

Frontiers
in
Artificial
Intelligence
and
Applications

MACHINE LEARNING AND ARTIFICIAL INTELLIGENCE

Proceedings of MLIS 2020

Edited by
Antonio J. Tallón-Ballesteros
Chi-Hua Chen

IOS
Press

MACHINE LEARNING AND ARTIFICIAL INTELLIGENCE

Machine learning and artificial intelligence are already widely applied to facilitate our daily lives, as well as scientific research, but with the world currently facing a global COVID-19 pandemic, their capacity to provide an important tool to support those searching for a way to combat the novel corona virus has never been more important.

This book presents the proceedings of the International Conference on Machine Learning and Intelligent Systems (MLIS 2020), which was due to be held in Seoul, Korea, from 25-28 October 2020, but which was delivered as an online conference on the same dates due to COVID-19 restrictions. MLIS 2020 was the latest in a series of annual conferences that aim to provide a platform for exchanging knowledge about the most recent scientific and technological advances in the field of machine learning and intelligent systems. The annual conference also strengthens links within the scientific community in related research areas.

The book contains 53 papers, selected from more than 160 submissions and presented at MLIS 2020. Selection was based on the results of review and scored on: originality, scientific/practical significance, compelling logical reasoning and language. Topics covered include: data mining, image processing, neural networks, human health, natural language processing, video processing, computational intelligence, expert systems, human-computer interaction, deep learning, and robotics.

Offering a current overview of research and developments in machine learning and artificial intelligence, the book will be of interest to all those working in the field.



ISBN 978-1-64368-136-8 (print)
ISBN 978-1-64368-137-5 (online)
ISSN 0922-6389 (print)
ISSN 1879-8314 (online)

MACHINE LEARNING AND ARTIFICIAL
INTELLIGENCE

Frontiers in Artificial Intelligence and Applications

The book series Frontiers in Artificial Intelligence and Applications (FAIA) covers all aspects of theoretical and applied Artificial Intelligence research in the form of monographs, selected doctoral dissertations, handbooks and proceedings volumes. The FAIA series contains several sub-series, including ‘Information Modelling and Knowledge Bases’ and ‘Knowledge-Based Intelligent Engineering Systems’. It also includes the biennial European Conference on Artificial Intelligence (ECAI) proceedings volumes, and other EurAI (European Association for Artificial Intelligence, formerly ECCAI) sponsored publications. The series has become a highly visible platform for the publication and dissemination of original research in this field. Volumes are selected for inclusion by an international editorial board of well-known scholars in the field of AI. All contributions to the volumes in the series have been peer reviewed.

The FAIA series is indexed in ACM Digital Library; DBLP; EI Compendex; Google Scholar; Scopus; Web of Science: Conference Proceedings Citation Index – Science (CPCI-S) and Book Citation Index – Science (BKCI-S); Zentralblatt MATH.

Series Editors:

Joost Breuker, Nicola Guarino, Pascal Hitzler, Joost N. Kok, Jiming Liu,
Ramon López de Mántaras, Riichiro Mizoguchi, Mark Musen, Sankar K. Pal,
Ning Zhong

Volume 332

Recently published in this series

- Vol. 331. A.J. Tallón-Ballesteros (Ed.), Fuzzy Systems and Data Mining VI – Proceedings of FSDM 2020
- Vol. 330. B. Brodaric and F. Neuhaus (Eds.), Formal Ontology in Information Systems – Proceedings of the 11th International Conference (FOIS 2020)
- Vol. 329. A.J. Tallón-Ballesteros (Eds.), Modern Management based on Big Data I – Proceedings of MMBD 2020
- Vol. 328. A. Utka, J. Vaičenonienė, J. Kovalevskaitė and D. Kalinauskaitė (Eds.), Human Language Technologies – The Baltic Perspective – Proceedings of the Ninth International Conference Baltic HLT 2020
- Vol. 327. H. Fujita, A. Selamat and S. Omatu (Eds.), Knowledge Innovation Through Intelligent Software Methodologies, Tools and Techniques – Proceedings of the 19th International Conference on New Trends in Intelligent Software Methodologies, Tools and Techniques (SoMeT_20)
- Vol. 326. H. Prakken, S. Bistarelli, F. Santini and C. Taticchi (Eds.), Computational Models of Argument – Proceedings of COMMA 2020

ISSN 0922-6389 (print)
ISSN 1879-8314 (online)

Machine Learning and Artificial Intelligence

Proceedings of MLIS 2020

Edited by

Antonio J. Tallón-Ballesteros

*Department of Electronic Engineering, Computer Systems and Automation,
University of Huelva, Spain*

and

Chi-Hua Chen

College of Mathematics and Computer Science, Fuzhou University, China

IOS
Press

Amsterdam • Berlin • Washington, DC

© 2020 The authors and IOS Press.

This book is published online with Open Access and distributed under the terms of the Creative Commons Attribution Non-Commercial License 4.0 (CC BY-NC 4.0).

ISBN 978-1-64368-136-8 (print)

ISBN 978-1-64368-137-5 (online)

Library of Congress Control Number: 2020949553

doi: 10.3233/FAIA332

Publisher

IOS Press BV

Nieuwe Hemweg 6B

1013 BG Amsterdam

Netherlands

fax: +31 20 687 0019

e-mail: order@iospress.nl

For book sales in the USA and Canada:

IOS Press, Inc.

6751 Tepper Drive

Clifton, VA 20124

USA

Tel.: +1 703 830 6300

Fax: +1 703 830 2300

sales@iospress.com

LEGAL NOTICE

The publisher is not responsible for the use which might be made of the following information.

PRINTED IN THE NETHERLANDS

MLIS 2020

Preface

The COVID-19 pandemic which began in late 2019 has posed a historical challenge to society. At the time of writing, scientists, clinicians, and healthcare experts around the globe continue to search for new technologies to support those tackling the pandemic. The evidence deriving from the application of Machine Learning (ML) and Artificial Intelligence (AI) to previous epidemics encourages researchers by providing a new angle with which to fight the novel corona virus outbreak. ML and AI are widely applied to facilitate our daily lives, as well as in scientific research.

The International Conference on Machine Learning and Intelligent Systems (MLIS) is an annual conference that aims to provide a platform for exchanging knowledge about the most recent scientific and technological advances in the field of ML and intelligent systems. It also strengthens links within the scientific community in related research areas. MLIS 2020 was initially scheduled to be held in Seoul, Korea from 25–28 October 2020, but was changed to an online conference on the same dates due to COVID-19 restrictions. Scientific topics covered at MLIS 2020 included data mining, image processing, neural networks, human health, natural language processing, video processing, computational intelligence, expert systems, human-computer interaction, deep learning, robotics, etc.

This book contains 53 papers selected from more than 160 submissions and presented at MLIS 2020. Selection was based on the results of review scored on: originality, scientific/practical significance, compelling logical reasoning and language. Rigorous peer review was conducted by members of the Editorial Committee led by Guest Editors Prof. Antonio J. Tallón-Ballesteros and Prof. Chi-Hua Chen, as well as active reviewers like Prof. Dheyaa Jasim Kadhim, Prof. Saeid Seyedi, Prof. Ivan Izonin, Prof. Behrouz Pourghebleh, Prof. Tahereh Eftekhari, Prof. Ilaria Cacciari, Prof. Ogundokun Roseline Oluwaseun, who engaged authors to answer stimulating questions for the camera-ready versions of their papers. We believe the 53 papers selected will provide some insights in the field of ML and AI.

On behalf of the conference organizers, we would like to take this opportunity to express our sincere thanks to the Guest Editor and all the Reviewers for their tremendous efforts and dedication to the conference. We also wish to thank all the authors for their relevant contributions to the conference, as well as all our colleagues at the publishers IOS Press for their support and tireless efforts in preparation for the publication of the conference proceedings. We believe that with such earnest support and contributions, future MLIS Conferences will also scale new heights.

Guest Editors:

Antonio J. Tallón-Ballesteros
University of Huelva (Spain)
Seville city, Spain

Chi-Hua Chen
Fuzhou University (China)
Fuzhou City, Fujian, China

This page intentionally left blank

Contents

Preface	v
<i>Antonio J. Tallón-Ballesteros and Chi-Hua Chen</i>	
Part I. Machine Learning and Artificial Intelligence	
Genetically Driven Optimal Selection of Opinion Spreaders in Complex Networks	3
<i>Alexandru Topîrceanu</i>	
Behavior Features for Automatic Detection of Depression from Facebook Users	12
<i>Siranuch Hemtanon, Saifon Aekwarangkoon and Nichnan Kittphattanabawon</i>	
Survey on Formal Concept Analysis Based Supervised Classification Techniques	21
<i>Hayfa Azibi, Nida Meddouri and Mondher Maddouri</i>	
Cross-Based Adaptive Guided Filtering	30
<i>Ming Yan, Yueli Hu, Kai Li and Jianeng Zhao</i>	
Anisotropic Diffusion with Deep Learning	40
<i>Hyun-Tae Choi, Yuna Han, Dahye Kim, Seonghoon Ham, Minji Kim, Yesol Park and Byung-Woo Hong</i>	
A Clinical Decision Support Tool to Detect Invasive Ductal Carcinoma in Histopathological Images Using Support Vector Machines, Naïve-Bayes, and K-Nearest Neighbor Classifiers	46
<i>Kyra Mikaela M. Lopez and Ma. Sheila A. Magboo</i>	
Energy Disaggregation Using Principal Component Analysis Representation	54
<i>Pierre V. Dantas, Waldir Sabino S. Júnior and Celso B. Carvalho</i>	
Research on Optimization Method of Tool Path in Five-Axis Process Singular Region	61
<i>Yujie Wang, Xin Shen, Yu Peng and Lixin Zhao</i>	
Big Data System for Medical Images Analysis	71
<i>Janusz Bobulski and Mariusz Kubanek</i>	
Research and Application of High Precision Machining Technology for Super-Large Integral Frame Parts of Aluminum Alloy	80
<i>Hong Ji, Yujie Wang, Yu Peng, Lixin Zhao and Song Huang</i>	
AutoSynPose: Automatic Generation of Synthetic Datasets for 6D Object Pose Estimation	89
<i>Heiko Engemann, Shengzhi Du, Stephan Kallweit, Chuanfang Ning and Saqib Anwar</i>	
Solving a Fuzzy Linear Equation with a Variable, Using the Expected Interval of a Fuzzy Number	98
<i>Richard Abramonte, Eder Escobar, Antenor Aliaga and Flabio Gutierrez</i>	

Analysis of the Software Most Used by Hackers to Carry Out Penetration Testing in Public Organizations <i>Segundo Moisés Toapanta Toapanta, Raúl Francisco Pérez González, Máximo Giovanni Tandazo Espinoza and Luis Enrique Mafla Gallegos</i>	107
Deep Robot-Human Interaction with Facial Emotion Recognition Using Gated Recurrent Units & Robotic Process Automation <i>Suchitra Saxena, Shikha Tripathi and Sudarshan Tsb</i>	115
Research on the Construction of Cutting Edge Technology Monitoring System Based on Multi-Source Heterogeneous Data <i>Qiang Xiao, Siming Tan and Shengfeng Yu</i>	127
An Octave Package to Perform Qualitative Analysis of Nonlinear Systems Immersed in R4 <i>Eder Escobar, Richard Abramonte, Antenor Aliaga and Flabio Gutierrez</i>	136
Detecting Similar Versions of Software by Learning with Logistic Regression on Binary Opcode Information <i>Hyun-il Lim</i>	146
Explainable AI: Using Shapley Value to Explain Complex Anomaly Detection ML-Based Systems <i>Jinying Zou and Ovanes Petrosian</i>	152
A Simple Ensemble Learning Knowledge Distillation <i>Himel Das Gupta, Kun Zhang and Victor S. Sheng</i>	165
Using Conversation Analysis for Examining Social Media Interactions <i>Matteo Farina</i>	172
A Study on the Effect of DropConnect to Control Overfitting in Designing Neural Networks <i>Hyun-il Lim</i>	178
The Evolution from Digital Mock-Up to Digital Twin <i>Sheng Dai, Gang Zhao, Yong Yu and Qiangwei Bao</i>	184
Adapted NRC Based Sentiment Analysis in Event Distraction <i>Zeling Wang, Bing Wei and Yibing Li</i>	192
Intelligence Accounting Information Fusion System: Theory, Model and Framework <i>Xinsheng Duan</i>	200
Modeling Multi-Order Adaptive Processes by Self-Modeling Networks <i>Jan Treur</i>	206
 Part II. Communication and Networks	
Application of Twin Objective Function SVM in Sentiment Analysis <i>Qiaoman Yang and Chunyu Liu</i>	221

Robust H-infinity State Estimation of Uncertain Neural Networks with Two Additive Time-Varying Delays	229
<i>Xiaoping Hu, Yajun Wang, Jiakai Ding and Dongming Xiao</i>	
Dynamic Space-Covered Broadcast Algorithm Based on Neighbor-Degree in Mobile Ad Hoc Wireless Networks	241
<i>Qiaoqiao Lou and Zhijin Zhao</i>	
Secure Server Key Management Designs for the Public Cloud	248
<i>Kevin Foltz and William R. Simpson</i>	
A Routing Protocol Based on Both of Density Variation and Distance-Aware for WSNs	254
<i>Dongmei Xing</i>	
Real Time Digital Filter for a Front-End Electronics in Dark Matter and Neutrino Measurements	266
<i>Alejandro D. Martinez R. on behalf of DarkSide Collaboration</i>	
Research on Fixture Design of NC Machine Tool Based on UG 3D Modeling	272
<i>Yanxia Li</i>	
Design and Implementation of Substation Operation Safety Monitoring and Management System Based on Three-Dimensional Reconstruction	278
<i>Yingyi Yang, Hao Wu, Fan Yang, Xiaoming Mai and Hui Chen</i>	
Modeling of a Leg and Knee System for the Analysis of Human Gait by Means of State Feedback Control	288
<i>P.A. Ospina-Henao, César H. Valencia, Marcelo Becker, Zuly A. Mora P and S.M. Vásquez</i>	
Two-Dimensional Spatial Resolution in Plane Temperature Monitoring Based on Raman Distributed Temperature Sensor	302
<i>Junfan Chen, Ning Sun and Zhongxie Jin</i>	
Analysis for the Adoption of Security Standards to Improve the Management of Securities in Public Organizations	310
<i>Segundo Moisés Toapanta Toapanta, Madeleine Lilibeth Alvarado Ronquillo, Luis Enrique Mafla Gallegos and Alberto Ochoa Zezzatti</i>	
Simple Receiving Scheme Based on Balanced Detection for Half-Cycled SSB DD-OFDM Signal	322
<i>Pengfei Yang and Xue Chen</i>	
Design, Simulation and Verification of a 7-DOF Joint Motion Simulation Platform	331
<i>Ze Cui, Saishuai Huang, Zenghao Chen, Hongxin Yang and Danjie Zhu</i>	
3D Air-Ground Channel Model Reconstruction Based on the Experimental Data and Q-D Method	337
<i>Alexander Maltsev, Olesya Bolkhovskaya, Vadim Sergeev, Anton Elokhin and Alexander Kokarev</i>	

Cost-Efficient RAA Technology for Development of the High-Gain Steerable Antennas for mmWave Communications <i>Olesya Bolkhovskaya, Alexander Maltsev, Valentin Seleznev and Ilya Bolotin</i>	346
Detection of Specific Components in a PCA Mixture <i>Jin Xie, Zian Zheng and Jian Gao</i>	354
Investigation of the Effect of Different Covariance Estimation Methods on the Performance of Least Squares Compressive Covariance Sensing <i>Nuha A.S. Alwan</i>	363
Ad-Hoc Network Access Control System and Method for Edge NFC Terminal <i>Yiqin Bao and Zhengtang Sun</i>	374
Joint User Clustering and Power Allocation for Vehicular Communications <i>Qian Chen, Gongbin Qian, Chunlong He, Rujun Zhao and Yuping Zheng</i>	383
Study of Work-Function Variation on Performance of Dual-Metal Gate Fin Field-Effect Transistor <i>Tianyu Yu, Liang Dai, Zhifeng Zhao, Weifeng Lyu and Mi Lin</i>	392
A Novel Approach to Radiometric Identification <i>Raoul Nigmatullin, Semyon Dorokhin and Alexander Ivchenko</i>	400
Order-Preserving and Efficient One-to-Many Search on Encrypted Data <i>Dongping Hu, Aihua Yin, Huaying Yan and Tao Long</i>	406
The Application of Improved Sparrow Search Algorithm in Sensor Networks Coverage Optimization of Bridge Monitoring <i>Yueqi Peng, Yunqing Liu and Qi Li</i>	416
Navigation Signal Design and Ranging Performance Evaluation of Cn Band Based on Satellite-to-Ground Link <i>Xiaofei Chen, Xue Wang, Xiaochun Lu, Jing Ke and Xia Guo</i>	424
Prototype Design of a Ship Intelligent Integrated Platform <i>Fuxin Wang, Hao Luo, Yuan Yu and Liyong Ma</i>	435
The Bounded Traveling Wave Solutions of a (3+1) Dimensions mKdv-ZK Equation <i>Yuzhong Zhang</i>	442
A Survey on Ship Intelligent Cabin <i>Shuli Jia, Fuxin Wang, Mengchao Dong and Liyong Ma</i>	453
An Identity-Based Directed Proxy Multi-Signature Scheme <i>Weiping Zuo</i>	459
Subject Index	465
Author Index	469

Part I
Machine Learning and Artificial Intelligence

This page intentionally left blank

Genetically Driven Optimal Selection of Opinion Spreaders in Complex Networks

Alexandru TOPÎRCEANU^{a,1}

^a*Department of Computer and Information Technology, Politehnica University Timișoara, Romania*

Abstract. The problem of influence maximization (IM) represents a major challenge for modern network science, with direct applicability in political science, economy, epidemiology, and rumor spreading. Here, we develop a novel computational intelligence framework (GenOSOS) based on genetic algorithms with emphasis on the optimal layout of spreader nodes in a network. Our algorithm starts with solutions consisting of randomly selected spreader nodes; then, by defining custom original *crossover* and *mutation* operators, we are able to obtain, in a short number of genetic iterations, nearly optimal solutions in terms of the nodes' topological layout. Experiments on both synthetic and real-world networks show that the proposed GenOSOS algorithm is not only a viable alternative to the existing node centrality approach, but that it outperforms state of the art solutions in terms of spreading coverage. Specifically, we benchmark GenOSOS against graph centralities such as node degree, betweenness, PageRank and k-shell using the SIR epidemic model, and find that our solution is, on average, 11.45% more efficient in terms of diffusion coverage.

Keywords. genetic algorithm, influence maximization, computational intelligence, complex networks

1. Introduction

Finding individuals with high social influence is one of the fundamental challenges for network science [1,2,3], and represents a critical issue for better understanding of the market [4], and for predicting political preference [5] as well. In its simplest formulation, IM sets out to select the initial spreader nodes which may influence a maximal number of users in a given network [1]. An important demand faced by IM algorithms is obtaining a balanced trade-off between the accuracy of the solution and the time/memory cost, especially over large-scale networks. Consequently, developing efficient algorithms for IM still represents a challenging research topic.

In terms of the more recent IM state of the art research, we first note the works of Zareie *et al.* [6,7]. Similar to our approach, in [6] the authors suggest that distances between spreaders should be taken into consideration to ensure minimum overlap and maximum coverage of a wider area of the network. Similar in scope, the goal in [7] is to maximize the distance between spreader nodes with the use of gray wolf optimization.

¹Corresponding Author: Alexandru Topîrceanu, Department of Computer and Information Technology, Politehnica University Timișoara, Timișoara, 300223, Romania; E-mail: alex@cs.upt.ro

Other evolutionary methods used to tackle the IM problem are found in [8,9]. Gong *et al.* [8] propose a local influence criterion for a reliable estimation of the influence propagation in cascade models and use particle swarm optimization (PSO) to optimize local influence criteria. Tang *et al.* [9] use a discrete shuffled frog-leaping algorithm (DSFLA) that combines both deterministic and random walk strategies.

Indeed, compared to our computational intelligence framework, we find also similar genetic methodologies proposed to solve the IM problem in complex networks. Bucur *et al.* [10] define a genetic algorithm approach in which the independent cascade (IC) model is used as a fitness function for nodes. Gong *et al.* [11] make use of a genetic algorithm for community-based influence maximization in social networks. Their idea is to optimize the 2-hop influence spread of nodes to find the most influential nodes. Finally, Cui *et al.* [12] propose degree-descending search evolution (DDSE); this strategy generates a node set whose influence spread is comparable to the degree centrality. The results claimed by the authors are obtained 5x faster than for greedy algorithms.

In this article, we develop a novel computational intelligence framework, called *Genetically driven Optimal Selection of Opinion Spreaders* (GenOSOS), to engage in the IM problem, and provide both qualitative and quantitative means of evaluating the performance of GenOSOS. We first apply state of the art methodology in selecting spreaders based on node centralities (degree, betweenness, PageRank and k-shell), then run the SIR epidemic model [13], and measure the diffusion coverage. The SIR diffusion simulations show that the potential of our solution exceeds expectations by offering superior quantifiable results compared to the state of the art. Compared to the analyzed related work [6,7,8,9,2,10,11,12], this paper brings several important contributions:

- We propose GenOSOS, a genetic algorithm approach for the IM problem, which represents an original attempt for dealing with the trade-off between spreader spacing and diffusion coverage.
- We propose a problem-specific modeling of the population and chromosome representation. Furthermore, we implement the fitness function based on a graph coloring algorithm, which can accelerate the convergence of the spreading process.
- We define an individual (chromosome) as a unique spreader set, bringing along custom implementations of *crossover* and *mutation*.
- We estimate the effectiveness of GenOSOS on synthetic and real-world networks. The experimental results show that our algorithm has competitive performances to similar centrality-based node selection methods.

2. Methods

2.1. Problem Definition

We consider a complex network modeled as an undirected graph $G = \{N, E\}$, where $N = \{n_i\}$ is the node set and $E = \{e_{ij} | n_i, n_j \in N\}$ is the edge set in the network. Nodes represent individual entities and edges represent social relationships between any two nodes. A node can be marked as a spreader if it has already adopted an opinion, or inactive otherwise.

Thus, the problem of IM is defined as follows: given network G and a number p , determine subset $N^* \subset N$ consisting of p spreaders (i.e., nodes) such that these nodes can

spread their influence to other nodes $N \setminus N^*$ in G by maximizing the influence coverage and minimizing the time taken by these spreaders.

In terms of centrality based spreader assignment, we select several of the most popular and robust node centralities to serve as comparison for our proposed selection method, namely, node degree, betweenness, PageRank, and k-shell centrality [14,15]. As a baseline serving for comparison, we also use random spreader assignment.

The degree Deg of a node n_i is defined as the sum of all incident edge weights to that node's vicinity N_i as $Deg(n_i) = \sum_{n_j \in N_i} w_{ij}$, with $w_{ij} = 1$ in an unweighted context. Betweenness centrality Btw is defined as the fraction of shortest paths between all node pairs that pass through a specific node n_i [16]. The PageRank algorithm, which is used at the core of Google's search engine [17], interprets an edge e_{ij} as a vote by node n_i to node n_j . Finally, with k-shell centrality, for every node a k-shell index gets assigned based on its topological location; to this end, nodes that are closer to the network core have higher k-shells. Nodes with greater k-shells are considered as more influential nodes. [15].

2.2. Network Datasets

In order to run and test GenOSOS we first create models for the four fundamental complex network topologies [14]: a random Erdős-Rényi network ER , a regular mesh network Me , a Watts-Strogatz small-world network SW , and a Barabási-Albert scale-free network SF . Then, we generate four complex synthetic models: Holme-Kim HK [18], covert cellular networks $Cell$ [19], Watts-Strogatz networks with degree distribution WD [20], and genetically optimized Genosian social networks Gen [21]. Finally, we include in our study the following four real-world datasets: a co-authorship network CoA [22], an online social network OSN [23], a scientific collaboration network Geo [24], and an email communication network Em [25].

We measure a standard set of network properties, for each dataset, which are given in Table 1. Here we include the network size (number of nodes N), number of edges E , average degree $avgD$, maximum degree $maxD$, average path length APL , average clustering coefficient ACC , network modularity Mod , and diameter Dmt [14].

Table 1. Network measures for the validation datasets divided into fundamental synthetic topologies, complex synthetic topologies, and real-world networks.

<i>Dataset</i>	<i>N</i>	<i>E</i>	<i>avgD</i>	<i>maxD</i>	<i>APL</i>	<i>ACC</i>	<i>Mod</i>	<i>Dmt</i>
<i>ER</i>	5000	25061	5.012	26	3.994	0.002	0.362	7
<i>Me</i>	5000	26948	5.390	44	11.515	0.148	0.821	30
<i>SW</i>	5000	19999	4.000	13	6.738	0.298	0.739	12
<i>SF</i>	5000	15762	3.152	294	5.378	0.007	0.64	13
<i>HK</i>	1000	3330	3.330	85	3.553	0.506	0.488	7
<i>Cell</i>	1041	6012	5.775	95	4.428	0.258	0.885	10
<i>WD</i>	1178	9048	7.681	58	15.419	0.659	0.93	32
<i>Gen</i>	1063	6915	6.505	25	4.765	0.498	0.882	9
<i>CoA</i>	1589	2742	3.451	34	5.823	0.878	0.955	17
<i>OSN</i>	1899	20296	10.688	339	3.055	0.138	0.338	8
<i>Geo</i>	3621	9461	5.226	102	5.316	0.679	0.743	14
<i>Em</i>	12625	20362	3.226	576	3.811	0.577	0.684	9

3. The GenOSOS Framework

3.1. Chromosome Representation

Each genetic generation consists of a solution population $S^j = \{s_1^j, s_2^j, \dots, s_i^j, \dots, s_n^j\}$ of n individual solutions s_i^j (chromosomes). The algorithm loop further consists of $1 \leq j \leq k$ iterative generations. As such, we represent a chromosome i from generation j as a solution network $s_i^j = \{N, N^*, E\}$. Here, N is the same set of nodes from the original network N , N^* is a subset of marked spreader nodes of size $|N^*| = p$, and E is the same set of edges which remain unchanged throughout the algorithm. Thus, each chromosome s_i^j differs through its custom selection of spreaders N^* .

Each initial chromosome is initialized by randomly marking $p \ll N$ nodes as spreaders. For a deterministic approach, we would have C_N^p possible combinations, which can be approximated by N^p solutions. While it is mathematically possible to obtain two initial equivalent chromosomes (with equivalent N^*), given the usual size of real-world networks $N > 1000$, and the limited number of selected spreaders $n < 100$, this probability is extremely small and is not a concern for our study.

3.2. Fitness Calculation

For every chromosome we need to be able to quantify its spreading efficiency. This efficiency is calculated by adopting a classic graph coloring algorithm starting from the marked nodes. Figure 1 exemplifies the fitness calculation on a network of $N = 20$ nodes. In step 1, we consider the $N^*(t = 1)$ spreaders from the chromosome representation as sources for coloring. Next, we repeat the graph coloring and mark all neighbors of the spreaders, obtaining a larger spreader set $N^*(t + 1)$. We keep track of the growing set of marked nodes until $N^*(t) \geq 95\%N$ of the network is covered. Once this stop condition is met, the fitness of chromosome s_i^j is expressed as the number of nodes successfully colored divided by the number of steps required $f(s_i^j) = N^*(t)/t$.

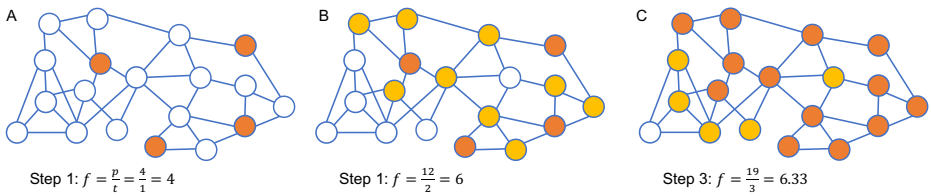


Figure 1. Example of computing fitness f on a small network with $N = 20$ nodes. (A) A number of $p = 4$ spreaders (dark orange) originate from the chromosome representation, leading to a fitness $f = 4$. (B) All adjacent nodes to the original spreaders are also marked as spreaders (yellow orange); at this stage $f = 6$ and only 60% of the network is colored. (C) The process continues until at least 95% of the network is colored; at this stage $f = 6.33$ and the graph coloring algorithm stops.

3.3. Defining Custom Genetic Operators

Elitism implies that a proportion r_e of the highest fitness (best) solutions from the n chromosomes are copied over to the next generation. This approach ensures that the

fitness scores of the top $r_e n$ of the next generation will be at least as good as the current generation.

Crossover takes a pair of two randomly selected chromosomes from the pool of elite solutions, merges them together, and returns two new chromosomes. A crossover index in each of the chromosomes is randomly selected, and all the spreader nodes of the chromosomes after that selection index are exchanged between the two chromosomes. We symbolize r_c as the crossover rate.

Mutation on a chromosome is implemented by randomly selecting a spreader node n_i from the marked spreaders N^* , and swapping it with a random unmarked node from the remaining graph $N \setminus N^*$. The mutation operator is repeated given the mutation rate r_m .

3.4. Algorithm Implementation

The genetic algorithm of GenOSOS, shown in Figure 2, relies on three genetic operators – elitism, crossover, and mutation – and runs according to the following steps:

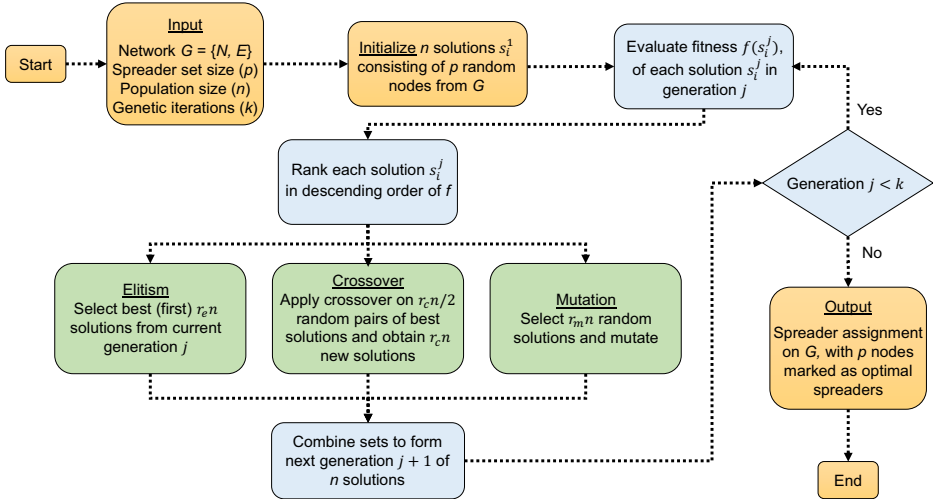


Figure 2. Flowchart of GenOSOS emphasizing the main algorithmic steps: input/output (orange), generation control (blue), and genetic operators (green). According to the flowchart, the algorithm finds an optimal solution s_i^j for placing p spreader nodes in a graph G , and runs k genetic iterations consisting of three operators that are used to generate n new solutions, from generation j , for the next generation $j + 1$. The output consists of a set of p nodes marked as spreaders in graph G .

1. Initialize n solutions (chromosomes), each with p randomly marked spreaders.
2. Compute fitness $f(s_i^j)$ of each chromosome s_i^j in generation j .
3. Sort chromosomes in descending order of fitness f .
4. Copy the first (best) $r_e n$ of the chromosomes to the next generation $j + 1$.
5. Pick $r_c / 2 \cdot n$ randomly selected pairs of chromosomes from the best chromosomes and apply crossover, resulting in $r_c n$ new chromosomes.
6. Pick $r_m n$ randomly selected chromosomes and apply mutation on them.
7. Combine sets $r_e n$, $r_c n$, and $r_m n$ to form the next generation $j + 1$ of size n .

8. Repeat steps (2-7) for $1 \leq j \leq k$ generations.

When solving NP-hard problems with heuristic methods (e.g., genetic algorithms), multiple combinations of model parameters can be feasible. As a trade-off between algorithmic speed and result precision, we simulate with a population size of $n = 1000$ chromosomes, a number of $k = 10$ generations, an elitism rate of $r_e = 0.5$, a crossover rate of $r_c = 0.3$, and a mutation rate of $r_m = 0.2$.

4. Results

4.1. Diffusion Coverage

We start by analyzing the diffusion coverage obtained by varying $p = 1 - 100$ spreaders on the random *ER*, mesh *Me*, small-world *SW*, and scale-free *SF* topologies, based on the SIR epidemic model. Spreaders are selected according to each of the six discussed selection methods: random *Rand*, degree *Deg*, betweenness *Btw*, PageRank *PR*, k-shell *KS*, and the proposed GenOSOS (*GOS*) method. Figure 3 displays the results for increasing p , and given in Table 2, are the values of diffusion coverage for $p = 10$ and $p = 50$. Each represented measurement is obtained after 10 repeated simulations.

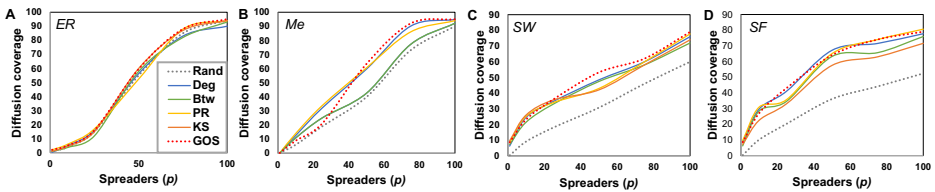


Figure 3. Diffusion coverage with increasing spreader set size $p = 1 - 100$ on the four fundamental complex topologies: (A) random *ER*, (B) mesh *Me*, (C) small-world *SW*, and (D) scale-free *SF*. The coverage obtained by GenOSOS is highlighted with the red dotted line (*GOS*).

Table 2. Diffusion coverage on the fundamental synthetic topologies, expressed in percentage (%), for the scenarios with $p = 10$ spreaders and $p = 50$ spreaders.

Topology	p	<i>Rand</i>	<i>Deg</i>	<i>Btw</i>	<i>PR</i>	<i>KS</i>	<i>GOS</i>
<i>ER</i>	10	4.03	5.22	3.78	6.12	4.56	5.16
<i>Me</i>	10	5.98	12.64	10.90	14.02	10.87	8.83
<i>SW</i>	10	9.05	24.50	21.42	25.21	25.94	23.15
<i>SF</i>	10	10.10	29.64	28.43	30.85	22.47	27.03
<i>ER</i>	50	55.03	56.32	57.88	52.46	60.02	59.54
<i>Me</i>	50	39.80	59.72	42.55	60.50	42.59	63.24
<i>SW</i>	50	30.72	48.45	47.43	42.98	42.04	53.22
<i>SF</i>	50	36.10	67.23	63.27	65.03	57.54	64.11

The simulation results over the synthetic topologies show that spreaders placed according to GenOSOS are capable of achieving similar, and superior diffusion performance compared to the state of the art centrality approach. Specifically, with GenOSOS we obtain the highest spreading coverage on the mesh and small-world (for $p = 50$). Our genetic algorithm approach outperforms state of the art graph centralities in 2 out of 4 cases on the fundamental topologies. On average, *Rand* is 32.67% lower, *Deg* is 3.49% lower,

Btw is 12.06% lower, *PR* is 7.97% lower, and *KS* is 15.79% lower in terms of diffusion coverage.

Next, we analyze the diffusion coverage on the Holme-Kim *HK*, cellular *Cell*, Watts-Strogatz with degree distribution *WD*, and Genosian *Gen* synthetic topologies. The same amount of spreaders is increased from $p = 1$ to $p = 100$ in the network, according to each of the six selection centralities. Table 3 presents the best results after 10 independent repetitions for each simulation scenario, obtained when $p = 10$, and $p = 50$. Figure 4 displays the increasing diffusion coverage for all values of p .

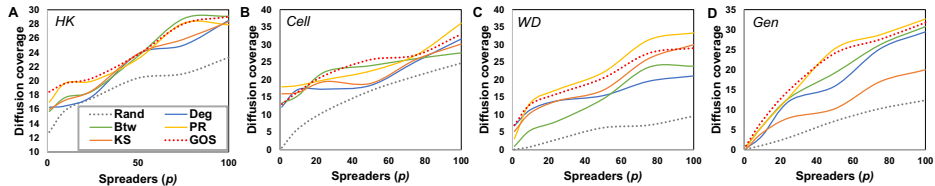


Figure 4. Diffusion coverage with increasing spreader set size $p = 1 - 100$ on the four complex synthetic networks: (A) Holme-Kim *HK*, (B) cellular *Cell*, (C) Watts-Strogatz with degree distribution *WD*, and (D) Genosian *Gen*. The coverage obtained by GenOSOS is highlighted with the red dotted line (GOS).

Table 3. Diffusion coverage on the complex synthetic topologies, expressed in percentage (%), for the scenarios with $p = 10$ spreaders and $p = 50$ spreaders.

Topology	p	<i>Rand</i>	<i>Deg</i>	<i>Btw</i>	<i>PR</i>	<i>KS</i>	<i>GOS</i>
<i>HK</i>	10	15.93	16.51	17.75	19.78	17.33	19.65
<i>Cell</i>	10	6.03	17.19	15.46	18.34	16.33	16.56
<i>WD</i>	10	0.84	11.46	5.62	13.49	10.55	13.04
<i>Gen</i>	10	1.12	3.57	5.92	5.92	4.25	7.24
<i>HK</i>	50	20.40	23.86	23.54	23.19	23.85	23.71
<i>Cell</i>	50	16.81	18.35	24.09	22.48	18.91	25.73
<i>WD</i>	50	6.28	15.44	14.51	22.07	16.97	20.43
<i>Gen</i>	50	7.15	15.89	19.19	25.21	10.24	24.10

The simulation results on the complex synthetic topologies show that spreaders placed according to GenOSOS achieve, again, a diffusion performance comparable to the centrality approach. Namely, GenOSOS scores the highest spreading coverage on the *HK* and *Gen* networks for $p = 10$, respectively *HK* and *Cell* for $p = 50$, being roughly on par with the other centralities on the other networks.

Based on the presented measurements, our genetic approach outperforms the state of the art on 2 out of 4 networks. In terms of diffusion coverage, the spreaders selected according to GenOSOS achieve higher coverage rates, 46.08% more than *Rand*, 21.77% more than *Deg*, 13.55% more than *Btw*, 1.13% more than *PR*, and 25.52% more than *KS*.

Finally, we measure the diffusion coverage on the real-world co-authorship network *CoA*, online social network *OSN*, Geometry scientific collaboration *Geo*, and Email *Em* networks. The same amount of $p = 1 - 100$ spreaders are selected according to each of the six selection methods. In Figure 5 we display the measured diffusion coverage for all values of p . Each entry in Table 4 represents the best measurement obtained after 10 independent simulation repetitions, for $p = 10$ and $p = 50$ spreaders.

Overall, we notice that spreaders placed according to GenOSOS achieve high diffusion performance compared to the state of the art. Specifically, GenOSOS scores the

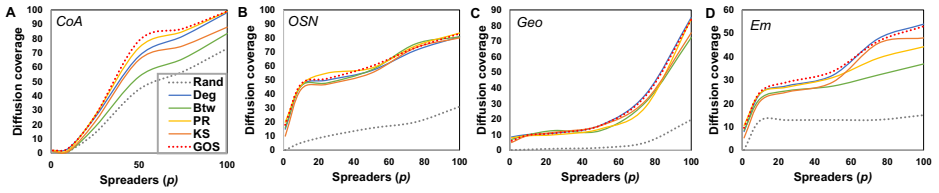


Figure 5. Diffusion coverage with increasing spreader set size $p = 1 - 100$ on the four real-world networks: (A) co-authorship network *CoA*, (B) online social network *OSN*, (C) Geometry scientific collaboration *Geo*, and (D) Emails *Em*. The coverage obtained by GenOSOS is highlighted with the red dotted line (GOS).

Table 4. Diffusion coverage on the real-world networks, expressed in percentage (%), for the scenarios with $p = 10$ spreaders and $p = 50$ spreaders.

Topology	p	<i>Rand</i>	<i>Deg</i>	<i>Btw</i>	<i>PR</i>	<i>KS</i>	<i>GOS</i>
<i>CoA</i>	10	1.06	3.52	1.25	2.39	1.46	3.69
<i>OSN</i>	10	5.22	45.86	45.81	46.18	43.33	47.10
<i>Geo</i>	10	0.46	10.08	9.91	9.36	9.38	9.75
<i>Em</i>	10	12.76	24.25	22.12	24.47	21.04	24.93
<i>CoA</i>	50	44.60	67.91	53.50	73.77	65.32	78.67
<i>OSN</i>	50	15.73	56.73	57.26	58.51	55.02	59.85
<i>Geo</i>	50	1.61	15.96	12.62	13.71	15.78	15.90
<i>Em</i>	50	12.95	32.06	27.43	31.11	29.15	33.72

highest spreading coverage on the co-authorship *CoA*, *OSN*, and email *Em* networks for $p = 10$. When $p = 50$, the results remain consistent, with GenOSOS scoring the highest coverage on the same networks. Based on the analyzed simulation results, we conclude that our genetic approach outperforms the state of the art on 3 out of 4 cases networks in terms of diffusion coverage. Specifically, the spreaders selected according to GenOSOS achieve higher coverage rates, namely 60.19% more than *Rand*, 8.22% more than *Deg*, 19.84% more than *Btw*, 5.86% more than *PR*, and 15.15% more than *Ks*.

5. Conclusion

In this paper we present a novel computational intelligence approach of selecting spreaders in complex networks based on genetic algorithms. We introduce the GenOSOS framework and compare it against state of the art methodology in selecting spreaders based on node centralities. SIR simulation results are quantified through diffusion coverage achieved on both synthetic and real-world datasets. The detailed analysis on three categories of network datasets show that the potential of our proposed solution is not only viable, but offers superior results compared to the state of the art centrality approach. Specifically, GenOSOS obtains a 11.45% higher coverage, averaged over all 12 datasets. In essence, our solution is superior to the state of the art on 7 out of 12 datasets (58.3%) in terms of diffusion coverage.

Overall, we have achieved to goal of this study, namely to: (i) investigate the alternative of optimal spreader selection using genetic algorithms, and (ii) also show that the genetic alternative can be, often, equal or superior in diffusion performance in comparison to the state of the art. Consequently, we have developed an important alternative spreader selection method without the need to estimate nodes centrality.

Acknowledgments

Author A.T. was supported by the Romanian National Authority for Scientific Research and Innovation (UEFISCDI), project number PN-III-P1-1.1-PD-2019-0379.

References

- [1] Chen W, Wang Y, Yang S. Efficient influence maximization in social networks. In: Proc. of 15th ACM SIGKDD Int. Conf. on Knowledge discovery and data mining. ACM; 2009. p. 199–208.
- [2] Kempe D, Kleinberg J, Tardos É. Maximizing the spread of influence through a social network. In: Proc. ACM SIGKDD Int. Conf. on Knowledge discovery and data mining. ACM; 2003. p. 137–146.
- [3] Topirceanu A. Competition-Based Benchmarking of Influence Ranking Methods in Social Networks. *Complexity*. 2018;2018.
- [4] Easley D, Kleinberg J. *Networks, crowds, and markets: Reasoning about a highly connected world*. Cambridge University Press; 2010.
- [5] Golbeck J. *Analyzing the Social Web*. Access Online via Elsevier; 2013.
- [6] Zareie A, Sheikahmadi A, Khamforoosh K. Influence maximization in social networks based on TOP-SIS. *Expert Systems with Applications*. 2018;108:96–107.
- [7] Zareie A, Sheikahmadi A, Jalili M. Identification of influential users in social network using gray wolf optimization algorithm. *Expert Systems with Applications*. 2020;142:112971.
- [8] Gong M, Yan J, Shen B, Ma L, Cai Q. Influence maximization in social networks based on discrete particle swarm optimization. *Information Sciences*. 2016;367:600–614.
- [9] Tang J, Zhang R, Wang Pea. A discrete shuffled frog-leaping algorithm to identify influential nodes for influence maximization in social networks. *Knowledge-Based Systems*. 2020;187:104833.
- [10] Bucur D, Iacca G. Influence maximization in social networks with genetic algorithms. In: *European conference on the applications of evolutionary computation*. Springer; 2016. p. 379–392.
- [11] Gong M, Song C, Duan C, Ma L, Shen B. An efficient memetic algorithm for influence maximization in social networks. *IEEE Computational Intelligence*. 2016;11(3):22–33.
- [12] Cui L, Hu H, Yu Sea. DDSE: A novel evolutionary algorithm based on degree-descending search strategy for influence maximization in social networks. *J of Net and Comp App*. 2018;103:119–130.
- [13] Pastor-Satorras R, Castellano C, Van Mieghem P, Vespignani A. Epidemic processes in complex networks. *Reviews of modern physics*. 2015;87(3):925.
- [14] Wang XF, Chen G. *Complex networks: small-world, scale-free and beyond*. IEEE circuits and systems. 2003;3:6–20.
- [15] Kitsak M, Gallos LK, Havlin S, Liljeros F, Muchnik L, Stanley HE, et al. Identification of influential spreaders in complex networks. *Nature physics*. 2010;6(11):888.
- [16] Topirceanu A, Udrescu M, Marculescu R. Weighted betweenness preferential attachment: A new mechanism explaining social network formation and evolution. *Scientific Reports*. 2018;8(1):1–14.
- [17] Brin S, Page L. Reprint of: The anatomy of a large-scale hypertextual web search engine. *Computer networks*. 2012;56(18):3825–3833.
- [18] Holme P, Kim BJ. Growing scale-free networks with tunable clustering. *Phys Rev E*. 2002;65(2).
- [19] Tsvetovat M, Carley KM. *Generation of realistic social network datasets for testing of analysis and simulation tools*. DTIC Document; 2005.
- [20] Chen Y, Zhang L, Huang J. The Watts–Strogatz network model developed by including degree distribution: theory and computer simulation. *J of Physics A: Math and Theo*. 2007;40(29):8237.
- [21] Topirceanu A, Udrescu M, Vladutiu M. Genetically Optimized Realistic Social Network Topology Inspired by Facebook. In: *Onl Soc Media Anal and Vis*. Springer; 2014. p. 163–179.
- [22] Newman MEJ. Finding community structure in networks using the eigenvectors of matrices. *Phys rev E*. 2006;74(3):036104.
- [23] Opsahl T, Panzarasa P. Clustering in weighted networks. *Social networks*. 2009;31(2):155–163.
- [24] Batagelj V, Mrvar A. Pajek-program for large network analysis. *Connections*. 1998;21(2):47–57.
- [25] Leskovec J, Lang KJ, Dasgupta A, Mahoney MW. Community structure in large networks: Natural cluster sizes and the absence of large well-defined clusters. *Internet Math*. 2009;6(1):29–123.

Behavior Features for Automatic Detection of Depression from Facebook Users

Siranuch HEMTANON ^{a,1}, Saifon AEKWARANGKOON ^b
and Nichnan KITTPHATTANABAWON ^a

^a*School of Informatics, Walailak University, Nakorn Si Thammarat, 80160 Thailand.*

^b*School of Nursing, Walailak University, Nakorn Si Thammarat, 80160 Thailand.*

Abstract. Major depressive disorder is one of common mental disorders globally. It is best to be early detected and cured. This work introduces a method to detect depressive disorder at risk via a behavior made on Facebook platform. A set of features related to Facebook main functions including amount of posting, sharing, commenting and replying is designed to represent users' activities in a numerical value form. The collected data with periodic and consecutive aspects are gathered without interpreting content. Thus, the data are easier to be collected with less privacy issue. To distinct between positive and negative depression-at risk, PHQ-9 questionnaire, a standard tool commonly used to screen depression patient in Thailand, was used. These features hence are used in supervised learning classification algorithm for detecting a risk of being depressive disorder. From the experiment of 160 Thai Facebook users, the statistical result indicated that depression-positive users tend to do consecutive actions and rarely reply to other comments. Moreover, they often have activities in late night. The classification experiment shows that the designed features based on users' activities from Facebook with deep learning algorithm yields about 87% in terms of F-measure. After analyzing the data, we thus split data regarding users' gender and removed obviously low active data, and the F-measure from classification went up to 91.4 which improves for 4 points.

Keywords. Depression, Detection, Behavior features, Social network activity, Facebook

1. Introduction

Depression (also called clinical depression or major depressive disorder) is a mood disorder causing a persistent negative thought, physical disturbance, behaviors change, and feeling of sadness and loss of interest [1]. It is one of the globally most common mental disorders [2, 3]. Depression can happen at any age. It causes severe symptoms affecting how ones feel, think, and handle daily activities including sleeping, eating, or working. To be diagnosed with clinical depression, the symptoms must be present for at least two weeks [4]. The study [5] shows that the earlier that treatment begins, the more effective it is. Thus, it is best for patients to be early diagnosed. Although it is impossible to physically examine all people at risk, surveillance screening test should be done.

¹ Corresponding Author: Nichnan KITTPHATTANABAWON, School of Informatics, Walailak University, Nakorn Si Thammarat, Thailand. 80160; E-mail: knichcha@wu.ac.th.

Since depression symptoms can alter patient's thinking process, their behaviors are changed in several aspects such as having less activities from losing interest to do thing and being uneasy easily from feeling hopeless. Though it is difficult to monitor every physical action for depression screening for every person, it is possible to find a sign in behavior on digital based media. The digital-based media refers to interact to ones' mobile phone, actions in an electronic game and behavior in social network applications. The signs of the disorder shown in social media can be easier to be recorded and analyzed.

In this work, we aim to analyze users' behavior via social network application to realize what the signs of depressive disorder. Since Facebook is a famous and widely used social network in Thailand as there are 26.9 million Thai users in 2019, we focus our feature design based on Facebook functions including posting, sharing, commenting and replying. The designed features are coped to be measurable as a numeric value without interpreting content; thus, it is easier to be gather, possibly automatically, with less privacy. We expect that the designed features regarding users' behavior on Facebook will help in detecting non-aware cases of depressive disorder and lead them for earlier treatments as well as prevention of a possible suicidal attempt.

2. Related Works

There are several researches on screening mental disorder using information technology. A summary of these works is provided in Table 1. From the summary, we found that most of the work handled a text related to mental disorder and suicidal risk. There are two main techniques including natural language processing to handle text and machine learning to create a decision-making model for determining the result. Their focused texts are from online sources such as twitter posts and news. Most of the works applies text-based classification techniques, such as Naïve-Bayes (NB) [6, 7, 8, 9], Decision-tree (DT) [7, 8, 9, 10], Support Vector Machine (SVM) [6, 8], Genetic algorithm (GA) [11] and Deep learning [12], to detect a post that is risky for expressing suicide attempt or shows a sign of mental disorder symptom. These works provide great help to reduce the risk of mental disorder; however, they require patients to make a post about what they have in mind. In a view of natural language processing (NLP), complexity of a target language also plays a crucial role in accuracy performance. Thus, the same technique applied to different language may result in noticeably different performance, and some languages including Thai still lack necessary resources and theoretical studies to improve its NLP performance.

In addition, the mental disorder may affect the patients to be unstable and result in rarely making a post, especially depressive disorder which makes one losing interest in doing anything [5]. Thus, solely focusing on detecting the mental risks from the text may be insufficient to detect all the at-risk people. Detecting people with depression from online sources may require additional information such as how they interact with other people in their social network.

Table 1. Related work on detection of mental disorder and suicide risk

Research Title	Techniques		Dataset
	NLP	ML	
An Automatic Screening for Major Depressive Disorder from Social Media in Thailand (2019) [6]	Word vector	NB, SVM	Facebook posts

Research Title	Techniques		Dataset
	NLP	ML	
Detection of Suicide-Related Posts in Twitter Data Streams (2018) [7]	Lexicon-based	NB, DT	Twitter posts
Google and suicides: what can we learn about the use of internet to prevent suicides? (2018) [8]	-	Regression analysis	Suicide data mortality from all American state
Detecting suicidality on Twitter (2015) [10]	-	SVM, Logistic Regression	Twitter posts
Analyzing the connectivity and communication of suicidal users on twitter (2016) [12]	-	TF-IDF, Unigram	Twitter posts
Natural Language Processing of Social Media as Screening for Suicide Risk (2018) [13]	Word vector	Deep Learning	Social media data
Mining Twitter for Suicide Prevention (2014) [14]	grammatical structures	NB, DT, K-NN	Twitter posts
Multi-class machine classification of suicide-related Communication on Twitter (2017) [15]	-	NB, DT	Twitter posts
The association of suicide-related Twitter use with suicidal behavior: A cross-sectional study of young internet users in Japan (2015) [16]	-	Association	Twitter posts

3. Methodology

This work focuses on Depression at risk screening via users' behavior in Face-book. We first gather users' behavior on Facebook for both positive and negative in depression into two groups. Statistics of their behavior are analyzed and calculated following our designed features. Then, we use the features to learn a classification model to detect depression-positive users. An overview of the process is illustrated in Figure 1.

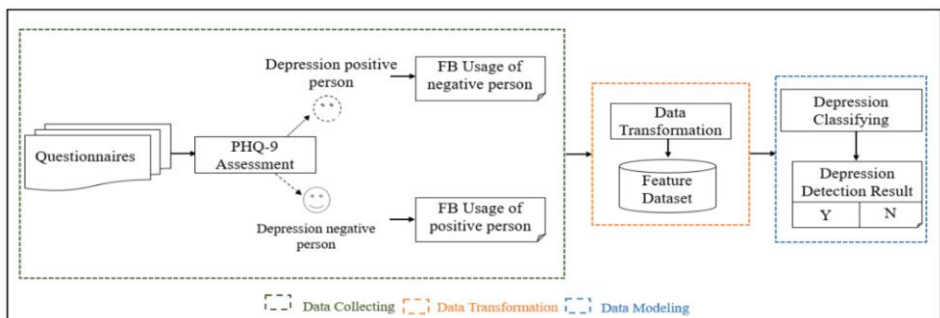


Figure 1. An overview of the proposed method.

3.1. Data Collection

To gather volunteers and differentiation between positive and negative in risk of depression, we design a questionnaire for asking for anonymous personal information and a test for depression. The chosen depression test in this work is PHQ-9 questionnaire [4] which is widely used and accepted for detecting signs of depression in Thailand. The details of questionnaire and other required information are provided in Table 2.

Table 2. The acquired information from participants

Question type	Question	Note
Personal Information	<ul style="list-style-type: none"> - Gender (Male, Female, Sexual Oriental, Alternative gender) - Age (year-olds) - Occupation 	
Depression testing	<ul style="list-style-type: none"> - PHQ-9 (0-3 scale answer where 0 means not at all, and 3 refers to often): <ol style="list-style-type: none"> 1. Little interest or pleasure in doing things 2. Feeling down, depressed, or hopeless 3. Trouble falling or staying asleep, or sleeping too much 4. Feeling tired or having little energy 5. Poor appetite or overeating 6. Feeling bad about yourself or that you are a failure or have let yourself or your family down 7. Trouble concentrating on things, such as reading the newspaper or watching television 8. Moving or speaking so slowly that other people could have noticed - Or the opposite being so fidgety or restless that you have been moving around a lot more than usual 9. Thoughts that you would be better off dead, or of hurting yourself 	These questions [5] are to capture symptoms of depression that affect ones' mind emotional, cognition, physical, and behaviors.
Facebook Account	<ul style="list-style-type: none"> - Consensus for monitoring their account for their public activities including post, share and reply 	Monitoring as seeing their actions as their friend

With PHQ-9 questionnaire, depression result is defined into five classes as minimal depression, mild depression, moderate depression, moderately severe depression and severe depression. In this work, volunteers are split into two groups for positive depression (from mild depression, moderate depression and severe depression) and negative depression (no depression) based on their depression result. We ask them to allow us to monitor their Facebook accounts and collect data according to our designed features. The data collection is conducted for at least 14 days according to depression guideline [5].

3.2. Data Transformation

Since each social network platform provides different functions, collectible features regarding users' behavior should be specified differently. In this work, we focus on Facebook's users; thus, the designed features are according to Facebook functions specifically. The features are about what and how a user uses and interacts to other accounts, especially posting, commenting, sharing and replying. The amount of these actions is counted and stored. Moreover, the timing of these actions is also collected to

calculate for a gap between each action. We though do not collect the content of these actions in this phase since they are private information. There are three aspects as Facebook functions, time and consecutive action. The details of the collected features are given in Table 3.

Table 3. The Features based on Facebook usage behavior for detecting depression at risk

Aspect	Features	Details
Facebook functions	Posting (posts made by an account owner)	- Average number of daily posts - Standard deviation of daily posts
	Sharing (posts made by sharing other account's post)	- Average number of daily shares - Standard deviation number of daily shares
	Commenting (comment as replying in another person's post and share)	- Average number of daily comments - Standard deviation number of daily comments
	Replying (comment as replying in another person's comments in own post and share)	- Average percentage of daily replies - Standard deviation number of daily replies
Time	Weekday - Weekend	- Number of actions done in weekday (Mon-Fri) - Number of actions done in weekend (Sat-Sun)
	Period	- Number of actions done in morning (06:01-12:00) - Number of actions done in afternoon (12:01-18:00) - Number of actions done in evening (18:01-24:00) - Number of actions done in late night (00:01-06:00)
Consecution	Consequence	- Number of actions done consecutively (activities within 5-minute gap)

Since the main functions of Facebook are posting, commenting, replying to a comment and sharing a post, our features are designed regarding the functions. The activity for these functions includes posting a text, an image and a video or replying with a text and a sticker/emoticon. Differently, we do not count the number of replying since there is a vary number of obtained comments for each person. Thus, we choose the percentage of replying from a total number of obtained comments from other users. This work excludes the chatting function provided with Facebook messenger because we are allowed to monitor only the activities that users set as public or friend only. The designed features in different aspects are to combine such as an amount of posting (function) done in weekday (period) and in consequence (consecution).

Data of the features are collected as counting number representing their frequency; hence, it is less complex and explicitly comparable. This work collects only actions via Facebook of Thai users; thus, we ignore the content of the posts and the replies. Although the posts and replies may contain the keywords obvious for showing the sign of depression, we decide to focus solely on interaction between users in this work since it breaches less privacy and is simpler in terms of collection. Moreover, analyzing Thai text requires a lot more techniques of natural language processing in which is complex and remains challenge.

We can use the features for analyzing the behavior of users who are positive in being depression. Moreover, these features can be used to automatically identify users who may be at risk of being depression using an automatic classification method.

3.3. Data Modeling

Once data are transformed, they are used to generate a model for an automated classification as a depression detection model. To generate the model, this work applies supervised learning which requires labeled dataset for training. The training dataset includes input data and their response value. Supervised Learning is to map an input to a particular class from a label of an output, and the mapping becomes a model for classification. Supervised Learning algorithms thus help to make predictions for new unseen data by referring to the generated model.

In this work, a set of features for behavior usage of Facebook users is designed for supervised learning input while designate classes of the output are positive-depressive and negative-depressive state of a user. The supervised learning hence generates a model as a depress detection model to map the significant features that have an impact towards the depressive status classes. There are several available supervised learning algorithms such Naïve-Bayes (NB) [6, 7, 8, 9] Support Vector Machine (SVM) [6, 8] and Deep learning [12]. They all have their own advantages and limitations from a different training approach. Without adjusting their default algorithm, we will apply the feature set and compare a performance of the generated models.

4. Experiments

We designed an experiment to explore the potentials of the designed features. We provided results that we extracted from the dataset using the features. Moreover, we used them in an automated classification with supervised machine learning techniques to see how they perform in detection of depression.

4.1. Data and Settings

This study has been granted with a Certificate of Ethical Approval from the Human Research Ethics Committee of Walailak University with the Project No. WU-EC-IN-0-187-62. Participants were volunteers who willingly applied to our experiment. They were first required to take a PHQ-9 test to determine their depression state. After selection, we had 160 Thai participants who commonly use Facebook and allowed us to monitor their account for 14 days. All participants are over 18-year-olds. The questionnaire results showed that there are 120 participants who are positive in depression-at risk and the rest are negative.

4.2. Outstanding Features for Detecting Depression

In this section, we provided results that we extracted from the dataset using the designed features. The results of extracted data of Facebook usage regarding depression are given in Figure 2-4.

From the data, the highly different statistics are from the number of late-nights actions, number of afternoon actions and standard deviation (SD) of daily posts, respectively. In details from function-based data, SD values of daily posts and shares were noticeably different while average (AVG) values of posts and shares were slightly different. This finding is understandable since depression-negative users tentatively had

consistent numbers of daily posts and shares while depression-positive users may have unstable mood that affects their inconsistent behavior. The depression-positive users had more likely post because it is a process of releasing a state of negative feelings or suffering. In the contrary, AVG values of comments and reply percentage are more usable than their SD values. This is because depression-positive users scarcely conducted a reply or comment, but depression-negative users are more likely to interact to other users. From time-based features, we can see the time when the two groups are active. The depression-positive users had more activity numbers during late night since it is a time when depression is more likely to occur, including loneliness and insomnia. As a result, people with depression spend more time online. Last, consecutive actions from depression positive users were also significantly higher than those of depression-negative users.

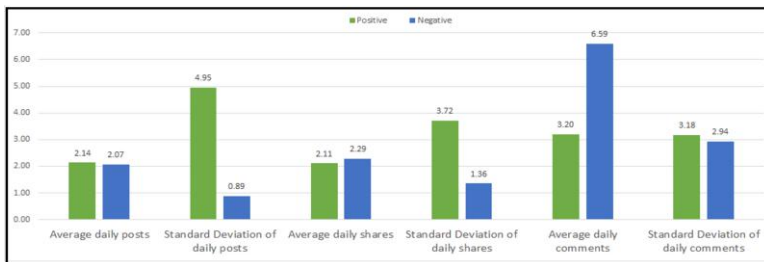


Figure 2. Data results of Facebook usage regarding depression according to the features of Facebook functions including posts, shares and comments.

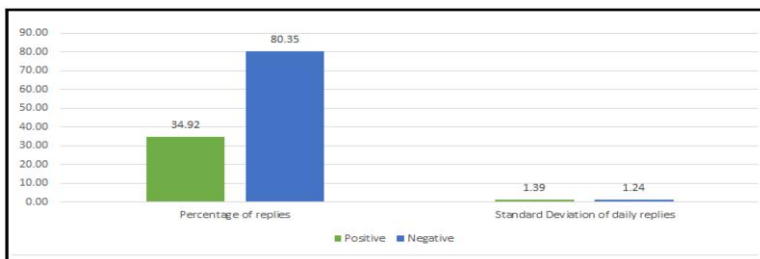


Figure 3. Data results of Facebook usage regarding depression of reply function

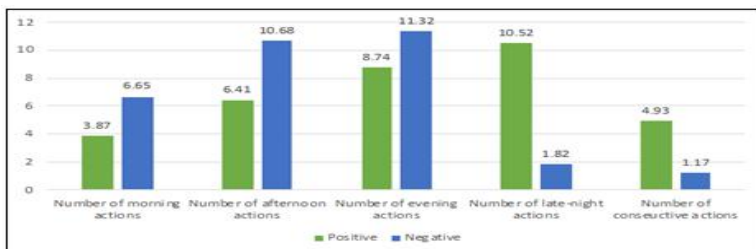


Figure 4. Data results of Facebook usage regarding depression according to the designed time and consecutiveness features

4.3. Performance of Classification Models

With the designed features given in Table 3, the data are in a form of numerical value and can be easily used in supervised machine learning. Four classification algorithms as SVM, NB, DT and deep learning with default setting are used to train the classifying model. With small set of data, 5-fold cross validation was applied. The evaluation results are shown in Table 4.

Table 4. Classification results from designed features

Algorithm	Precision	Recall	F-measure
SVM	81.4	88.1	84.6
NB	80.8	86.8	83.7
DT	77.9	80.3	79.1
Deep learning	84.4	89.7	87.0

From the results, deep learning performed the best in all measurements among the chosen algorithms. The results though were still low; thus, we decided to split training data based on participant gender and removed the data with less than 5 activities in total. There were 102 and 33 remained participants for female and male, respectively. Then, we experimented with the gender-based model and gained evaluation results as given in Table 5.

Table 5. Classification results from designed features splitting model based on participant's genders

Algorithm	Precision	Recall	F-measure
SVM	86.5	90.3	88.4
NB	84.6	89.9	87.2
DT	80.6	83.7	82.1
Deep learning	90.2	92.6	91.4

The results of classification were improved with splitting model based on participant's genders. This shows that gender of participants is an important factor for differentiating behavior of Facebook users.

5. Conclusion and Future Works

This paper proposes a method to detect depression at risk person via a behavior of Facebook users. A set of features related to Facebook functions interact able to other users is designed to represent users' activities in Facebook platform in a numerical value form. The designed features include posting, sharing, commenting and replying amount with periodic focus without interpreting its content. The limitation of the acquired data from Facebook user account is the number of activities as frequency of actions in a specific periodic time and consequential manner. Therefore, the data are easier to be collected with less privacy issue. To distinct between positive and negative depression-at risk, PHQ-9 questionnaire, commonly used to screen depression patient in Thailand, was used. The collected data according to the designed features with a depression result are used to classify a likelihood of being depress disorder. These features hence are used in supervised learning classification algorithm for detecting a risk of being depress disorder.

From the experiment of 160 Thai Facebook users, we found that the designed features based on users' activities from Facebook with deep learning algorithm yields about 87% in terms of F-measure. After analyzing the data, we thus split data regarding users' gender and removed obviously low active data, and the F-measure from classification went up to 91.4 which improves for 4 points. From the data, the highly significant features in differentiation between depression-positive and negative users are from the number of late-nights actions, percentage of replies, number of afternoon actions, standard deviation (SD) of daily posts and consecutive actions, respectively. In the future, we plan to include text analysis in detection of depression including sentimental analysis of a post, vulgarity degree of a language and keyword detection that depress person often used.

References

- [1] DSM-5 American Psychiatric Association. Diagnostic and statistical manual of mental disorders. Arlington: American Psychiatric Publishing, 2013.
- [2] 6 statistics about "Depression" that you may never know. Cited: 2017, 16 May. Available from: <https://www.sanook.com/health/7137/>.
- [3] Mike J. 1985. Classification algorithms. Wiley-Interscience.
- [4] Depression and suicide assessment form (Including 2q 9q and 8q assessment). DEPARTMENT OF MENTAL HEALTH Ministry of Public Health. (in Thai) Retrieved November 18, 2018, from: <https://www.dmh.go.th/test/download/view.asp?id=22>.
- [5] Halfin A. Depression: the benefits of early and appropriate treatment. *Amer. J. Manag. Care*, 2007, 13(4 Suppl), S92-7.
- [6] Hemtanon S, Kittiphattanabawon N. An automatic screening for major depressive disorder from social media in Thailand. The 10th National & International Conference "Global Goals, Local Actions: Looking Back and Moving Forward" Suan Sunandha Rajabhat University, Bangkok, Thailand, 2019 1(10), 103-113.
- [7] Vioulès MJ, Moulahi B, Azé J, Bringay S. Detection of suicide-related posts in Twitter data streams. *IBM Journal of Research and Development*, 2018, 62(1), 7-1.
- [8] Chandler V. Google and suicides: what can we learn about the use of internet to prevent suicides? *Public health*, 2018, 154, 144-150.
- [9] Desmet B, Hoste V. Online suicide prevention through optimised text classification. *Information Sciences*, 2018 439, 61-78.
- [10] O'dea B, Wan S, Batterham PJ, Calear AL, Paris C, Christensen H. Detecting suicidality on Twitter. *Internet Interventions*, 2015 2(2), 183-188.
- [11] Wu C, Landgrebe DA, Swain PH. The decision tree approach to classification, 1975.
- [12] Colombo GB, Burnap P, Hodorog A, Scourfield J. Analysing the connectivity and communication of suicidal users on twitter. *Computer communications*, 2016, 73, 291-300.
- [13] Coppersmith G, Leary R, Crutchley P, Fine A. Natural language processing of social media as screening for suicide risk. *Biomedical informatics insights*, 2018, 10, 1178222618792860.
- [14] Abboute A, Boudjeriou Y, Entringer G, Azé J, Bringay S, Poncelet P. Mining twitter for suicide prevention. In *International Conference on Applications of Natural Language to Data Bases/Information Systems*, Springer, Cham, 2014, June, pp. 250-253.
- [15] Burnap P, Colombo G, Amery R, Hodorog A, Scourfield J. Multi-class machine classification of suicide-related communication on Twitter. *Online Social Networks and Media*, 2017, 2, 32-44.
- [16] Sueki H. The association of suicide-related Twitter use with suicidal behavior: a cross-sectional study of young internet users in Japan. *J. Affective Disorders*, 2015, 170, 155-160.

Survey on Formal Concept Analysis Based Supervised Classification Techniques

Hayfa AZIBI^{a,1}, Nida MEDDOURI^b and Mondher MADDOURI^c

^aLIPAH, Faculty of Sciences of Tunis, University of Tunis El Manar, Tunisia

^bGREYC-CNRS UMR 6072, University of Caen Normandy, France

^cCollege of Business, University of Jeddah, Kingdom of Saudi Arabia

Abstract. Classification is a data mining task and which is a two-phase process: learning and classification. The learning phase consists of constructing a classifier or a model from a labeled set of objects. The classification phase consists classifying new objects by using the generated classifier. Different approaches have been proposed for supervised classification problems through Formal Concept Analysis, and which is a mathematical theory to build upon hierarchies of formal concepts. The proposed approaches in literature rely on the use of single classifier and ensemble methods. Single classifier methods vary between them according to different criteria especially the number of formal concepts generated. We distinguish overall complete lattice methods, sub-lattice methods and concept cover methods. Methods based on ensemble classifiers rely on the use of many classifiers. Among these methods, there are methods based on sequential training and methods based on parallel training. However, with the large volume of data generated from various sources, the process of knowledge extraction with traditional methods becomes difficult. That's why new methods based on distributed classifier have recently appeared. In this paper, we present a survey of many FCA-based approaches for classification by dividing them into methods based on a mono-classifier, methods based on ensemble classifiers and methods based on distributed classifiers. Different methods are presented and compared within this paper.

Keywords. Artificial intelligence, Data mining,, Machine learning, Supervised classification, Formal concepts analysis, Ensemble methods, Cloud

1. Introduction

The exploding volume and speed of data growth have triggered several challenges in many learning problems in real world. Classification is one of the most important tasks in Machine Learning. The classification problem aims to predict a class to which new data will fall under. In fact, the supervised classification analyzes the attributes and develops an accurate description or model for each class using descriptions submitted by attributes. Several classification algorithms were proposed in the literature and widely applied in practice. As references in the fields, we can highlight the Artificial Neural Network, Association Rule Mining, Formal Concept Analysis, Induction of Decision

¹Corresponding Author: Hayfa Azibi, Faculty of Sciences of Tunis, University of Tunis El Manar, Tunisia.
E-mail: hayfa.azibi@fst.utm.tn

Trees, Naïve Bayes and Support Vector Machine. Each supervised classification method is characterized by some features that can be fitted to some classification tasks.

Formal Concept Analysis (FCA) [1] is a popular method of Machine Learning methods [2]. It is a mathematical theory which builds upon hierarchies of formal concepts. Also, FCA is a theoretical framework which structures a group of objects and their attributes. The classification approach based on FCA is divided into two steps: learning step and classification step. In the learning step, a classifier is built by means of analysis of objects described by attributes in the training set. Each object is assumed to belong to a pre-defined class represented by a particular attribute in the training set. In the classification step, the model built in the first stage is used to classify the new objects.

This article provides a comparative study of FCA-based classification methods. In literature, several studies had carried out this comparative study. The authors in [3] carried out a theoretical and experimental comparative studies of some classification methods based on FCA. New methods have appeared in [4] which are based on a single classifier. The authors classified the classification methods based on the FCA by evoking the notions of complete lattice, sub-lattice and cover of the concept. Other algorithms are based on the taxonomy of [2] of existing supervised classification methods. This taxonomy is divided into two categories: exhaustive methods and combinatory methods. The first category is characterized by the use of a single classifier. The second contains the learning methods which exploit the paradigms of ensemble learning. Hence, this article presents a comparison of the FCA-based classification methods cited in existing work with those that have recently emerged. We introduce a new category of methods which is based on a distributed classifier. The paper is organized as follows: in section 2 we introduce methods based on mono-classifier. Section 3 introduces methods based on ensemble classifiers. Section 4 presents the methods based on distributed classifiers. In section 5, we discuss such methods. Finally, we present our conclusion.

2. Methods based on mono-classifier

Several classification methods based on FCA, were presented in literature. Many FCA-based classification algorithms that generate a complete lattice have been developed, we can cite GRAND [5,6], GALOIS [7,8], RULEARNER [9] and NAVIGALA [10].

GRAND builds a complete lattice. All concepts are presented in the lattice except the supremum or the infimum each one of them is an empty set. To update the lattice, for each new object that has attributes shared lattice nodes, nodes with common attributes will be inserted, in the meantime, all redundant connections will be removed. It induces the most accurate rules in order to be applied on each new object.

GALOIS is a system that provides an incremental aspect of lattice construction. It uses two different methods to determine classes of new objects. The system performs a similarity calculation between the new object and each concept. This similarity is the concept properties verified by the object. Finally, GALOIS assigns to the object the class of the most similar concept.

RULEARNER constructs a full-concept lattice. In the learning phase, this system builds a set of rules that overlap the object nodes of the lattice. Besides, to classify new objects with an ordered list, RULEARNER uses these rules by keeping the order. However, it uses majority voting for unordered list.

NAVIGALA is a recognition system that was developed to recognize noisy graphical objects and symbol images by navigating through the complete lattice like navigating in a classification tree [10].

In [11], the authors proposed an incremental learning method for mining sequential patterns to find different human behavioral patterns in non-stationary smart environments. The input data are labeled sequences that gradually arrive from a sensor. If a lattice is not found before, the presented method makes a lattice initialization. For each element of the new training data set, an iteration of the lattice checks whether the iterated concept should be updated, created or ignored [11].

FCA-based query expansion was discussed in [12]. This study is based on the extraction of description topics from documents. In fact, a set of retrieved documents is obtained based on a query against a set of documents to perform the expansion. The description topics defined as intrinsic concepts in a document are extracted from the recovered documents. Using the retrieved documents as objects and the description topics as attributes, a lattice is constructed as the possible expansion space. The expanded query will be generated by the selected lattice nodes.

Despite the several systems of lattice concept-based classification, their problem remains in time and space complexities. This common limit is due to the navigation in the whole space search. To solve this issue, many researches presented approaches based on sub-lattice classification. A sub-lattice is a partial part of Galois lattice. The classification process is the same for a complete lattice and sub-lattice methods but the major difference between them is about how many formal concepts are generated. LEGAL [8], CIBLe [13], CLNN & CLNB [14] and CLANN [15] build a sub-lattice. The sub-lattice generation contributed drastically to the reduction of theoretical complexities and execution time.

To build a sub-lattice, LEGAL applies two learning parameters. The objects of the initial formal context will be divided into positive objects and negative objects. For each new node, LEGAL begins by constructing its sub-nodes. Valid nodes are then retrieved using learning parameters. These valid nodes are characterized by a great number of positive objects. The algorithm ends when there are no valid nodes.

CIBLe starts with the construction of a sub-lattice. It gives a numerical re-description to the training data. In its classification step, CIBLe performs a similarity calculation to classify new objects. In practice, it uses three different measures: Manhattan distance, Mahalanobis distance and Euclidean distance.

CLNN & CLNB integrate respectively two classifiers, the Naïve Bayes classifier and the Nearest Neighbors classifier, in the lattice concept. They also use the majority vote to classify new objects.

CLANN builds a sub-lattice in the training phase and only data with two classes are handled. The obtained sub-lattice will be used to construct neural networks that perform classification.

The authors in [16] proposed a classification method based on FCA which applies the minimum description length (MDL) to select concepts. Target class code tables are used individually to get compression objects. For classification, the attributes of an object are covered by sets of elements found in code tables of classes. Finally, coding lengths are calculated for all classes. The class that has minimum coding length is chosen.

A concept cover is defined as a part of lattice which contains only relevant concepts. To build cover concepts, IPR [17] resorted to the greedy algorithm. For classification,

IPR looks for rules with the premise that matches attributes of the new object. Rules that were applied represent the most weighed ones for the involved object. CITREC [18] is another cover—based classification method. The first step of CITREC is to convert numeric and nominal attributes to binary ones. Then, the creation of a new context (the reduced context) where the objects and the classes of the different objects of the training set are equal. Next, the lattice is built using the reduced context. To classify new objects, CITREC uses the majority vote.

Different supervised classification methods based on FCA were presented in this section: complete lattice methods, sub-lattice methods and concept cover methods. Concept lattice and sub-lattice proceed similarly. But, using sub-lattice is feasible due to its running time compared to the concept lattice. In fact, this feature leads to the generation of the relevant rules but this causes a loss of information. The problems in the presented methods remain in the use of a single classifier, the high complexity and the type of handled data which is binary for almost all systems. As a result, many researches in literature oriented to the combination of classification methods based on the ensemble methods the best known of which are boosting and bagging.

3. Methods based on ensemble classifiers

There is a growing realization that the use of ensemble classifiers can be more effective than the use of single classifiers. Why rely on the best single classifier, while we can obtain the most accurate and reliable result from a combination of several? This is the reasoning behind the idea of ensemble methods. Several classifiers based on ensemble methods were developed in literature. There are two categories: methods based on sequential training and methods based on parallel training. The difference between the two categories is that the first one generates classifiers sequentially but the second method generates parallel classifiers.

In this context, BFC [19] and BNC [20] are two methods based on sequential training that were proposed in literature. BFC is a method based on FCA and also benefits from boosting algorithms. The basic idea of BFC consists in selecting a group of data from the learning set after assigning equal weights to the training objects. Then, BFC extracts the relevant formal concepts within the subset. To classify the learning data, the BFC method uses the training objects weights. This process is repeated until getting the final classifier. For BNC, it proceeds in the same way as BFC. However, what differs between BFC and BNC is the data type and attribute selection. BFC makes the learning from binary data but BNC handles nominal data. For attribute selection, BFC uses Shannon's Entropy while BNC uses informational gain.

FPS-FCA [21], DNC [22], RMCS [23] and B-RCL [24] are based on parallel training. FPS-FCA divides the training set into many subsets. FPS-FCA uses the obtained subsets to generate formal contexts in order to extract classification rules. DNC builds several parallel classifiers. In this case, each classifier is constructed using the same learning algorithm. DNC creates disjoint and stratified samples. On each sample, CNC (Classifier Nominal Concept) is then constructed [22]. The classifiers' outputs are finally combined by a majority vote. RMCS is also a method based on parallel training. RMCS begins with the construction of a classification table using a formal context. Then, RMCS assigns classifiers to the objects that exist in the context. After matching, it searches the

test set object neighbors by means of a similarity metric. The classifiers that are selected for classification are those which have more neighbors that were found [23]. The author of [24] proposed the fusion of Random Conceptual Coverage Learner with bagging paradigm. RCL differs from other FCA coverage methods in attribute selection. RCL performs the selection randomly from the training set. B-RCL was proposed to reduce the variance caused by RCL.

4. Methods based on distributed classifiers

In recent decades, the volume of data generated from different sources flows continuously. Hence, the extraction of knowledge from numerous data sources using mono-classifier methods and ensemble learning methods becomes a difficult task. The existing algorithms are not scalable to the huge new and larger datasets for knowledge extraction and representation. To solve this issue, frameworks for big data applications are developed [25]. However, these frameworks are based on a distributed environment like Cloud Computing [26]. In this field, several distributed data mining tools were developed. In [27], the authors introduced a cloud-based framework to implement home diagnostic service. The user submits a query which contains the disease information. A dispatcher selects nodes to determine the medical records corresponding to the request. The dispatcher then merges the search results and passes them on to a data analysis cluster. A lattice will be constructed according to the medical records retrieved and reveals the relationships between diseases with common symptoms [27].

The authors of [28] presented a Multi-Cloud Service Composition approach which based on FCA. In fact, from each lattice the requested services are extracted. Then, the use of the lattice for filtering candidate clouds according to providers that were selected [28]. Finally, there is a selection of the appropriate and optimal cloud set from which the best services are selected.

The work presented in [29] is a Distributed Classifier Nominal Concept. This method is a distributed version of CNC, which handles nominal data. Dist-CNC was implemented using Distributed Weka Spark which is a distributed framework for weka. During the learning phase, the master node divides the input data into partitions and then distributes the training task and the partitions obtained to the slave nodes. Slave nodes apply CNC on each partition and return results to master node. To evaluate the model, the master node distributes the model formed to the slave nodes. Then, each slave node uses its partitions to evaluate the model. The final results are merged and then returned.

5. Discussion

Tables 1 and 2 show a comparison between supervised classification methods based on FCA by category. The comparison criteria chosen show the characteristics of each method. As shown in tables 1 and 2, these methods handle various data types as binary, numeric and nominal data. In table 1, the methods construct complete lattices, sub-lattice or cover concepts. These methods classify the datasets which contain several classes with the exception of LEGAL, RULEARNER and CLANN. For lattice construction, these methods use algorithms to generate concept lattices. These algorithms can be incremen-

Table 1. Comparison between mono-classifier based methods

Methods	GRAND	GALOIS	RULEARNER	NAVIGALA	LEGAL	CIBLe	CLNN & CLNB	CLANN	IPR	CITREC
Data type	Binary	Nominal	Nominal	Binary	Binary	Numeric	Numeric	Binary	Binary	Binary
Number of classes	multi-class	multi-class	2 classes	multi-class	2 classes	multi-class	multi-class	2 classes	multi-class	multi-class
Construction lattice algorithm	Oosthuizen	Carpineto and Romano	Oosthuizen	Bordat extension	Bordat	Modified Bordat	Top-down approach	Modified Bordat	Heuristic approach	Godin
Concepts structure	Complete lattice	Complete lattice	Complete lattice	Complete lattice	Sub-lattice	Sub-lattice	Sub-lattice	Sub-lattice	Cover concepts	Cover concepts
Concept selection	Maximally complete concepts	Maximally complete concepts	Maximally complete concepts	Distance measure	Maximally complete concepts	Lattice level Entropy	Support Precision	Heuristic algorithms	Entropy	Support
Incremental	Yes	Yes	Yes	Yes	No	No	No	No	Yes	Yes
Learned knowledge	Rules	Relevant concepts	Ordered and unordered rules	Concepts	Relevant concepts	Relevant concepts	Rules	Relevant concepts	Rules	Rules
Classification	Vote	Similarity or vote	General rule	Navigation in a GALOIS lattice like navigation in the decision tree	Vote	k-nearest neighbors algorithm	Verified rule + vote	Neural networks	weighted rules	Vote
Complexities	$O(2^l \times l^4)$ with 1 is the minimum between n and m.	$O(3^m \times 2^n \times n) < O(3^{2m} \times n)$	Idem to GRAND.	$O(L \times n^3) + O(nm^2)$ with L the number of concepts	$O((L \times n(1-\alpha))$ with L the number of concepts and α the validity criteria	$O(L \times m^3)$ with L the number of concepts of the sub-lattice	$O(L \times n \times (1-\alpha))$	$O(2^{\min(n,m)})$	$O(n^2 \times m^2 \times (m+n))$	$O(2^m \times n)$

Table 2. Comparison between ensemble based methods

Methods	BFC	BNC	FPS-FCA	DNC	RMCS	B-RCL
Concepts structure	Cover	Cover	Sub-lattice	Cover	Complete lattice	Sub-lattice
Data type	Binary	Nominal	Nominal	Nominal	Binary	Nominal
Concept selection	Shannon entropy	Informational gain	Relevant patterns	Informational gain	Distance	Euclidean Random coverage
Learned knowledge	Rule	Rules	Graph pattern structure	Rules	Concepts	Rules
Classification	Weighted vote	Weighted vote	Hypotheses	Majority vote	Maximum number of neighbors	Majority vote
Ensemble	Sequential	Sequential	Parallel	Parallel	Parallel	Parallel
Complexity	$O(n \log(n) + nm)$	$O(n \log(n) + nm)$ with m = nominal attribute	$O(nm/k)$ with k is the number of processors	$O(n')$ with n' = stratified samples	$O(nm \log(n))$	$O(N^3)$ with N is the number of base classifiers

tal or non-incremental. All methods in tables 1 and 2 use the concept selection to induce rules in order to classify new instances through these rules.

However, these rely on different selection measures such as the Informational Gain for BNC and DNC, the support for CITREC and the Shannon entropy for BFC. IPR and CLNN & CLNB uses support, precision and recall to obtain concepts.

Methods based ensemble classifiers use multiple classifiers that are combined by vote techniques. These methods choose to represent the learned knowledge by relevant concepts or rules. In the classification phase, each system uses its appropriate method to determine a class for each new object. The majority vote is applied by GRAND, CITREC, LEGAL, BFC and DNC. It may also be used for GALOIS that also applies the similarity calculation. CLNN & CLNB applies voting strategy and verified rules for prediction. To predict a class for a new object, RULEARNER makes a selection of rules by respecting the order of the antecedents, CIBLE applies K-Nearest Neighbors algorithm and CLANN utilizes neural networks algorithm for classification. IPR uses weighed rules and BNC uses the weighed vote. RMCS classifies new examples by looking for the maximum number of neighbors and FPS-FCA uses hypothesis.

Tables 1 and 2 also propose a comparison of the theoretical complexities of different classification methods based on FCA where n is the number of objects and m is the number of attributes. As shown in table 1, all methods have an exponential complexity. Sub-lattice methods reduce this complexity because the build is a part of the lattice. Cover concepts methods have minimal complexity thanks to the generation of only the most relevant concepts. However, for ensemble classifiers, parallel methods like DNC, RMCS and FPS-FCA have a linear complexity, a polynomial logarithmic complexity and a polynomial complexity, respectively. For BFC and BNC, there is a complexity optimization that reaches a polynomial logarithmic order. The extraction of knowledge from large data sets still a challenge and a difficult task for traditional data mining tool. Distributed classifiers constitute a solution to answer this problem.

6. Conclusion

In this paper, we presented several FCA-based classification methods. First, we introduced methods based on mono-classifier that regroup the methods based on full lattice, sub-lattice and cover concept. Second, we presented methods based on classifiers ensemble. They rely on the use of many classifiers by parallel or sequential training. Finally, we introduced methods based on distributed classifiers to answer the problem of knowledge extraction from large data sets.

References

- [1] Ganter B, Wille R. *Formal Concept Analysis: Mathematical Foundations*. Springer Verlag; 1999.
- [2] Trabelsi M, Meddouri N, Maddouri M. New taxonomy of classification methods based on Formal Concepts Analysis. In: *Proceedings of the 5th International Workshop "What can FCA do for Artificial Intelligence"?* co-located with the European Conference on Artificial Intelligence. vol. 1703; 2016. p. 113–120.
- [3] Fu H, Fu H, Njiwoua P, Nguifo EM. A Comparative Study of FCA-Based Supervised Classification Algorithms. In: *Proceeding of Second International Conference on Formal Concept Analysis*; 2004. p. 313–320.
- [4] Meddouri N, Maddouri M. Classification methods based on formal concept analysis. In: *Proceedings of the 6th International Conference on Concept Lattices and Their Applications*; 2009. p. 9–16.
- [5] Oosthuizen GD. *The use of a lattice in knowledge processing*. University of Strathclyde. Glasgow, Scotland, UK; 1988.
- [6] Oosthuizen G. *The application of concept lattice to machine learning*. Dept Comput Sci, Univ Pretoria, Pretoria, South Africa, Tech Rep CSTR. 1996;94(01).
- [7] Carpineto C, Romano G. GALOIS: An Order-Theoretic Approach to Conceptual Clustering. In: *Proceedings of the Tenth International Conference on Machine Learning*. Morgan Kaufmann; 1993. p. 33–40.
- [8] Njiwoua P, Nguifo EM. Forwarding the choice of bias LEGAL-F: using feature selection to reduce the complexity of LEGAL. *Proceedings of BENELEARN-97, ILK and INFOLAB*. 1997;p. 89–98.
- [9] Sahami M. *Learning Classification Rules Using Lattices (Extended Abstract)*. In: *Proceedings of the 8th European Conference on Machine Learning*. vol. 912 of Lecture Notes in Computer Science. Springer; 1995. p. 343–346.
- [10] Visani M, Bertet K, Ogier J. Navigala: an Original Symbol Classifier Based on Navigation through a Galois Lattice. *International Journal of Pattern Recognition and Artificial Intelligence*. 2011;25(4):449–473.
- [11] Hao J, Bouzouane A, Gaboury S. An incremental learning method based on formal concept analysis for pattern recognition in nonstationary sensor-based smart environments. *Pervasive and Mobile Computing*. 2019;59:101045.
- [12] Yu H, Shi C, Bai Y, Zhang C, Hearne R. Query Expansion Based on Formal Concept Analysis From Retrieved Documents. *Journal of Internet Technology*. 2019;20(2):409–421.
- [13] Njiwoua P, Nguifo EM. Améliorer l'apprentissage à partir d'instances grâce à l'induction de concepts: le système CIBLe. *Revue d'intelligence artificielle*. 1999;13(2):413–446.
- [14] Xie Z, Hsu W, Liu Z, Lee ML. Concept lattice based composite classifiers for high predictability. *Journal of Experimental and Theoretical Artificial Intelligence*. 2002;14(2-3):143–156.
- [15] Tsopzé N, Mephu Nguifo E, Tindo G. CLANN: Concept Lattice-based Artificial Neural Network for Supervised Classification. In: *Proceedings of the 5th International Conference on Concept Lattices and Their Applications*. vol. 331; 2007. .
- [16] Makhhalova T, Kuznetsov SO, Napoli A. A First Study on What MDL Can Do for FCA. In: *CLA 2018 - The 14th International Conference on Concept Lattices and Their Applications*; 2018. .
- [17] Maddouri M. Towards a machine learning approach based on incremental concept formation. *Journal of Intelligent Data Analysis*. 2004;8(3):267–280.
- [18] Douar B, Latiri C, Slimani Y. Approche hybride de classification supervisée à base de treillis de Galois: application à la reconnaissance de visages. In: *Actes des 8èmes Journées Francophones en Extraction et*

Gestion des Connaissances. vol. E-11 of *Revue des Nouvelles Technologies de l'Information*. Cépaduès-Éditions; 2008. p. 309–320.

- [19] Meddouri N, Maddouri M. Boosting Formal Concepts to Discover Classification Rules. In: *Proceeding of the 22rd International Conference on Industrial, Engineering & Other Applications of Applied Intelligent Systems*. vol. 5579 of *Lecture Notes in Computer Science*. Springer; 2009. p. 501–510.
- [20] Meddouri N, Maddouri M. Adaptive learning of nominal concepts for supervised classification. In: *Proceedings of the 14th International Conference on Knowledge-Based and Intelligent Information and Engineering Systems*. vol. 6276 of *Lecture Notes in Computer Science*. Springer; 2010. p. 121–130.
- [21] Kuznetsov SO. Fitting pattern structures to knowledge discovery in big data. In: *International Conference on Formal Concept Analysis*. Springer; 2013. p. 254–266.
- [22] Meddouri N, Khoufi H, Maddouri M. Parallel learning and classification for rules based on formal concepts. In: *Proceedings of the 18th International Conference on Knowledge-Based and Intelligent Information and Engineering Systems*. *Procedia Computer Science*. Elsevier; 2014. p. 358–367.
- [23] Kashnitsky Y, Ignatov DI. Can FCA-based Recommender System Suggest a Proper Classifier? *arXiv preprint arXiv:150405473*. 2015;.
- [24] Ali MAT. Bagged Randomized Conceptual Machine Learning Method. *College of Engineering*. Qatar; 2018.
- [25] Inoubli W, Aridhi S, Mezni H, Maddouri M, Nguifo EM. An experimental survey on big data frameworks. *Future Generation Computer Systems*. 2018;86:546 – 564.
- [26] Mezni H, Abdeljaoued T. A cloud services recommendation system based on Fuzzy Formal Concept Analysis. *Data and Knowledge Engineering*. 2018;116:100 – 123.
- [27] Lin W, Dou W, Zhou Z, Liu C. A cloud-based framework for Home-diagnosis service over big medical data. *Journal of Systems and Software*. 2015;102:192 – 206.
- [28] Mezni H, Sellami M. Multi-Cloud Service Composition Using Formal Concept Analysis. *Journal of Systems and Software*. 2017;134(C):138152.
- [29] Fray R, Meddouri N, Maddouri M. Cloud Implementation of Classier Nominal Concepts using Distributed WekaSpark. In: *Supplementary Proceedings of ICFCA 2019 Conference and Workshops*. vol. 2378 of *CEUR Workshop Proceedings*. CEUR-WS.org; 2019. p. 125–136.

Cross-Based Adaptive Guided Filtering

Ming YAN^{a,b}, Yueli HU^{a,b, 1}, Kai LI^{a,b} and Jianeng ZHAO^{a,b}

^a *Shanghai University Microelectronic Research and Development Center, Shanghai 200072, China*

^b *School of Mechatronic Engineering and Automation, Shanghai University, Shanghai 200072, China*

Abstract. Edge-preserving and structure-preserving smoothing filtering has attracted much interest in the last decades. A conventional linear filter effectively smoothens noise in homogeneous regions but blurs the edges of an image. This study aimed to present an adaptive guided filter using a cross-based framework. The proposed method outperformed many other algorithms in terms of sharpness enhancement and noise reduction. Moreover, the cross-based adaptive guided filter had a fast and nonapproximate linear-time algorithm as the guided filter.

Keywords: Adaptive denoising, cross-based framework, edge-preserving smoothing, guided filter

1. Introduction

Image enhancement and image noise reduction are two opposing technologies. The former aims to improve the high-frequency components of the image and enhance the edge and texture features, while the latter aims to smooth the signal and remove high-frequency noise. These two technologies have been the research hotspots in the field of image processing. Nowadays, many filtering techniques, such as bilateral filtering [1], anisotropic diffusion [2-5], guided filtering [6], and so on are available, which can remove high-frequency details while removing noise in smooth areas.

Anisotropic diffusion can be used to preserve and strengthen large edges, but its oversmoothness filters out noise and small textures at the same time. Bilateral filtering has been widely used [7], which achieves image noise reduction according to spatial correlation and gray correlation. However, it has certain limitations due to the complexity of the adaptive value of the coefficients. At the same time, it has a staircase effect during smooth operations.

The effect of guided filtering is similar to that of bilateral filtering. However, its algorithm complexity is much lower than that of bilateral filtering, and it has unique advantages in terms of detail enhancement and high dynamic range compression [8]. Different from bilateral filtering, guided filtering is essentially a local multi-point estimation [9]; that is, the calculation of multiple observation points in the neighborhood is needed to estimate a certain point. Bilateral and guided filtering can be combined to make an overall combined estimation because the points in these neighborhoods are also

¹ Corresponding Author: Yueli Hu, ^a *Shanghai University Microelectronic Research and Development Center, Shanghai 200072, China*; ^b *School of Mechatronic Engineering and Automation, Shanghai University, Shanghai 200072, China*; E-mail: huyueli@shu.edu.cn.

used for other points. The estimates are worthy of calculation. This type of combined estimation of multiple points has a great improvement over previous operations for domain window estimation [10-12].

Edge enhancement leads to clearer subjective perception of the image and more comfortable visual effect for people. However, the a forementioned algorithm is only for the noise reduction and smoothing effect in the neighborhood. Therefore, the selection of the domain window is crucial in edge enhancement. Compared with global stereo matching, local stereo matching has more advantages in terms of calculation amount, parameter setting, and implementation efficiency [13]. For video noise reduction, the real-time requirement cannot be ignored. A previous study focused on real-time matching and introduced speed comparison in stereo matching [14]. The principle of bilateral filtering and stereo matching was described in previous studies for the effectiveness of dynamic stereo processing architecture [15-16].

In response to the need for edge enhancement, this study proposed a cross-based adaptive guided filter. The cross-based neighborhood was introduced into a guided filter, and a locally adaptive variable was added [17]. The adaptive change in the neighborhood might guide the smoothing effect of filtering.

2. Guided filtering

A previous study proposed guided filtering to make a linear transformation of the guided image, which had edge protection characteristics like those of bilateral filtering. Compared with bilateral filtering, guided filtering avoided the gradient reversal effect [6].

2.1. Definition

The main idea of guided filtering is to have a linear relationship between output image Q and guided image I , that is:

$$Q_i = a_k I_i + b_i, \forall i \in w_k \quad (1)$$

w_k is a square filter window with radius r , and (a_k, b_k) is a linear coefficient. Within a certain range w_k , their values are fixed. Eq. (1) shows that if image I has edges, image Q also retains the edges, and $\nabla Q = a \nabla I$.

The value of (a_k, b_k) is determined by minimizing the difference between input image P and filtered output image Q as shown in Eq. (2):

$$E(a_k, b_i) = \sum_{i \in w_k} ((a_k I_i + b_i - P_i)^2 + \varepsilon a_k^2) \quad (2)$$

ε is a normalization coefficient used to adjust a_k , preventing its value from being too large. (a_k, b_k) is calculated using Eq. (3) and Eq. (4):

$$a_k = \frac{\frac{1}{|w|} \sum_{i \in w_k} I_i P_i - \mu_k \bar{P}_k}{\sigma_k^2 + \varepsilon} \quad (3)$$

$$b_k = \bar{P}_k - a_k \mu_k \quad (4)$$

μ_k and σ_k^2 are the mean value and variance value of guided image I in the filter window, respectively. $|w|$ is the number of pixels in w_k , and $\bar{P}_k = \frac{1}{|w|} \sum_{i \in w_k} P_i$ is the mean value of input image P in the filter window w_k .

The proposed linear model was used for all partial windows of the entire image. The occurrence of pixel i in different windows Q_i leads to its uncertain values in different windows. Therefore, the average value of Q_i is calculated using Eq. (5):

$$Q_i = \frac{1}{|w|} \sum_{k: j \in w_k} (a_k I_i + b_k) = \bar{a}_i I_i + \bar{b}_i \quad (5)$$

$$\text{In Eq. (5), } \bar{a}_i = \frac{1}{|w|} \sum_{k \in w_i} a_k, \quad \bar{b}_i = \frac{1}{|w|} \sum_{k \in w_i} b_k.$$

After such adjustments, the linear coefficient (\bar{a}_i, \bar{b}_i) refers to the spatial mean value, and a direct linear relationship no longer exists between ∇Q and ∇I . (\bar{a}_i, \bar{b}_i) is the output of an average filter, and its gradient is much smaller near the strong edge than I . However, $\nabla Q \approx \bar{a} \nabla I$ is still used here, indicating that the gradient changes in I can still be largely retained in Q .

A function that contains the guided image I , input image P , and output image Q is obtained as follows:

$$Q_i = \sum_j W_{ij}(I) P_j \quad (6)$$

W_{ij} is a function that guides the image and is independent of input image P .

a_k can be written as a weighted sum of P : $a_k = \sum_j A_{kj}(I) P_j$, which is relevant to only the weights of. If Eq. (5) is rewritten as Eq. (6), then:

$$W_{ij}(I) = \frac{1}{|w|^2} \sum_{k: (i,j) \in w_k} \left(1 + \frac{(I_i - \mu_k)(I_j - \mu_k)}{\sigma_k^2 + \varepsilon} \right) \quad (7)$$

2.2. Edge Protection Filtering

If $\varepsilon = 0$, Eq. (2) has an optimal solution when $a_k = 1, b_k = 0$. If $\varepsilon > 0$, two situations occur:

Flat area: If the pixel value of image I is constant in the window, then the solution of Eq. (2) should be $a_k = 0, b_k = \bar{P}_k$.

Fluctuation area: If the image pixel value of image I changes greatly within window w_k , that is, its variance value is large, then the value of a_k tends to 1 and the value of b_k tends to 0.

\bar{a}_i and \bar{b}_i are the average values of a_k and b_k , respectively, which can be obtained using Eq. (5): If a pixel is a point in the undulating region, its value does not change. If it is in a flat region, its value is replaced by the average value of the surrounding pixels. The difference between the flat area and the undulating area is defined by \mathcal{E} . For image blocks with large variance values, that is, when σ^2 is larger than \mathcal{E} , the pixel values in the image blocks are retained, and conversely, the pixel values are replaced by the mean value. The effect of \mathcal{E} is similar to the gray variance in bilateral filtering.

2.3. Summary

The edge protection smoothing characteristics of guided filtering can be illustrated by a 1D signal as an example. As shown in Figure 1, if I_i and I_j are on the same side of the edge, $I_i - \mu_k$ and $I_j - \mu_k$ have the same +/- sign. However, if they are on

$$1 + \frac{(I_i - \mu_k)(I_j - \mu_k)}{\sigma_k^2 + \mathcal{E}}$$

different sides, they have opposite signs. Then, the value of in this case is much smaller than the value in the case of the same side, which tends to 0. This means that the pixel values on both sides of the edge are not averaged; when

$\sigma_k^2 \ll \mathcal{E}$ (flat area), $W_{ij}(I) = \frac{1}{|w|^2} \sum_{k(i,j) \in w_k} 1$, it is equivalent to a low-pass filter.

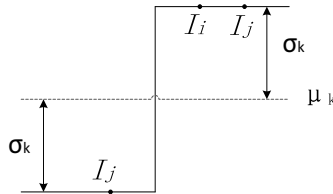


Figure 1. A 1D example of an ideal step edge.

3. Improved cross-based adaptive guided filtering

3.1. Offset ξ

This study introduced an adaptive offset ξ in the guided filtering to enhance the ability of bilateral filtering to strengthen edges [18]. The improved adaptive guided filtering was found to have better filtering effects at the edges and better detail retention.

In the filter window of guided image I , the maximum value MAX_k , minimum value MIN_k , and average value $MEAN_k$ can be used to replace the pixel values in the image. Supposing that $\Delta_{ki} = I_i - MEAN_k$, Δ_{kj} is the difference between each pixel in the image and the average value. ξ is defined as Eq. (8):

$$\xi_i = \begin{cases} MAX_k - I_i, & \text{if } \Delta_{ki} > 0 \\ MIN_k - I_i, & \text{if } \Delta_{ki} < 0 \\ 0, & \text{if } \Delta_{ki} = 0 \end{cases} \quad (8)$$

$(I_i + \xi_i - \mu_k)$ is used instead of the $(I_i - \mu_k)$ in the original guided filter. If the pixel value I_i to be obtained is greater than the average value, its value is replaced with MAX_k ; if the pixel value is less than the average value, it is replaced with MIN_k ; and it is not replaced if it is equal to the average value. Substituting the original value with the lowest value helps strengthen the edges, thereby making the border texture clearer.

3.2. Cross-Based Adaptive Guided Filtering Algorithm

Scharstein and Szeliski proposed a stereo matching algorithm [19]. Compared with global stereo matching, local stereo matching is more efficient and easier to implement [20-22]. A cross-based local stereo matching algorithm was proposed to work out the size and shape of the adaptive window of the pixel [17]. A cross-based adaptive guided filtering algorithm is more practical through orthogonal integral image technology. The present study introduced this method into guided filtering and adjusted the size and shape of the neighborhood window.

The local cross-based adaptive guided algorithm is divided into two steps: First, a cross-based with a base radius is established with the pixel to be measured as the center, and the central pixel can be called a base pixel P . As shown in Figure 2, the color similarity is calculated for the pixel values in its four directions to adjust the arm length $\{h_p^-, h_p^+, v_p^-, v_p^+\}$. Using h_p^- , for example, to calculate the difference between consecutive pixels in this direction and P , and to find the maximum arm length r^* similar to the base pixel value, r^* is calculated as shown in Eq. (9):

$$r^* = \max_{r \in [1, L]} (r \prod_{i \in [1, r]} \delta(p, p_i)) \tag{9}$$

$p_i = (x_p - i, y_p)$, L is the preset maximum arm length value, $\delta(p_1, p_2)$ is the color similarity indicator function between pixels P_1 and P_2 , and Eq. (10) provides its definition:

$$\delta(p_1, p_2) = \begin{cases} 1, & \max_{c \in \{R, G, B\}} (|I_c(p_1) - I_c(p_2)|) \leq \tau \\ 0, & \text{otherwise} \end{cases} \tag{10}$$

τ is the pixel color difference threshold between p_1 and p_2 . When the maximum arm length r^* is obtained, it equals the value of h_p^- .

Second, the determination of the cross-based shows that each pixel only needs to store four arm length values, and the field determination of the base pixel is the superposition of the horizontal basis on its vertical basis, that is:

$$U(p) = \bigcup_{q \in V(p)} H(q) \tag{11}$$

q is the pixel on the vertical basis of P , as shown in Figure 2, $H(q)$ is the horizontal basis of q , and $U(p)$ is the adaptive neighborhood of pixel p .

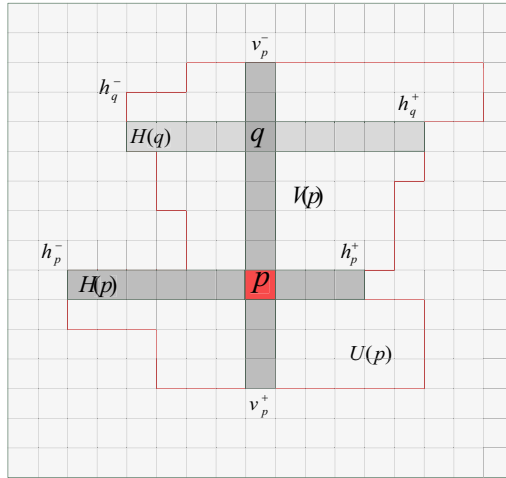


Figure 2. Configuration of a cross-based frame.

4. Experimental results

The value of ε in the ζ was determined. The three images in Figure 3 are the processing results of waterfall images under different ε values. Through the combined results of peak signal to noise ratio and subjective vision, $\varepsilon = 0.8\tau$ was finally selected, where τ is the threshold value of the difference in the cross-based pixel.

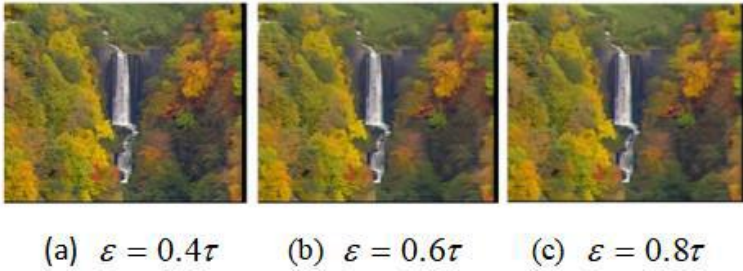


Figure 3. Effect of different coefficients of guided filtering.

Standard sequences such as Akiyo and Foreman were selected for experimental simulation to compare the advantages and disadvantages of the proposed algorithm and the original guided filtering algorithm.

Figure 4 is the 20th frame of the Akiyo sequence. Random noise with a magnitude of 0.05 was added. Bilateral filtering, guided filtering, and cross-based adaptive guided filtering proposed in this study were used to reduce the noise. Figure 4a and 4d are original images of noise reduction, and Figure 4e to 4h are corresponding enlarged views. Figures 4a and 4e are images with random noise added, Figure 4b and 4f are images processed with bilateral filtering on the noise map, and Figure 4c and 4g are directed noise filtering with guided filtering, Figure 4d and 4h are the effect diagrams of noise reduction after cross-based adaptive guided filtering. The cross-based adaptive guided filtering had better filtering effects at the edges and had obvious advantages in terms of detail processing.

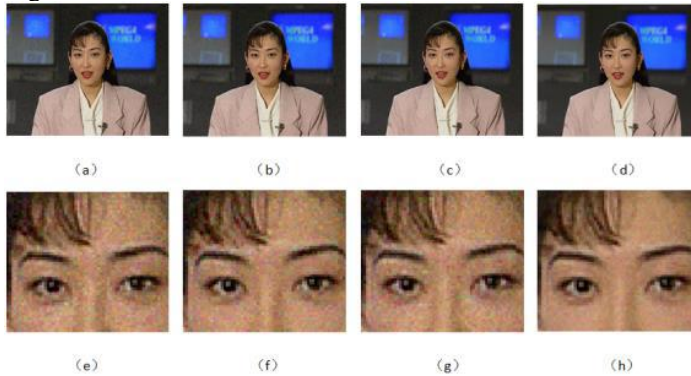


Figure 4. Denoising comparison of Akiyo.

Figure 5 is an effect diagram of noise reduction processing on a night-time real shot image. From left to right are an original image, a bilateral filter image, and a cross-based improvement guided filter image. At night, the noise of the image is large due to the insufficient exposure, which causes problems such as large color spots and blurred borders. The figure shows that the bilateral filtering and the improved cross-based guided filtering in this study both had a better smoothing effect and a better effect on removing speckles, but the over-smoothing of the bilateral filtering caused the distortion of the picture. The detailed enlarged view of Figure 5(d-f) reveals that the smoothing effect of bilateral filtering blurred the boundary at the texture and edges, while the cross-based improved guided filtering still had clear boundaries.

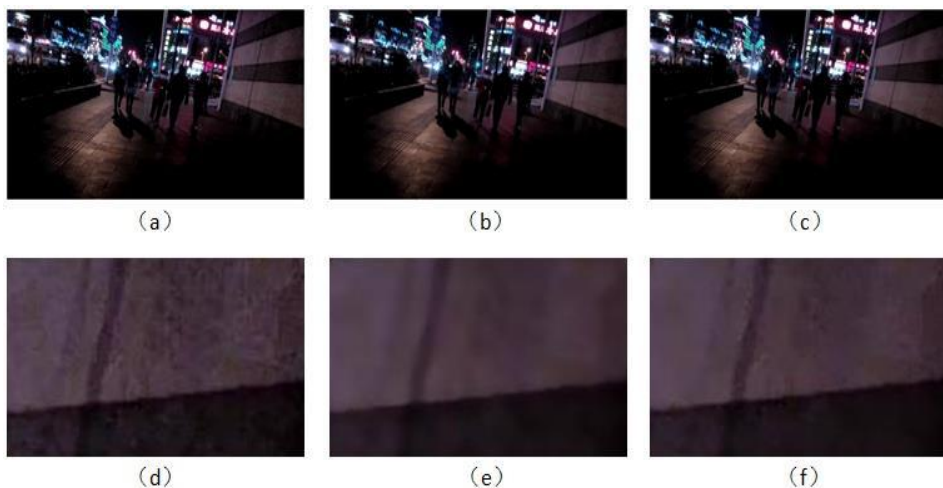


Figure 5. Denoising process on the picture of real objects.

Table 1 shows the results of the peak signal-to-noise ratio, and the structural similarity of standard sequence noise reduction based on previous studies [1, 6, 18] and the cross-based adaptive guided filtering algorithm.

Table 1. PSNR and SSIM comparison of four methods

	GF [6]	BF [1]	ABF [7]	Proposed method
PSNR comparison (dB)				
Akvio	27.0561	27.105	28.4098	29.3644
Foreman	26.7908	26.8976	27.549	28.7079
Waterfall	24.9809	25.4488	25.9767	27.6231
Bus	28.4903	29.099	29.9811	30.9946
Mobile	28.3293	28.0561	29.3002	30.724
Bottle	25.3579	25.6792	26.0593	27.4684
Flower	26.1684	26.5467	27.4918	28.3674
Luggage	27.6492	27.7843	28.8139	28.4627
Tempete	31.0575	31.1183	32.0243	33.0527
Bridge-close	26.8309	26.8976	27.9832	28.2753
SSIM comparison				
Akvio	0.827	0.835	0.842	0.860
Foreman	0.834	0.841	0.856	0.873
Qaterfall	0.825	0.834	0.844	0.865
Bus	0.838	0.846	0.862	0.884
Mobile	0.819	0.836	0.854	0.868
Bottle	0.805	0.826	0.848	0.861
Flower	0.821	0.837	0.853	0.875
Luggage	0.839	0.854	0.876	0.902
Tempete	0.828	0.842	0.865	0.887
Bridge-close	0.827	0.840	0.849	0.868

Compared with GF, BF, and ABF, the proposed method achieved better performance in terms of PSNR and SSIM when denoising the image. At the same time, the proposed algorithm still retained the advantages of guided filtering in computing time and implementation complexity.

5. Conclusions

This study proposed a cross-based adaptive guided filter. In this method, the adaptive selection of the filtering neighborhood was added to the framework of the guided filter. Filter blocks were generated whose size could be adjusted automatically and their shape could be adjusted adaptively based on gray similarity and stereo matching, which was beneficial to the smoothing effect of the filter. The improvement of the offset was added to the processing of the guided image, which enhanced the slope of filtering at the edges, reduced the blur, and made the image sharper. The selected value of the offset changed proportionally with the gray threshold of the filter window, which had a better effect than the hard threshold. The improved guided filtering still retained the original characteristics of easy hardware implementation with low algorithm complexity, and had good prospects in engineering.

References

- [1] Tomasi C, Manduchi R. Bilateral filtering for gray and color images. In: Proceedings of the IEEE International Conference on Computer Vision. 1997 23(1): 45-78.
- [2] Buades A, Coll B and Morel J M. A review of image denoising algorithms, with a new one. Multiscale Modeling Simul. 2005 4(2):490-530.
- [3] Buades A, Coll B, Morel J M. A non-local algorithm for image denoising. IEEE Computer Society Conference on Computer Vision and Pattern Recognition. 2005 2: 60-5.
- [4] Coup'e P, Yger P, Barillot C. Fast non-local means denoising for 3D MR images. MICCAr06, Copenhagen. 2006.
- [5] Buades A, Coll B, Morel J. Nonlocal image and movie denoising. Int. J. Comp. Vision. 2008:123-139.
- [6] He K, Sun J, Tang X. Guided image filtering. International Proceeding of ECCV, 2010. 1-6, 8.
- [7] Paris S, Kornprobst P, Tumblin J and Durand F. Bilateral filtering: theory and applications. Foundations Trends Comp. Graphics Vis. 2008 4(1):1-73.
- [8] Rhemann C, Hosni A, Bleyer M, Rother C and Gelautz M. Fast cost-volume filtering for visual correspondence and beyond. In *Proc. CVPR* 2011, 1, 2, 6, 7.
- [9] Katkovnik V, Foi A, Egiazarian K, Astola J. From local kernel to nonlocal multiple-model image denoising. Int J Comp Vis, 2010 86(1):1-32.
- [10] Hosni A, Bleyer M, Rhemann C, Gelautz M and Rother C. Realtime local stereo matching using guided image filtering. IEEE International Conference on Multimedia and Expo .2011 pp. 1- 6.
- [11] De-Maeztu L, Mattoccia S, Villanueva A and Cabeza R. Linear stereo matching. International Conference on Computer Vision, 2011, pp. 1708-1715.
- [12] Rhemann C, Hosni A, Bleyer M, Rother C and Gelautz M. Fast cost-volume filtering for visual correspondence and beyond. *2011 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2011, pp. 3017-3024.
- [13] Lu J, Lafruit G and Catthoor F. Anisotropic local high-confidence voting for accurate stereo correspondence in *Proc. SPIE-IS&T Electron. Imaging*, Jan. 2008, 6812: 605822-1-605822-10.
- [14] Gong M, Yang R, Wang L and Gong M. A performance study on different cost aggregation approaches used in real-time stereo matching. Int. J. Comp. Vis., 2007 75(2):283-296.
- [15] Mattoccia S, Giardino S, Gambini A. Accurate and efficient cost aggregation strategy for stereo correspondence based on approximated joint bilateral filtering. *Computer Vision-ACCV 2009*, ed: Springer, 2010, pp. 371-380.
- [16] Wang L, Liao M, Gong M, Yang R and Nister D. High-quality realtime stereo using adaptive cost aggregation and dynamic programming. *Third International Symposium on 3D Data Processing, Visualization, and Transmission*, 2006 pp. 798-805.
- [17] Zhang K, Lu J and Lafruit G. Cross-based local stereo matching using orthogonal integral images. IEEE Trans. CSVT, July 2009. 19(7):1073-1079.
- [18] Zhang B, Allebach J P. Adaptive bilateral filter for sharpness enhancement and noise removal. IEEE Transaction on Image Process 2008 17(5), 664-678.
- [19] Scharstein D, Szeliski R. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms Int. J. Comput. Vision, May 2002 47(1): 7-42.f

- [20] Tombari F, Mattoccia S, Di Stefano L and Addimanda E. Classification and evaluation of cost aggregation methods for stereo correspondence. 2018 IEEE Conference on Computer Vision and Pattern Recognition, 2018 pp. 1–8.
- [21] Gong M, Yang R, Wang L, Gong M. A performance study on different cost aggregation approaches used in real-time stereo matching. *Int. J. Comp. Vision* 2007, 75(2):283-296.
- [22] Hirschmuller H and Scharstein D. Evaluation of cost functions for stereo matching. 2007 IEEE Conference on CVPR, 2007 pp. 1–8.

Anisotropic Diffusion with Deep Learning

Hyun-Tae CHOI¹, Yuna HAN, Dahye KIM, Seonghoon HAM, Minji KIM, Yesol PARK, and Byung-Woo HONG

Computer Science Department, Chung-Ang University, Seoul, Korea

Abstract. We propose a deep learning framework for anisotropic diffusion which is based on a complex algorithm for a single image. Our network can be applied not only to a single image but also to multiple images. Also by blurring the image, the noise in the image is reduced. But the important features of objects remain. To apply anisotropic diffusion to deep learning, we use total variation for our loss function. Also, total variation is used in image denoising pre-process.[1] With this loss, our network makes successful anisotropic diffusion images. In these images, the whole parts are blurred, but edge and important features remain. The effectiveness of the anisotropic diffusion image is shown with the classification task.

Keywords. anisotropic diffusion, total variation, classification, deep learning

1. Introduction

Many approaches to object classification make essential use of machine learning methods. To improve the accuracy, we collect larger datasets, make more complicated models, and pre-process the training, testing dataset. But from the pre-processing point of view, the noise of image is a big problem to train and test a deep learning network model. There is a lot of research to reduce noise in images.[2] Like other denoising with deep learning, we try to decrease the noise in the image by diffusing using our deep learning network. In our experiments, we define noise as not important information for the classification task. For image denoising, we make a network that executes **Anisotropic Diffusion**. Filters in anisotropic diffusion can outperform isotropic diffusion to certain applications such as denoising of highly degraded edges.[3] And anisotropic diffusion reduces noise in the image without removing significant parts of the image content, typically edges[4]. With anisotropic diffusion, we can reduce the noise of the image and blur unnecessary parts of the image except the edge. It resembles the process that creates a scale space[3] where an image generates a parameterized family of successively more and more blurred images based on the diffusion process. This anisotropic diffusion is a generalization of this diffusion process. To observe the effect of our anisotropic diffusion image result, we check the classification accuracy.

Section 2 reviews anisotropic diffusion and its traditional algorithm. Section 3 describes the training process of our network. Section 4 includes the experiment results of our experiment. Finally, section 5 finish the paper by summarizing.

¹Corresponding Author: E-mail: hyuntae@image.cau.ac.kr

2. Anisotropic diffusion

Anisotropic diffusion reduces noise in the image without removing significant parts such as the edge. To make this image, anisotropic diffusion produces a family of parameterized images. Each resulting image is a combination of original images and a filter that depends on the local content of the original image. So anisotropic diffusion is a non-linear and space-variant transformation of the original image[5]. To produce a family of parameterized images, anisotropic diffusion iterates the algorithm under the assumption that 'Inside of edge will be the same area'. This is the simple algorithm of anisotropic diffusion.

Algorithm 1

for number of training iterations **do**

1. Calculate edge gradient in four directions based on the current pixel.
 2. In each direction, generalize it using generalization constant and sum it.
 3. Output sum of the above-calculated values and weights(Lambda).
-

The traditional anisotropic diffusion algorithm works with Eq. (1).[4] In Eq. (1). ∇ denotes gradient, Δ denotes the Laplacian, div is the divergence operator and $c(x, y, t)$ is the diffusion coefficient that controls the rate of diffusion. It is usually chosen as a function of the image gradient. Because of this coefficient, anisotropic diffusion can preserve edges in the image.

$$\frac{\delta I}{\delta t} = \text{div}(c(x, y, t)\nabla I) = \nabla C \cdot \nabla I + c(x, y, t)\Delta I \quad (1)$$

But the traditional anisotropic diffusion algorithm is just for a single image and the speed is too slow. So, we apply anisotropic diffusion to deep learning framework for generalization and faster than the traditional algorithm.

3. Training

3.1. Anisotropic Diffusion Loss

To calculate the gradient of each direction, we try to calculate the total variation of anisotropic diffusion image which is the output of our network. Given an image I , the output of the anisotropic diffusion image is U . The loss function is defined as Eq. (2):

$$\text{Loss} = \|U - I\|^2 + \lambda \|\nabla U\| \quad (2)$$

$$\|\nabla U\| = \|\nabla U_x\| + \|\nabla U_y\| \quad (3)$$

λ is the regularization parameter. ∇U_x is the gradient in the x-axis direction of the anisotropic diffusion image and ∇U_y is the gradient in the y-axis direction of the anisotropic diffusion image. In this loss function, λ term makes the output image is not too blurred. If this value is too big, the output image is too blurred to recognize the objects in the image. And if this value is too small, the output image is the same as the original image. So we set this value to 0.01 through several experiments. To calculate $\|\nabla U\|$ in Eq. (2), we calculate Eq. (3). To get the value of $\|\nabla U_x\|$ and $\|\nabla U_y\|$, we use Central-Difference with Dirichlet boundary condition. With this $\|\nabla U\|$, we can compute the total variation of network output. For this output images, our network is Auto-encoder network with 4-layers for encoding, and 4-layers.[6] We use the Auto-encoder structure for decoding to force the learned representations of the input to have useful properties.[7] In this architecture, the regularization parameter is fixed to 0.01, the filter size is 3×3 , and the input, output image size is 256×256 .

3.2. Classification

To compare the accuracy of classification with the original image and anisotropic diffusion image, we train three networks with the same structure. The first network is trained with the original images and tested with the original images. The second network is trained with anisotropic diffusion images and tested with anisotropic diffusion images. And the third network is trained with CIFAR-10(Canadian Institute For Advanced Research) [8] and tested with anisotropic diffusion images of the CIFAR-10 test dataset. We use VGG-19(Visual Geometry Group)[9] network for classification network to compare the results.

4. Experiment Result

4.1. Anisotropic Diffusion Network

We train our network with the celeba dataset[10]. Celeba dataset is a collection of celebrity faces with a simple background and their faces are clear to recognize. So it is not too complicated. And face has distinct features such as eyes, nose, and mouth. So with this celeba dataset, we can compare the original image and anisotropic diffusion image to see whether important features remain after the denoising process. Figure 1 illustrates the original celeba image and its anisotropic diffusion celeba image from our anisotropic diffusion network. With anisotropic diffusion, the whole image is blurred except the edge of objects in the image. Although the image is blurred, we can still observe the eyes, nose, and mouth. So, noise in the image can be reduced with anisotropic diffusion by blurring. Also, distinct features remain, so there is no problem in recognizing the object. Additionally, to see whether the network works on more complex images, we use the BSDS dataset(Berkeley Segmentation Dataset) which is usually used for segmentation and the PASCALVOC dataset(PASCAL Visual Object Classes) which is usually used for object detection.[11][12] They are more complicated than celeba dataset. Figure 2 illustrates the result of the BSDS dataset and Figure 3 illustrates the result of the PASCALVOC dataset. With this Figure 2 and Figure3, we can see that our anisotropic diffusion network also works on a more complex image dataset. In each pair of images,

the left side is original images and the right side is anisotropic diffusion images. The anisotropic diffusion image is blurred than the original images. As we can see with BSDS and PASCALVOC anisotropic diffusion images results, the whole images are blurred but the shape of objects in images remains. Despite the data becoming more complex, the anisotropic diffusion result images still have their object's shape. Therefore, we could see that even complex images could be applied to our network.



Figure 1. In each pair of images, the left side is the original celeba image and the right side is the anisotropic diffusion celeba image.

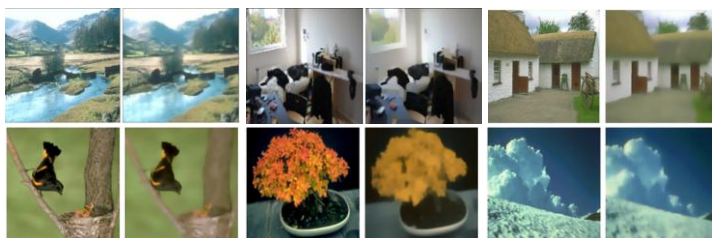


Figure 2. In each pair of images, the left side is the original BSDS image and the right side is the anisotropic diffusion BSDS image.

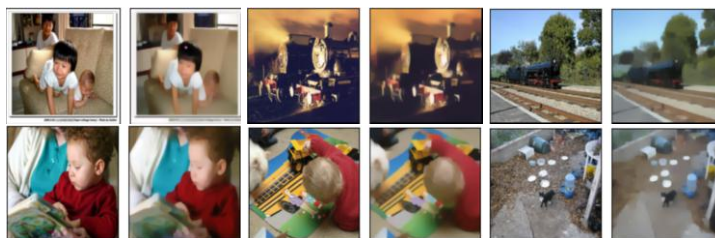


Figure 3. In each pair of images, the left side is the original PASCALVOC image and the right side is the anisotropic diffusion PASCALVOC image.

4.2. Classification

To see the anisotropic diffusion image is effective, we use the result of anisotropic diffusion for three experiments with the CIFAR-10 dataset and VGG-19 network. Table 1 illustrates the classification accuracy(%) results. With case 1 and case 2, we can see that the anisotropic diffusion image of the original is slightly more effective for classification. Also with case 1 and case 3, even the network is trained with the original image, test accuracy with anisotropic diffusion image is higher than the original image. With this result and Figure 1, anisotropic diffusion reduces noise in the image by blurring the original image but remains important features. So important information about the objects in the image is preserved by this anisotropic diffusion method.

Table 1. Train and Test accuracy(%) with CIFAR-10 and anisotropic diffusion image of CIFAR-10. Case 1 train and test with CIFAR-10. Case 2 train and test with anisotropic diffusion image. Case 3 train with CIFAR-10 and test with anisotropic diffusion image of the CIFAR-10 test set.

	Train Accuracy	Test Accuracy
Case 1	88.16	77.56
Case 2	90.24	83.27
Case 3	88.16	80.94

5. Conclusion

The traditional algorithm iterates its process with a single image.[13] So the diffusion coefficient and other parameters are suitable for only one image. And the speed of the algorithm is too slow. But with this our deep learning anisotropic diffusion network, we can apply it to other datasets not just for a single image. Also, by blurring the image, unnecessary information is reduced but the important features of the object remain. We show the effect of our result with classification accuracy using anisotropic diffusion image, that the difference from the accuracy of the original image is not so big and even better. With this result, the anisotropic diffusion image can be used to classification tasks instead of the original image. Even, although we use the VGG-19 network that is a very simple network for classification, test accuracy is slightly higher than the original image. It proves that anisotropic diffusion images still have important features and can be used for the classification task. Our result is simple, but it shows that anisotropic diffusion image can be used in other areas such as object tracking. Because object tracking does not need whole information of the image. It just needs important features that can represent the object. Our experiments show that anisotropic diffusion with deep learning reduces noise in the images but still have important features of objects.

Acknowledgement

This research was supported by the MSIT(Ministry of Science and ICT), Korea, under the National Program for Excellence in SW(20170001000041001) supervised by the IITP(Institute of Information & communications Technology Planning &Evaluation)

References

- [1] Rudin LI, Osher S, Fatemi E. Nonlinear total variation-based noise removal algorithms. 1992.
- [2] Tian CW et al. Deep learning on image denoising: An overview. arXiv preprint arXiv:1912.13171, 2019.
- [3] Weickert J Anisotropic diffusion in image processing. Stuttgart: B.G. Teubner; 1998:5-8.
- [4] Perona P and Malik J. Scale-space and edge detection using anisotropic diffusion. 1990. pp. 629-639
- [5] Joachim W. A review of nonlinear diffusion filtering. *Scale-Space Theory in Computer Vision*. Springer, LNCS 1252; 1997. pp. 1–28.
- [6] Baldi P. Autoencoders, unsupervised learning, and deep architectures. In: *Proceedings of ICML workshop on unsupervised and transfer learning*; 2012. pp.37-49.
- [7] Goodfellow I, Bengio Y, Courville A. *Deep Learning*. MIT Press; 2016. 499 p.
- [8] Krizhevsky A, Hinton G. Learning multiple layers of features from tiny images. Technical report, University of Toronto; 2009. 3 p.
- [9] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition.2014.
- [10] Liu ZW et al. Large-scale celebfaces attributes (celeba) dataset. Retrieved August, 2018, 15: 2018.
- [11] Arbelaez P, Fowlkes C, Martin D. The Berkeley segmentation dataset and benchmark. Computer Science Department, Berkeley University. 2007.
- [12] Everingham M, Eslami SMA, Van Gool L, Williams CKI, Winn J and Zisserman A. *Int J Comp Vision*, 2015, 111(1), 98-136.
- [13] Yu YJ, Acton ST. Speckle reducing anisotropic diffusion. *IEEE Transactions on image processing*, 2002, 11.11: 1260-1270.

A Clinical Decision Support Tool to Detect Invasive Ductal Carcinoma in Histopathological Images Using Support Vector Machines, Naïve-Bayes, and K-Nearest Neighbor Classifiers

Kyra Mikaela M. LOPEZ and Ma. Sheila A. MAGBOO
University of the Philippines Manila, Manila, Philippines

Abstract. This study aims to describe a model that will apply image processing and traditional machine learning techniques specifically Support Vector Machines, Naïve-Bayes, and k-Nearest Neighbors to identify whether or not a given breast histopathological image has Invasive Ductal Carcinoma (IDC). The dataset consisted of 54,811 breast cancer image patches of size 50px x 50px, consisting of 39,148 IDC negative and 15,663 IDC positive. Feature extraction was accomplished using Oriented FAST and Rotated BRIEF (ORB) descriptors. Feature scaling was performed using Min-Max Normalization while K-Means Clustering on the ORB descriptors was used to generate the visual codebook. Automatic hyperparameter tuning using Grid Search Cross Validation was implemented although it can also accept user supplied hyperparameter values for SVM, Naïve Bayes, and K-NN models should the user want to do experimentation. Aside from computing for accuracy, the AUPRC and MCC metrics were used to address the dataset imbalance. The results showed that SVM has the best overall performance, obtaining accuracy = 0.7490, AUPRC = 0.5536, and MCC = 0.2924.

Keywords. Invasive ductal carcinoma (IDC), oriented FAST and Rotated BRIEF (ORB), Support Vector Machines, Naïve-Bayes, K-Nearest Neighbors.

1. Introduction

Breast cancer is one of the most common types of cancer worldwide with over two million new cases of breast cancer diagnosed in 2018 [1]. This represents around 12.3% of the total new cancer cases that year. In the Philippines, breast cancer had the highest number of new cases in 2015, representing 19% of the overall new cancer cases in both men and women [2]. The most common subtype of breast cancer is called Invasive Ductal Carcinoma (IDC) which makes up 80% of all invasive breast cancer cases [3, 4]. At present, there is no definite main cause of breast cancer. Aside from genetics, there are still several risk factors that have been known to influence a person's susceptibility to breast cancer [5]. Therefore, the key to improving breast cancer survival is early detection and screening [6]. In most cases, IDC can manifest as micro-calcifications or thickening of breast tissues [7, 8]. For further confirmation, doctors may recommend a

breast biopsy, the only diagnostic procedure that can determine and verify the presence of cancer [9, 10].

Research has shown that the application of artificial intelligence and machine learning during diagnosis has helped further improve cancer detection and staging [11, 12]. Computer-aided diagnosis has helped with automating labor-intensive steps and reducing reader bias [13-15]. Several studies have shown that accurate breast cancer predictions may depend on the right combination of feature selection and/or ML techniques [16, 17].

2. Methodology

A literature review was conducted to determine the performance of traditional machine learning as well as deep learning approach for classification of histopathological images particularly for invasive ductal carcinoma. The dataset used in this study, the “Breast Histopathology Images”, consisted of 162 whole mount slide images of Breast Cancer specimens scanned at 40x magnification. From there, 277,524 patches of size 50 x 50 were extracted (198,738 IDC negative and 78,786 IDC positive). This dataset was originally described by Cruz-Roa [18] and is now hosted in Kaggle [19]. Due to hardware limitations, for this particular study, only 54,811 images were selected of which 48,848 was used for training and 10,963 was used for training. This new set maintained the original 28:72 or approximately 3:7 ratio of positive to negative images.

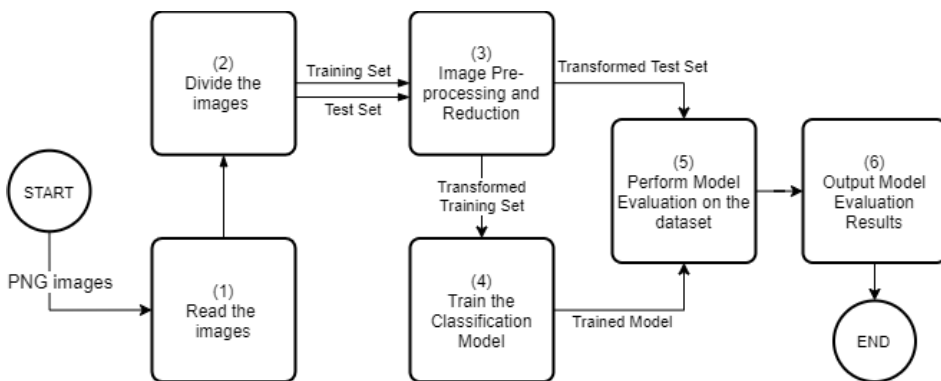


Figure 1. General Workflow

The general workflow is common to all ML applications and is illustrated in Figure 1. The images are loaded and then split into two for training and testing. The model is built after undergoing a series of pre-processing steps which includes feature extraction via ORB, feature scaling via Min-Max normalization, then clustering via K-Means over the training set. The model is then evaluated using the test set.

To train the classifier, hyperparameter values for each of the classifiers (Support Vector Machines, Naïve-Bayes, and k-Nearest Neighbors) must be specified. The user may experiment by providing these values or they may opt to use automatic hyperparameter tuning. This uses GridSearchCV, which runs a 10-fold cross validation to determine which hyperparameter values will produce the best performing estimator (classifier), ranked by the mean test score. The trained model is then saved using Joblib

to be used for image classification of new biopsy images. This process will result in a trained model and the average classification accuracy attained by that model.

The machine learning classifiers are implemented via the scikit-learn libraries: SVC for SVM classifier, MultinomialNB for Naïve-Bayes classifier, knn for the k-Nearest Neighbor classifier. The performance metrics to be used is also implemented using scikit-learn and include the accuracy, precision, recall, average precision, and Matthew's Correlation Coefficient (MCC).

The ORB features of the training and test sets were extracted on a local machine while the remaining steps (codebook generation, training, testing, and performance evaluation) were performed in Google Colaboratory.

2.1. Bag of Visual Words

The Bag-of-Visual-Words (BoVW) is a technique used in image classification [20]. It represents an image as a set of features consisting of corner points, edges, and flat regions. Since an image can have multiple features, for each image, features are extracted then a visual dictionary or bag of visual words is generated using k-means clustering [21]. This visual dictionary, represented by a collection of histograms, will be used in training machine learning algorithms in order to classify a new input image [22-24]. In this study, we extracted the features of each image in the training set using ORB (Oriented FAST and Rotated BRIEF), applied normalization on each feature and then reduced the number of features via KMeans clustering. The clusters are then collected into a visual dictionary or codebook. The generated codebook is then used to build a histogram of features for each of the training images. These histograms will serve as the input for training the classifier.

2.1.1. ORB (Oriented FAST and Rotated BRIEF) for Feature Extraction

ORB, also known as Oriented FAST Rotated BRIEF, was first presented in 2011 by Ethan Rublee et. al. [25] for computer vision tasks such as object recognition, detection, and matching. ORB was developed by OpenLabs as an open source alternative to SIFT and SURF. ORB uses FAST or Features from Accelerated Segment Test to create a sequence of images, all of which are versions of the image at different resolutions [26-28]. Next, it extracts keypoints or regions in the image which are points of interest. BRIEF or Binary Robust Independent Elementary Feature then takes all keypoints found by the FAST algorithm then converts each keypoint into a binary feature vector [29]. BRIEF uses a randomly-selected distribution of point-pairs relative to a central point to create the descriptor [25, 30]. Since BRIEF is sensitive to rotation, ORB used the rBRIEF (rotation aware BRIEF) in order to make it invariant to rotation.

2.1.2. Normalization of Extracted Features

Normalization is a feature scaling technique used to ensure that each feature contributes approximately proportionately to a measure. In this study, we used the min-max normalization scaled to unit range to linearly transform the extracted features to values ranging from 0 to 1. This will ensure that each data point will be on the same scale making each feature equally important while preserving the distribution of the original data. [31, 32]

2.1.3. Feature Reduction through K-Means Clustering then Codebook Generation

K-Means clustering is an algorithm used to find groups in data where k represents the number of groups or clusters. It is typically used as an unsupervised learning algorithm but is also commonly used as a vector quantization step in codebook generation in the BoVW Model [24]. Each cluster contains a centroid, a data point at the center of a cluster representing a multi-dimensional average of the cluster [32]. Now that the clusters are formed, the next step is to quantify and represent an image as a histogram by counting the number of times each visual word appears. This histogram is our actual bag of visual words. In this study, $k = 500$ is the number of clusters used.

2.2. Model Training

2.2.1. Support Vector Machines

Support Vector Machine or SVM is a supervised machine learning algorithm that aims to determine the optimal separating hyperplane to correctly classify the data in a given space [33, 34]. SVM has parameters that may be tuned to increase accuracy, especially if given non-linearly separable data points. These parameters include regularization parameter, gamma, and the kernel. The regularization parameter, known as the lambda (λ) parameter, represents the degree of importance that is given to misclassifications. The higher the value of λ , the smaller the max-margin and the lesser incorrect classifications are allowed. A lower λ value allows the classifier to find a larger max-margin but with a greater tendency to misclassify data points. The gamma (γ) parameter describes the degree of influence a single data point has over the decision boundary. For a higher gamma value, the closer data points to the hyperplane are considered which can help handle more complexity in data, but if it is considerably high, it may have a tendency to overfit the data. A lower gamma value considers farther data points but may lead to underfitting to the data, making less stable classification. The kernel trick makes use of a kernel function ϕ , another parameter in SVM which transforms the data into a higher dimensional feature space so that a linear separation is possible [35]. The different types of SVM kernels include: Linear kernel, Sigmoid kernel, and Radial Basis Function (RBF) kernel [36]. In this study, we used the RBF kernel with $\lambda = 1$, and gamma = 0.01 and resulting to accuracy = 0.7490, Precision = 0.6991, and Recall = 0.2135.

2.2.2. Naïve-Bayes Classifier

The Naïve-Bayes algorithm is commonly used for classification problems and is suitable for high dimensional input. Based on Bayes' Theorem of Probability [37], the goal of Naïve -Bayes is to maximize the posterior probability from the training data to formulate a decision rule for new data [38]. For variables that have categories not observable in the training set, the Naïve-Bayes model may use the alpha (α) value, also known as the additive smoothing parameter or the Laplace correction. This hyperparameter is more commonly applied in histogram steps of text classification. Certain instances that are not encountered in the training set have zero frequency, thus having zero probability. The smoothing parameter prevents the model from assigning this null probability by converting the instance count into a "pseudo-count". In this study, we used $\alpha = 1.0$ resulting to accuracy = 0.6485, Precision = 0.421, and Recall = 0.6135.

2.2.3. K-Nearest Neighbor Classifier

K-Nearest Neighbor or K-NN is another supervised learning algorithm known for its simple implementation and low calculation time. This is commonly used in statistical estimations and pattern recognition. This algorithm stores the entire training dataset, making use of all the data while classifying a new data point or instance [39].

The value of K affects the shape of the decision boundaries and is usually an odd numbered integer if the number of classes is even. A small K results in a flexible but less stable decision boundary having low bias and high variance with a tendency to overfit data. When K is relatively large, the classifier is more resilient to outliers, making smoother decision boundaries but can consequently have higher bias. Some methods like ten-fold cross-validation can be used to estimate the optimal K value. In this study K = 3 resulting to Accuracy = 0.7071, Precision = 0.2563, and Recall = 0.0131.

2.3. Performance Metrics

Although the accuracy score is reliable, it may not always be relevant to diagnosis, especially given an imbalanced dataset. Instead, the following metrics, Area Under the Precision-Recall Curve (AUPRC) or Average Precision, and Matthews Correlation Coefficient (MCC) were applied as these are the most commonly used metrics when dealing with imbalanced data [40, 41].

To know if the classifier performance is good, the performance of the random classifier must be computed first using Eq (1).

$$\text{Performance of Random Classifier} = \frac{\text{Total Positive}}{\text{Total Positive} + \text{Total Negative}} \quad (1)$$

For the dataset used in this study, the baseline performance of the random classifier is 0.2839.

Precision refers to the percentage of results that are relevant and is a good measure to determine when the cost of false-positives is high. High precision relates to the low false positive rate which is important in diagnostics so as not to subject patients without a disease to expensive and even invasive procedures. The best value for precision is 1 and the worst value is 0. Recall, on the other hand, expresses the ability to find all relevant instances in a dataset. This is commonly referred to as the true positive rate or sensitivity.

A precision-recall curve shows the relationship between precision and recall for every possible cut-off. This focuses on the minority class making it an effective measure whenever there is class imbalance [42]. The resulting score, AUPRC, also called average precision, can be used to compare performance of different classifiers. A classifier's performance is rated good if the average precision is higher than the performance of the random classifier.

The Matthew's Correlation Coefficient or MCC measures the correlation between the predicted and observed binary classification of a **sample** and can be directly computed from the confusion matrix. An MCC score of +1 describes a perfect prediction, a 0 is no better than a random prediction, and a -1 score represents a complete disagreement between prediction and outcome.

MCC is generally regarded as a balanced measure which can be used even if there is a class imbalance problem. In some studies, MCC is regarded as the most informative single score to establish the quality of a binary classifier prediction in a confusion matrix

context since its score is high only if the classifier does well on both the *negative* and the *positive* elements [41].

3. Results and Discussion

Although deep learning using convolutional neural networks and its variants are popular techniques for image classifications, there are still a number of recent studies that still use traditional machine learning techniques such as support vector machines, Naïve-Bayes, logistic regression, k-nearest neighbors, and random forest, among others.

In some studies, the performance of traditional ML specifically SVM, is comparable to the performance of deep learning approach but requiring less resources (less number of parameters to consider, less number of training samples, less number of iterations to reach convergence, effectiveness of application of active learning techniques [43-45]. The number of available high-quality annotated images, the pre-processing techniques, the feature selection methods, and the classification algorithms employed including the selection of the best hyperparameters are factors that can affect the performance of the classification model using traditional approaches. The search for the best combination of these factors for the given dataset is the challenge ML experts are working on.

Table 1. Hyperparameter Values for Each Machine Learning Classifier

Model	Hyperparameter	Value
SVM (RBF kernel)	regularization (λ)	1
	gamma (γ)	0.01
Naïve-Bayes	additive smoothing (α)	1.0
KNN	Neighbors (K)	3

Table 1 shows the hyperparameter values used in each model. The hyperparameter values chosen were obtained by performing 10-fold cross validation on 2,500 images, a separate set of images from the training and testing dataset. K-means clustering was implemented with $k = 500$.

As summarized in Table 2, the AUPRCs of both SVM and Naïve Bayes are way above the baseline of the performance of the random classifier computed as 0.2839 based from the formula in [42] with SVM as higher among the two. K-NN was the worst in all aspects, even obtaining a negative value for MCC.

Table 2. Summary of Model Performances Based on Different Evaluation Metrics

Model	Accuracy	AUPRC	Precision	Recall	MCC
SVM	0.7490	0.5536	0.6991	0.2135	0.2924
Naïve-Bayes	0.6485	0.5089	0.421	0.6135	0.2538
KNN	0.7071	0.2827	0.2563	0.0131	-0.0139

SVM's precision at 69.91% means the SVM model is good at finding relevant results. MCC is also good (29.24%) indicating that the SVM model is doing well on both the *negative* and the *positive* IDC cases.

Although a simple and traditional method has been presented, it is worth experimenting with newer methods such as [46, 47] that have shown to perform better.

4. Conclusion

This paper describes a method for Invasive Ductal Carcinoma (IDC) breast cancer classification. This is based on the Bag-of-Visual-Words (BOVW) model as a general approach and utilized the Oriented FAST and Rotated BRIEF (ORB) descriptors for feature extraction, Min-Max Normalization for feature scaling, and K-Means Clustering on the ORB descriptors to generate the visual codebook before feeding it to the SVM, Naïve-Bayes and K-Nearest Neighbor machine learning classifiers.

After evaluating the three machine learning models on various performance metrics, it was found that SVM obtained the best results in terms of accuracy (74.90%), AUPRC (55.36%), and MCC score (29.24%).

References

- [1] Worldwide cancer data Global cancer statistics for the most common cancers. [Online]. Available: <https://www.wcrf.org/dietandcancer/cancer-trends/worldwide-cancer-data>. [Accessed: 10-Jan-2020].
- [2] Laudico AV, Mirasol-Lumague MR, Medina V, Mapua CA, Valenzuela FG, Pukkala E. 2015 Philippine Cancer Facts and Estimates, 2015.
- [3] Types of breast cancer. [Online]. Available: <https://www.breastcancer.org/symptoms/types>.
- [4] Invasive ductal carcinoma (IDC). [Online]. Available: <https://www.breastcancer.org/symptoms/types/idc>.
- [5] What are the risk factors for breast cancer? [Online]. Available: https://www.cdc.gov/cancer/breast/basic_info/risk_factors.htm.
- [6] Breast cancer: prevention and control. [Online]. Available: <https://www.who.int/cancer/detection/breastcancer/en/>.
- [7] Breast cancer signs and symptoms. [Online]. Available: <https://www.acrf.com.au/support-cancer-research/types-of-cancer/breast-cancer/>.
- [8] Understanding breast calcifications. [Online]. Available: https://www.breastcancer.org/symptoms/testing/types/mammograms/mamm_show/calcifications.
- [9] Tests for diagnosing IDC. [Online]. Available: <https://www.breastcancer.org/symptoms/types/idc/tests/diagnosing>.
- [10] Breast biopsy. [Online]. Available: <https://www.cancer.org/cancer/breast-cancer/screening-tests-and-early-detection/breast-biopsy.html>.
- [11] Asri H, Mousannif H, Al Moatassime H, Noel T. Using machine learning algorithms for breast cancer risk prediction and diagnosis, *Procedia Comput. Sci.*, 2016, 83:1064–1069.
- [12] Nindrea RD, Aryandono T, Lazuardi L, Dwiprahasto I. Diagnostic accuracy of different machine learning algorithms for breast cancer risk calculation: a meta-analysis. *Asian Pac. J. Cancer Prev.*, Jul. 2018, 19(7): 1747–1752.
- [13] Dhahri H, Al Maghayreh E, Mahmood A, Elkilani W, Faisal Nagi W. Automated breast cancer diagnosis based on machine learning algorithms, *J. Healthc. Eng.*, Nov. 2019, 2019: 1–11.
- [14] Negrão de Figueiredo G, Ingrisich M, Fallenberg EM. Digital analysis in breast imaging. *Breast Care (Basel)*, Jun. 2019, 14(3): 142–150.
- [15] Kaushal C, Bhat S, Koundal D, Singla A. Recent Trends in Computer Assisted Diagnosis (CAD) system for breast cancer diagnosis using histopathological images, *IRBM*, Aug. 2019, 40(4): 211–227.
- [16] Kourou K, Exarchos TP, Exarchos KP, Karamouzis MV, Fotiadis DI. Machine learning applications in cancer prognosis and prediction. *Comput. Struct. Biotechnol. J.*, 2015, 13: 8–17.
- [17] Ferroni P, Zanzotto FM, Riondino S, Scarpato N, Guadagni F, Roselli M. Breast cancer prognosis using a machine learning approach. *Cancers (Basel)*, Mar. 2019, 11(3).
- [18] Cruz-Roa A et al. Automatic detection of invasive ductal carcinoma in whole slide images with convolutional neural networks. 2014, p. 904103.
- [19] Breast histopathology images. [Online]. Available: <https://www.kaggle.com/paultimothymooney/breast-histopathology-images>.
- [20] Magoulas GD, Prentza A. Machine learning in medical applications, *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, LNAI, 2001, 2049: 300–307.
- [21] Davida B. Bag of visual words in a nutshell. 2018. [Online]. Available: <https://towardsdatascience.com/bag-of-visual-words-in-a-nutshell-9ceea97ce0fb>.

- [22] Green K. Generating and applying a bag of visual words model for image classification, 2017. [Online]. Available: <http://www.deepcore.io/2017/04/18/generating-and-applying-a-bag-of-visual-words-model-for-image-classification/#page-content>.
- [23] Al Chanti D, Caplier A. Improving bag-of-visual-words towards effective facial expressive image classification, Proceedings of the 13th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications, 2018, pp. 145–152.
- [24] Karim AAA, Sameer RA. Image classification using bag of visual words (BoVW), Al-Nahrain J. Sci., Dec. 2018, 21(4): pp. 76–82.
- [25] Tyagi D. Introduction to ORB (Oriented FAST and Rotated BRIEF). 2019. [Online]. Available: <https://medium.com/analytics-vidhya/introduction-to-orb-oriented-fast-and-rotated-brief-4220e8ec40cf>.
- [26] Rosten E, Drummond T. Fusing points and lines for high performance tracking. Tenth IEEE International Conference on Computer Vision (ICCV'05) Volume 1, 2005, 2:1508-1515.
- [27] Viswanathan DG. Features from Accelerated Segment Test (FAST) Deepak Geetha Viswanathan 1., 2011.
- [28] Rublee E, Rabaud V, Konolige K, Bradski G. ORB: An efficient alternative to SIFT or SURF, 2011 International Conference on Computer Vision, 2011, pp. 2564–2571.
- [29] Rosten E, Porter R, Drummond T. Faster and better: a machine learning approach to corner detection, IEEE Trans. Pattern Anal. Mach. Intell., Jan. 2010, 32(1): 105–119.
- [30] Calonder M, Lepetit V, Strecha C, Fua P. BRIEF: binary robust independent elementary features, Springer Berlin Heidelberg, 2010, pp. 778–792.
- [31] KumarSingh B, Verma K, Thoke AS. Investigations on impact of feature normalization techniques on classifier performance in breast tumor classification, Int. J. Comput. Appl., Apr. 2015, 116(19): 11–15.
- [32] Kallipolitis A, Maglogiannis I. Creating visual vocabularies for the retrieval and classification of histopathology images. in 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 2019, pp. 7036–7039.
- [33] Evgeniou T, Pontil M. Support vector machines: theory and applications, 2001, pp. 249–257.
- [34] Gandhi R. Support vector machine — introduction to machine learning algorithms, 2018. [Online]. Available: <https://towardsdatascience.com/support-vector-machine-introduction-to-machine-learning-algorithms-934a444fca47>.
- [35] Schölkopf B. The kernel trick for distances. Adv. Neural Information Processing Systems. 2001.
- [36] Wang LP ed. Support vector machines: theory and applications. Springer Science & Business Media, 177: 29-45, 2005.
- [37] Joyce J. Bayes' theorem. 2003.
- [38] Karim M, Rashedur MR. Decision tree and naive bayes algorithm for classification and generation of actionable knowledge for direct marketing. 2013.
- [39] Peterson LF. K-nearest neighbor. Scholarpedia 4.2 (2009): 1883.
- [40] Magboo MSA, Coronel, A. 30-day hospital readmission prediction model for diabetic patients within the 30-70 age group. Proceedings of Academics World 130th International Conference, Madrid, Spain, 10th -11th June, 2019.
- [41] Chicco D. Ten quick tips for machine learning in computational biology. BioData Min., Dec. 2017, 10(1): 35.
- [42] Introduction to the precision-recall plot, 2017 [Online] Available: <https://classeval.wordpress.com/introduction/introduction-to-the-precision-recall-plot/>
- [43] Liu P, Choo KKR, Wang L, Huang F. SVM or deep learning? A comparative study on remote sensing image classification. Soft Comput., Dec. 2017, 21(23): 7053–7065.
- [44] Huynh BQ, Li H, Giger ML. Digital mammographic tumor classification using transfer learning from deep convolutional neural networks. J. Med. Imaging, Aug. 2016, 3(3): 034501.
- [45] Borges Sampaio W, Moraes Diniz E, Corrêa Silva A, Cardoso de Paiva A, Gattass M. Detection of masses in mammogram images using CNN, geostatistic functions and SVM. Comput. Biol. Med., Aug. 2011, 41(8): 653–664.
- [46] Izonin I, Trostianchyn A, Duriagina Z, Tkachenko R, Tepla T, Lotoshynska N. The combined use of the wiener polynomial and SVM for material classification task in medical implants production, Int. J. Intell. Syst. Appl. 2018, 10(9): 40-47. DOI: 10.5815/ijisa.2018.09.05
- [47] Bodyanskiy YE, Perova I, Vynokurova O, Izonin I. Adaptive wavelet diagnostic neuro-fuzzy system for biomedical tasks, 14th International Conference on Advanced Trends in Radioelectronics Telecommunications and Computer Engineering (TCSET), February 20–24, 2018, pp. 299-303.

Energy Disaggregation Using Principal Component Analysis Representation

Pierre V. DANTAS¹, Waldir SABINO S. Júnior² and Celso B. CARVALHO³

Federal University of Amazonas (UFAM), Amazonas, Brazil

Abstract. The main purpose of disaggregation is to decompose a signal into a set of other signals that together constitute it. This approach could be applied to audio signals, health care, home automation, ubiquitous systems and energy systems. It may be unworkable to individually measure the energy consumption of loads in a system simultaneously and, through disaggregation, we can make an inference using a main meter. The main contribution of this work is to use PCA to extract representativeness of an energy consumption signal we want to disaggregate, identifying its most relevant characteristics. The field of study is relevant because it allows information to be obtained in a simpler and cheaper way about the individual consumption of loads that make up a system. This opens up perspectives for other approaches such as smart grids and IoT. We demonstrate that when compared to other techniques, the proposal produces more accurate disaggregation results.

Keywords. energy disaggregation, principal component analysis (PCA), representativeness.

1. Introduction

Energy disaggregation consists of infer individual consumption of equipment that compose an amount of energy consumption, knowing only the latter. In this case, only one main meter is sufficient, which can make the measurement simpler and cheaper. In several situations, when we work with a database of great magnitude, may be more appropriate to work with the representativeness of a data set, which is a common condition in the area of signal processing [1].

In this work, we use *principal component analysis* (PCA) to extract representativeness from an energy consumption database and, thus, process disaggregation. We choose PCA in this work because it is widely used in multivariate analysis, used to transform a data set, with supposedly correlated variables, into a data set with uncorrelated variables and reduced dimensionality [2]. There are applications in data compression, extraction detection and pattern recognition [3, 4] and

¹Corresponding author: pierre.dantas@gmail.com

²waldirjr@ufam.edu.br

³ccarvalho_@ufam.edu.br

beyond [5, 6, 7]. PCA can be also analyzed as a projection issue where a vector in multidimensional space is projected in a vector subspace of smaller dimension, seeking to minimize the mean squared error or to maximize the variance of the projected vector. Although understanding that PCA has a higher computational cost operation, when compared to other techniques [8], we want to analyze the accuracy of disaggregation when dealing with representativeness. The increase in the accuracy of the disaggregation, in this case, may justify the use of a more complex previous processing.

We construct a new combination of initial variables in such a way that the new variables (i.e., principal components) are uncorrelated and most of the information is compressed into the first components, that is, we obtain representativeness from energy consumption database to process disaggregation. So, the idea is PCA tries to put maximum possible information in the first component, then the maximum remaining information in the second and so on. Thus, the main idea is to assume that components of the highest energy of an energy signal corresponds to a *signature* (pattern). In this work, we observed that the six first principal components contain more than 99.9% of the signal energy, considering this the criterion for delimiting the dimension of the signal decomposition.

Regarding the database, we use the public Reference Energy Disaggregation Data Set (REED) database [9], prepared by Massachusetts Institute of Technology (MIT) for research purposes in energy disaggregation.

This article is organized as follows: first we present a related work section, detailing the advantages and disadvantages of techniques used in disaggregation. Then, we present a methodology section, detailing the procedures used to disaggregate an energy consumption signal using PCA, how the data were acquired and how we disaggregate the signal. Then, we detail the experiments, followed by a results section and, in the end, the discussions and conclusions of the work.

2. Related Work

There are several approaches to deal with energy disaggregation such as analyzing transitions in time domain, factorial hidden Markov models (FHMM), neural networks and machine learning algorithms. For a comprehensive survey and more details on non-intrusive load monitoring methods and techniques for energy disaggregation problem, reader should refer to a survey presented in [10]. We present next a summary of the main approaches.

2.1. Event and State-Based Approach

The approach to the problem of disaggregation had its first work published in 1992 [11], where the transitions of energy consumption signals in the time domain were analyzed. The event-based strategies center on the transition generated by appliances and use change detection algorithm to identify start and end of an event. Simple mean [12] calculates the mean squared error between the estimated energy consumption of each appliance and the consumption measured. Powerlets [12] uses a set of electrical energy consumption data from different appliances for training and identifies energy consumption patterns. Then, created a dictionary with the identified patterns and executed the disaggregation.

2.2. Hidden Markov Model

Hidden Markov model (HMM) is a technique whose each state is represented by a probability distribution function, modeling the observation corresponding to that state. Regardless of how HMM-based approaches were used in disaggregation [13, 14], specialized a priori knowledge is required. Consequently, its presentation is constrained by the way in which the models obtained seem approximate to true. In addition, HMM-based approaches have better success in controlling multi-state appliances, but their performance degrades for multi-state and variable uncontrolled appliances.

2.3. Graph Signal Processing

Graph signal processing (GSP) is an emerging field that extends the classical theory of signal processing to general graph indexed data. GSP is a powerful, scalable, and flexible approach to signal processing suited to machine learning and data mining issues. GSP-NILM disaggregates active power dataset without prior knowledge and relies on GSP to accomplish adaptive threshold, signal clustering and pattern matching [15, 16]. It works well if the average load of each appliance is sufficiently distinct from that of the other appliance load and if the power of each load does not fluctuate much. GSP approach requires that appliances be manually labeled after disaggregation.

2.4. Deep Learning

Deep learning is an approach to machine learning that has drawn on the human brain knowledge, statistics and applied mathematics. This is the artificial neural networks (ANN) that are made up of several layers. Recently, various deep learning architectures have been applied to the issue of energy disaggregation [17, 18, 19]. Some methods have been trained using few data, but in order to be well generalized, deep learning models need a lot of data. SSI-SD methodology [18] address disaggregation using data mining techniques and elaborate clusters (dictionaries) considering that the states of operation of the appliances and the signal of main consumption are dependent instances.

3. Methodology

The first stage of the methodology consists of elaborating a dictionary formed by the combining the energy consumption of the devices and the respective main consumption. This is a supervised learning where we use around 150,000 samples of energy consumption from each device of the energy consumption database [9]. In order to apply the PCA to extract representativeness, we performed the singular value decomposition (SVD) [2] of an array arranged from the energy consumption data via Hankel transformation [12]. With this transformation, we make it possible to treat a one-dimensional signal as a two-dimensional signal, in order to apply matrix operations. At the end of this operation, we obtain a matrix of eigenvectors arranged according to the energy of their eigenvalues.

Thus, we consider the eigenvalues with the highest signal energy to apply the concept of selection of principal components, that is, those that correspond to more than

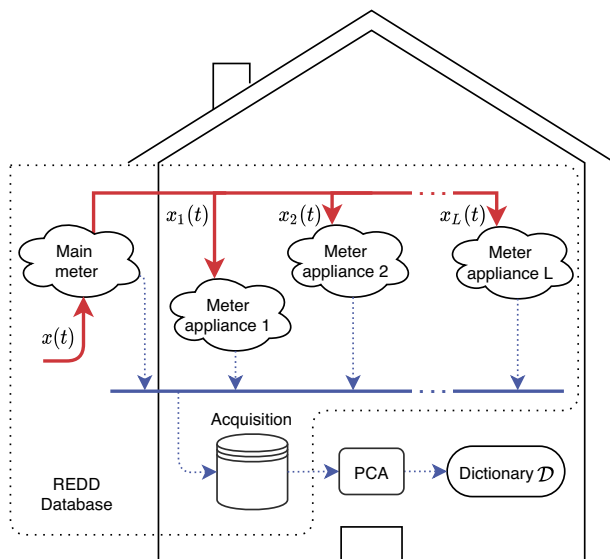


Figure 1. Framework for obtaining dictionary using PCA. The red line means electric power flow and the blue line means data flow. The blocks in the form of a cloud mean that we have a set of an appliance associated with an energy consumption meter.

99.9% of the total energy. Thus, we developed a dictionary composed of the eigenvalues and eigenvectors with the highest signal energy. We illustrate this situation in Figure 1.

After the training stage, we perform the second step of the signal disaggregation. At this point, we applied the PCA to a set of samples different from those selected for the training stage, in this case, considered as test samples, also around 150,000 samples. In this case, we only use the main electricity consumption data. We performed the PCA for each sample in the test set, obtaining a representation that allows comparison with the dictionary content obtained in the first step. Thus, we try to find, in the dictionary, the most similar set of this representation, called *matching* procedure [20], that consists on calculating dissimilarity, inferring about the solution of the disaggregation. With this, we estimate the most likely disaggregation set of energy appliances consumption. The estimated energy consumption signals of the L appliances are designated by $\hat{x}_{\{1,2,\dots,L\}}(t)$. We illustrate this situation in Figure 2.

4. Experiments

To evaluate the generalization of proposed method, we used *holdout validation*. We split up the data set into a *training* and *test* set. The training set is what the model is trained on, and the test set is used to see how well that model performs on unseen data. We divide the total data set in the ratio of 50% for training and 50% for testing. We decided to work with this proportion because we understand that, given the nature of the energy consumption signal, the number of samples is sufficient to allow a training stage with this proportion.

In the first stage of disaggregation, which consists in the elaboration of the dictionary, we use 150,000 samples of the main electrical energy consumption signal

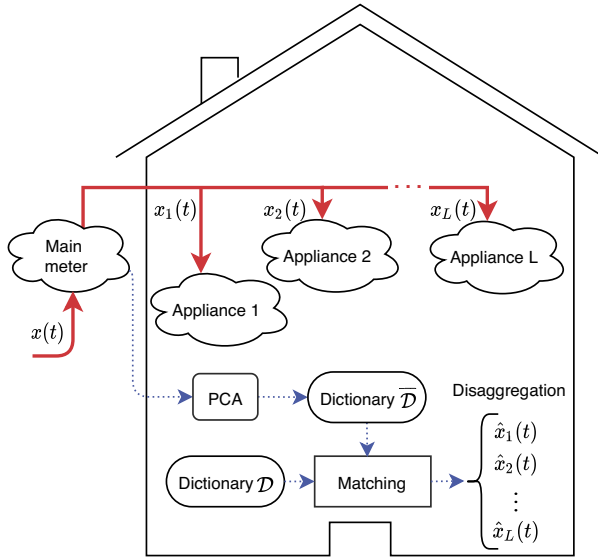


Figure 2. Framework for disaggregation of energy signal using PCA. The red line means electric power flow and the blue line means data flow. The blocks in the form of a cloud mean that we have an appliance.

and the same quantity for the electrical energy consumption signal of the appliances. Each signal was divided into 15,000 segments of dimension $w = 10$ to allow the application of Hankel's transformation. We apply this transformation of dimension $\ell = 5$. Then we apply singular value decomposition (SVD). We consider using the six principal components on PCA because we observed that this amount is sufficient to represent at least 99.9% of the signal energy. Then, for each segment of energy consumption signals, we select the first principal component and we elaborate the dictionary.

In the second stage of the disaggregation, we use 150,000 samples of a main electrical energy consumption signal only, divided into 15,000 segments of dimension $w = 10$, and we extracted the dictionary. Then, for each principal component, we look for a similarity in dictionary obtained in training step. Then, we recover the electricity consumption signals of the appliances, obtaining the estimates $\hat{x}_{\{1,2,\dots,L\}}(t)$.

This procedure was developed in MATLAB, after importing the energy consumption database. For the experiments, we used a computer with a 1.7GHz Intel Core i5 processor, 2 cores, 4GB memory, 1,600MHz DDR3, macOS Sierra operating system version 10.12.5.

5. Results

In Table 1, we present the accuracy of disaggregation obtained with the application of PCA. Note that the accuracy obtained is greater than the other methodologies used in comparison. We were careful to compare the results with other methodologies that also used the REDD database to implement the disaggregation. Likewise, the metric for calculating the disaggregation is the same in all works. Note that there are several

parameters that influence the accuracy of the disaggregation and we found that, given the same methodology, the result of the accuracy may be different.

Note that the accuracy is calculated at the end of the disaggregation of all samples in the test data set. In order to determine accuracy, it is necessary to know the true solution of the disaggregation. In this case, the database used allows this possibility, given that this database was specifically designed for these purposes. In a different case, it would not be possible to calculate the disaggregation accuracy and we would have to use the performance parameter of the test step to rate the effectiveness of the training. In these cases, it may be necessary to use unsupervised learning methods [21, 22] given that it does not require any labeled data.

Table 1. Accuracy of methodologies of energy disaggregation.

Methodology	Accuracy (%)
PCA	97,2
SSI-SD	87,3
Powerlets	79,0
FHMM	59,6
Simple Mean	39,0

6. Conclusion

A new approach for disaggregating energy consumption signals by a non-intrusive method is proposed. The estimate of the individual energy consumption of devices in a household, given the knowledge of general consumption only, is based on the extraction of representativeness, the elaboration of a dictionary and the search for a solution of disaggregation by metrics of dissimilarity analysis. The approach is based on the application of principal component analysis (PCA) to develop dictionaries that represent the operating states of the devices that make up the general energy consumption. This allows the elaboration of the dictionary based on the analysis of the magnitude of the signal energy.

The proposed algorithm was tested using a data set specifically made for the purpose of studies on the disaggregation. The result showed better performance, on average, for other methodologies that used the same database and the same accuracy calculation metrics. Emphasizing that this estimation method is supervised. The advantages of the proposed algorithm include obtaining greater disaggregation accuracy at the expense of a higher computational cost, given the required matrix decomposition operations. This method is, therefore, a new highly accurate disaggregation method.

7. Acknowledgment

This research, as provided for in Article 48 of Decree No. 6008/2006, was funded by Samsung Electronics of Amazonia LTDA, under the terms of Federal Law No. 8387/1991, through agreement No. 004, signed with the Center for R&D in Electronics and Information from the Federal University of Amazonas - CETELI/UFAM.

References

- [1] Elhamifar E, Sapiro G, Sastry SS. Dissimilarity-based sparse subset selection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2016;38(11):2182–2197.
- [2] Syms C. Principal components analysis. In: *Encyclopedia of Ecology*. Elsevier; 2018. p. 566–573.
- [3] Harris T, Yuan H. Filtering and frequency interpretations of singular spectrum analysis. *Physica D: Nonlinear Phenomena*. 2010;239(20-22):1958–1967.
- [4] Roy Chowdhury M, Tripathi S, De S. Adaptive Multivariate Data Compression in Smart Metering Internet of Things. *IEEE Transactions on Industrial Informatics*. 2020;PP:1–1.
- [5] Qi J, Jiang G, Li G, Sun Y, Tao B. Surface EMG hand gesture recognition system based on PCA and GRNN. *Neural Computing and Applications*. 2019;p. 6343–6351.
- [6] Jafarzadegan M, Safi F, Beheshti Z. Combining Hierarchical Clustering approaches using the PCA Method. *Expert Systems with Applications*. 2019;137.
- [7] Ait-Sahalia Y, Xiu D. Principal Component Analysis of High-Frequency Data. *Journal of the American Statistical Association*. 2017;114(525):287–303.
- [8] Tkachenko R, Izonin I. Model and Principles for the Implementation of Neural-Like Structures Based on Geometric Data Transformations. In: *Advances in Computer Science for Engineering and Education*. Springer; 2019. p. 578–587.
- [9] Kolter J, Johnson M. REDD: A Public Data Set for Energy Disaggregation Research. *Artif Intell*. 2011;25.
- [10] Faustine A, Mvungi N, Kaijage S, Kisangiri M. A Survey on Non-Intrusive Load Monitoring Methodies and Techniques for Energy Disaggregation Problem. *CoRR*. 2017;abs/1703.00785.
- [11] Hart G. Nonintrusive Appliance Load Monitoring. *Proceedings of the IEEE*. 1992 01;80(12):1870 – 1891.
- [12] Elhamifar E, Sastry S. Energy Disaggregation via Learning ‘Powerlets’ and Sparse Coding. In: *Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence*. AAAI Press; 2015. p. 629–635.
- [13] Kolter JZ, Jaakkola T. Approximate inference in additive factorial HMMs with application to energy disaggregation. In: *Journal of Machine Learning Research*; 2012. .
- [14] Yang C, Wu Z. Research on Non-intrusive Load Decomposition Based on FHMM. *IOP Conference Series: Materials Science and Engineering*. 2020;768:62046.
- [15] Stankovic V, Liao J, Stankovic L. A graph-based signal processing approach for low-rate energy disaggregation. In: *2014 IEEE symposium on computational intelligence for engineering solutions (CIES)*. IEEE; 2014. p. 81–87.
- [16] Li D, Dick S. A graph-based semi-supervised learning approach towards household energy disaggregation. In: *2017 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*. IEEE; 2017. p. 1–7.
- [17] Shin C, Rho S, Lee H, Rhee W. Data requirements for applying machine learning to energy disaggregation. *Energies*. 2019;12:201–213.
- [18] Dantas P, Junior W. Energy Disaggregation via Data Mining. In: *Brazilian Technology Symposium*. Springer; 2019. p. 541–546.
- [19] Kelly J, Knottenbelt W. Neural NILM: Deep neural networks applied to energy disaggregation. In: *BuildSys 2015 - Proceedings of the 2nd ACM International Conference on Embedded Systems for Energy-Efficient Built*; 2015. .
- [20] Hakak S, Kamsin A, Palaiahnakote S, Gilkar G, Khan W, Imran M. Exact String Matching Algorithms: Survey, Issues, and Future Research Directions. *IEEE Access*. 2019;PP:1–1.
- [21] Holweger J, Dorokhova M, Bloch L, Ballif C, Wyrsh N. Unsupervised algorithm for disaggregating low-sampling-rate electricity consumption of households. *Sustainable Energy, Grids and Networks*. 2019;19:100244.
- [22] Hosseini S, Kelouwani S, Agbossou K, Cardenas A, Henao NF. Adaptive on-line unsupervised appliance modeling for autonomous household database construction. *International Journal of Electrical Power & Energy Systems*. 2019;112:156–168.

Research on Optimization Method of Tool Path in Five-Axis Process Singular Region

Yujie WANG¹, Xin SHEN, Yu PENG and Lixin ZHAO

CAC Chengdu Aircraft Industrial (Group) Co. Ltd., Chengdu 610092, China

Abstract. For the five-axis machine into the singular region in the process of parts processing, resulting in a discontinuous and rapid rotation of the axis of rotation of large angles. Based on the analysis of the cause of the obvious ripple on the machined surface and the influence on the machining precision, a mathematical model of the singular region is established, and an optimization method of the tool path in the singular region is proposed. The simulation and practical machining results show that the method can effectively overcome the problem of excessive movement of the rotating shaft in the singular region of 5-axis machine tool, and solve the surface corrugated defects caused by the problem, while improving the processing efficiency.

Keywords. Digital control process, Singular region, Tool path optimization

1. Introduction

The shape of aircraft structural parts is related to the aerodynamic layout of the aircraft, and the machining quality of the shape is closely related to the aerodynamic performance. Therefore, most of the contour surfaces of aircraft structural parts are curved, which need five-axis NC machine tools to complete.

The five-axis machining adds two rotating shafts to the three-axis machining, which makes the machining mode more flexible, the material removal rate is higher, the processing time is shorter, and the more complex parts can be dealt with. Therefore, five-axis machining has been widely used in aviation, aerospace, automobile, ship and other industrial fields. But the motion of rotation axis also makes the attitude control of tool more complex, which introduces many special problems of five axis machining. The singular point problem is an important one. When the tool passes through the region near the singular point, the rotation axis will produce discontinuous and rapid rotation, which greatly increases the non-linear error, and it is easy to destroy the workpiece, and even damage the machine parts.

Because of the existence of singular region, when the five-axis NC machining is carried out, it is easy to appear in a very short cutting length, the rotating axis rotates rapidly and discontinuously, and the variation is very large, which will produce obvious ripples on the machined surface, resulting in a great increase in the amount of grinding and grinding difficulty of the subsequent fitters, which seriously affects the

¹ Corresponding Author: Yujie Wang, CAC Chengdu Aircraft Industrial (Group) Co., Ltd., Chengdu China; E-mail: wyj410245590@163.com.

machining quality of the parts and reduces the machining efficiency of the parts [1], as shown in Figure 1.

Therefore, it is very important to optimize the tool path in the singular region for improving the machining accuracy and efficiency [2]. In this paper, the cutter axis vector in the singular region is taken as the research object to solve the problem that the motion of rotation axis in the singular region is too large. In order to improve the machining accuracy and efficiency of five axis machine tool, the relevant research is carried out by means of theoretical analysis, mathematical calculation, computer programming and experimental verification. Finally, the tool path optimization software is formed, and the optimization method is verified by simulation analysis and trial cutting. The processing quality and efficiency are improved synchronously.

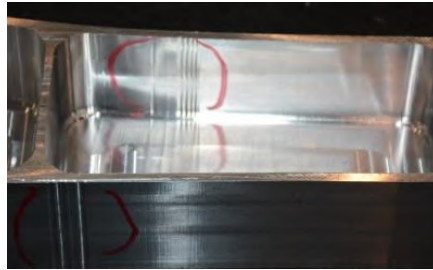


Figure 1. Surface ripple caused by excessive movement of the axis of rotation.

2. Theoretical study on the optimization method of tool path in singular region

Singular region is characterized by very large changes in the axis of rotation when the tool-axis vector changes very little one-dimensional angle [3]. Therefore, through the establishment of singular region tool path optimization mathematical model, analysis of the surface profile error as a guide to calculate the tool axis variation tolerance, The optimization vector is found in the tolerance region to minimize the angle between the starting vector and the target vector on the projection plane, so that the motion of the rotation axis in the singular region is minimized.

2.1. Singularities in 5 Axis Machine Tools

In 5-axis NC machining the so-called singular problem arises when there is inverse kinematics singular point in the machining space [4]. The inverse kinematics problem of machine tool can be expressed as solving the coordinates of each moving axis of the machine tool under the condition of known coordinate value of tool tip in workpiece coordinate system and vector direction of cutter axis. Take the AC double turntable five-axis machine tool as an example. The solution of A and C angles is obtained from the inverse kinematical transformation as follows [5]:

$$\begin{cases} A = k_A \cdot \alpha \cos(k), k_A = \pm 1 \\ C = -\arctan\left(\frac{i}{j}\right) + k_C \cdot \pi, k_C = 0,1,2 \end{cases} \quad (1)$$

where i, j, k is tool vector, A is in $[-90^\circ, 90^\circ]$, C is in $[0^\circ, 360^\circ]$. The value of A, C in Table 1 and Table 2 based on Eq. (1).

Table 1. AC corner numerical value($A \in [-90^\circ, 0^\circ]$).

$A \in [-90^\circ, 0^\circ]$	i	>0	<0	<0	>0	$=0$
	j	>0	>0	<0	<0	>0
	C	$2\pi - \arctan(i/j)$	$-\arctan(i/j)$	$\pi - \arctan(i/j)$	$\pi - \arctan(i/j)$	2π
	i	$=0$	>0	<0	$=0$	
	j	<0	$=0$	$=0$	$=0$	
	C	π	$3\pi/2$	$\pi/2$	Can't be sure.	

Table 2. AC corner numerical value($A \in (0^\circ, 90^\circ]$).

$A \in (0^\circ, 90^\circ]$	i	>0	<0	<0	>0	$=0$
	j	>0	>0	<0	<0	>0
	C	$\pi/2 - \arctan(i/j)$	$\pi - \arctan(i/j)$	$2\pi - \arctan(i/j)$	$-\arctan(i/j)$	π
	i	$=0$	>0	<0	$=0$	
	j	<0	$=0$	$=0$	$=0$	
	C	2π	$\pi/2$	$3\pi/2$	Can't be sure.	

From the Table 1, it can be seen that there is no singular problem in the solution of angle A , but when the solution of angle c is in $i=0$ and $j=0$, the solution of $c = \pi - \arctan(0/0)$ or $C = -\arctan(0/0)$ has no solution, so it is impossible to determine the value of c , that is to say, it is singular point. At this time, the turntable of the machine tool and the tool shaft are vertical, no matter what the value of the C angle does not affect the direction of the tool axis of the point, that is, the c axis can swing at any angle to obtain the tool axis vector at the singular point position, so for the AC double turntable five-axis machine, the point of the tool axis vector is the singularity point, and the normal axis of the C turntable is the singular axis.

In fact, When the angle between the tool axis vector and the singular axis is less than a certain angle, the angle of the rotation axis will change more and more, resulting in large error [6]. This angle is singular value, and the area formed by it is a cone area in space, as shown in the Figure 2.

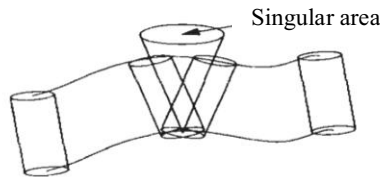


Figure 2. Schematic diagram of singular regions.

2.2. Mathematical Modeling of Singular Regions

According to the definition of singular region, the geometric model of the singular region is established, as shown in Figure 3.

In Figure 3, V_1 and V_2 represent the starting vector and target vector of tool axis motion, respectively, the angle between the starting vector and the target vector on the

projection plane, O for the origin, V_p for the polar axis, V_p as the axis, V_1 and V_2 vectors as the busbar as the conical surface, and projection to the plane as circle A.

When the angle between V_1 , V_2 and polar axis V_p is very small, even if the angle between V_1 and V_2 is very small, But the angle $\Delta\beta$ of their projection on the plane of the polar axis V_p as the normal vector may be very large. This is the reason why the tool axis vector on the same surface changes very little, but the machine tool rotation axis movement still needs to swing substantially.

Secondly, after understanding the problems existing in the singular region, in order to reduce the movement of the rotation axis, the tool path in the singular region must be optimized. The optimization principle is to establish the mathematical model of the tool path optimization, as shown in Figure 4. The results show that the variation tolerance of the tool axis is calculated by using the machining surface profile error as the guide, and the optimal vector is found in the allowable region to minimize the $\Delta\beta$, so as to minimize the movement of the rotation axis in the singular region.

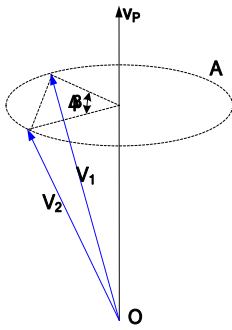


Figure 3. The geometric model of singular region.

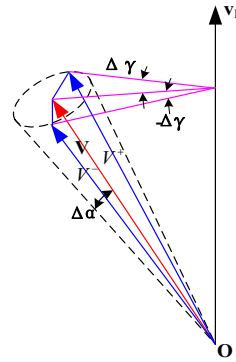


Figure 4. Geometric model of tool path optimization.

In Figure 4, the vector V to be optimized is used as the center line of rotation, the origin O is taken as the item point, and the tolerance $\Delta\alpha$ as the half-top angle as the conical face. The angles between the bus and the vector V of the conical surface is $\Delta\alpha$. The super polar axis V_p does two planes tangent to the conical surface, and the tangent vectors are represented by vectors V^- and V^+ .

The projection angles of V^- and V^+ on the plane where the polar V_p is the normal vector are $\Delta\gamma$ and $-\Delta\gamma$, respectively. V and V^-, V^+ make plane p' and p'' respectively. The previous tool axis vector is used as the optimization objective vector of the current tool axis, and the included angle θ of the projection between the plane where the polar axis V_p is the normal vector is calculated.

Assuming that V is the vector optimized by the cutter axis, it can be divided into four cases by geometric judgment

Case 1: $\theta < -\Delta\gamma$ is $V_y = V^-$;

Case 2: $\theta > \Delta\gamma$ is $V_y = V^+$;

Case 3: $0 < \theta < \Delta\gamma$, V_y is p'' with the intersection of the target vector V_{TARGET} and the plane of the polar axis.

Case 4: $-\Delta\gamma < \theta < 0$, V_y is p' with the intersection of the target vector V_{TARGET} and the plane of the polar axis.

2.3. Tool Path Optimization Method

Generally, the motion path of five axis CNC machine tool (tool path) is obtained by CAM software, and the running NC program can be recognized by post-processing software. We can optimize the tool path through post-processing software to solve the problem of singular region. According to the mathematical model of tool path optimization in singular region, the flow diagram of tool path optimization is shown in Figure 5.

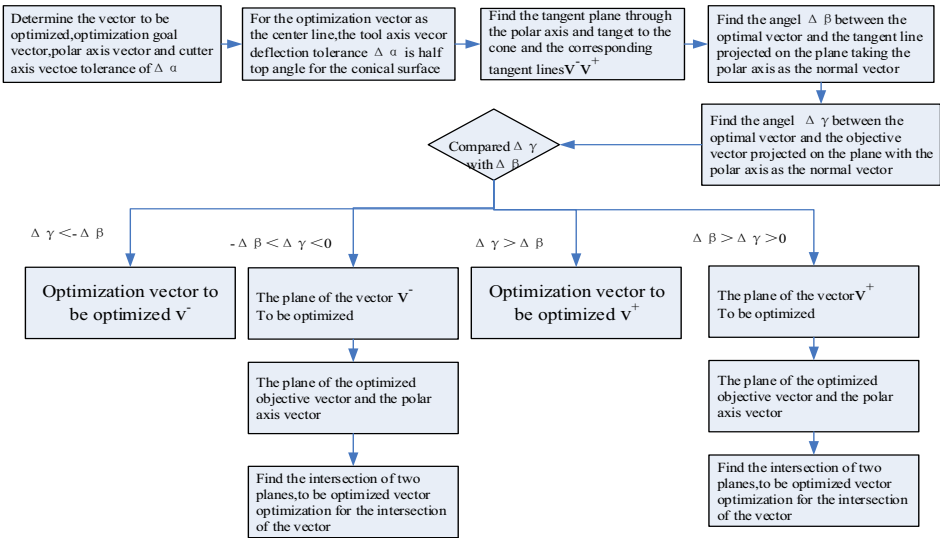


Figure 5. the flow diagram of tool path optimization.

Know: Optimize tool vector $v_1 = (i_1 \ j_1 \ k_1)$, optimizing vector $v = (i \ j \ k)$, axis vector $v_p = (i_p \ j_p \ k_p)$, tolerance $\Delta\alpha$

Find: Optimize the cutter axis vector $v_y = (i' \ j' \ k')$, make the angle between the tool axis vector and the vector to be optimized $(v_p \ v) < \Delta\alpha$, and projected on the plane with the polar vector as the normal vector of the objective optimization vector $Ang(v_y \ v_1 \ v_p)$.

By the above analysis, the model can be reduced to a three-dimensional quadratic system of equations.

$$\begin{cases} \begin{pmatrix} i \\ j \\ k \end{pmatrix} \cdot \begin{pmatrix} i' \\ j' \\ k' \end{pmatrix} = \cos \alpha \\ \begin{pmatrix} i \\ j \\ k \end{pmatrix} \times \begin{pmatrix} i' \\ j' \\ k' \end{pmatrix} \cdot \begin{pmatrix} i_p \\ j_p \\ k_p \end{pmatrix} \times \begin{pmatrix} i' \\ j' \\ k' \end{pmatrix} = 0 \\ \begin{pmatrix} i' \\ j' \\ k' \end{pmatrix} = 1 \end{cases} \quad (2)$$

Put Eq. (2) simplify

$$\begin{cases} ii' + jj' + kk' = \cos \alpha \\ (jk' - kj')i_p + (ki' - ik')j_p + (ij' - ji')k_p = 0 \\ i'^2 + j'^2 + k'^2 = 1 \end{cases} \quad (3)$$

From the Eq. (3) the optimized polar axis of the cutter axis vector can be obtained by the mathematical calculation $v_y = (i' \ j' \ k')$.

3. Software Development and Software Testing of Tool Path Optimization in Singular Region

In order to realize the automatic optimization of tool path, reduce the difficulty of calculation, improve the working efficiency, and facilitate the engineering application, this paper develops the tool path optimization processing software based on VB.NET language, and verifies that the function of the software meets the requirements of tool path optimization by optimizing the preprogram of typical parts.

3.1. Software Development

The software input is the tool path file processed by CAM. The tool axis vector in the file is identified line by line, and the tool axis vector (target vector) meeting the optimization conditions is identified as 0. Then read the tool axis vector marked as 0 for optimization. The tool axis vector after optimization is judged. If the tool axis vector meets the requirements of optimization objectives, the tool axis vector is identified as 1. If it does not meet the requirements of optimization objectives, the optimization cycle will be carried out until the requirements of optimization objectives are met. The software running framework is shown in the Figure 6.

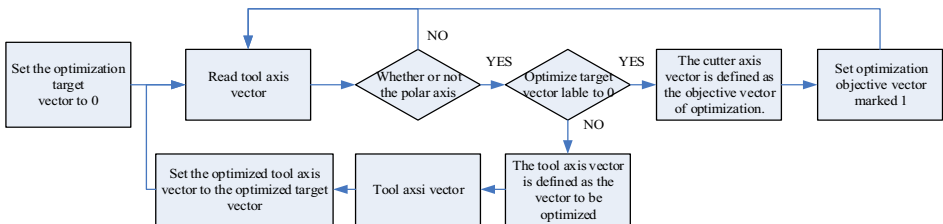


Figure 6. Block diagram of software.

3.2. Software Testing

After the completion of the software, the ability test is carried out to test whether the software can effectively optimize the cutter axis vector, reduce the amount of rotation axis movement, and realize the tool of tool-path optimization method.

The "S" specimen is taken as the specimen for testing the machining accuracy of the machine tool in five axes. Its structure is shown in Figure 7. In the label area, there is a typical singular region problem, so we choose "S" pre-program to optimize the test.

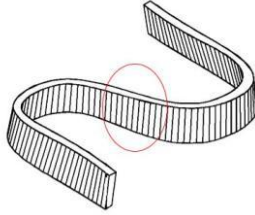


Figure 7. The schematic diagram of "S" specimen.

NC program of singular region before and after optimization are shown in Figure 8. Before the optimization, C angle enters and leaves from the singular region of s specimen, and then decreases from 113° to -34°. The C angle is transported in a short time. After optimization, the C angle remained almost unchanged at -230°, effectively reducing the C angle movement.

N1011 X0.335 Y38.697 Z30.256 A-6.995 C-229.351	N1011 X0.482 Y38.382 Z30.211 A-6.709 C113.982
N1012 X0.865 Y37.351 Z30.227 A-6.794 C-229.654	N1012 X1.009 Y37.034 Z30.182 A-6.509 C113.174
N1013 X1.432 Y35.871 Z30.194 A-6.565 C-229.935	N1013 X1.574 Y35.553 Z30.149 A-6.283 C112.278
N1014 X1.954 Y34.477 Z30.161 A-6.345 C-230.162	N1014 X2.092 Y34.158 Z30.117 A-6.064 C111.420
N1015 X2.599 Y32.701 Z30.118 A-6.055 C-230.378	N1015 X2.734 Y32.381 Z30.075 A-5.779 C110.296
N1016 X3.209 Y30.977 Z30.075 A-5.766 C-230.517	N1016 X3.341 Y30.655 Z30.032 A-5.497 C109.159
N1017 X3.783 Y29.311 Z30.031 A-5.479 C-230.582	N1017 X3.913 Y28.988 Z29.990 A-5.217 C107.998
N1018 X4.288 Y27.811 Z29.991 A-5.218 C-230.585	N1018 X4.416 Y27.487 Z29.951 A-4.962 C106.887
N1019 X4.847 Y26.116 Z29.944 A-4.913	N1019 X4.973 Y25.793 Z29.906 A-4.670 C105.535
N1020 X5.309 Y24.691 Z29.904 A-4.653	N1020 X5.432 Y24.368 Z29.867 A-4.422 C104.297
N1021 X5.864 Y22.947 Z29.852 A-4.327	N1021 X5.983 Y22.627 Z29.819 A-4.117 C102.625
N1022 X6.409 Y21.200 Z29.799 A-3.995	N1022 X6.525 Y20.883 Z29.770 A-3.810 C100.729
N1023 X6.940 Y19.470 Z29.744 A-3.659	N1023 X7.053 Y19.157 Z29.720 A-3.506 C98.567
N1024 X7.435 Y17.836 Z29.691 A-3.338	N1024 X7.545 Y17.527 Z29.672 A-3.221 C96.188
N1025 X7.890 Y16.315 Z29.640 A-3.036	N1025 X7.997 Y16.011 Z29.627 A-2.959 C93.592
N1026 X8.379 Y14.664 Z29.582 A-2.704	N1026 X8.482 Y14.367 Z29.578 A-2.679 C90.234
N1027 X8.799 Y13.232 Z29.530 A-2.415	N1027 X8.899 Y12.940 Z29.535 A-2.445 C86.727
N1028 X9.301 Y11.504 Z29.465 A-2.063	N1028 X9.397 Y11.220 Z29.483 A-2.174 C81.536
N1029 X9.801 Y9.771 Z29.396 A-1.707	N1029 X9.894 Y9.494 Z29.430 A-1.923 C74.929
N1030 X10.290 Y8.063 Z29.323 A-1.355	N1030 X10.380 Y7.794 Z29.378 A-1.707 C66.650
N1031 X10.783 Y6.338 Z29.242 A-0.998	N1031 X10.870 Y6.077 Z29.327 A-1.533 C56.166
N1032 X11.305 Y4.501 Z29.146 A-0.618	N1032 X11.389 Y4.250 Z29.273 A-1.420 C42.715
N1033 X11.850 Y2.580 Z29.033 A-0.220	N1033 X11.931 Y2.339 Z29.218 A-1.394 C27.216
N1034 X12.394 Y0.760 Z28.937 A0.000	N1034 X12.451 Y0.511 Z29.167 A-1.463 C12.928
N1035 X12.933 Y-1.141 Z28.924 A0.401	N1035 X12.990 Y-1.378 Z29.116 A-1.619 C0.320
N1036 X13.457 Y-2.977 Z28.903 A0.782	N1036 X13.513 Y-3.204 Z29.068 A-1.830 C-9.301
N1037 X13.997 Y-4.854 Z28.875 A1.168	N1037 X14.051 Y-5.071 Z29.019 A-2.085 C-16.827
N1038 X14.528 Y-6.687 Z28.843 A1.542	N1038 X14.582 Y-6.895 Z28.972 A-2.361 C-22.425
N1039 X15.071 Y-8.538 Z28.809 A1.914	N1039 X15.123 Y-8.737 Z28.926 A-2.654 C-26.760
N1040 X15.598 Y-10.377 Z28.790 A2.286	N1040 X15.664 Y-10.556 Z28.880 A-2.954 C-30.080
N1041 X16.147 Y-12.200 Z28.753 A2.746	N1041 X16.213 Y-12.372 Z28.834 A-3.258 C-32.677
N1042 X16.701 Y-14.010 Z28.714 A3.088	N1042 X16.766 Y-14.176 Z28.789 A-3.561 C-34.715
N1043 X17.263 Y-15.809 Z28.674 A3.419	N1043 X17.327 Y-15.971 Z28.744 A-3.863 C-36.319
N1044 X17.832 Y-17.597 Z28.634 A3.740	N1044 X17.896 Y-17.755 Z28.701 A-4.160 C-37.576
N1045 X18.409 Y-19.372 Z28.595 A4.051	N1045 X18.474 Y-19.528 Z28.658 A-4.453 C-38.550
N1046 X18.996 Y-21.134 Z28.555 A4.350	N1046 X19.062 Y-21.289 Z28.616 A-4.737 C-39.286
N1047 X19.593 Y-22.881 Z28.515 A4.636	N1047 X19.662 Y-23.036 Z28.575 A-5.014 C-39.817
N1048 X20.202 Y-24.611 Z28.476 A4.909	N1048 X20.273 Y-24.766 Z28.535 A-5.282 C-40.161
N1049 X20.823 Y-26.324 Z28.438 A5.168	N1049 X20.898 Y-26.486 Z28.497 A-5.542 C-40.318

Pre-optimization

Optimization

Figure 8. Comparison of Pre-optimization program and Optimization program.

Statistical optimization program C Angle travel range and C Angle total movement, the data as shown in Table 3, we can see, the range of C angle travel is -234° to 277° before optimization and -230° to -82° after optimization.

Table 3. "S" specimen simulation of the cutting results.

Contrast items.	Before optimization	After optimization
The range of C	-234°~277°	-230°~ -82°
C angle exercise	10892.916°	4669.791°

Therefore, after testing, the software can achieve tool path optimization algorithm, to reduce the rotation axis of the target to be.

4. Application and Verification of Tool-Path Optimization Method in Singular Region

The tool path optimization software is used to optimize the typical parts. Through simulation analysis and trial cutting, the machining states of the parts before and after optimization are compared to verify the optimization effect of the tool path optimization method in the singular region.

4.1. Tool Path Optimization Simulation Analysis

Select a typical part of the tool path optimization test verification, the simulation diagram of trial cutting part model formed by Vericut (A simulation software for NC machining) as shown in Figure 9.

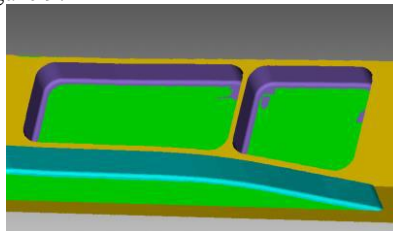


Figure 9. Simulation diagram of trial cutting part model.

The part outline program is optimized before and after the comparison, as shown in Figure 10. The tool path file processed by CAM was imported into the software for optimization. The N2172-N2185, a typical C angle with a large variation, is selected for case study.

2176 N2170 X169.336 Y41.895 Z-12.524 A3.838	2176 N2171 X169.099 Y37.287 Z-12.600 A3.237 C-182.881
2176 N2171 X169.104 Y37.287 Z-12.600 A3.237 C-181.287	2177 N2172 X168.748 Y30.311 Z-12.714 A2.323 C-182.881
2177 N2172 X168.751 Y30.311 Z-12.714 A2.323 C-181.287	2178 N2173 X168.395 Y23.294 Z-12.828 A1.401 C-182.881
2178 N2173 X168.397 Y23.294 Z-12.828 A1.401 C-181.287	2178 N2174 X168.031 Y16.061 Z-12.945 A0.451 C-182.879
2179 N2174 X168.031 Y16.061 Z-12.945 A0.451 C-181.287	2180 N2175 X168.032 Y15.638 Z-12.952 A0.405 C-172.340
2180 N2175 X168.028 Y15.637 Z-12.952 A0.405 C-181.287	2181 N2176 X168.114 Y15.223 Z-12.959 A0.362 C-157.525
2181 N2176 X168.106 Y15.221 Z-12.959 A0.362 C-181.287	2182 N2177 X168.300 Y14.843 Z-12.967 A0.314 C-155.663
2182 N2177 X168.292 Y14.842 Z-12.967 A0.314 C-181.287	2183 N2178 X168.566 Y14.515 Z-12.975 A0.281 C-161.654
2183 N2178 X168.561 Y14.514 Z-12.975 A0.281 C-181.287	2184 N2179 X168.885 Y14.237 Z-12.982 A0.229 C-160.292
2184 N2179 X168.881 Y14.236 Z-12.982 A0.229 C-181.287	2185 N2180 X169.243 Y14.010 Z-12.988 A0.115 C-133.171
2185 N2180 X169.243 Y14.010 Z-12.988 A0.000	2186 N2181 X169.637 Y13.854 Z-12.993 C-53.396
2186 N2181 X169.637 Y13.854 Z-12.993	2187 N2182 X170.057 Y13.801 Z-12.997 A0.162 C-20.737
2187 N2182 X170.054 Y13.801 Z-12.997 A-0.162 C-181.286	2188 N2183 X170.473 Y13.870 Z-12.999 A0.081 C-13.773
2188 N2183 X170.474 Y13.870 Z-12.999 A0.000	2189 N2184 X170.867 Y14.027 Z-13.000 A0.001 C0.000
2189 N2184 X170.867 Y14.027 Z-13.000	2190 N2185 X176.164 Y16.433 Z-13.000 A0.000 C-63.435
2190 N2185 X176.164 Y16.433	2191 N2186 X181.635 Y18.817
2191 N2186 X181.635 Y18.817	2192 N2187 X186.339 Y20.794
2192 N2187 X186.339 Y20.794	2193 N2188 X191.062 Y22.710
2193 N2188 X191.062 Y22.710	2194 N2189 X195.814 Y24.569
2194 N2189 X195.814 Y24.569	2195 N2190 X200.603 Y26.374
2195 N2190 X200.603 Y26.374	2196 N2191 X205.429 Y28.125
2196 N2191 X205.429 Y26.125	2197 N2192 X210.294 Y29.822

postoptimality

before optimization

Figure 10. Comparison of actual optimization.

Before optimization, the C angle of the program changes rapidly from -182.8° to C0°. After optimization, the range of C angle variation is very small, and there is no

change at all. Therefore, the optimization of C angle variation in the optimization process is basically the same as that in C -63.4. The effect is obvious as shown in Figure 11.

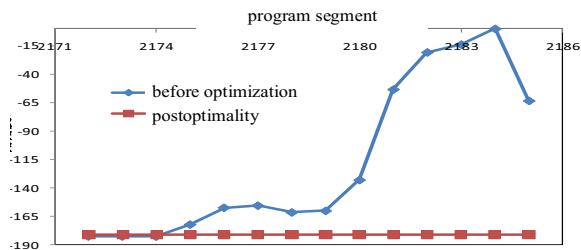


Figure 11. C angle change before and after optimization.

4.2. Verification and Application of Cutter Path Optimization Test

In order to verify the effectiveness of the tool path optimization method, some typical parts are tested and verified on the spot. The real-time machining time of the pre optimization and post optimization of the inner contour of the test cutting part is accurately collected by the DNC real-time monitoring module of the machine tool. The specific values are shown in Table 4.

Table 4. Comparison of actual cutting time before and after optimization.

	pre-process work time	after optimization work time	Shorten processing time	Improve processing efficiency
Contour finishing process	270s	187s	83s	44.38%
Inner form finishing process	221s	123s	98s	79.67%

As can be seen from Table 4, the processing time before shaper optimization is 270s. After optimization, the efficiency is 44.38%, the pre-process time is 221s, after optimization 123s, improve processing efficiency is 79.67%.

The actual cutting effect is shown in Figure 12, 13 and 14. Before the optimization of the surface of the surface are corrugated, seriously affecting the surface quality, the need for benchwork follow-up grinding, but also for the final delivery of parts of the quality of filling the hidden dangers.



Figure 12. Comparison before and after optimization.

After the optimization, the surface waviness has disappeared completely. The effect is good. Not only the machining efficiency is improved, but also the surface quality is greatly improved. It is verified that the optimization method of tool path in singular region can effectively solve the problem of workpiece surface waviness caused by oversized rotation axis of machine tool, and improve machining efficiency at the same time.



Figure 13. Surface Before Optimization

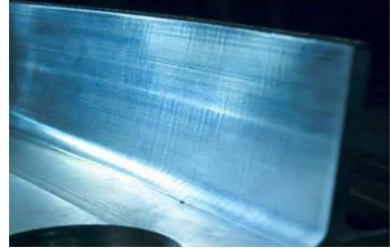


Figure 14. Surface After Optimization

5. Conclusions and Prospects

Based on this method, the tool path optimization software of singular region is developed. Through the simulation comparison and trial test, it is verified that this method can effectively solve the problem of workpiece surface waviness caused by the excessive rotation axis of the machine tool, and can also improve the machining efficiency significantly.

Looking ahead, as the new research project cycle is getting shorter and shorter, aircraft junction component quality requirements are getting higher and higher, five-axis NC machining applications will also be more and more, The engineering application and popularization of tool path optimization software can effectively improve the quality and efficiency of zero-part machining, so as to alleviate the pressure of production delivery and quality. We will continue to study the tool path optimization of five-axis machine tool in the future to make the tool path optimization method more reasonable and more effective, while optimizing and promoting tools software, strengthen engineering application, solve practical problems in production.

References

- [1] Wang F, Lin H, Zheng LM et al. A control algorithm for rapid movements near the radius compensation singularity in five-axis end milling. *J. Syst. Sci. Complexity*. 2013, 26(5):718-734.
- [2] Lin ZW, Fu JZ, Shen HY et al. Non-singular tool path planning by translating tool orientations in C-space. *Int. J. Adv. Manufacturing Tech*. 2014, 71(9-12):1835-1848.
- [3] Knut S. Inverse kinematics of five-axis machines near singular configurations. *Int. J. Machine Tools Manufacture*. 2006, 47(2):299-306.
- [4] Affouard A, Duc E, Lartigue C et al. Avoiding 5-axis singularities using tool path deformation. *Int. J. Machine Tools Manufacture*. 2003, 44(4):415-425.
- [5] Yang JX, Yusuf A. Generalized kinematics of five-axis serial machines with non-singular tool path generation. *Int. J. Machine Tools Manufacture*. 2013,75:119-132
- [6] Lin ZW, Fu JZ, Yao XH, Sun YF. Improving machined surface textures in avoiding five-axis singularities considering tool orientation angle changes. *Int. J. Machine Tools Manuf.*, 2015, 41-49.

Big Data System for Medical Images Analysis

Janusz BOBULSKI¹ and Mariusz KUBANEK

Czestochowa University of Technology, Department of Computer Science, Poland

Abstract. Big Data in medicine contains conceivably fast processing of large data volumes, alike new and old in perseverance associate the diagnosis and treatment of patients' diseases. Backing systems for that kind activities may include pre-programmed rules based on data obtained from the medical interview, and automatic analysis of test diagnostic results will lead to classification of observations to a specific disease entity. The current revolution using Big Data significantly expands the role of computer science in achieving these goals, which is why we propose a computer data processing system using artificial intelligence to analyse and process medical images. We conducted research that confirms the need to use GPUs in Big Data systems that process medical images. The use of this type of processor increases system performance.

Keywords. Big data, intelligent systems, deep learning, multi-data processing.

1. Introduction

The demand for solutions offering effective analytical tools has been increasing in recent years. This trend can also be observed in the field of Big Data analysis. Almost all sectors are interested in these solutions, although undoubtedly business organizations excel in the use of this type of analysis. However, it can be seen that the healthcare sector is also more likely to use these solutions. Theory and practice show that Big Data analytics in this sector can contribute, among others to: improve patient care, designate and implement appropriate paths (methods) for patient treatment, support clinical treatment and diagnostic support [1]. However, Big Data is also associated with a certain type challenges in the form of complexity, threats to security and privacy as well as the demand for new technologies and human skills [2, 3].

In Big Data, the main problem is the integration and homogeneity of data, because they come from different sources and there are different formats, for example, the format of dates in the US and in Europe, whether upper and lower case letters in the name of their own. This results in the need to control their condition [4, 5]. The key factor here is the correctness and quality of the data, because in the case of their lack the results of the system will be unsatisfactory, according to the principle: garbage at the input-garbage at the output. Therefore, the system should be equipped with a data control and conditioning module, analogous to signal processing systems where the input signal is filtered to eliminate interference [6, 7].

¹ Corresponding Author: Janusz Bobulski, Czestochowa University of Technology, Department of Computer Science, 73 Dabrowskiego Str., 42-201 Czestochowa, Poland; E-mail: janusz@icis.pcz.pl.

In modern IT, we have two main trends, big data and deep learning. The development of these fields will significantly affect technological progress. By adding Computer Vision to this you can talk about the systems of the future. That is why we decided to develop a system for analysing large volumes of images using artificial intelligence for the needs of medicine [7].

Image labelling is crucial in big data image processing systems. Improving the efficiency of labelling is essential to the development and speed of operation of such systems. In this article, we conducted research that confirms the need to use GPUs in Big Data systems that process medical images. The use of this type of processor increases system performance.

2. Big Data architectures

2.1. Lambda

Lambda architecture is a popular architecture used in Big Data systems. It allows simultaneous access to large data sets and their parallel processing.

The main feature of the Lambda architecture is the presence of two identical data streams (Figure 1), where one is processed in real time and the other in batch mode [8].

In real time mode, data is processed continuously, and thanks to the short time of data access it is possible to quickly search for information. In this case, access to historical data is not possible and not all operations are possible. The quality and reliability of data in this mode is lower. Batch mode is more reliable, however, due to its longer processing time, the real mode allows you to process data in real time.

In batch mode, calculations are made for the entire data set and take much longer, but the data received is of high quality and contains a full history. This dataset has an indivisible form that only needs to be expanded without removing data from it. This ensures data consistency and access to historical data.

Data views are created based on real and batch data in access mode. Their aggregation causes them to bond in such a way that they are visible as a whole. Views give the opportunity to perform various ad hoc examines while providing quick access to data.

The idea of Lambda construction offers a compromise between real-time and batch processing. The biggest disadvantage of this solution is the necessity to maintain two sovereign subsystems - one for the real-time layer and second for the batch layer. The implements used in each layer are unlike, hence the need to use different solutions for each mode, which makes this structure more complicated and costly to preserve.

2.2. Kappa

Kappa architecture is an alternative without the main disadvantage of Lambda [8] architecture. Its idea results from four core assumptions:

- Everything is kind of a stream that any data source can generate.
- The data is immutable and can be reused at any time.
- KISS - Keep principle is short and simple.
- You can restore the data state at any time.

The data in the stream must remain original and unchanged, otherwise you will not be able to get consistent calculation results. In figure 1 we can see the structure of the Kappa architecture. There are a real-time mode and an access mode that implements the identical roles. There is no Batch layer mode here, which was redundant, for the reason that the story can be restored at any time. Because of these advantages, we chose this architecture for our system [9, 10].

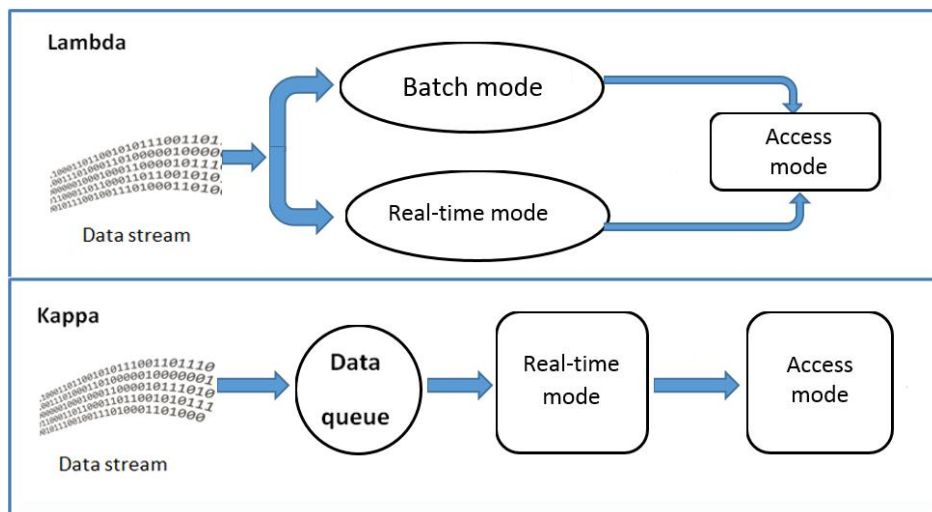


Figure 1. Comparison of Lambda and Kappa architectures.

3. Elements of the system

The broad structure of the system is exposed in the figure 2. Key parts of the structures are:

1. System for processing to some extent kind of data that is capable to find solutions to new problems; the objective of the system is to gather information, not data.
2. Users: a person, system or device capable to communicate with the system,
3. Data: a set of data stream or files.
4. External commands: questions from the external users, e.g.: “Find all images with red hat”.
5. Methods library: a set of procedures work on external data provided to the system, based on algorithms of artificial intelligence, deep learning, application algorithms, data processing and analysis; this library will be used by the processing unit.
6. Pre-processing: identification of input data; and analysis of correctness and error repair.
7. Language interpreter: convert external command to internal commands of system.
8. Internal commands list: a set of methods and instructions that can do system.
9. Knowledge database.
10. Analysis module: based on deep learning and artificial intelligence.

11. System output: results of data processing.

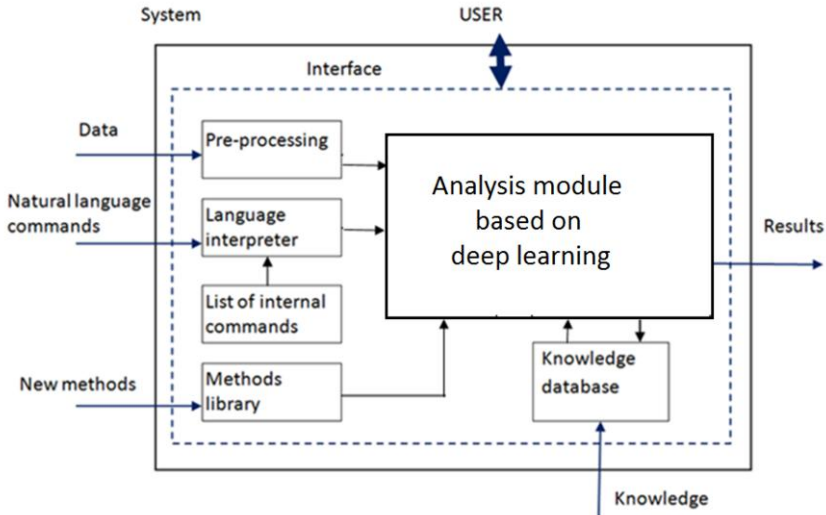


Figure 2. Architecture of propose system.

Preliminary data analysis has to control the data quality, for the reason that agreeing to the GIGO (“garbage in - garbage out”) principle - incoming incorrect data consequences in incorrect results. Deprived data makes it difficult to pull correct conclusions, and henceforward, decision making and knowledge mining. Pre-processing ought to include data cleaning for further investigation. It is estimated that the preliminary data processing is 70-80 % of the information finding procedure. Data afterward verification e.g. a range, will be transformed to the correct internal layout. Then in the system appears a fused data structure. An next profit of this stage is data standardization.

4. Data model for knowledge database

Database model is understood as a set of rules that describe the structure of data in a given database. Allowed operations or data structure are defined by specifying representations of entities and relationships allowed in a given model. In the context of Big Data processing, NoSQL (Not only Structured Query Language) databases that have no SQL database (Structured Query Language) restrictions apply.

The graph model of data representation is a model with labelled and directed multi-graphs, which contains attributes. Labelled, because it has labels for all edges. Directed, because it has edges with certain direction (from the source to the end node). The presence of attributes in the graph causes the attribute list variable to be assigned to each node and edge. The attribute is the value associated with the name.

A multigraph can have multiple edges between two nodes. Therefore, different edges can connect two nodes many times, without paying attention to whether these edges have the same source, end node and label.

In GBD, graph is a native form of information storage. It is often stated that GBD provides index-free neighbourhoods. In GBD, information about vertex neighbours is

stored locally. For comparison, in RBD there would be a reference to an index that would show the vicinity of the vertices. Therefore, in the graph option, the neighbourhood search time depends on the number of vertices neighbours. In the relational option, however, on the number of all edges [11, 12].

Semi-structured data is often modelled as large tables, with lots of columns, blank for most rows, which significantly reduces performance. The alternative in the form of modelling this data, with the participation of many table connections, is also not an efficient solution, considering the costs of joins during queries. In addition, normalization of graph structures in RBD degrades performance during queries. Such a decrease in performance is related to the recursive nature of e.g. file trees or social networks, as well as to the form of recording data as relationships. Each operation performed on the edge of the graph gives the effect of joining between tables in a relational database. This is a very slow operation, which in addition is not scalable [13].

GBD support a graph model that allows direct storage of specific objects and relationships between them in a database. GDB should allow access to query methods that cope not only with stored objects, but with the graph structure itself. The best-known example is traverse, which in its simplest form can be used to obtain adjacent vertices in the graph. The use of the graph database in the propose System has the additional advantage of being able to search graphs in depth and breadth [14].

The use of GDB technology will result in the creation of a semantic data network. The semantic net is a graph in which individual elements have their meaning. A semantic graph is not a specific product, specification or standard. It's more of an idea or vision than technology. The goal of the semantic graph project is to make data available for processing: people and machines so that they can be used not only for display purposes, but also for automation, integration and reuse in many different applications, such as intelligent agents. They will use distributed databases in the form of semantic networks. This in turn will allow the creation of an automatic infrastructure that, if properly designed, will make a significant contribution to the evolution of human knowledge [15]. In addition, the semantic web enriches the current web with annotations written in a machine-process able language, which can also be associated with each other [16].

5. Analysis module on the base deep learning

In connection with the growing need to provide remote diagnostics systems, new methods are being sought to enable automatic performance of tasks that until now were only possible for people, e.g. speech understanding, object recognition or the entire image context. One of these techniques is deep learning. This rapidly growing field of machine learning, based on deep (having many hidden layers) neural networks, has become an indispensable tool that allows computers to solve problems of perception of the world around us. Deep learning is a technique known for many years. The earliest deep learning models, consisting of many layers of nonlinear features, date back to the 1960s. In 1965, Ivakhnenko and Lapa [17] published the architecture of the first deep one-way (feedforward) neural network based on polynomial activation functions. Unfortunately, in those days models based on deep learning were not effective. Only recently has deep learning become a key approach used, among others, in image recognition. One of the reasons for this progress is the ability to perform calculations on larger models, thanks to the availability of faster processors and graphics cards. In addition, better regularization methods are now known that allow the training of large

networks that respond to complex problems. Another important aspect is the ability to provide models with the right amount of data (Big Data) necessary to successfully train a neural network. The Big Data trend has made machine learning much simpler because the significance of the main limitation of estimation techniques (incorrect reasoning based on a small data set) has been reduced [18].

Many aspects have contributed to the success of the deep neural networks observed over the past few years. First of all, thanks to the development of the Big Data trend and much greater availability of various databases, neural networks using deep learning algorithms can draw appropriate conclusions. It turned out that these algorithms become very effective when we provide models with the right amount of data necessary to make proper observations, min. 5000 tagged data in each class. Another reason for developing deep learning is the ability to perform calculations on larger models. Networks consisting of a small number of neurons are not able to model complex problems. However, the more neurons, the more intelligent our system becomes. The use of large models is now possible due to the availability of much faster processors and graphics cards, as well as better methods of preventing over fitting of the network, such as early termination, L2 regularization (adding an additional factor to the error function that imposes a penalty on high weights obtained by the network), or dropout (removing a selected portion of the activation function at random). The greater availability of data used to train models and much better hardware resources meant that deep learning could compete with other algorithms in the world-wide image recognition professions 'ImageNet Large Scale Visual Recognition'. In 2012, convolutional neural networks [19] achieved first place for the first time, reducing the error of detecting the correct category in the first five results (top-5 error) from 26.1% to 15.3%. Since then, deep neural networks have been winning this competition every year, and now the top-5 error has been reduced by up to 3.6%. Due to the rapid development of the deep learning technique, several models developed on large data sets have been created, which allow classification of objects with great efficiency - Inception model trained using ImageNet [20] containing 1000 categories of images is subject to a top-5 error of 5.6%). In addition, using transfer learning [21 - 23], it is possible to re-train existing proven and effective models for new tasks by using existing network weights for all layers except the last one, which is removed and re-trained using a new data set. Transfer learning opens up a number of options for using deep learning algorithms when the training data set is too small to learn full deep representation or when we have limited network training time. Recently, several papers have been devoted to the issue of transfer learning used for the purpose of medical diagnostics in order to improve and improve the quality of services provided, e.g. celiac disease diagnosis and detection of anomalies in duodenal endoscopic images [24]. In both works, the revolutionary neural networks trained using the ImageNet image database [25] were re-trained on a new smaller set of images. The applied transfer learning procedure allowed to significantly reduce the time of network training while maintaining high accuracy of classification. In addition, the same model has been adapted to two completely different tasks, giving very good results in both cases. On this basis, we can conclude that known effective models can be used for virtually any object classification task, as long as we provide the appropriate input data to the networks. In this regard, it becomes very important to create publicly available collections of medical images that could be used to overtrain deep neural networks. The trained models can then be used for remote medical diagnostics.

6. Experiments

Image labelling is crucial in big data systems that process medical images. Improving the efficiency of labelling is essential to the development and speed of operation of such systems. This study looked at the main aspects of image labelling, namely the time needed for labelling. Reducing the required time will allow for an increase in the number of processed images, which in turn will improve the efficiency of the system. We conducted research on the time needed for image operations using the CPU and GPU. The same artificial neural network was used in both cases. The results show a significant reduction in computation time during labelling. The time efficiency is based on the average of a hundred runs of a test program in which a thousand photos are labelled. Four series of tests were performed in which labelling and image processing were performed:

- only labelling
- Labelling and improvement of image contrast
- Labelling and image sharpening
- Labelling and histogram equalization

The results of the measurements are presented in Table 1. As can be seen, the use of GPU processors allows to reduce the computation time by 35%.

Table 1. Times executions for 1000, 500, 100 example images [s], Comparison CPU and GPU.

		Labeling +			
		-	contrast	unsharp	histogram equalization
CPU	1000 images	761	758	775	767
	500 images	381	383	391	391
	100 images	76	77	78	78
		Labeling +			
		-	contrast	unsharp	histogram equalization
GPU	1000 images	502	498	520	501
	500 images	251	253	253	254
	100 images	50	51	51	51

7. Conclusion

In the article, we conducted research that confirms the need to use GPUs in Big Data systems that process medical images. The use of GPU processor increases system performance and reduce computing time. Additional advantage is decreasing power consumption.

Thanks to the development of deep learning algorithms, image analysis can be performed in real time on generally available desktop computers while maintaining high accuracy of obtained results. However, it should be remembered that models based on deep learning become effective only when we provide them with the necessary amount of data to make the right observations. The development of the Big Data trend has meant

that we have much more information at our disposal, thanks to which we can create universal deep neural networks and adapt them to our needs using the transfer learning technique. In order to be able to use them also in medicine and remote medical diagnostics, it is very important to create further databases of medical photos with open accesses, which will be used to create neural networks and appropriate inference when analysing new images.

Big Data significant tasks of the modern digital world. Big Data as a complex of IT issues requires the creation of new data analysis techniques and technological solutions that will allow to extract valuable and valuable information from them. New technologies of data gathering and processing force interdisciplinary exploration and the need to combine existing solutions. Next generation huge computer systems will be grounded on performances that use Deep Learning, Computer Vision, Big Data and others [26], so new technologies should be developed that may process large volumes of data and mine suitable data for medical world.

References

- [1] Dash S, Shakyawar SK, Sharma M et al. Big data in healthcare: management, analysis and future prospects. *J Big Data* 6, 54 (2019).
- [2] Buhl H, Rooglinger M, Moser F, Heidemann J. *Big Data, Business & Information Systems Engineering*, April 2013, Volume 5, Issue 2, pp 65-69.
- [3] Zikopoulos P, Eaton C, deRoos D, Deutsch T, Lapis G. *Understanding big data: analytics for enterprise class hadoop and streaming data*, McGraw Hill, USA, 2012.
- [4] Jinchuan C, Yueguo C, Xiaoyong D, Cuiping L, Jiaheng L, Suyun Z, Xuan Z. *Big data challenge: a data management perspective*, *Frontiers of Computer Science*, SP Higher Education Press, 2013, 7(2): 157-164.
- [5] Katal A, Wazid M, Goudar RH. Big data: issues, challenges, tools and good practices, 2013 Sixth International Conference on Contemporary Computing (IC3), IEEE, Noida, 2013, pp. 404-409.
- [6] Doug L: *Data Management: Controlling Data Volume, Velocity, and Variety*, Application Delivery Strategies, META Group, Gartner, 2011.
- [7] Bobulski J, Kubanek M. Design of the BLINDS System for Processing and Analysis of big data - a pre-processing data analysis module, *Adv. Intel. Syst. Comp.*, 2019, Volume 889, pp. 132-139.
- [8] Marz N, Warren J. *Big data principles and best practices*, Manning Publications Co., 2015.
- [9] Zadeh LA. *Computing with words: principal concepts and ideas*. Springer Publishing Company, Incorporated, 2012.
- [10] Schank RC. *Conceptual information processing*, Yale University, New Haven, Connecticut, 1975
- [11] Sadalage JP, Fowler M. t. l. Hubisz J. *NoSQL Distilled: A Brief Guide to the Emerging World of Polyglot Persistence*, Pearson Education, 2013.
- [12] Barbierato E, Gribaudo M, Iacono M. Performance evaluation of NoSQL bigdata applications using multi-formalism models, *Future Generation Computer Science*, 2014, Vol.37, pp. 345-353.
- [13] Kolomicenko V. *Analysis and experimental comparison of graph databases*. Masters's Thesis, Charles University in Prague, 2013.
- [14] Robinson I, Webber J, Eifrem E. *Graph Databases*. O'Reilly Media, 2013.
- [15] Slotwinski D. *Graph databases - technology review*, AGH, 2010.
- [16] Berners-Lee T, Hendler J, Lassila O. The semantic web. *Scientific American*, 2001, 284(5), pp.34-43.
- [17] Euzenat J, Shvaiko P. *Ontology Matching*, Springer Science & Business Media, 2007.
- [18] Ivakhnenko AG, Lapa VG, *Cybernetic predicting devices*, CCM Information Corporation, 1965.
- [19] Goodfellow I, Bengio Y, Courville A. *Deep Learning*, Cambridge, MA: MIT Press 2017.
- [20] Krizhevsky A, Sutskever I, Hinton GE. *ImageNet classification with deep convolutional neural networks*, NIPS'2012: Neural Information Processing Systems, Lake Tahoe, Nevada, 2012
- [21] Szegedy C. *Going deeper with convolutions*, 2015 IEEE Conference on Computer Vision and Pattern Recognition CVPR, Boston, MA, 1-9, 2015.
- [22] Hunter J, Hsu CH. *Formal Acknowledgement of Citizen Scientists' Contributions via Dynamic Data Citations*. In: Allen R., Hunter J., Zeng M. (eds) *Digital Libraries: Providing Quality Information*. ICADL 2015. *Lecture Notes in Computer Science*, vol 9469. Springer, Cham., 2015, pp 64-75.

- [23] Wimmer G, Veecei A, Uhl A. CNN transfer learning for the automated diagnosis of celiac disease, The 6th International Conference on Image Processing Theory, Tools and Applications (IPTA), Oulu, 2016, pp. 1-6.
- [24] Ianiro G, Bibbo S, Pecere S, Gasbarrini A, Cammarota G. Current technologies for the endoscopic assessment of duodenal villous pattern in celiac disease. *Comput. Biol Med.* 1. October 2015, 65, 308-314.
- [25] ImageNet database, <http://www.image-net.org>, Last access 20 Jan 2019.
- [26] Bobulski J, Kubanek M. CNN use for plastic garbage classification method, 25TH ACM SIGKDD Conference on Knowledge Discovery and Data Mining, Workshop on Data Mining and AI for Conservation, 4-8 August 2019, Anchorage, Alaska, USA, 2019.

Research and Application of High Precision Machining Technology for Super-Large Integral Frame Parts of Aluminum Alloy

Hong JI¹, Yujie WANG, Yu PENG, Lixin ZHAO and Song HUANG

CAC Chengdu Aircraft Industrial (Group) Co., Ltd., Chengdu 610092, China

Abstract. Compared with ordinary large-scale structural parts, super-large aircraft aluminum alloy integral frame parts have the characteristics of large size, high ribs and thin-walled, which lead to the difficulty of deformation control and dimensional accuracy assurance in the machining process, and the problems of spring knife and broach are easy to occur. In this paper, the research on super-large aluminum alloy integral frame parts is carried out, and a set of methods with part deformation control and coordinate drift error control are proposed, and the processing programming strategy is further optimized. This method has been successfully applied to a super-large aircraft aluminum alloy integral frame part, which greatly reduces the deformation of parts, improves the processing stability, and improves the processing efficiency by about 30%.

Keywords. super-large integral frame part, deformation control, drift error reduction, thin-walled high rib

1. Introduction

With the rapid upgrading of modern aircraft and the continuous improvement of aircraft performance, more and more integral frame parts are used in aircraft, and the development trend is towards super-large and complex structure. Figure 1 and Figure 2 shows a large integral frame part of Eurofighter Typhoon and a wing rib of Airbus plane, respectively [1].

The size of super-large aluminum alloy integral frame is generally larger than 3000 * 1500 * 150 mm. The minimum thin-walled thickness is 1.5mm. To achieve high precision and high efficiency machining of this kind of parts is one of the difficult problems to be solved in the field of NC machining [2].

Zhang *et al.* have put forward effective vibration method in deformation control technology, which can solve the deformation problem of small and medium-sized wall plate thin-walled parts [3]. Dong and Cheng and others put forward a finite element correction scheme for parts deformation correction, which can predict the deformation trend of specific parts in advance and adjust accordingly [4]. However, the above

¹ Corresponding Author: Hong JI, CAC Chengdu Aircraft Industrial (Group) Co., Ltd., Chengdu China;

methods cannot solve the deformation problems such as side bending, warping and shrinkage existing in the machining of super-large aluminium alloy integral frame, and cannot effectively guarantee the machining accuracy of super-large aluminium alloy integral frame. Therefore, this paper takes a super-large aluminium alloy integral frame as the application object, and puts forward a new process method which integrates deformation control, drift error control and machining strategy optimization, which can improve the deformation control ability and precision guarantee ability of super-large aluminium alloy integral frame.

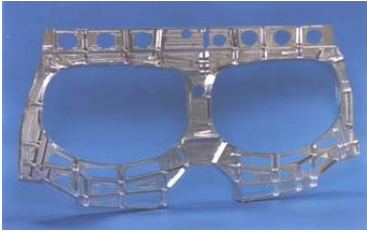


Figure 1. A large integral frame part of Eurofighter Typhoon.



Figure 2. A wing rib of Airbus plane (15V2).

2. Analysis of Machining Difficulties

At present, the main problems existing in super-large aluminium alloy integral frame are as follows:

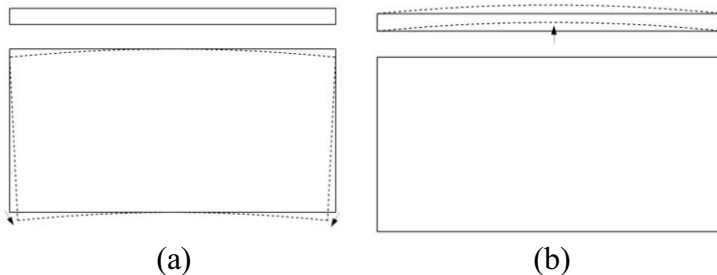
2.1. Difficulties of Deformation Control

The major influential factors leading to processing deformation are summarized as:

- Super large size and asymmetric structure.
- Unevenly release of residual stress and internal stress caused by large amount of material removal in the stage of rough machining.

The typical deformation trend of large-scale integrated structural component in machining is illustrated as follows (as shown in Figure 3):

- Lateral bending along the length with deformation of about 0.4 mm;
- Warping along the thickness with deformation of about 3 mm;
- Shrinking along the length with deformation of about 0.4 mm.



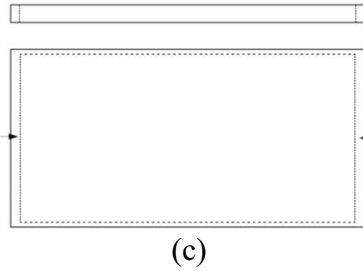


Figure 3. Schematic diagram of typical deformation trend for large integrated frame part machining: (a) lateral bending; (b) warping; (c) shrinking.

2.2. Difficulties of Drift Error Control

The coordinate drift exists in all axes of large-scale CNC machine tools. The maximum stroke of the coordinate axes of the machine tools is about 4000mm when machining. Statistics of the coordinate axes drift of large-scale CNC machine tools are usually about 0.2-0.4mm by field data collection, which seriously affects the accuracy of parts and the quality of products [5].

2.3. Poor Stability of Machining Process

Generally, the external profile of super-large integral frame part is designed with 1.5~2.5 mm ribs along outer profile, accompanied by asymmetric structure, such as the depth of cavity on one side approaching 70 mm, that on other side reaching 95 mm. On strong consideration for light weighting, web plates and high-ribs are empirically designed as 1.5 mm, except for the key areas where are susceptible to fatigue damage. In addition, the lug for mounting flange and some lightening holes are also used to saving in weight. In the high-speed milling process, such structures, like thin wall and high rib (as shown in Figure 4), make it easy to rebound or scrape the cutter. Because cutting force on the cutter is complicated, especially when it simultaneously contacts with three sides, the bottom and the lateral of lug boss as well as the web plate [6].



Figure 4. Typical structure of high ribs and thin web plates.

2.4. Long Period of Machining Process

A typical super-large aluminum alloy integral part weighs about 3500 kg. After rough machining, the weight of the part is about 500 kg, the amount of blank removed from rough machining exceeds 3000 kg, and the material removal rate is about 85%. The manufacturing cycle of the whole part is 21 days. It can be imagined that such a high material removal rate and such a long-time span will pose a serious challenge to the deformation control of the part [7].

3. High Efficiency and High Precision Strategy

Based on the analysis of machining difficulties and production experiences on large integral frame parts, present work investigated the CNC technology with high machining precision from two aspects: distortion control and product quality control.

3.1. Multi-Dimensional Deformation Control by Improving Process Scheme

Technology of distortion control proposed for super large integrated frame parts is shown in Figure 5.

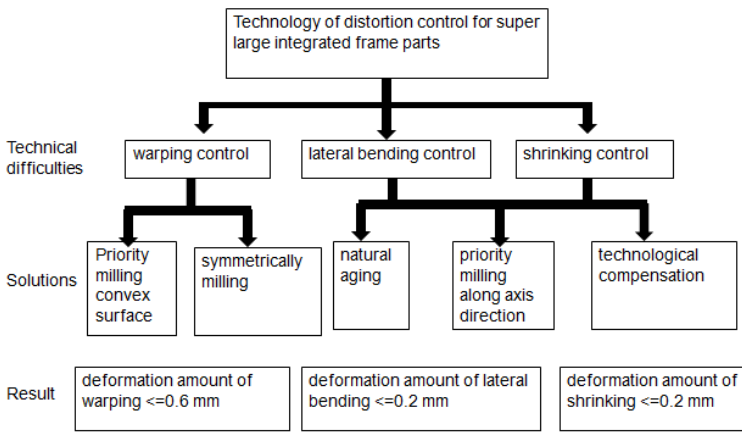


Figure 5. Technology roadmap of high-precision machining on large integral frame parts.

3.1.1. Priority milling convex surface and symmetrically machining to reduce warping Increase natural aging, combined with axial priority process compensation to reduce side bending

Forging piece as mentioned in section 2.4 (size of 3800×1800×180 mm) itself has warping deformation of 10 mm, but that can be even worse in case of improper distortion control during machining. In order to overcome this problem, present study proposed a method illustrated in Figure 6: first mill the convex surface on workblank that caused by itself warping after forging, and the milling depth is half of the forging blank. Warping deformation occurring in machining opposites to that from itself through all processing steps. As a result, the amount of warping deformation of final product can be reduced.

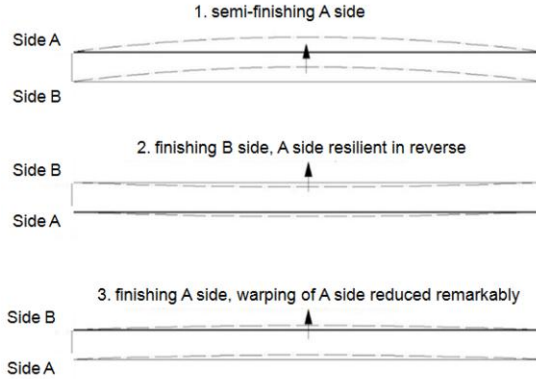


Figure 6. Method of symmetrically machining to reduce warping deformation.

3.1.2. Increase natural aging, combined with axial priority process compensation to reduce side bending

Due to the huge amount of material removal (sometimes even reach to 97%) in the CNC operation of large aluminum alloy integral frame, lateral bending and shrinking often happen resulting from the effect of residual stress and cutting force on material. Generally, residual stress releases constantly as the process proceeds; then larger material removal, more distortion in parts. Thus, some approaches, such as priority milling along axis direction and natural aging, should be adopted to assist the release of residual stress.

However, machining distortions are so complicated that cannot be eliminated. According to the actual data collected about lateral bending and shrinking, aluminum alloy component with size of 1000 mm undergoes undesirable deformation of around 0.1 mm during machining. For the aluminum alloy integral frame part with profile of 3600×1600×165 mm, it can be up to 0.4 mm that much larger than the tolerance requirement. To solve these problems, the method of technological compensation (as shown in Figure 7) is often used in practice. It can help not only minimize the distortion during milling, but also homogenize deformation in the frame parts.

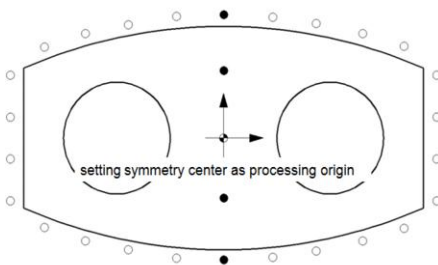


Figure 7. The origin is set at the center line of symmetry.

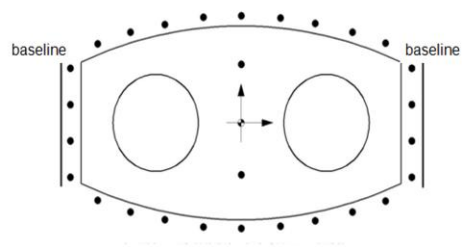


Figure 8. Set baseline.

3.2. Drift Error Distribution and Calibration

Statistics show that, the drift error along X-axis on large CNC machine is around 0.05 mm at intervals of 1000 mm, and that on aged CNC machine is about 0.1 mm. Through setting the symmetry center as the origin for processing, it can be reduced to $(0.1-0.1/2/2)=0.075$ mm. In addition, with the help of baseline on both sides of workpiece (Figure 8), the drift error along X-axis can be further decreased to 0.05 mm. Based on this method, the drift error from machining large integral structure components of aluminum alloy can be smaller than 0.1 mm.

3.3. Optimization of Dynamic Rigid Machining Strategy for Thin-Walled High-Ribs Strip

1. Rough machining

The material removal rate of rough machining is about 85%, which is the main factor of long manufacturing cycle. Therefore, more efficient cutting parameters should be used in rough machining. The cutting parameters before and after optimization are shown in the table 1.

Table 1. Rough machining parameters of previous and optimized program.

Roughing	Cutter	Speed (S)	Feed (F)	Depth of cut (AP)	Width of cut (AE)
Previous program	D32*65R3	9000	4500	5 mm	24 mm
Optimized program	D32*65R3	9000	4500	5 mm	18 mm

2. Semi finishing

In this case, high-ribs was divided into two parts for processing. During the process, the cutting depth was changed layer by layer. Finally, the upper of ribs were tapered as shown in Figure 9.

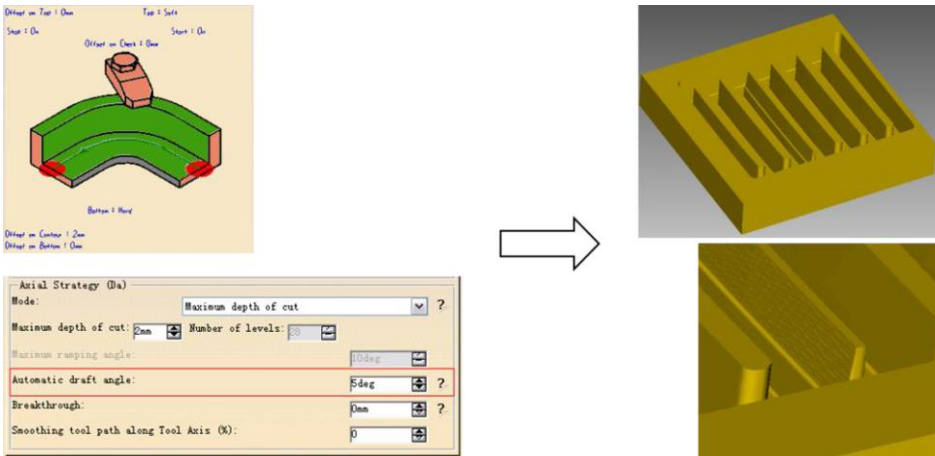


Figure 9. Semi-finishing approach: layer-by-layer machining taper.

3. Finishing

As shown in Figure 10, the finishing process of upper part was divided into two steps. In this approach motion, spiral machining method was applied for the first part due to its poorer stiffness. But the conventional method that is single-side layer-by-layer machining, was used at the second step, because the lower part has greater stiffness and processing stability.

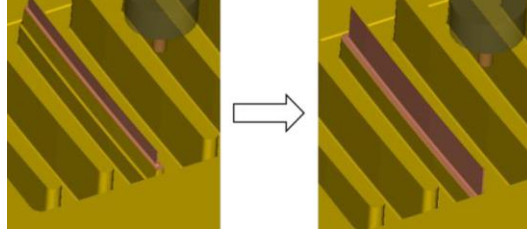


Figure 10. Finish machining approach motion: dividing finishing process for the upper part into two steps.

4. Process Validation and Application Evaluation

Above machining strategy was applied to a representative super-large integrated frame part of aluminum alloy. The results show that above methods possess certain processing accuracy and can be adopted in similar parts.

4.1. Results of Distortion Control

According to Figure 11, the deformation amount of the final product is sampled. The test results (Figure 12) indicate that distortion caused by residual stress and inhomogeneous machining stress has been controlled well and the final product is qualified. It can be concluded that the machining strategy proposed in this article can effectively solve the distortion problem in super large integrated frame parts of aluminum alloy.

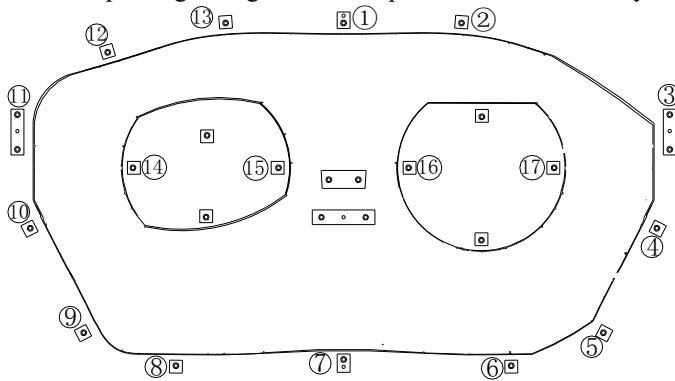


Figure 11. Sampling points used for collecting data of distortion deformation.

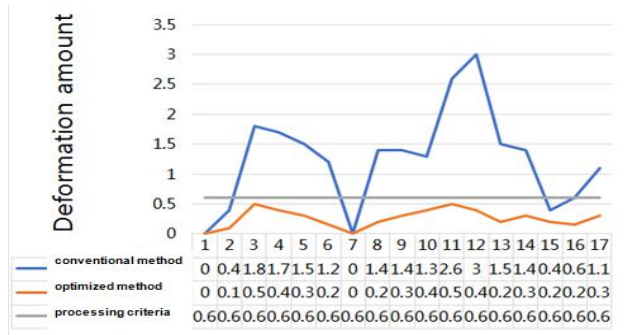


Figure 12. Comparison of deformation amount between previous and optimized program.

4.2. Results of Surface Quality Control

According to requirements for large structural components, surface roughness in key area for preventing fatigue damage is 1.6, that of other areas is 3.2; dimensional tolerance of ribs, grooves and web plates is ± 0.2 mm. Data from the actual measurements indicates that, after applying above machining strategy in the frame part, machining precision and surface quality were all qualified. As shown in Figure 13, there no evident trail from cutter rebounding and scraping is observed.

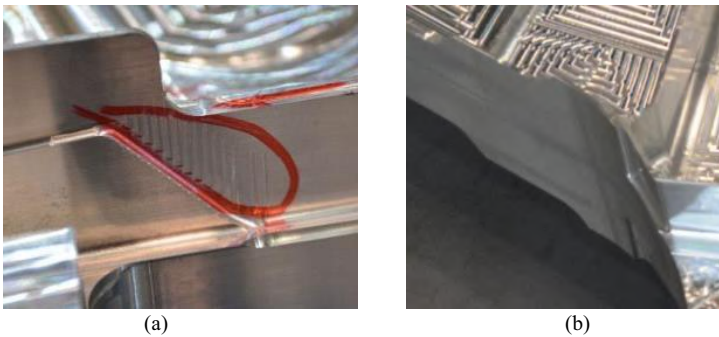


Figure 13. Surface quality in ribs (a) before and (b) after optimization.

4.3. Results of Efficiency Improvement

Statistics for actual processing time show that, it usually took about 21 days to finish a large integral frame part before. But, nowadays, the whole process only took 15 days by using efficiency strategy mentioned above. The processing efficiency for each product has increased by 30%. In combination with the comparison of processing time in every operation between conventional method and optimized method (Figure 14), it can be concluded that the methods of efficiency improvement in present work is useful and can be implemented in super large integral frame parts.

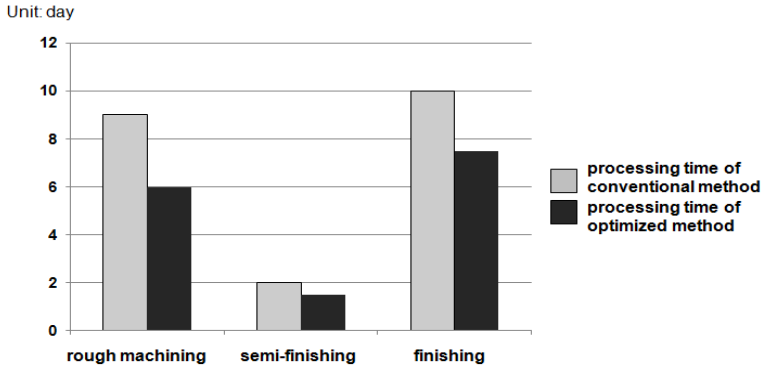


Figure 14. Processing time before and after optimization.

5. Conclusions

The machining strategy for representative super-large integrated frame parts of aluminum alloy was investigated. Based on analysis of technical difficulties in such parts, this study discussed the factors affecting machining deformation and dimension accuracy, proposed key technological solutions for deformation control and quality control. These machining methods were successfully implemented and promoted in NC factory of AVIC CAC. Finally, through the mass production, the qualified rate reached 99.5%, and the processing efficiency increased by about 30%. In summary, technological solutions of this paper are applicable to similar parts or structures with same technological characteristics. Studies in present work lay a solid foundation for super large integral frame part in aviation and accelerate the development of aviation manufacturing technology.

References

- [1] Yu FF, Du BR, Ren WJ et al. Slicing recognition of aircraft integral panel generalized pocket. *Chin. J. Aeronautics*. 2008, 21(6):585-592.
- [2] Lasemi A, Xue DY, Gu PH. Recent development in CNC machining of freeform surfaces: A state-of-the-art review. *Computer-Aided Design*. 2010, 42(7):641-654.
- [3] Zhang DJ, Cui ZS, Ruan XY, et al. An analytical model for predicting springback and side wall curl of sheet after U-bending. *Comp. Mater. Sci*. 2007, 38(4):707-715
- [4] Dong ZC, Cheng HB. Study on removal mechanism and removal characters for SiC and fused silica by fixed abrasive diamond pellets. *Int. J. Machine Tools Manuf*. 2014, 85:1-13.
- [5] Michael G, Josef M, Nils F et al. High precision grey-box model for compensation of thermal errors on five-axis machines. *CIRP Annals - Manufacturing Tech*. 2014, 63(1):509-512
- [6] Masaomi T, Shintaro T, Noriyuki K et al. Enhancement of geometric accuracy of five-axis machining centers based on identification and compensation of geometric deviations. *Int. J. Machine Tools Manuf*. 2013, 68:11-20.
- [7] Guo Q, Zhao B, Jiang Y, et al. Cutting force modeling for non-uniform helix tools based on compensated chip thickness in five-axis flank milling process. *Precision Eng*. 2018, 51:659-681.

AutoSynPose: Automatic Generation of Synthetic Datasets for 6D Object Pose Estimation

Heiko ENGEMANN ^{a,b}, Shengzhi DU ^{a,1}, Stephan KALLWEIT ^{a,b,2},
Chuanfang NING ^c and Saqib ANWAR ^b

^a*Tshwane University of Technology (TUT), Pretoria, South Africa*

^b*FH Aachen – Aachen University of Applied Sciences, Aachen, Germany*

^c*Ecole polytechnique fédérale de Lausanne (EPFL), Lausanne, Switzerland*

Abstract. We present an automated pipeline for the generation of synthetic datasets for six-dimension (6D) object pose estimation. Therefore, a completely automated generation process based on predefined settings is developed, which enables the user to create large datasets with a minimum of interaction and which is feasible for applications with a high object variance. The pipeline is based on the Unreal 4 (UE4) game engine and provides a high variation for domain randomization, such as object appearance, ambient lighting, camera-object transformation and distractor density. In addition to the object pose and bounding box, the metadata includes all randomization parameters, which enables further studies on randomization parameter tuning. The developed workflow is adaptable to other 3D objects and UE4 environments. An exemplary dataset is provided including five objects of the Yale-CMU-Berkeley (YCB) object set. The datasets consist of 6 million subsegments using 97 rendering locations in 12 different UE4 environments. Each dataset subsegment includes one RGB image, one depth image and one class segmentation image at pixel-level.

Keywords. synthetic dataset, 6D pose estimation, domain randomization

1. Introduction

Computer vision based on deep neural networks (DNNs) enables robots to perceive their environment in a human like manner. Intelligent robots depend on robust perception strategies to perform their key tasks: autonomous navigation [1,2] and adaptive manipulation [3,4]. State-of-the-art approaches for 6D pose (object position and orientation) estimation [5–9] as well as object tracking [10–12] and novel grasping techniques [13–16], enable adaptive task execution and closed loop control of high precision manipulation tasks.

One major factor closely related to the model performance is the quality and quantity of the training dataset [17]. Traditionally, the generation of a dataset for object detection is time-consuming and costly. It is a two-step task, which is divided into: data acquisition and data annotation, including manually executed operations [18]. The difficulty is

¹ Corresponding Author: Shengzhi Du, E-Mail: dus@tut.ac.za

² Corresponding Author: Stephan Kallweit, E-Mail: kallweit@fh-aachen.de

increased in case of 6D pose estimation datasets, caused by the need of object pose information in each dataset subsegment.

In contrast to data sets obtained from the real world, the generation of synthetic data can be fully automated, making it suitable for computer vision tasks with high object variations. The major challenge using synthetic data is closing the gap between the simulated data and its real-world counterpart, i.e. the reality gap. For this reason, different techniques were developed, divided into two basic categories: domain adaptation and reality match. Domain adaptation describes the problem to generalize a trained model from its source space to an unknown target space [19]. Reality match approaches simulate the feature rich real-world data as realistic as possible. It was proved that a model, fully trained on synthetic data, can provide state-of-the-art performance for 6D pose estimation [7]. This supports the hypothesis that synthetic data is an effective alternative when the acquisition and annotation of real-world data is not feasible.

In this paper, we present a fully automated pipeline for the generation of a 6D object pose dataset. The result is a significant reduction in preparation time and a minimization of necessary human interaction. The pipeline can be easily adapted to different objects and environments. It provides multiple improvements to comparable approaches, such as various parameter settings and detailed metadata for each dataset subsegment. An exemplary dataset is generated including six million subsegments. Each subsegment consists of an RGB, depth and class segmentation image at pixel level. Two metadata files provide the 6D object pose and a corresponding 3D bounding box as well as the parameter settings for randomization. Based on the detailed metadata, the data set can be divided into smaller sub-datasets, allowing the study of different parameter combinations.

2. Related Work

Methods to generate synthetic data for training neural networks can be split into two general categories: superimposing and rendering.

The main idea of superimposing is to project object images onto background images. In [20] a method is presented to generate synthesized data from existing real-world datasets. Object images are extracted by cropping the images according to the provided metadata. The object images are projected onto another set of background images, using support surfaces estimated by semantic segmentation and plane fitting. The dataset in [21] is generated using rendering models and real-world images. This method was improved in [22] by adding variations like random illumination, noise and blurring of the object images before superimposing. The implementation of superimposing is straightforward and easy to automate. However, the missing interaction between the environment and the object results in a reality gap due to the absence of important real-world features, e.g. shadows.

Rendering refers to the generation of an image from 2D or 3D models as well as complete environments [23]. Typical rendering software provide numerous setting options for e.g. the ambient lighting, the object appearance, the shape and the poses. Due to the numerous setting options, the generation of datasets in detail is a very time-consuming process. Various methods were developed to provide synthetic data - as realistic as possible - based on rendering. Domain randomization techniques [4] increase the ability of generalizing the trained model to real-world data. An optical flow dataset [24] was presented derived from an animated short film by modifying the motion blur pipeline of the 3D creation suite Blender [25]. The resulting Sintel dataset has been

extended to depth and segmentation images. In [26] the Unity development platform [27] was used to create a virtual urban world, populated with cars, vans, pedestrians and cyclists. The resulting Synthia dataset was generated using a dynamic illumination engine and provides different scene appearances related to the four seasons. The approach proposed in [28] creates a unique virtual world for each dataset subsegment, instead of generating the whole dataset based on a single virtual world. The resulting Synscape dataset [29] is a photorealistic synthetic dataset for street scene parsing. The dataset SceneNet-RGBD [30] provides indoor scenes including household objects. It is an extension of the work presented in [31], using an automatic random scene generator based on 174683 potential 3D objects. In contrast to the previous mentioned approaches, the synthetic dataset Falling things (FAT) [32] is focusing the field of robotic manipulation. Therefore, it uses objects of the real-world Yale-CMU-Berkeley (YCB [33]) object set. In addition to RGB, depth and segmentation images, each dataset subsegment also includes metadata in form of 6D pose information. The SIDOD dataset [34] is closely related, since both datasets were generated using the Unreal Engine 4 (UE4 [35]) and the open-source custom plugin NVIDIA DeepLearning Data Synthesizer (NDDS [36]), a tool to extract view data during rendering.

3. Method

Our approach differentiates from others by focusing a high automation level to enable a wide parameter range for randomization and less human interaction. The pipeline was developed based on the game engine UE4, providing functionalities to predefine parameter settings for environments and objects, including the manual chosen rendering locations, later referred to as training spots. The actual rendering process is highly automated, which simplifies the generation of large object-related datasets and the usage of synthetic data in real world applications. An exemplary dataset was generated using a RTX 2080 Ti, where each rendering process per dataset subsegment takes around 50ms. The UE4 project, including one open-source environment is available at: <http://autosynpose.fh-aachen.de>.

3.1. Domain Randomization

The presented pipeline simulates the following aspects with an user adjustable variability: environment, training spot, object appearance, camera-object transformation, ambient lighting, number and size of geometric primitive distractors, presence of complex object distractors, and multiple instances of the object of interest.

The object appearance can be randomized by an adjustable percentage for the base color. In addition, the appearance parameters roughness, metallic and specular are configurable. The camera describes a motion alongside the surface of a sphere during the automated rendering processes (see Figure 1). The camera motion is reduced to a circular motion in case of training spots on flat surfaces, e.g. tables. The origin of the sphere coordinate system S is located at the manual selected training spot. The optical axes of the camera, which is represented by the z -axis of the camera coordinate system C , points towards the origin of the coordinate system S during the rendering processes.

The object of interest is randomly positioned in a 3D cube with adjustable dimensions. The so-called object cube describes a linear motion alongside the z -axis of the coordinate system C during the rendering processes. Therefore the location of the

object-cube origin can be described as a vector $\mathbf{o}_{oc} = (0, 0, z_{oc})^T$ in reference to the coordinate system C , with $0.2 \text{ m} \leq z_{oc} \leq 1.5 \text{ m}$. The pivot point of the vector \mathbf{o}_{oc} is located at the origin of the coordinate system S , during the motion of the camera and the object-cube. This approach generates a large number of dataset subsegments with different perspectives. An additional 3D cube encloses the object-cube. Inside this distractor-cube, the distractors are positioned randomly around the object of interest, including geometric primitive and object distractors. The distractor-cube will be partly outside the field of view of the camera at close distances. Thus, not all distractors will be present in the corresponding image data. Therefore, we introduce the concept of distractor density, which depends on the size of the distractor-cube, the size of the distractors itself and the number of distractors present.

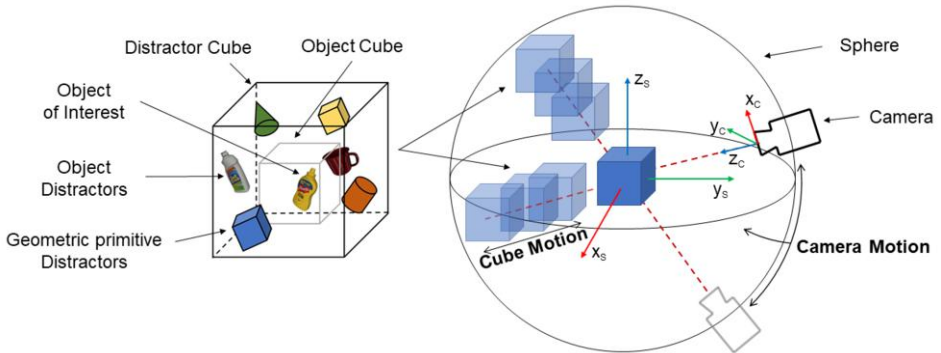


Figure 1. Spherical camera motion around the object of interest.

3.2. Metadata

The presented pipeline provides two metadata files in JSON format for each dataset subsegment. The first one is generated by NDDS, including the 6D object pose and its bounding box in the coordinate system C . The second file contains detailed information of the parameter configuration for domain randomization as shown in table 1.

Table 1. Description of the metadata file for domain randomization.

Category	Name	Description
Generating Info	EnvironmentNo	Environment Number
	SpotNo	Spot Number
Appearance	HueShiftPercent	Shift in object hue in percent
	Roughness	Object roughness value
	Metallic	Object metallic value
Distractors	WithDistractors	Presence of distractors
	DistractorSize	Size of primitive distractors in 10 mm
	DistractorCount	Number of distractors in scene
	WithObjectDistractors	Presence of object distractors
Configurations	WithLightningVariations	Presence of lightning variations
	MultipleInstances	Presence of multiple object instances
	InstanceCount	Number of object instances
	MultipleObjects	Presence of additional objects

4. Dataset Analysis

The proposed pipeline was used to generate the exemplary dataset *AutoSynPose*, which can be downloaded here: <http://autosynpose.fh-aachen.de>. The dataset includes six million subsegments rendered at 97 training spots spread over 12 environments. The environments vary from realistic indoor to game-like outdoor scenes (see Figure 2) in order to bypass the machine learning specific problem that the factors of variation cannot always be directly observed [37].



Figure 2. Environments of the AutoSynPose dataset.

As object of interest, the mustard bottle of the YCB dataset was chosen. In addition, the YCB objects: toy plane, meat can, power drill and hammer are in 17% of the dataset present. Each subsegment consist of one RGB image, one depth image and one segmentation image providing class instance information at pixel level. The subsegment does not include object instance information to reduce the size of the dataset. Figure 3 shows one subsegment including the object of interest and geometric primitive as well as object distractors.



Figure 3. Subsegment of the AutoSynPose dataset.

Figure 4 shows the statistics of the dataset, across the different environments and manually selected training spots. Furthermore, Figure 4 shows the composition of the dataset regarding the presence of distractors, lightning variations, multiple instances of the object of interest and multiple objects.

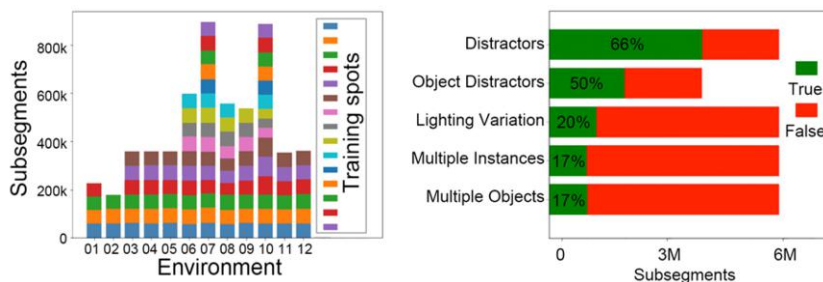


Figure 4. Dataset statistics. Left: Distribution of subsegments over environments and training spots. Right: Dataset composition.

Figure 5 shows a statistical analysis of the six million dataset subsegments. The edge length of the object-cube was set to 0.3m and the edge length of the distractor-cube was set to 0.75m.

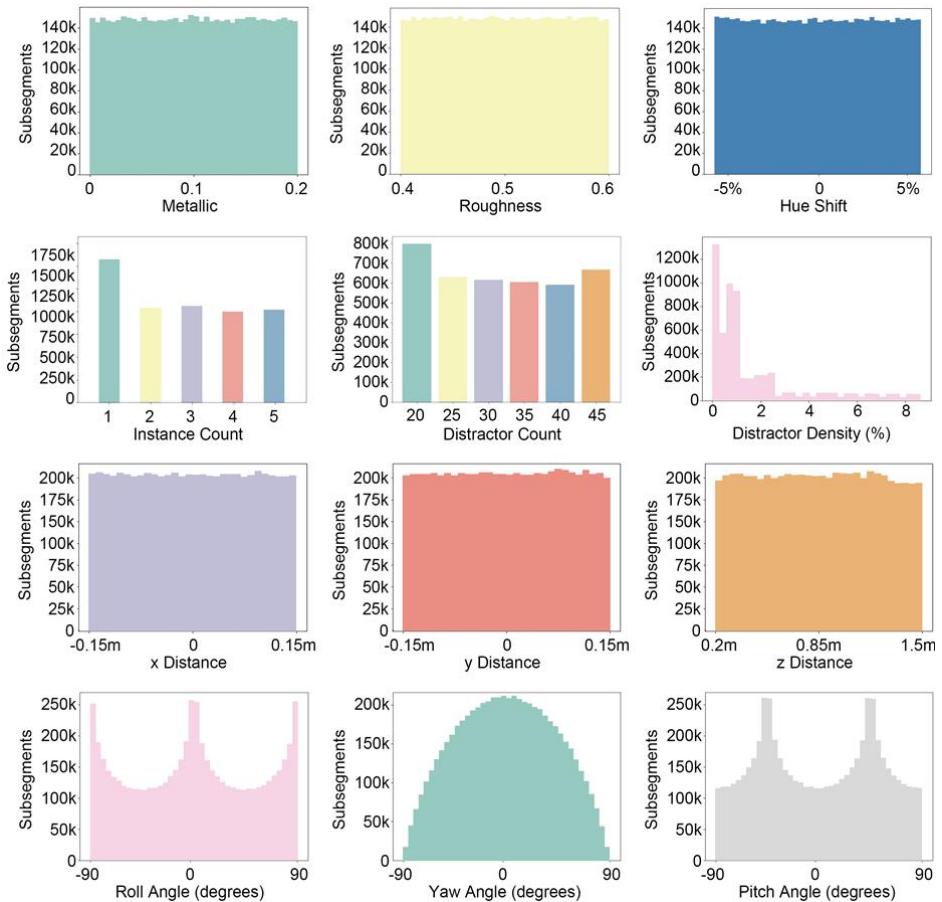


Figure 5. Statistical analysis of the dataset: Object appearance, object instance count, geometric primitive distractors, and camera-object transformation.

The first column of Figure 5 shows the randomization of the object appearance, which is equally distributed around the selected parameter settings. The second column shows the distribution of the instance count and the geometric primitive distractors. The distractor density inside the distractor-cube varies from 0% up to 8%, with a peak in the range between 0% and 1%. The third column shows the distribution of the 3D object position in the camera coordinate system, which is almost uniformly distributed. The distributions of the Euler angles reflect the orientation of the object in the camera coordinate system.

Table 2 shows our dataset in contrast to comparable state-of-the-art datasets. To our knowledge the AutoSynPose dataset is the only 6D object pose dataset providing detailed information about parameter settings used for domain randomization at subsegment level.

Table 2. Comparison of synthetic datasets (Masks* IS – instance segmentation, CS – class segmentation; FD* – Flying Distractors, DI* – Domain Randomization Information).

Dataset	#Segments	Scenes	Masks*	Poses	B-Box	FD*	DI*
Sintel (2015) [24]	1.6 k	movie	IS	3D	×	×	×
Synthia (2016) [26]	200 k	urban	CS	×	×	×	×
SN-RGBD (2016) [31]	5 M	household	IS CS	3D	×	×	×
Synscape (2018) [29]	25 k	urban	IS CS	3D	2D	×	×
FAT (2018) [32]	60 k	household	IS CS	6D	3D	×	×
SIDOD (2019) [34]	115 k	household	IS CS	6D	3D	✓	×
Ours (2020)	6 M	various	CS	6D	3D	✓	✓

The SIDOD and the FAT dataset are closely related to our dataset. Table 3 shows a detailed comparison of the three datasets.

Table 3. Detailed comparison with the synthetic datasets FAT and SIDOD.

Parameter	FAT (2018) [32]	SIDOD (2019) [34]	Ours (2020)
No. of objects	21	21	5
No. of environments	3	3	12
No. of training spots	15	18	97
Object appearance variations	×	×	✓
Lightning variations	✓	✓	✓
Object position distribution	normal	normal	uniform
Object orientation distribution	mixed	mixed	mixed

We were able to significantly increase the number of environments and training spots, because of the high automation level of the proposed pipeline. In addition to the lightning domain, our pipeline also provides configurable settings for the object appearance domain. Another improvement is the uniform distribution of the object position in the camera coordinate system. The compared datasets provide a normal distribution for x and y, whereby the mean is located at the optical axis of the camera. The same applies for the camera-object distance, whereby the mean is located at the mid of the intended range. A uniform distribution represents a larger number of different perspective viewpoints and reduces overfitting problems at test time. The focus of our dataset is to enable studies on the impact of different parameter settings for domain randomization. Therefore, the AutoSynPose dataset only includes five different objects, whereby the mustard bottle of the YCB object set is prioritized.

5. Conclusion

We developed an automated synthetic dataset generating pipeline, based on the game engine UE4 and the open-source plugin NDDS. Adjustable parameters for domain randomization and the rendering process allow a simple change between different objects of interests and environments. The required knowledge about game and rendering engines is reduced. The pipeline generates a uniform distribution of the 3D object position in reference to the camera coordinate system. In addition, the presented camera-object motion concept covers different perspective viewpoints per training spot. To our knowledge, it is the first approach providing detailed information about the parameter settings used for domain randomization at subsegment level. In addition, we introduced the concept of distractor density as a metric for distractors. We hope that the proposed

pipeline makes it easier for other researchers to generate synthetic data from 3D models and environments.

The presented pipeline can be used to synthesize a variety of different datasets. However, the identification of the important factors of variation, which cannot directly be observed, has the potential to decrease the reality gap of future synthetic dataset generation. Therefore, the exemplary generated dataset AutoSynPose aims to study the effect of randomization parameters in detail.

6. References

- [1] Milioto A, Mandtler L, Stachniss C. Fast instance and semantic segmentation exploiting local connectivity, metric learning, and one-shot detection for robotics. In *International Conference on Robotics and Automation (ICRA)*, 2019, pp. 5481–5487.
- [2] Zou Q, Jiang H, Dai Q, Yue Y, Chen L, Wang Q. Robust lane detection from continuous driving scenes using deep neural networks. In *IEEE Transactions on Vehicular Technology*, 2019, 69(1), pp. 41–54.
- [3] Schwarz M, Behnke S. Semantic RGB-D perception for cognitive service robots. In *RGB-D Image Analysis and Processing*, Springer, Cham, 2019, pp. 285–307.
- [4] Tobin J, Fong R, Ray A, Schneider J, Zaremba W, Abbeel P. Domain randomization for transferring deep neural networks from simulation to the real world. In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2017, pp. 23–30.
- [5] Hu Y, Hugonot J, Fua P, Salzmann M. Segmentation-driven 6D object pose estimation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2019, pp. 3385–3394.
- [6] Xiang Y, Schmidt T, Narayanan V, Fox D. Posecnn: A convolutional neural network for 6d object pose estimation in cluttered scenes. arXiv preprint arXiv:1711.00199 (2017).
- [7] Tremblay J, To T, Sundaralingam B, Xiang Y, Fox D, Birchfield S. Deep object pose estimation for semantic robotic grasping of household objects. arXiv preprint arXiv:1809.10790 (2018).
- [8] Deng X, Mousavian A, Xiang Y, Xia F, Bretl T, Fox D. Poserbpf: A rao-blackwellized particle filter for 6d object pose tracking. arXiv preprint arXiv:1905.09304 (2019).
- [9] Wang C, Xu D, Zhu Y, Martín-Martín R, Lu C, Fei-Fei L, Savarese S. Densefusion: 6d object pose estimation by iterative dense fusion. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2019, pp. 3343–3352.
- [10] Choi C, Christensen H I. 3D textureless object detection and tracking: An edge-based approach. In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2012, pp. 3877–3884.
- [11] Crivellaro A, Rad M, Verdie Y, Yi K M, Fua P, Lepetit V. A novel representation of parts for accurate 3D object detection and tracking in monocular images. In *IEEE International Conference on Computer Vision (ICCV)*, 2015, pp. 4391–4399.
- [12] Yuan D, Kang W, He Z. Robust visual tracking with correlation filters and metric learning. *Knowledge-Based Systems* (2020), 105697.
- [13] Kumra S, Kanan C. Robotic grasp detection using deep convolutional neural networks. In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2017, pp. 769–776.
- [14] Morrison D, Corke P, Leitner J. Closing the loop for robotic grasping: A real-time, generative grasp synthesis approach. arXiv preprint arXiv:1804.05172 (2018).
- [15] Morrison D, Corke P, Leitner J. Learning robust, real-time, reactive robotic grasping. *The International Journal of Robotics Research* 39 (2020), 183–201.
- [16] Mahler J, Matl M, Satish V, Danielczuk M, DeRose B, McKinley S, Goldberg K. Learning ambidextrous robot grasping policies. *Science Robotics* 4 (2019).
- [17] Voulodimos A, Doulamis N, Doulamis A, Protopoulos E. Deep learning for computer vision: A brief review. *Computational intelligence and neuroscience* 2018 (2018).
- [18] Weiss G M, Provost F. Learning when training data are costly: The effect of class distribution on tree induction. *Journal of Artificial Intelligence Research* 19 (2003), 315–354.
- [19] Wang M, Deng W. Deep visual domain adaptation: A survey. *Neurocomputing* 312 (2018), 135–153.
- [20] Georgakis G, Mousavian A, Berg A C, Kosecka J. Synthesizing training data for object detection in indoor scenes. arXiv preprint arXiv:1702.07836 (2017).
- [21] Josifovski J, Kerzel M, Pregizer C, Posniak L, Wermter S. Object detection and pose estimation based on convolutional neural networks trained with synthetic data. In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2018, pp. 6269–6276.

- [22] Hinterstoisser S, Lepetit V, Wohlhart P, Konolige K. On pre-trained image features and synthetic images for deep learning. In *Proceedings of the European Conference on Computer Vision (ECCV)*, 2018, pp. 682–697.
- [23] Akenine-Möller T, Haines E, Hoffman N. *Real-time rendering*. Crc Press, 2019.
- [24] Butler D J, Wulff J, Stanley G B, Black M J. A naturalistic open source movie for optical flow evaluation. In *European conference on computer vision*, 2012, pp. 611–625.
- [25] Blender Online Community, Blender - a 3D modelling and rendering package, <http://www.blender.org> [cited 2020 July 18].
- [26] Ros G, Sellart L, Materzynska J, Vazquez D, Lopez A M. The synthia dataset: A large collection of synthetic images for semantic segmentation of urban scenes. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 3234–3243.
- [27] Hass J K. A history of the unity game engine (2014).
- [28] Tsirikoglou A, Kronander J, Wrenninge M, Unger J. Procedural modeling and physically based rendering for synthetic data generation in automotive applications. *arXiv preprint arXiv:1710.06270* (2017).
- [29] Wrenninge M, Unger J. Synscapes: A photorealistic synthetic dataset for street scene parsing. *arXiv preprint arXiv:1810.08705* (2018).
- [30] McCormac J, Handa A, Leutenegger S, Davison A J. SceneNet RGB-D: 5M photorealistic images of synthetic indoor trajectories with ground truth. *arXiv preprint arXiv:1612.05079* (2016).
- [31] Handa A, Patraucean V, Badrinarayanan V, Stent S, Cipolla R. Understanding real world indoor scenes with synthetic data. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016, pp. 4077–4085.
- [32] Tremblay J, To T, Birchfield S. Falling things: A synthetic dataset for 3d object detection and pose estimation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 2018, pp. 2038–2041.
- [33] Calli B, Singh A, Walsman A, Srinivasa S, Abbeel P, Dollar A M. The YCB object and model set: Towards common benchmarks for manipulation research. In *International Conference on Advanced Robotics (ICAR)*, 2015, pp. 510–517.
- [34] Jalal M, Spjut J, Boudaoud B, Betke M. SIDOD: A synthetic image dataset for 3D object pose recognition with distractors. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 2019, pp. 475–477.
- [35] Epic Games. Unreal Engine, <http://www.unrealengine.com> [cited 2020 July 18].
- [36] To T, Tremblay J, McKay D, Yamaguchi Y, Leung K, Balanon A, Cheng J, Hodge W, Birchfield S. NDDS: Nvidia Deep Learning Dataset Synthesizer, 2018.
- [37] Goodfellow I, Bengio Y, Courville A. *Deep Learning*, MIT Press, 2016.

Solving a Fuzzy Linear Equation with a Variable, Using the Expected Interval of a Fuzzy Number

Richard ABRAMONTE^a, Eder ESCOBAR^a, Antenor ALIAGA^b, Flabio GUTIERREZ^{a1}

^aDepartment of Mathematics, ^bDepartment of Electronics, Universidad Nacional De Piura, Piura, Perú.

Abstract. In this work, a fuzzy linear equation $AX + B = 0$, is solved, were A, B y C are triangular diffuse numbers, could also be trapezoidal. For this type of equations there are several solution methods, the classic method that does not always obtain solutions, the most used is the method of alpha cuts and arithmetic intervals that although it always finds a solution, as a value is taken closer to zero (more inaccurate), the solution satisfies less to the equation. The new method using the expected interval, allows to obtain a smaller support set where the solutions come closer to satisfying the equation, also allows to find a single interval where the best solutions for decision making are expected to be found. It is recommended to study the incorporation of the concept of the expected interval in the methods to solve systems of fuzzy linear equations

Keywords. Fuzzy number, expected interval, fuzzy arithmetic, alpha cut, fuzzy equations.

1. Introduction

The fuzzy logic created by Lofti Zadeh in 1965 is an adequate means to represent and model a type of uncertainty such as imprecision, which appears in various reality problems [1].

Fuzzy linear equations are applied in various areas such as optimization, control systems, physics, economy [2-5]. In this work, we focus on the case of the linear equation of the form;

$$\tilde{A}\tilde{X} + \tilde{B} = \tilde{C} \quad (1)$$

where \tilde{A} , \tilde{B} y \tilde{C} are inaccurate data; since the imprecision is represented with fuzzy triangular numbers, the equation is called the fuzzy linear equation (FLE).

The problem of solving a FLE is not trivial, even for cases such as $\tilde{A}\tilde{X} = \tilde{B}$, were \tilde{A} y \tilde{B} are triangular fuzzy numbers (TFN), i.e., for some TFN \tilde{A} y \tilde{B} , you can't find a fuzzy set \tilde{X} , such that, using fuzzy arithmetic, $\tilde{A}\tilde{X}$ is exactly equal to \tilde{B} [6].

¹ Corresponding Author: Flabio Gutierrez, Department of Mathematics, Universidad Nacional de Piura, Campus Universitario, Urb. Miraflores s/n, Castilla, Piura, Perú; E-mail: flabio@unp.edu.pe.

In [7], a strategy is proposed for the solution of a FLE: First, to seek the solution with the classic method \tilde{X}_c , because it is the most accurate solution, but it does not always exist; second, if there was no classic solution, look for the solution by the principle of extension \tilde{X}_e whose construction process is quite complicated (very little used); and finally the solution by the method of alpha cuts and interval arithmetic \tilde{X}_α that gets an approximate solution, but with the condition that always exists. The relationship is true

$$\tilde{X}_c \leq \tilde{X}_e \leq \tilde{X}_\alpha \tag{2}$$

While the alpha-cuts and arithmetic interval method always gets a solution for Eq. (1), of the form $x(\alpha) = [x_1(\alpha), x_2(\alpha)]$, $\alpha \in [0, 1]$; has a problem, according to the ratio (2), the solution support set is much wider than the other methods, i.e. for the alpha cut $\alpha = 1$, you get the most accurate (exact) solution, which completely satisfies the equation; as the equation's tends to 0 (increases inaccuracy) the solution of the equation tends to satisfy every single thing much less to the equation.

That is why, in this work it is proposed to incorporate the concept of the expected interval to the method of alpha cuts and interval arithmetic to obtain a solution with a smaller support set and find a single interval where the best solutions for decision making are found.

The method of alpha cuts and interval arithmetic is the basis of several current methods for solving systems of fuzzy linear equations [8-10].

2. Material and Methods

2.1. Fuzzy Numbers

Definition 1: A generalized LR fuzzy number \tilde{A} with the membership function $\mu_{\tilde{A}}(x)$, $x \in \mathbb{R}$ can be defined as [11].

$$\mu_{\tilde{A}}(x) = \begin{cases} L(x), & a \leq x \leq b, \\ 1, & b \leq x \leq c, \\ R(x), & c \leq x \leq d, \\ 0, & \text{otherwise} \end{cases}$$

Where $L(x)$ is the left membership function that is an increasing function on $[a, b]$ and $R(x)$ is the right membership function that is a decreasing function on $[c, d]$ such that $L(a) = R(d) = 0$ and that $L(b) = R(c) = 0$.

Definition 2: A triangular fuzzy number (TFN) (see Figure 1) is represented by $\tilde{A} = (a_1, a_2, a_3)$.

Definition 3: A TFN $\tilde{A} = (a_1, a_2, a_3)$ is positive if and only if $a_1 > 0$.

Definition 4: Given a fuzzy set $\tilde{A} = \{(x, \mu_{\tilde{A}}(x))\}$, the support of \tilde{A} is defined as the ordinary set.

$$Sop(\tilde{A}) = \{x \in \mathbb{R}, \mu_{\tilde{A}}(x) > 0\}$$

Definition 5: The *alpha cut* set of a fuzzy set \tilde{A} is given by:

$$A_\alpha = \{x \in \mathbb{R} / \mu_{\tilde{A}}(x) \geq \alpha\}, \quad \alpha \in [0, 1]$$

For a fuzzy set with a triangular type membership function (see Figure 1) the alpha cut is given by: $A_\alpha = [a_1 + \alpha(a_2 - a_1), a_3 - \alpha(a_3 - a_2)]$

The alpha cut is perhaps the most important concept of fuzzy sets, because by adjusting the value α , the range or set of values that satisfy a certain degree of belonging, the level of satisfaction, precision of the result or robustness of the model can be determined [12].

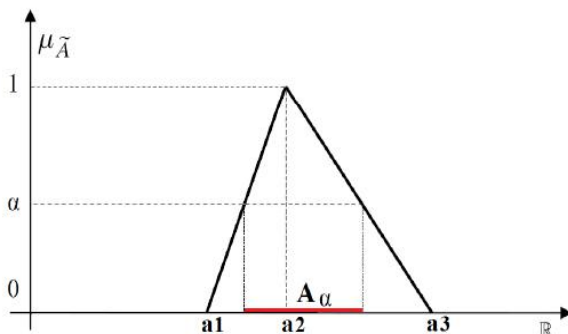


Figure 1. Triangular fuzzy Number.

Theorem 1: (Representation theorem [13]) If \tilde{A} is a fuzzy set and A_α its alpha cuts, $\alpha \in [0, 1]$, it is verified that:

$$\tilde{A} = \bigcup_{\alpha \in [0,1]} \alpha A_\alpha$$

This formal notation should be understood as the equality between the membership functions of both sets, where

$$\mu_{A_\alpha}(x) = \begin{cases} 1 & \text{si } y \text{ solo si } x \in A_\alpha \\ 0 & \text{otherwise} \end{cases}$$

On the other hand,

$$\mu_{\tilde{A}}(x) = \sup_{\alpha \in [0,1]} \min(\alpha, \mu_{A_\alpha}(x))$$

2.2. Interval arithmetic operations

Definition 6: The four basic arithmetic operations in closed intervals are: addition, difference, multiplication and division, they are defined as follows [14]:

$$[a, b] + [c, d] = [a + c, b + d]$$

$$[a, b] - [c, d] = [a - d, b - c]$$

$$[a, b] \cdot [c, d] = [\min \{ac, ad, bc, bd\}, \max\{ac, ad, bc, bd\}]$$

$$\frac{[a,b]}{[c,d]} = [\min \left\{ \frac{a}{c}, \frac{a}{d}, \frac{b}{c}, \frac{b}{d} \right\}, \max \left\{ \frac{a}{c}, \frac{a}{d}, \frac{b}{c}, \frac{b}{d} \right\}] , \text{ provided that } 0 \notin [c, d]$$

2.3. Expected Interval of a Fuzzy Number

Given a fuzzy number \tilde{N} , if we denote by $[N_\alpha^L, N_\alpha^R]$ with $\alpha \in [0,1]$ the sets of level α or α - cuts, we define the expected interval of \tilde{N} by [15].

$$EI(\tilde{N}) = \left[\int_0^1 N_\alpha^L d\alpha , \int_0^1 N_\alpha^R d\alpha \right]$$

The expected interval of a fuzzy number is an interval that concentrates the best information of a fuzzy number.

2.4. Fuzzy Linear Equation

The imprecision of the parameters of the equation, we will represent it with TFN of the form $\tilde{A} = (a_1, a_2, a_3)$, that is, the value of the parameter with the most possibility of occurrence is a_2 but it can occur between a_1 and a_3 .

Given the TFN $\tilde{A} = (a_1, a_2, a_3)$, $\tilde{B} = (b_1, b_2, b_3)$ and $\tilde{C} = (c_1, c_2, c_3)$, then the FLE is defined by:

$$\tilde{A}\tilde{X} + \tilde{B} = \tilde{C} \tag{3}$$

One of the methods proposed to solve this equation was that of alpha cut and interval arithmetic [7]. To this method we incorporate the concept of the expected interval.

2.5. Alpha Cut Method and Interval Arithmetic + Expected Interval

This method consists of obtaining the solution from a classical linear equation, that is, without imprecision, the alpha cuts of the fuzzy numbers representing the imprecise parameters are then replaced, then simplified by interval arithmetic [2].

Given the equation (3) and assuming that $0 \notin \text{Sop}(\tilde{A})$, to find the solution, follow these steps:

Step 1: Determine the cut of alpha coefficients imprecise entities: \tilde{A} , \tilde{B} y \tilde{C} .

Step 2: Replace the alpha cuts in the equation: $\tilde{x} = \frac{\tilde{c}-\tilde{b}}{\tilde{a}}$

$$x(\alpha) = \frac{c[\alpha] - b[\alpha]}{a[\alpha]} = \frac{[c_1(\alpha), c_2(\alpha)] - [b_1(\alpha), b_2(\alpha)]}{[a_1(\alpha), a_2(\alpha)]}$$

Step 3: Simplify using the definitions of interval arithmetic.

The interval difference is made

$$x(\alpha) = \frac{[c_1(\alpha) - b_2(\alpha), c_2(\alpha) - b_1(\alpha)]}{[a_1(\alpha), a_2(\alpha)]}$$

And then the interval division.

$$x1(\alpha) = \min \left\{ \frac{c_1(\alpha) - b_2(\alpha)}{a_1(\alpha)}, \frac{c_1(\alpha) - b_2(\alpha)}{a_2(\alpha)}, \frac{c_2(\alpha) - b_1(\alpha)}{a_1(\alpha)}, \frac{c_2(\alpha) - b_1(\alpha)}{a_2(\alpha)} \right\}$$

$$x2(\alpha) = \max \left\{ \frac{c_1(\alpha) - b_2(\alpha)}{a_1(\alpha)}, \frac{c_1(\alpha) - b_2(\alpha)}{a_2(\alpha)}, \frac{c_2(\alpha) - b_1(\alpha)}{a_1(\alpha)}, \frac{c_2(\alpha) - b_1(\alpha)}{a_2(\alpha)} \right\}$$

$$\forall \alpha \in [0,1]$$

The solution of (3) with a degree of precision $\alpha \in [0,1]$, is an interval $x(\alpha) = [x_1(\alpha), x_2(\alpha)]$; these intervals are the alpha sets of the fuzzy number (not necessarily a TFN) that will be the solution of the (3) [14].

According to Theorem 1 (see Subsection 2.1), the number that is the solution of the FLE is given by:

$$\tilde{S} = \bigcup_{\alpha \in [0,1]} \alpha \mu_{x(\alpha)}$$

where $\mu_{x(\alpha)}$ is the membership function of $x(\alpha)$.

The most used values are $\alpha = \{0.0, 0.1, 0.2, \dots, 0.9, 1.0\}$; $\alpha = 0$ gives the most robust (most imprecise) solution; $\alpha = 1$ gives the most precise (exact) solution.

The concept of the expected interval is incorporated into the alpha cut method.

Step 4: Calculate the expected interval of $x(\alpha)$

$$IE(\tilde{S}) = [\int_0^1 x_1(\alpha) d\alpha, \int_0^1 x_2(\alpha) d\alpha,]$$

In this interval, we find most solutions of the FLE. This interval corresponds to a unique $\alpha \in [0,1]$.

3. Results and Discussion

3.1. Case study: let be the linear equation (3).

Given de TFN $\tilde{A} = (1,2,3)$, $\tilde{B} = (5,6,7)$, $\tilde{C} = (7,11,15)$. To find the solution, we apply steps 1 to 4 of Section 2.5.

Step 1: We find the alpha cut of the triangular fuzzy numbers:

$$A[\alpha] = [1 + \alpha, 3 - \alpha], \quad B[\alpha] = [5 + \alpha, 7 - \alpha], \quad C[\alpha] = [7 + 4\alpha, 15 - 4\alpha]$$

Step 2: We replace the alpha cut in the solution:

$$x(\alpha) = \frac{c[\alpha] - b[\alpha]}{a[\alpha]} = \frac{[7 + 4\alpha, 15 - 4\alpha] - [5 + \alpha, 7 - \alpha]}{[1 + \alpha, 3 - \alpha]}$$

Step 3: Apply the interval difference in the numerator

$$x(\alpha) = \frac{[(7 + 4\alpha) - (7 - \alpha), (15 - 4\alpha) - (5 + \alpha)]}{[1 + \alpha, 3 - \alpha]} = \frac{[5\alpha, 10 - 5\alpha]}{[1 + \alpha, 3 - \alpha]}$$

Interval division is performed

$$x(\alpha) = [\min \{ \frac{5\alpha}{1 + \alpha}, \frac{5\alpha}{3 - \alpha}, \frac{10 - 5\alpha}{1 + \alpha}, \frac{10 - 5\alpha}{3 - \alpha} \}, \max \{ \frac{5\alpha}{1 + \alpha}, \frac{5\alpha}{3 - \alpha}, \frac{10 - 5\alpha}{1 + \alpha}, \frac{10 - 5\alpha}{3 - \alpha} \}]$$

$$x(\alpha) = [\frac{5\alpha}{3 - \alpha}, \frac{10 - 5\alpha}{1 + \alpha}] \quad , \quad \alpha \in [0,1]$$

The solution to the fuzzy linear equation is the fuzzy number,

$$\tilde{S} = \bigcup_{\alpha \in [0,1]} \alpha x(\alpha)_{\alpha} = \bigcup_{\alpha \in [0,1]} \alpha \left\{ \begin{array}{l} 1 \quad x \in [\frac{5\alpha}{3 - \alpha}, \frac{10 - 5\alpha}{1 + \alpha}] \\ 0 \quad \text{otherwise} \end{array} \right.$$

$$\tilde{S} = \bigcup_{\alpha \in [0,1]} \left\{ \begin{array}{l} \alpha \quad x \in [\frac{5\alpha}{3 - \alpha}, \frac{10 - 5\alpha}{1 + \alpha}] \\ 0 \quad \text{in another case} \end{array} \right.$$

The graph of the fuzzy number $\tilde{S} = \tilde{S}_L \cup \tilde{S}_R$ where

$$\tilde{S}_L = \left\{ \left(\frac{5\alpha}{3 - \alpha}, \alpha \right), \alpha \in [0,1] \right\}, \quad \tilde{S}_R = \left\{ \left(\frac{10 - 5\alpha}{1 + \alpha}, \alpha \right), \alpha \in [0,1] \right\}$$

The intersection point of the curves \tilde{S}_L y \tilde{S}_R occurs at: $\frac{5\alpha}{3 - \alpha} = \frac{10 - 5\alpha}{1 + \alpha}$

Solving for α you have $\alpha = 1$, which corresponds to a value of $x = \frac{5(1)}{3-1} = 2.5$

The graph of the fuzzy set \tilde{S} can be seen in Figure 2. The fuzzy number \tilde{S} in Cartesian form is given by

$$\tilde{S} = \begin{cases} \frac{3x}{5+x}, & 0 < x \leq 2.5 \\ \frac{10-x}{5+x}, & 2.5 \leq x < 10 \\ 0, & x \leq 0 \text{ and } x \geq 10 \end{cases}$$

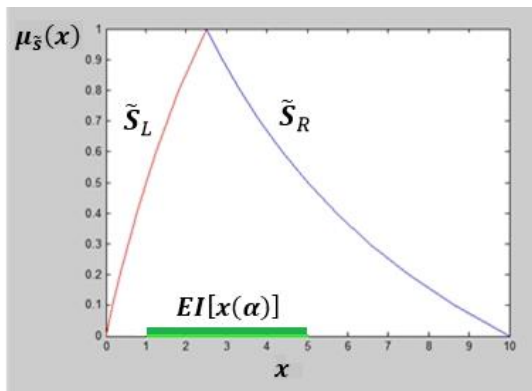


Figure 2. Solution of the first-degree linear fuzzy equation.

The fuzzy set is a fuzzy number (See Definition 1). The solution of the FLE will most likely occur at 2.5, but it can occur between 0 and 10.

The solution of the FLE with different degrees of precision is shown in Table 1.

The most robust (most imprecise) solution is the interval [0, 10], obtained with $\alpha = 0$. The most precise (exact) solution is the interval [2.50, 2.50], obtained with $\alpha = 1$; this interval is the real number $x = 2.5$. For $\alpha = 0.5$, the solution is the interval [1, 5]. In other words, you have different alternatives to choose the solution interval with the precision you want.

Table 1. Alfa cuts for the solution of the fuzzy linear equation of the first degree

Precision degree α	Alpha Cuts
0.0	[0, 10]
0.1	[0.17, 8.64]
0.2	[0.36, 7.50]
0.3	[0.56, 6.54]
0.4	[0.77, 5.71]
0.5	[1.00, 5.00]
0.6	[1.25, 4.38]
0.7	[1.52, 3.82]
0.8	[1.82, 3.33]
0.9	[2.14, 2.89]
1.0	[2.50, 2.50]

Step 4: Evaluate the expected interval of \tilde{S}

$$EI[\tilde{s}] = \left[\int_0^1 \frac{5\alpha}{3-\alpha} d\alpha, \int_0^1 \frac{10-5\alpha}{1+\alpha} d\alpha \right] = [1.081, 5.317]$$

The solution intervals with different degrees of precision $\alpha \in [0,1]$ are infinite, if an interval where the majority of solutions, of the fuzzy linear equation possibly occur is desired, this is the expected interval [1,081, 5,317] (see Figure 2).

3.2. Comparison of Solutions

For $\alpha = 1$ (total precision)

The solution to (3) with the method of alpha cutting and interval arithmetic is the TFN $\tilde{X} = (2.5, 2.5, 2.5)$, replacing in (3) and performing interval operations you get $11 = 11$, i.e. the equation is satisfied.

For $\alpha = 0$ (maximum inaccuracy)

The solution to equation (3) with the method of alpha cut and interval arithmetic is the fuzzy number $\tilde{X} = (0, 2.5, 10)$; replacing in (3) and performing interval operations gets $[5, 37] \neq [7, 15]$, that is, the equation is not satisfied, and what is worse, the interval of the first member is very different from that of the second member.

The solution to (3) incorporating the concept of the expected interval to the method of alpha cutting and interval arithmetic, is the fuzzy number $(1,081, 2.5, 5.317)$; replacing in (3) and performing interval operations results in $[6.0811, 22.951] \neq [7, 15]$, which also does not satisfy the equation, but is closer to equality than the solution obtained with the alpha cuts and interval arithmetic method.

It is noted that for $\alpha = 0$, the expected interval $IE(\tilde{s}) = [1.081, 5.317] \subset [0,10]$, and decreases the inaccuracy, which is the difficulty of the solution by arithmetic interval according to the ratio (2).

4. Conclusions

The new method of solving a fuzzy linear equation of a variable using the expected interval, allows to obtain a smaller support set where the solutions come closer to satisfying the fuzzy linear equation, also allows to find a single interval where the best solutions for decision making are expected to be found.

This method can be applied to problems where inaccuracy can be represented by triangular fuzzy or trapezoidal numbers.

It is recommended to study the incorporation of the concept of the expected interval in the methods to solve systems of fuzzy linear equations.

References

- [1] Zadeh LA. Fuzzy sets. Information and control. 1965, 8(3):338-353.
- [2] Abbasbandy S, Otadi M, Mosleh M. Minimal solution of general dual fuzzy linear systems, Chaos, Solitons and Fractals. 2008. 37: 1113-1124.
- [3] Ghomasshi A, Salahshour S, Hakimzadeh A. Approximating solutions of fully fuzzy linear systems: A financial case study. Journal of Intelligent & Fuzzy Systems. 2014. 26: 367-378.

- [4] Skalna I, Rao MR, Skalna I, Rao MR, Pownuk A. Systems of fuzzy equations in structural mechanics. *Journal of computational and Applied Mathematics*. 2008. 218(1): 149-156.
- [5] Wu CC, Chang NB. Grey input–output analysis and its application for environmental cost allocation, *European Journal of Operational Research*. 2002. 145:175–201.
- [6] Sevastjanov P, Dymova L. A new method for solving interval and fuzzy equations: linear case. *Information Sciences*. 2009. 179(7): 925-937.
- [7] Buckley JJ, Eslami E, Feuring T. Solving fuzzy equations. En *Fuzzy mathematics in economics and engineering*. Physica, Heidelberg. 2002:19-44.
- [8] Allahviranloo T, Babakordi F. Algebraic solution of fuzzy linear system as: $AX + BX = Y$. *Soft Computing*. 2017. 21(24): 7463-7472.
- [9] Lodwick WA, Dubois D. Interval linear systems as a necessary step in fuzzy linear systems. *Fuzzy sets and systems* 281. 2015: 227-251.
- [10] Chakraverty S, Behera D. Fuzzy system of linear equations with crisp coefficients. *Journal of Intelligent & Fuzzy Systems*. 2013. 25(1): 201-207.
- [11] Dubois D, Prade H. *Fuzzy sets and systems: theory and application*. Academic Press, New York. 1980.
- [12] Gutierréz F. et al. Un modelo de optimización difusa para el problema de atraque de barcos. *Investigación operacional*. 2018. 38(2): 160-168.
- [13] Negoită CV, Ralescu DA. *Applications of fuzzy sets to systems analysis*. Basel, Switzerland: Birkhäuser, 1975: 187.
- [14] Klir GJ, Yuan B. *Fuzzy sets and fuzzy logic*. Prentice hall. 1995.
- [15] Heilpern S. The expected value of a fuzzy number. *Fuzzy sets and Systems*, 1992. 47(1):81-86.

Analysis of the Software Most Used by Hackers to Carry Out Penetration Testing in Public Organizations

Segundo Moisés Toapanta TOAPANTA^{a1}, Raúl Francisco Pérez GONZÁLEZ^a, Máximo Giovani Tandazo ESPINOZA^a and Luis Enrique Mafla GALLEGOS^b

^aDepartment Computer Science, Salesian Polytechnic University of Ecuador (UPS), Chambers 227, Ecuador

^bFaculty of Systems Engineering, National Polytechnic School (EPN), Ladrón the Guevara E11-253, Ecuador

Abstract. Currently, the software handled by hackers is the main one to tackle a series of empirical knowledge, with this software attacking and helping organizations. The main objective is to analyze and systematize the software that is detected by hackers and crackers, in order to prevent risks and study the tactical levels and strategies for a given process. The analytical method is used in this investigation, for the study or analysis of the offensive software structure in public organizations. The results obtained from this research were an attack launching algorithm, software prototype taken by hackers, massive obfuscation model, and quantitative encryption model. It was concluded that piracy tools are used for preventive prevention and systems aggression, that is, to be defensive or offensive for a period, throughout an attack cycle.

Keywords. hacker, hacking tools, hacking, malicious program, public organizations.

1. Introduction

Within the penetration processes towards organizations or establishments of each type are based on a consecutive row of manipulation techniques for each specific case, these courses have enlarged many organizations[1]. Organizations choose hackers to manipulate these events in neighboring establishments and thus begin a launch of competitions between organizations or companies, they are generally having a very high level of appointment, while hackers are the most benefited today for its great ability to manipulate technology[2]. Old penetration software has been very useful for attack situations, exploits in a more practical way for individuals who know about these techniques and their vulnerabilities[3]. The software may have poorly developed code, which will allow ineffective functionality within an attack process, it is recommended that a check be performed before starting or violating a system, this refers to the high vulnerability of mobile phones[4]. Tracking the signals by means of an attack between two distant nodes, it will be possible to carry out a respective analysis on the different

¹Corresponding Author: Segundo Moisés Toapanta Toapanta, Department Computer Science, Salesian Polytechnic University of Ecuador (UPS), Chambers 227, Ecuador. Email:stoapanta@ups.edu.ec

types of software that come onto the market, due to automation and execution problems within a software, a large part of the structure will be found running on one level. low[5]. For the processes of a comparative evaluation to the tests that these software undergo, they have a lot to do with the hardware parts, as indicated by the strategies of the components retracted by the intangible processes of each stage in their permanence[6].

As a result of the above analysis, the following possible risks can be concluded within public organizations by each hacking software.

- Public institution provides access control at occupational levels.
- The security that is equated is extremely low, giving continuity to an incidence management.
- Physical preparation is mitigated by risk management presented in processes and activities of public services.

Why is it necessary to analyze penetration software in public organizations?

It is necessary to analyze the software, because there is a high risk of violation creating malicious functions that will reproduce and stagnate within a system.

The main objective is to analyze and systematize the software that is obtained by hackers and crackers, in order to prevent risks and study the tactical levels and strategies for a certain process.

The reviewed articles are: PsyBoG: A scalable botnet detection method for large-scale DNS traffic[1], An application of linear algebra theory in networked control systems: Stochastic cyber-attacks detection approach[2], Structural analysis of packing schemes for extracting hidden codes in mobile malware[3], An Empirical Evaluation of the Effectiveness of Attack Graphs and Fault Trees in Cyber-Attack Perception[4], Reinforcement learning for efficient network penetration testing[5], Hybrid detection of intermittent cyber-attacks in networked power systems[6], Effect of anti-malware software on infectious nodes in cloud environment[7], An Approach to Optimize the Management of Information Security in Public Organizations of Ecuador[8], A Malicious Web Site Identification Technique Using Web Structure Clustering[9], Andro-Dumpsys: Anti-malware system based on the similarity of malware creator and malware centric information[10], Analysis and Findings of Social Engineering Industry Experts Explorative Interviews: Perspectives on Measures, Tools, and Solutions[11], Smart Contract: Attacks and Protections[12], Security Verification via Automatic Hardware-Aware Exploit Synthesis: The CheckMate Approach[13], Detection and elimination of spyware and ransomware by intercepting kernel-level system routines[14], A Comprehensive Overview of Government Hacking Worldwide[15], Definition of parameters to perform audit in cybersecurity for public one organization of ecuador[16], Creation of a DDOS attack using HTTP-GET Flood with the Cyber Kill Chain methodology[17].

The analytical method is used in this investigation, for the study or analysis of the offensive software structure in public organizations.

The results obtained from this research are disclosed in each structural phase that is developed with the hacking tools, these tools are expanded by each system that composes it.

It is concluded that the phases revealed with piracy software are part of a rigorous prevention and exploitation of computer processes in public organizations, as well as the security and management of a business structure.

2. Materials and Methods

The research resources and treatments are based on the scientific articles investigated, with a sequential aspect in their results.

2.1. Materials

It is considered a formal analysis taking into account the research references that are presented with the research material.

It was verified that for a meticulous penetration environment it is specified that, within the programming of the software, the implementation of a subsystem is necessary in order to integrate remote access within the social and strategic localities, such as public organizations[1]. Hackers evaluate their network before launching an attack, with the respective software they manipulate the functions of the public organization side, apart from quantitatively analyzing each and every risk within their own network, engineers continually determine the flaws of these software, starting from a structural analysis to give the appropriate answers to all the systems that surround the organization or institution[2]. Domain threats, it is required for most security within dynamic routes, improving the resistance of self-defense security is the most useful way for a static software course. Hackers run in parallel the critical analyzes your infrastructure needs when planning an assault on your public organization[3]. By performing different types of algorithms, it would help the software establish a design protocol for the analysis of vulnerabilities that have been run in the system, this helps not to decrease the development or progress that was established in the regulations[4]. The software and techniques that hackers have used in many public and private organizations. These have been very helpful for the security of the organization[5]. Indirect penetrations are observed indirectly, since their components are similar to the functions of attacks with defects to their launches[6]. It is analyzed that, with the technology and tools specified for a certain piracy purpose, it is advisable to study the receiving points[7]. Given that in a computer attack it is intentional and it is not possible to consider all the counter attack scenarios, as much as possible that the target is declaring a tracking attack in self-defense, as well as activating an alarming signal where it has defense support[8].

Social-engineer was used software with the participation of Wireshark, an extended process of exploit launches occurs for full access of internal nodes through social engineering[9]. It should be recognized that the level of compatibility with the end user is unavoidable for this analysis, without this knowing what would happen within the source services in their environment[10]. The participation of tcp, http, udp and dns traffic is classified and detected, by identifying them within the network, producing better precision and the data transfer rate improving[11]. Using the Hoffman code, it will be possible to understand the valuation of the data received, if the strict way of the exact values of that information will be represented[12]. In itself, the use of software in a public organization is automated in its application profile since employees do not have the corresponding knowledge in coding[13]. The permanencies of the current states of the software operations are:

1. The probability of software decay with its encoding is low.
2. The standard that is set to the software has a length of responses the same as the arrival information.
3. Diagram the coding structure must be probable.

Also to the states, the servers and / or complete infrastructure of the public organization will be verified [14]. Sharing a remote connection to this destination, we proceed to investigate areas and check each point inside and outside the organization. [15]. A method that is very misleading for individuals with a lack of knowledge is analyzed when entering a website, which is classified as adware [16]. Therefore, this method is a considerable achievement for the general staff, the function of this method is to add badware while the person navigates and fulfills the needs that are within their reach [17].

2.2. Methods

In this study the analytical method was carried out, to make known the criteria that each software process analyzed by Institute for Security and Open Methodologies (ISECOM).

2.2.1. Manipulation of Hackers

To carry out the secret writing of a system or a file in general, one must focus on its encryption, developing an difficult function for the user in general, with this an initial space is made in the encoding of messages, where it requests the different characters for a secure structure within a system. In cryptographic systems the vast majority are made up of different settings, among them we have:

- Transmission of the encrypted message.
- Algorithms to decrypt the element.
- Data transmission rate combinations.

2.2.2. Systems with Gnu / Linux

Publicize this free system, which has allowed hackers to distribute distros in the development processes for each specific task, thus achieving the operation of generating different types of files, uniform to Table 1.

Table 1. The impact of each file extension generated in the GNU / LINUX system.

Impact	
Level	Malicious Extensions
Low	.exe
Low	.pif
Half	.com
Half	.scr
Half	.sys
Half	.ovl
High	.drv
High	.bin
High	.dll
High	.bat

2.2.3. The Adware

Assimilating Table 2, the function for the structure of a malicious program is applied in the presence of browsing a website. This program can be hidden as a backdoor, without disclosing the structure to the end user.

Table 2. Ability level to denote the structure of a malicious program.

Technique in a Malicious Program		
Level	Structure	Capacity
1	Directives	Very low
2	Main Features	Low
3	Coding	Low
4	Instruction	Half
5	Organization Loops	High
6	Root Element	Very High
7	Depuration	Critical

2.2.4. Using the Web Audit

To get vulnerabilities (XSS) complying with your command chain. On sites crossed with customer penetration on that website.

- Attack target vulnerable points of (DOM).
- Execution of the attacker script by the client.
- Theft of cookies.
- Inject (SCRIPTS JAVA, SCRIPTS CODE) into the website.

3. Results

Below are the phases acquired for the analysis of the software most used by hackers, these are:

- Attack Launch Algorithm.
- Prototype of Software taken by Hackers.
- Massive Obfuscation Model.
- Quantitative Encryption Model.

3.1. Attack Launch Algorithm

In the algorithm of Figure 1, the presence of attack launches to a given medium can be manifested, to intercept security rings and penetrate the secondary systems of public organizations and spread their coding in them.

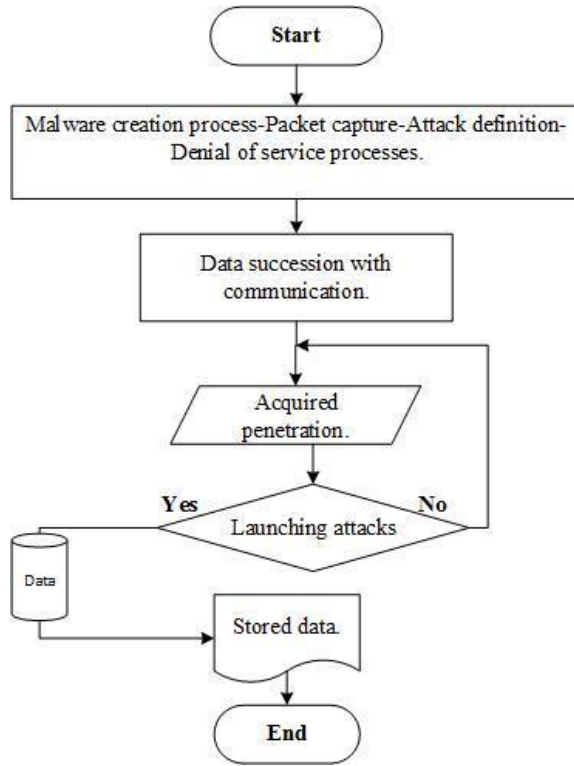


Figure 1: Algorithm for launching attacks on a secondary node.

3.2. Software prototype taken by Hackers

In this prototype the main computer security tools on the side of a hacker are enunciated, knowing the status of the Manual of the Open Method of Security Verification (OSSTMM) (Figure 2).

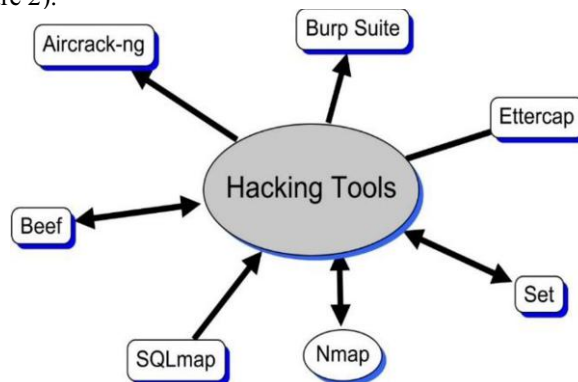


Figure 2: Prototype of Software taken by hackers.

3.3. Massive Obfuscation Model

In this equation a data template is expressed, with which information can be hidden within a process without limits to the amount of creation, data and malicious information of the software, thus giving rise to the exploitation knowledge of the Method Penetration Test (ISSAF).

$$x(C)r \approx \frac{x!}{(x-r)!r!} \in z...n \quad (1)$$

Where: Desired number of data (x), combinations of information taken from the data (Cr).

3.4. Quantitative Encryption Model

In this quantitative encryption model, a process restart is performed for encrypted attacks, which occurs on an attack path between two nodes (primary, secondary).

$$R_0 \approx f(x_0) \in z_1 \rightarrow f(R_1)_2 \quad (2)$$

Where: attack numbers (R), Primary node (f), Secondary node (z), Denial processes.

4. Discussion

The triumphs achieved for this investigation are related to the means of software transmission through an attack zone, in such a way that they can obfuscate or carry out the respective assault launches, both to prevent and attack an organization and perform these acts on an impassive system[13].

The information in this research is not limited to learning actions or facts, generally based on computer attacks using cybersecurity software.

The results contracted by other authors and researchers in this research were pleased by techniques and tactics of software management and executions inside and outside public institutions or organizations.

It was concluded that any result obtained with piracy software can be attacked or prevented by an act of cybersecurity in a public organization.

5. Future work and conclusion

Given the bad information that is extracted by hackers, it is advisable to carry out an intense test on hacking software, it should be noted that the software can leave signs or can carry out searches in the primary software.

It is concluded that the dominated results are ways of preventing an unwanted act before an organization, placing yourself on the side of the individual who defends or mitigates the supernatural laws of technology.

6. Acknowledgments

The authors thank to Salesian Polytechnic University of Ecuador, to the research group of the Guayaquil Headquarters “Computing, Security and Information Technology for a Globalized World” (CSITGW) created according to resolution 142-06-2017-07-19 and Secretaría de Educación Superior Ciencia, Technology e Innovation (Senescyt).

References

- [1] Kwon J, Lee J, Lee H, Perrig A. PsyBoG: A scalable botnet detection method for large-scale DNS traffic. *Comput. Networks*, 2016, vol. 97, pp. 48–73.
- [2] Li Y, Voos H, Darouach M, Hua C. An application of linear algebra theory in networked control systems: Stochastic cyber-attacks detection approach, *IMA J. Math. Control Inf.*, 2016 33(4), 1081–1102.
- [3] Lim J and Yi J H. Structural analysis of packing schemes for extracting hidden codes in mobile malware. *Eurasip J. Wirel. Commun. Netw.*, 2016, 2016(1).
- [4] Lallie H S, Debattista K, Bal J. An empirical evaluation of the effectiveness of attack graphs and fault trees in cyber-attack perception, *IEEE Trans. Inf. Conf., Forensics Secur.*, 2018, 13(5): 1110–1122.
- [5] Ghanem MC, Chen TM. Reinforcement learning for efficient network penetration testing, *Proc 3rd Int. Conf.*, 2020, 11(1): 1–23.
- [6] Kontouras E, Tzes A, Dritsas L. Hybrid detection of intermittent cyber-attacks in networked power systems, *Energies*, 2019, 12(24): 1–29.
- [7] Abazari F, Analoui M, Takabi H. Effect of anti-malware software on infectious nodes in cloud environment. 2016., *Int. Conf. Comput. Secur.*, 58: 139–148.
- [8] Moisés Toapanta Toapanta S, Enrique Mafla Gallegos L. An approach to optimize the management of information security in public organizations of Ecuador, *CITS 2019- 2019 Int. Conf. Comput Inf Telecommun*, 2020.
- [9] Nagai T, Kamizono M, Shiraishi Y, Member S. A malicious web site identification technique using web structure clustering, 9: 1665–1672, 2019.
- [10] Jang JW, Kang H, Woo J, Mohaisen A, Kim HK. Andro-Dumpsys: Anti-malware system based on the similarity of malware creator and malware centric information, 2015 *Int. Conf. Comput. Secur.*, 58: 125–138, 2016.
- [11] Aldawood H, Skinner G. Analysis and findings of social engineering industry experts explorative interviews: perspectives on measures, tools, and solutions. *IEEE Access*, 2020, 8: 67321–67329.
- [12] Sayeed S, Marco-Gisbert H, Caira T. Smart contract: attacks and protections. *IEEE Access*, 2020, 8: 24416–24427.
- [13] Trippel C, Lustig D, Martonosi M. Security verification via automatic hardware-aware exploit synthesis: the checkmate approach. *IEEE Micro*, 2019, 39(3): 84–93.
- [14] Javaheri D, Hosseinzadeh M, Rahmani AM. Detection and elimination of spyware and ransomware by intercepting kernel-level system routines. *IEEE Access*, 2018, 6: 78321–78332.
- [15] Li CY, Huang CC, Lai F, Lee SL, and Wu J. A comprehensive overview of government hacking worldwide. *IEEE Access, Int. Conf.* 2018, 6:55053–55073.
- [16] Toapanta SMT, Peralta NA, Gallegos LEM. Definition of parameters to perform audit in cybersecurity for public one organization of ecuador, *ACM Int. Conf. Proceeding Ser.*, no. 2019, pp. 91–96.
- [17] Martínez-Lozano JE, Atencio-Ortiz PS. Creation of a DDOS attack using HTTP-GET Flood with the Cyber Kill Chain methodology, *Iteckne, Int. Conf.* 2019, 16(1): 41–47.

Deep Robot-Human Interaction with Facial Emotion Recognition Using Gated Recurrent Units & Robotic Process Automation

Suchitra SAXENA¹, Shikha TRIPATHI and Sudarshan TSB
Faculty of Engineering, PES University, Bangalore, India

Abstract. This research work proposes a Facial Emotion Recognition (FER) system using deep learning algorithm Gated Recurrent Units (GRUs) and Robotic Process Automation (RPA) for real time robotic applications. GRUs have been used in the proposed architecture to reduce training time and to capture temporal information. Most work reported in literature uses Convolution Neural Networks (CNN), Hybrid architecture of CNN with Long Short Term Memory (LSTM) and GRUs. In this work, GRUs are used for feature extraction from raw images and dense layers are used for classification. The performance of CNN, GRUs and LSTM are compared in the context of facial emotion recognition. The proposed FER system is implemented on Raspberry pi3 B+ and on Robotic Process Automation (RPA) using UiPath RPA tool for robot human interaction achieving 94.66% average accuracy in real time.

Keywords. Facial Emotion Recognition, Deep learning; CNN, GRUs, Raspberry Pi II, Robotic Process Automation platform

1. Introduction

One of the prominent robotics research areas is the design of intelligent robots that can communicate and act as a human companion. With rapid advancements in hardware, computer graphics, robotic technology and artificial intelligence, more and more cobots and social robots have been designed. For a robot to act as a human companion, it should have emotional intelligence for effective human interaction. Emotions include cognitive evaluation, body language, expressions and feelings [1]. Emotion recognition is fascinating and a challenging task. Text, speech, facial expressions, biological signals and gestures can be used to recognize human emotions. Facial expressions have significant importance in the identification of human emotions in direct human to human interaction [2]. Researchers and engineers have attempted to design artificial intelligence frameworks, which are cognitively and/or physically similar to human behavior. The increase in computational power since a decade has largely contributed to the development of fast learning machines. In addition, the internet has generated a substantial amount of training data. These two development triggered research into smart

¹ Corresponding Author: Suchitra Saxena, Faculty of Engineering PES University, Bangalore, India; E-mail: suchitra@pes.edu.in, suchitrasaxena10@gmail.com

self-learning systems, with one of the most successful emerging techniques being deep learning networks. Progress in robot-human interaction over the past decade has contributed to several applications in robotic technology where robots need to comprehend human actions and emotions. With emotional intelligence robots, human action can be better predicted and performance can be increased in many applications such as Human Robot Interaction (HRI) for kid's therapy with autism and attention deficit hyperactivity disorder (ADHD), driver awareness alerting system, legal disciplines, medical guides, E-Learning, psychology and entertainment feedback system, emotional support as socially assistive robots for kids [3-8]. Developing an algorithm, which can identify emotions from facial images, is therefore desirable to improve HRI. The advancement of automated processes, tools such as the Robotic Process Automation (RPA) have been successful in many fields in improving operational accuracy and performance. RPA is the implementation of automated robots or bots that use artificial intelligence techniques to automate repetitive tasks in real time. Also, implementation of the proposed architecture on the RPA platform would be relevant for HRI or Industry 4.0 applications [9].

In this work, a robust and efficient GRUs Facial Emotion Recognition System (GFERS) is proposed for recognizing emotions from facial expressions. In real time, the proposed method achieves 94.66% accuracy. CMU Multi-PIE [10] and FER 2013 databases [11] are used for training. GFERS is developed using GRUs, which is found to be efficient and robust in recognizing emotions under constraints such as head pose, illumination variations and age differences. The proposed system is deployed on Raspberry Pi3 B+ [12] and also implemented on UiPath robotic process automation tool which can be used in HRI applications. The performance of the proposed system is also compared with CNN and LSTM based facial emotion recognition.

The paper overview is as follows: In Section II related work and contributions are discussed, Section III describes proposed architecture of GFERS, Section IV explains results and analysis. Conclusion and future directions are discussed in Section V.

2. Related Work

Since several years, considerable work on recognition of emotions for HRI applications has been reported. Deep learning techniques, Convolution Neural Networks (CNN) and Recurrent Neural Networks (RNN) as applied to facial emotion recognition have been explored in recent years. In 2014, Jung H, et al., used CNN to recognize emotion in real time. They have used three convolution layers and two fully connected (FC) layers. They have used CK+ and FER-2013 databases for training and achieved 72.78% & 86.54% respectively [13]. In 2017, Chen et al. also used CNN with four convolution layers and two FC layers. They have used CK+ and JAFFE databases [14] and achieved 98.15% recognition accuracy for frontal pose. In [15], Saxena et al. proposed an algorithm based on CNN with three convolution layers and two FC layers with batch normalization to address overfitting. They have used Softmax classifiers and CMU MultiPIE databases for training. They have achieved average accuracy of 95.8% with pose and illumination variations in real time with 25 subjects. In [16], Baddar et al. used LSTM to address the influence of mode variability on the encoded spatio-temporal attributes. They have shown that LSTM encoded spatio-temporal features and retains a bias due to different variations such as illumination variation or pose variations by using static sequences. They have used Oulu-CASIA, AFEW and KAISI facial expression datasets and achieved

85.185, 51.44% and 84.98% accuracy respectively for these databases. In [17], Hasani et al. used an Inception-ResNet 3D CNN with LSTM to extract the spatial-spectral relationships within images and between different frames in the videos. They have used CK+, MMI, FERA, DISFA and achieved 67.52%, 54.76%, 41.93% and 40.51% respectively for head pose variations. In [18], Kim et al. used a spatio-temporal representation of a hybrid combination of CNN and LSTM to recognize facial expression and achieved 78.61% and 60.98% for MMI, CASME II databases respectively. In 2018, Yan et al. proposed a framework Joint Convolutional Bidirectional LSTM (JCBLSTM) to jointly model discriminative facial textures and spatial relationships between different regions [19]. They achieved average accuracy of 90.89% and 71.99% for different pose and illumination variations using Multi-PIE, FER-2013 databases. In 2019, Ilyas et al. proposed a hybrid CNN and LSTM model to address complexities and limitations of Traumatic Brain Injured (TBI) human-robot interaction. They used TBI-patient database, which is a collection of multimodal data annotated by physiotherapists, caretakers, experts, and doctors [20]. They have used CK+ database and achieved accuracy of 86.16% for pose and illumination variations. In [21], Deng et al. used 3D CNN algorithm for FER in videos. They used the framework of 3D Inception-ResNets structure, Stem layer, RNN special type GRU layer, Island layer, Dropout layer and Softmax layer to capture spatial relationships in facial expression images and temporal relationships between different facial frames. They have achieved average accuracies of 68.73%, 58.76% and 43.56% for CK+, MMI, AFEW databases. The proposed system is pose and illumination invariant. In [22], Li et al. proposed an HRI emotion recognition system. They first used CNN model to extract features from static images and later LSTM to find the relationship between the transformation of facial image sequences. They used CK+ database and achieved 90.51% accuracy for the frontal face. In 2019, Kang et al. proposed a CNN-RNN hybrid model. In which, first they used VGG16 to extract features from video frames and then used convolutional GRU to decode motion features [23]. They achieved 47% accuracy with the AFEW database for pose and illumination variations. In 2020, An et al. used CNN-LSTM hybrid model for feature extraction and Support Vector Machine (SVM) for classification [24]. They achieved average accuracy for pose and illumination variation of 98.9%, 99.3%, 86.6%, 87.7% and 88.3% for CK+, JAFFE, FER-2013, BU-3DFE, Oulu-CASIA respectively.

Most of the work reported in literature uses CNN and a hybrid of CNN-RNN. Mostly LSTM and GRUs are used for text classification and speech translation or classification tasks where sequence of data are used. To reduce training time and to capture temporal information LSTM and GRU can be used. Limited work is reported on LSTM based facial expression recognition, whereas use of GRUs based RNN for facial emotion recognition tasks will be more efficient. In this work, a facial emotion recognition system using RNN with GRU is proposed for facial images. We believe that using GRU for facial emotion recognition would reduce training time as compared to CNN and LSTM. The design of proposed architecture for the framework is described in the next section.

3. Proposed Architecture of GFERS

In Facial emotion recognition there are three main steps: first step is face detection, second is feature extraction and third is emotion classification. In the proposed work, face detection is obtained by using Viola-Jones face detection technique [25] and facial emotion recognition is implemented using Gated Recurrent Units (GRU) deep learning

technique. In GRUs model algorithm, GRU is used for feature extraction and followed by two dense layers for emotion classification as shown in Figure. 1. Layers details are given in Table 1.

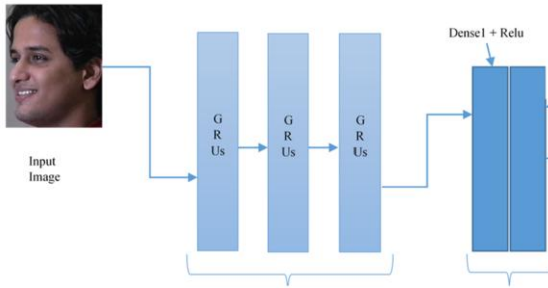


Figure 1. GFERS architecture



Figure 2. Few samples of Datasets used in GFERS

Table 1. Details of GFERS Layers

Input Image: 48x48x1	
GRUs Block	*GRU (512, Input_shape = (32,48,48), return_sequences = True)
	GRU (256, return_sequences=True)
	GRU (128, return_sequences=True)
	GRU (64, return_sequences=False)
Dense Block	Flatten: **FC1: 64 Activation function layer (ReLU)
	FC2: 6 Softmax

Notes: *GRU: Gated Recurrent Units Layer , **FC: Fully Connected layer

CMU MultiPIE and FER-2013 databases are used to train the layered model with six basic emotions (Happy, Anger, Neutral, Disgust, Sad and Surprise). The CMU MultiPIE database consists of 750,000 images of 5 basic emotions (Happy, Anger, Neutral, Disgust, and Surprise) posed by 337 subjects in different sessions. To add one more basic emotion (sad), we have used the FER-2013 database. FER-2013 database consists of 35685 images of 48x48 grayscale images with basic emotions (Happy, Sad, Neutral, Anger, Surprise, Anger and Disgust). The sample images used in training and validations are shown in Figure. 2. The algorithm of GFERS is shown in Table 2 and explained in the following section.

Table 2. GFERS Algorithm

Step1: Load training labeled data as
 $S = S^{(1)}, S^{(2)}, S^{(3)} \dots, S^{(N)}$; N is the total class

Step2: Define Layers to extract and learn features
Input image Layer (Dimensions)
GRU: {Units, Activation, Input_shape, return_seq}
GRU: {Units, Activation, return_seq}
GRU: {Units, Activation}
 ...
Dense: {Dimensional_vector}
Activation function layer
Dense: {Dimensional_vector = N }
Softmax layer

Step 3: Training process
 Set training options for training network Initial Weights $W = W_0$; Initial bias $\theta = \theta_0$
 options = Training_Options_Adamwithproperties:
 {Initial_Learning_Rate, beta_1, beta_2, Decay, MaxEpoch, MiniBatchSize}
 [network, info] = train_Network (labelled_data, layers, options)

Step4: Testing Process in real time
Capture Image (webcam)
 while (true)
 Face = detect_faces (Haar_Cascaded classifier)
 for face: Faces
 Predicted Class = classify (net, Face)
 classified_emotions (Display)
 end
 end

3.1. Feature Extraction

Gated Recurrent Units (GRU) layer is used for extracting features from images. GRUs are a new generation of Recurrent Neural Networks (RNN), which are a new gated mechanism introduced in 2014. GRU is equivalent to other RNN mechanisms like LSTM but has shown better performance on smaller datasets. GRU has only two gates; a reset gate and an update gate, they omit an output gate as shown in Figure. 3.

GRU's use the hidden state to pass information. The function of an update gate decides the set of data that can be retained for inclusion and remaining data to be excluded. The reset gate decides about how much of prior information should retain and discard. Equations for gated unit of GRU are shown below Eq. (1)-(3):

$$z_t = \sigma_g(W_z x_t + U_z h_{t-1} + b_z) \quad (1)$$

$$r_t = \sigma_g(W_r x_t + U_r h_{t-1} + b_r) \quad (2)$$

$$h_t = \{z_t \odot h_{t-1} + (1 - z_t) \odot \phi_h(W_h x_t + U_h (r_t \odot h_{t-1}) + b_h)\} \quad (3)$$

where, x_t , h_t , z_t and r_t are input, output, updated gate and reset gate vectors respectively. W , and U are parameter matrices and b vector respectively. ϕ_h and σ_g are hyperbolic tangent and sigmoid activation functions respectively. Initially, for $t = 0$, the output vector $h_0 = 0$.

In this work, after repeated experimentation with different layers of combinations and units, it was found that better training and validation accuracy could be obtained using the proposed layered architecture. In this architecture there are four GRU layers with the first layer of 512 units, second layer of 256 units, third layer of 128 units and fourth with 64 units with input image dimensions of 48x48x1. In the preprocessing stage CMU MultiPIE database is resized from 205x260x3 to 48x48x1 for five emotions

(Happy, Anger, Neutral, Disgust and Surprise) and for sad emotion FER-2013 database is used without any modification.

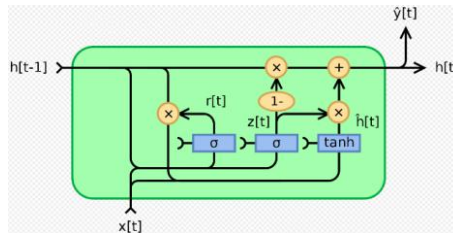


Figure 3. Gated Recurrent Unit

3.2. Classification

The GRUs output is fed to first dense layer as input for classification of emotion. The dimensional vector $N = 256$ with ReLu activation function is used for first dense layer output. The first dense layer output is given to second dense layer with $N = 6$ dimensional vector which is followed by Softmax classifier to get 6 emotion classes probability distribution using following Eq. (4):

$$S(p)_i = \frac{e^{p_i}}{\sum_{j=1}^N e^{p_j}} \quad (4)$$

where, p_i is the input vector, and since there are six emotions, p_j varies from 1,2,3...N=6.

4. Results and Analysis

In this section, GFERS performance evaluation and analysis has been discussed. 25 subjects are used for real-time performance testing, that resulted in an average accuracy of 94.66%. 2266 images per emotion were used (Happy, Neutral, Anger, Sad, Disgust, and Surprise), resulting in a total of 13596 images. The system's robustness and efficacy is measured under varying head pose and illumination conditions. It can recognize facial emotions from a distance between 0.30 m to 3 m approximately. The algorithm is implemented using Keras with Tensorflow as backend. GFERS is also implemented on RPA platform to make it easier to use in future robotic applications. The network is trained with adaptive moment estimation (Adam) with learning rate of 0.00008, a decay of 10⁻⁶ and beta1 and beta2 of 0.9 and 0.999 respectively for 359 epochs, batch size 32 and steps per epoch as 350 as shown in Table 3. The training accuracy and validation accuracy achieved by the model is 99.5% and 95.5% respectively. In real-time, the highest recognized probability value for each emotion is at-least 99%. GFERS achieved average recognition accuracy of 94.66% with 25 subjects in real-time as shown in Table 4. Single face and multi-face snapshots of results are shown in Figure 4 and 5 respectively under different illumination effect. GFERS is robust in various conditions of head pose and illumination and achieved good results in real time as shown in Figure 6. The system supported maximum head pose variation from frontal pose is $\pm 75^\circ$ in yaw. It is also observed that GFERS could not detect multiple faces in few cases. In some cases, neutral is recognized as disgust, sad and vice versa [16]. The results are shown in Figure 7. Also, the proposed system is implemented for RPA platform using UiPath RPA tool [26], which is considered one of the industry's fastest and robust solution for RPA

implementation, since it helps the robots to make certain real-time process changes depending on the requirements of the tasks. The snapshot for RPA implementation is shown in Figure 8.

Table 3. Hyper parameters for proposed system GFERS.

Hyper parameter	Values
Adam optimizer	
learning rate	0.00008
decay	10-6
beta1	0.9
beta2	0.999
Epoch	359
Batch size	32
Step size per epoch	350

Table 4. GFERS Recognition rate

Emotion	Recognition Rate (25 Subjects)
Anger	92%
Disgust	92%
Happy	96%
Neutral	96%
Sad	96%
Surprise	96%
Average Probability	94.66%



Figure 4. Results of GFERS for single face under different illumination and head pose



Figure 5. Results of GFERS for Multiple faces under different illumination and head pose



Figure 6. Results of GFERS for Pose (Maximum variation $\pm 75^\circ$ from frontal pose in yaw)



Figure 7. Failure results of GFERS

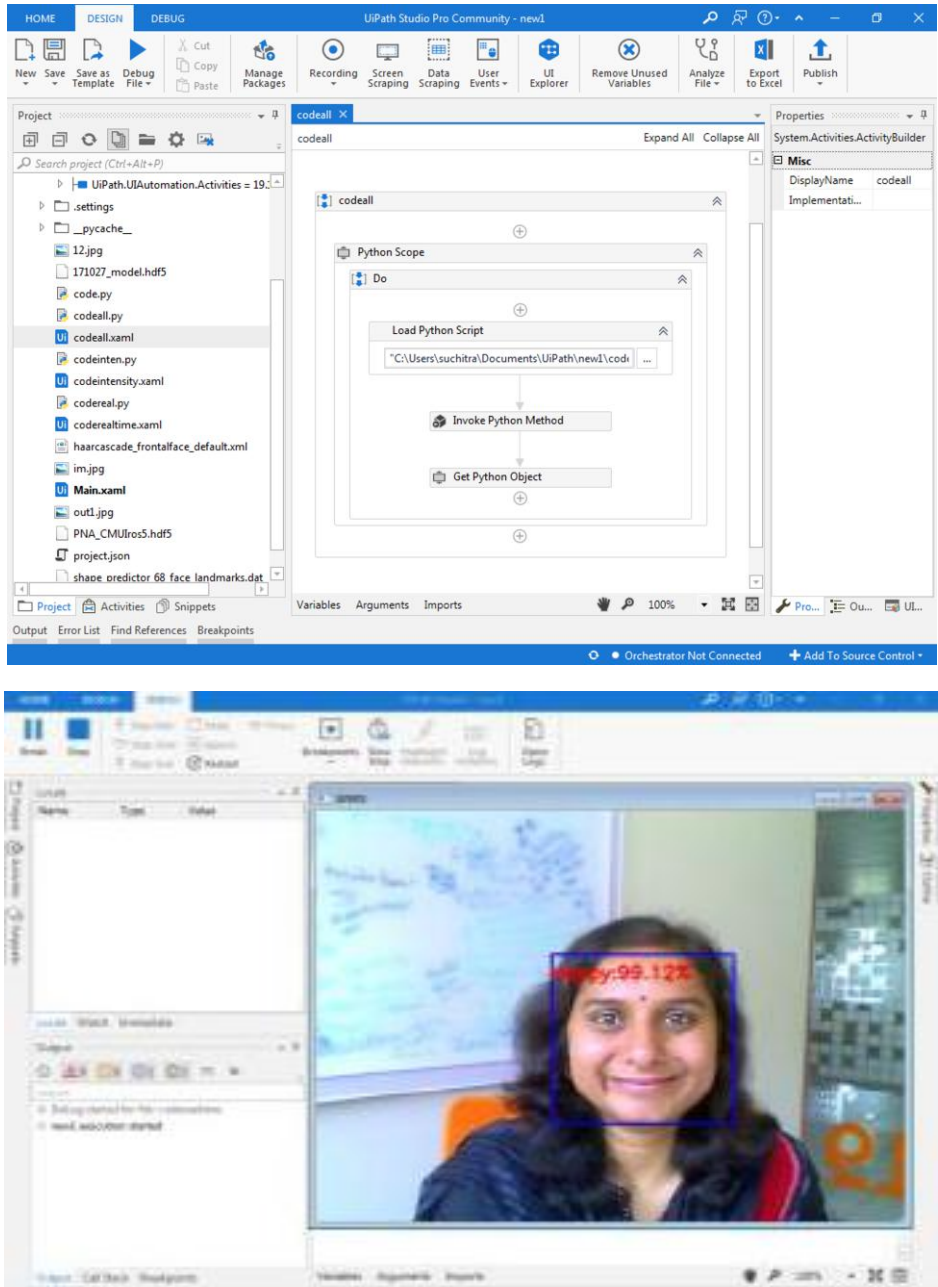


Figure 8. Snapshots of GFERS on UiPath RPA tool

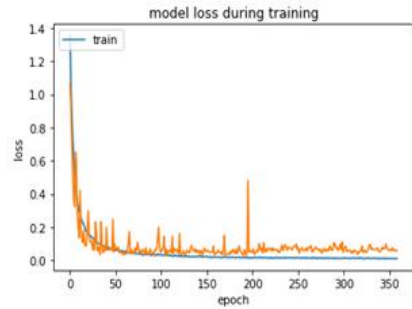
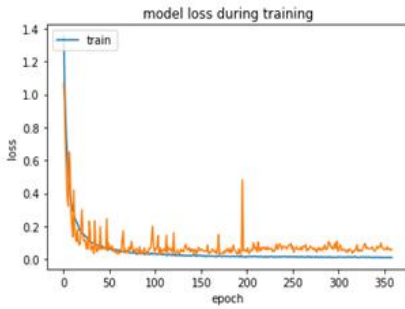
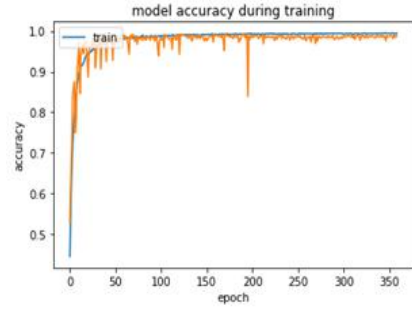
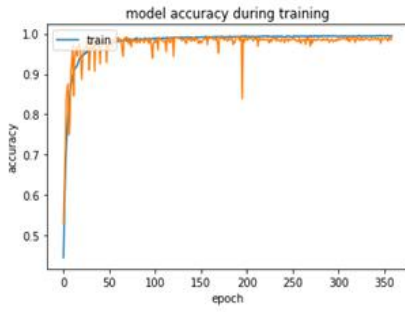
The GFERS results are compared with other related literature and a description of comparison is given in Table 5. The accuracy achieved by GFERS with various pose and illumination conditions is higher compared to the state-of-the-art techniques. In the existing literature, most of the reported work is based on CNN and CNN combined with

LSTM or GRU. Mostly LSTM and GRUs are used for text and speech classifications. For image classification, LSTM and GRUs are used for classification after feature extraction is done by using CNN. In the proposed architecture, standalone GRUs are used for feature extraction followed by dense and Softmax layers for classification achieving comparable accuracy with CNN and LSTM. Training time for GFERS, LSTM based model and CNN based model is 17.08 hours, 21.78 hours and 48 hours respectively. The training of all models is carried out by using same hardware platform (Intel i7-7700 CPU @ 3.60GHz and 16GBRAM). To train the model, use of GRUs took significantly lesser time compared to CNN models for the same number of epochs. The training time was reduced by a factor of approximately 30 hours making it feasible to carry out multiple experimentations with different combinations. The training process comparison of three model is as shown in Figure 9 (a)-(c).

Table 5. GFERS Results comparison with existing Literature (Real time)

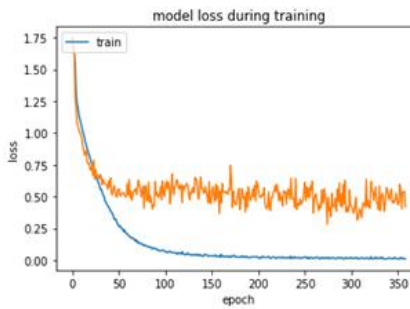
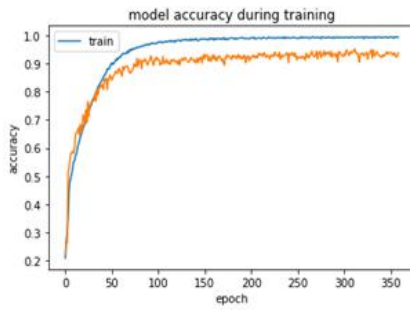
[Author, year]	Technique used	Data set used	Acc.*	PI*/II*/MF*
[Jung, 2014] [12]	DNN & CNN	CK, FER 2013	72.78% 86.54%	No/ No/No
[Chen, 2017] [13]	CNN	CK+, JAFFE	98.15%	No/ No/No
[Saxena, 2019][14]	CNN	CMU Multi PIE	95.8%	PI/II/MF
[Baddar, 2018][15]	Mode viriatonal LSTM	Oulu-CASIA, AFEW, KAISTFace MPMI	85.18%, 51.44%, 84.98%	PI/ II/No
[Hasani, 2017] [16]	3D CNN-LSTM	CK+, MMI, FERA, DISFA	67.52%, 54.76%, 41.93%, 40.51%	PI/No/No
[Kim, 2019] [17]	CNN-LSTM	MMI, CASME II	78.61%, 60.98%	No/No/No
[Yan, 2018][18]	Joint CNN-bidirectional LSTM	Multi-PIE, FER-2013	90.89%, 71.99%	PI/II/No
[Ilyas, 2019] [19]	CNN-LSTM	CK+	86.16%	PI/II/No
[Deng, 2019] [20]	CNN-GRU	CK+, MMI, AFEW	68.73%, 58.76%, 43.56%	PI/II/No
[Li, 2019] [21]	CNN-LSTM	CK+	90.51%	No/No/No
[Kang, 2019] [22]	VGG16, GRU	AFEW	47%	PI/II/No
[An, 2020] [23]	Hybrid CNN-LSTM, SVM	CK+, JAFFE, FER-2013, BU-3DFE, Oulu-CASIA	98.9%, 99.3%, 86.6%, 87.7%, 88.3%	PI/II/No
GFERS	GRU	CMU Multi PIE, FER-2013 (for Sad emotion)	94.66%	PI/II/MF

*Acc.: Average accuracy, PI: Pose Invariant, II: Illumination Invariant, MF: Multiple faces**NR: Not reported



(a)

(b)



(c)

Figure 9. Comparison of Training Process for (a) CNN (b) LSTM (c) GFERS

5. Conclusion and Future Work

In this work, we have proposed a methodology for facial emotion recognition using GRUs. The proposed GFERS works successfully for multiple facial emotion recognition under the constraints such as pose, illumination and age variations. Average accuracy of 94.66% is achieved by proposed GFERS in real time for 25 subjects, which is higher than the results reported in existing literature. The probability of recognized emotion is almost 100% in some cases. In real-time input images, the processing time is in between 0.032-0.037 sec for CPU and 0.200-0.350 sec for Raspberry pi implementation respectively. The GFERS is also implemented for UiPath Robotic Process Automation making it suitable for HRI applications.

Future work involves developing algorithms to solve the constraints of occlusion and to test these techniques on social robots as a real time application.

6. Acknowledgment

The authors would like to thank all the volunteers for the experimentation and also would like to thank the host organization for providing CMU Multi PIE database. We thank all other researchers for making other relevant databases available for such research experiments.

References

- [1] Dautenhahn K. Methodology & themes of human-robot interaction: A growing research field. *Int. J. Adv. Robotic System*. 2007; vol. 4, no. 1; p. 15.
- [2] Mehrabian A. Communication without words. *Psychology Today*, 1968, pp. 53-56.
- [3] Happy SL, et al. Automated alertness and emotion detection for empathic feedback during e-learning. *IEEE 5th Int. Conference on Technology for Education (T4E)*; 2013; India; pp. 47-50.
- [4] Coco MD, Leo M, Distanto C, Palestra G. Automatic emotion recognition in robot-children interaction for ASD Treatment. *IEEE Int. Conference on Computer Vision Workshop*; 2015; Santiago; pp.537-545.
- [5] Goodfellow IJ, et al. Challenges in representation learning: a report on three machine learning contests. *Workshop Challenges in Representation Learning (ICML2013)*; 2013; pp. 1-8.
- [6] Rosalind WP. *Affective computing*. MIT press, Cambridge. 2000.
- [7] Suchitra, Palaniswamy S, Tripathi S. Real-time emotion recognition from facial images using raspberry Pi II. *3rd International Conference on Signal Processing and Integrated Networks, (SPIN), IEEE*. 2016; Noida, India, pp. 666-670.
- [8] Romao M, Costa J, Costa CJ, *Robotic process automation: a case study in the banking industry*. 14th Iberian Conference on Information Systems and Technologies, 2019, Portugal, pp1-6.
- [9] Osman C, Ghiran A. When industry 4.0 meets process mining, *Procedia Computer Science*, 2019, Volume 159, pp. 2130-2136, ISSN1877-0509, <https://doi.org/10.1016/j.procs.2019.09.386>.
- [10] Gross R, Matthews I, Cohn J, Kanade T, Baker S. Multi-PIE. *Proc Int Conf Autom Face Gesture Recognit*. 2010;28(5):807-813.
- [11] Jung H, et al. Development of deep learning-based facial expression recognition system. 21st Korea-Japan Joint Workshop on Frontiers of Computer Vision (FCV). 2015; Mokpo; pp. 1-4.
- [12] Goodfellow IJ, et al. Challenges in representation learning: A report on three machine learning contests. *Neural Networks*, 64:59-63, 2015.
- [13] Warren Gay. 2014. *Raspberry Pi Hardware Reference* (1st. ed.). Apress, USA.
- [14] Chen X, Yang X, Wang M Zou J. Convolution neural network for automatic facial expression recognition. *International Conference on Applied System Innovation*. 2017; Sapporo; pp. 814-817.
- [15] Saxena S, Tripathi S, Sudarshan TSB. Deep dive into faces: pose & illumination invariant multi-face emotion recognition system. 2019 *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. 2019; Macau, China; pp. 1088-1093.
- [16] Baddar WJ, Ro YM. Mode Variational LSTM robust to unseen modes of variation: application to facial expression recognition. *AAAI* (2018).

- [17] Hasani B, Mahoor MH. Facial expression recognition using enhanced deep 3D convolutional neural networks. 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). 2017; Honolulu, HI; pp. 2278-2288.
- [18] Kim DH, Baddar WJ, Jang J, Ro YM. Multi-objective based spatio-temporal feature representation learning robust to expression intensity variations for facial expression recognition. in *IEEE Transactions on Affective Computing*. 1 April-June 2019; vol. 10: no. 2; pp. 223-236.
- [19] Yan J et al. A joint convolutional bidirectional LSTM framework for facial expression recognition. *IEICE Transactions on Information and Systems*. 2018;101(4); pp. 1217-1220.
- [20] Ilyas CMA, Schmuck V, Haque MA, Nasrollahi K, Rehm M, Moeslund TB. Teaching pepper robot to recognize emotions of traumatic brain injured patients using deep neural networks. 28th IEEE International Conference on Robot and Human Interactive Communication. 2019; India; pp. 1-7.
- [21] Deng L, Wang Q, Yuan D. Dynamic facial expression recognition based on deep learning.14th International Conference on Computer Science & Education (ICCSE). 2019; Toronto, Canada; pp. 32-37.
- [22] Li TS, Kuo P, Tsai T, Luan P. CNN and LSTM based facial expression analysis model for a humanoid robot. in *IEEE Access*, 2019, 7; pp. 93998-94011.
- [23] Kang K, Ma X. Convolutional gate recurrent unit for video facial expression recognition in the wild. Chinese Control Conference (CCC). 2019; Guangzhou, China; pp. 7623-7628.
- [24] An F, Liu Z. Facial expression recognition algorithm based on parameter adaptive initialization of CNN and LSTM. *The Visual Computer*. 2020; 36; pp 483–498.
- [25] Viola P, Jones MJ. Robust real-time face detection. *Int. Journal of computer vision*. 2004; 57,137-154.
- [26] <https://www.uipath.com/>

Research on the Construction of Cutting Edge Technology Monitoring System Based on Multi-Source Heterogeneous Data

Qiang XIAO¹, Siming TAN and Shengfeng YU

Qingdao Institute of Science and Technology Information, Qingdao, China

Abstract. With the rapid development of the Internet and big data, the data resources in various industries and technical fields are constantly emerging and growing. How to monitor and identify the effective data information in the massive big data has become one of the key contents of the current scientific and technological information work. This paper designs and implements the advanced technology monitoring system based on multi-source heterogeneous data. It comprehensively uses information collection technology, database technology and big data mining technology to realize the accurate monitoring, acquisition and analysis of multi-source heterogeneous data. It reveals the coupling relationship of technologies, people and institutions in different fields and the future technology development trend, and finally visualizes in various states. It provides a reference for the strategic decision-making of relevant government departments, and provides efficient and convenient research tools and methods for scientific research institutes and enterprises.

Keywords. Cutting edge technology, multi-source heterogeneous data, monitoring, system, construction

1. Introduction

Cutting edge technology is a leading and exploratory major technology in the core or key technology field. It represents the latest development trend of high-tech in the world. It not only plays an important leading role in the formation and development of emerging industries in the country and even the world in the future, but also plays a positive role in promoting the technological upgrading and R & D infrastructure construction of enterprises.

With the rapid development of science and technology, a new round of industrial technology revolution will be triggered in the fields of information, manufacturing, biology, new materials and energy. In the national science and technology innovation plan of the 13th five-year plan, China has clearly pointed out that it is necessary to strengthen the early warning of the trend of industrial change and major technologies, strengthen the prediction of the turning point of disruptive technology replacing

¹ Corresponding Author: Qiang Xiao, Qingdao Institute of Science and Technology Information, Qingdao, China; E-mail: qd82898286@163.com.

traditional industries, and timely lay out the research and development of frontier technologies in emerging industries. Therefore, it has become the focus of scientific and technological workers and managers to quickly and accurately understand and master the state-of-the-art technology situation, so as to help the government and enterprises to formulate science and technology development strategy, and improve the national high-tech research and development ability and the international competitiveness of the industry.

Some government departments or research institutions in some countries have established corresponding strategic decision support systems to provide intellectual support for their competitiveness and sustainable development in some technical fields. The U.S. Department of defense has established a number of information technology analysis centers (IACS) to provide information analysis services for managers and decision makers of the Department of defense by using databases and intelligence analysis tools. Japan's Nomura institute takes consulting and knowledge services, system integration services and decision-making services for the government as its core business, and takes the construction of information technology service platform as an important means to provide high-quality services to users. The British Institute of international strategy also has a corresponding Intelligence Analysis Department [1].

In recent years, scholars have begun to pay attention to data sources and identification methods of cutting-edge technologies. For example, scientific and technological media data are used as data sources of frontier technology identification, and neural network model is used in citation analysis process. This study holds that the biggest characteristic of Internet information is that there are many kinds of data sources and rich contents, and the data are not limited by countries and geographical space. These characteristics make the acquisition of data more convenient, and the data objects used for the monitoring of frontier technologies are more multi-source, which is more conducive to improving the accuracy of monitoring results of frontier technologies. The corresponding data analysis methods need to be more in-depth data fusion and systematic analysis methods [2].

In this study, we propose and design a cutting-edge technology monitoring system, which is a research method and tool platform. Using network information collection technology, database technology and data mining technology, it provides research services such as information automatic collection, data depth calculation, knowledge intelligent discovery and visualization, and realizes real-time monitoring, automatic collection and automatic collection of the latest technology trends and scientific research achievements. Intelligent analysis can accurately grasp or predict the development trend of frontier technology in specific fields, study and judge its development path, provide decision-making reference for the government's scientific research plan and major project layout in specific fields, and provide research methods and tools for scientific research institutes and enterprises to carry out scientific research.

2. Structure Design of Monitoring System

The overall structure design of advanced technology monitoring system is divided into three layers: data acquisition layer, analysis storage layer and application display layer. As shown in the figure 1 as below:

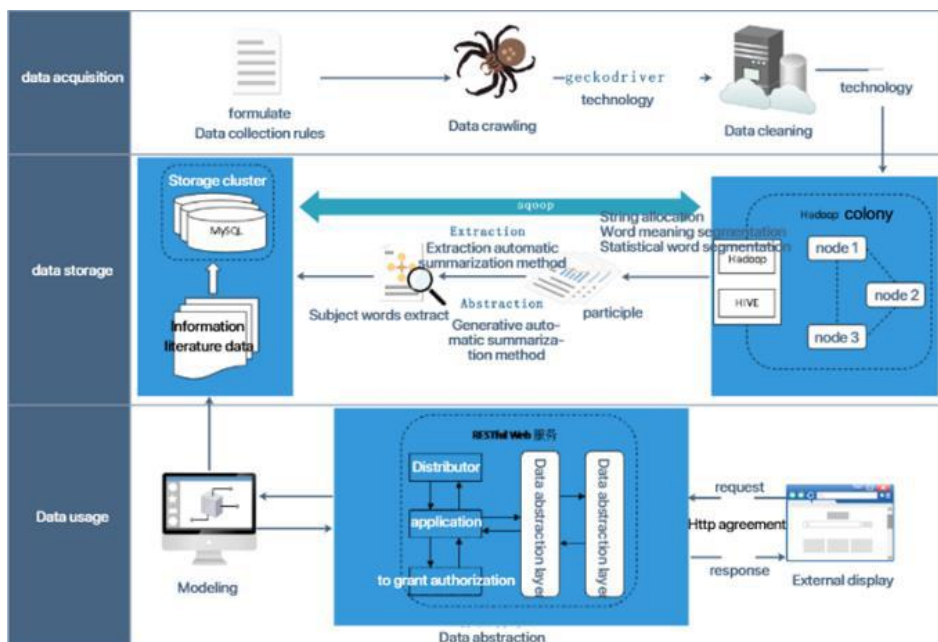


Figure 1. Overall design of the system.

2.1. Data Acquisition Layer

The data acquisition of this system should meet three characteristics: comprehensiveness, multidimensional and high efficiency. The system uses distributed web crawler, and the bottom layer of crawler is geckodriver technology. Geckodriver can complete remote web content crawling. The system can collect dynamic data related to cutting-edge technology according to the set rules.

Data cleaning includes: 1) data De duplication: only one copy of the same paper or patent obtained from different data sources can be retained. The paper title and the author can be used to judge whether the paper is duplicate, and the patent title and the inventor / inventor can be used to judge whether the patent is repetitive; 2) data filling: in the process of data De duplication, if it is found that the duplicate data of subsequent papers / patents contains the previous data (3) data standardization: if the abbreviation or nonstandard elements in the paper / patent data are found, standardization shall be carried out. For example, if the address in China is written as Chi, it shall be standardized as China; if the unit abbreviation is abbreviated, it shall be changed to the full name.

2.2. Analysis Storage Layer

The system needs to extract effective information from multi-source heterogeneous data. The entity extraction method based on the deep learning model BiLSTM-CRF is used to add the manually extracted features such as word segmentation, indexing and word frequency statistics into the model to describe the semantic information of each Chinese character more fully, and extract new technical keywords from the technical dynamic data as far as possible In the future, which technologies may become hot spots provide

data support. The dynamic data storage engine of the system implements permanent storage of the effective data collected by the web crawler, which can be used for various needs analysis of the system or shared to the third party. At the same time, the data storage should consider the periodic update of the data.

2.3. Application Display Layer

This system uses RESTful architecture to realize the access of display chart to statistical analysis data. In this way, the complexity of the system and the coupling degree of the system can be reduced.

This system has statistical analysis algorithm library, all of which can be independent of the line algorithm program, and the algorithm library records the meaning of each statistical analysis algorithm and the executable SQL statement. The client uses the general web request such as GET/ POST to visit the statistical analysis algorithm program based on HTTP protocol, and the server receives the request from the display side through the distributor Parameters automatically forward the request to the corresponding statistical analysis algorithm program in the algorithm library; the statistical analysis algorithm program calls the SQL statements stored in the database, and then executes to obtain the statistical analysis data. After obtaining the data, the statistical analysis results are submitted to the format converter, and the format converter converts the statistical analysis result set into JSON format and then transmits it to the display terminal. The application display end can use echarts graphic display component, which contains rich graphics such as fishbone diagram, hot word graph, relationship diagram and so on. After receiving the data in JSON format, echarts displays the conclusion through the specified graph.

3. Function Design of Monitoring System

The function design of advanced technology tracking and monitoring system is divided into four parts: (1) technology dynamic analysis; (2) expert dynamic analysis; (3) mechanism dynamic analysis; (4) large screen visualization centralized display.

3.1. Technical Dynamic Analysis

First of all, the dynamic analysis of technology is to display the dynamic information of the latest technology development in a specific field, and at the same time, it should have the translation function for the collected foreign language information. Secondly, it analyzes the development process of technology in specific fields, excavates important node figures and events, and displays the development trend of technology in various forms. Then there is dynamic information analysis, which is used to show the geographical distribution, source type and ranking of all collected information. The last is hot word analysis. Through clustering analysis of hot words (or key words), a high-frequency hot word atlas is formed, and the correlation between technical dynamic information and hot words is established, and relevant information such as hot word retrieval technology, organization and expert dynamics is realized.

3.2. Dynamic Analysis of Experts

According to the matching degree of the given domain keywords, the domain expert database is formed. According to the data collection, the basic information of the experts' institutions, research hotspots, technical cooperation and academic achievements is displayed. At the same time, according to information sources, journals, patents and other information, the distribution of key experts in countries, industries and disciplines is analyzed.

3.3. Dynamic Analysis of Mechanism

According to the matching degree of keywords in the given field, the Organization database can be classified according to the attributes of universities, scientific research institutes and enterprises. Display the research direction, research hotspot, key experts, cooperation network and research dynamic information of each institution. Cooperation network analysis and regional distribution of all scientific research institutions in specific fields can be carried out.

3.4. Large Screen Visual Display

According to the actual needs, it can realize various functional requirements, such as dynamic tracking of new technology, hot word map, new technology trend prediction, expert dynamic, organization dynamic, regional distribution of technology dynamic, etc., and conduct statistical analysis and display in the form of column chart, broken line chart, fan chart, relationship diagram, fishbone diagram, hot word chart, etc.

4. Application of Analysis and Mining Technology

There are two main tasks for data analysis and mining in this system: one is to extract knowledge from the acquired new technology information to form a new technology industry knowledge map, which includes the relationship among new technology, institutions, scholars and events. Second, according to the knowledge map of new technology industry, we can extract the new rules or relationships among the elements of the industry, and provide data basis for the trend analysis, statistical analysis and other modules.

So in general, information analysis and mining includes two sub modules: knowledge extraction and knowledge mining. In addition, keyword extraction is also a key function of the system, which needs to capture new technology keywords according to the information captured. Its function flow is shown in Figure 2

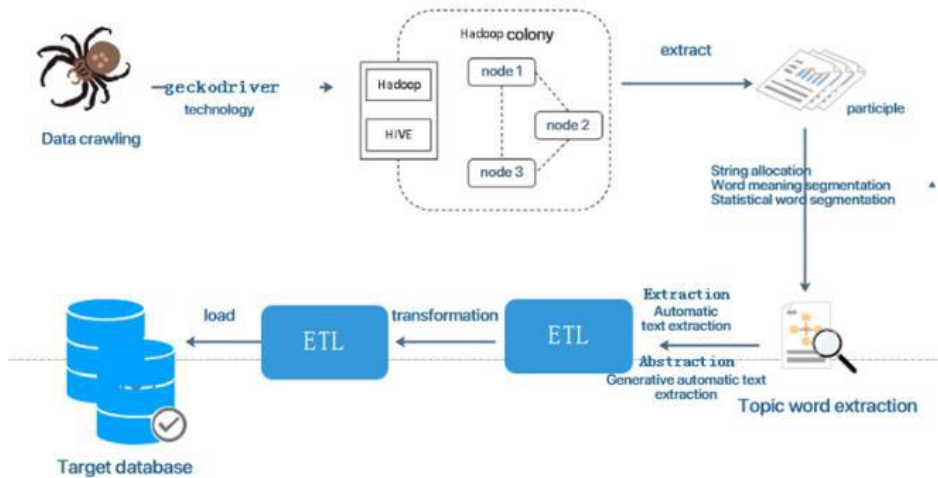


Figure 2. System function flow chart

4.1. Knowledge Extraction

The advanced technology tracking and monitoring system needs to extract effective information from structured, semi-structured and pure text data from the Internet and various thematic databases to form knowledge (structured data) and store it in the knowledge map. A lot of hidden information can be extracted from the dynamic data of advanced technology.

This system uses the method of entity extraction based on the deep learning model BiLSTM-CRF. New named entities will appear in the dynamic information of cutting-edge technology, such as information source type, organization name, expert name, etc., and the core of rule-based naming extraction method is rule-making. Therefore, once a new naming entity appears, it will consume a lot of time and energy to update the rules manually, in other words, its formulation and the portability of the rules is poor. Therefore, in order to solve the deficiency of rule-based name extraction, the system adopts the entity extraction method based on the deep learning model BiLSTM-CRF. On the basis of this model, this scheme adds the manually extracted features such as word segmentation, indexing and word frequency statistics into the model to describe the semantic information of each Chinese character more fully. Through the above process, BiLSTM-CRF algorithm can extract new technical keywords from the captured dynamic data to the maximum extent, and provides data support for predicting which technologies may become hot technologies in the future.

4.2. Extraction of Structured and Semi-Structured Data

In the cutting-edge technology monitoring system, structured data mainly refers to the archived literature data and patent data, which can be transformed into RDF or other forms of knowledge base content. For example, a common W3C recommended mapping language is R2RML (RDB2RDF).

In the cutting-edge technology monitoring system, semi-structured data mainly refers to the data captured from major science and technology news websites, official websites of authoritative scientific research institutions, microblogs and science and

technology columns. This kind of data itself has a certain structure, but it needs further extraction and sorting to get the data. Web page data extraction is generally generated by wrapper, and the process of extracting information is shown in the following figure 3:

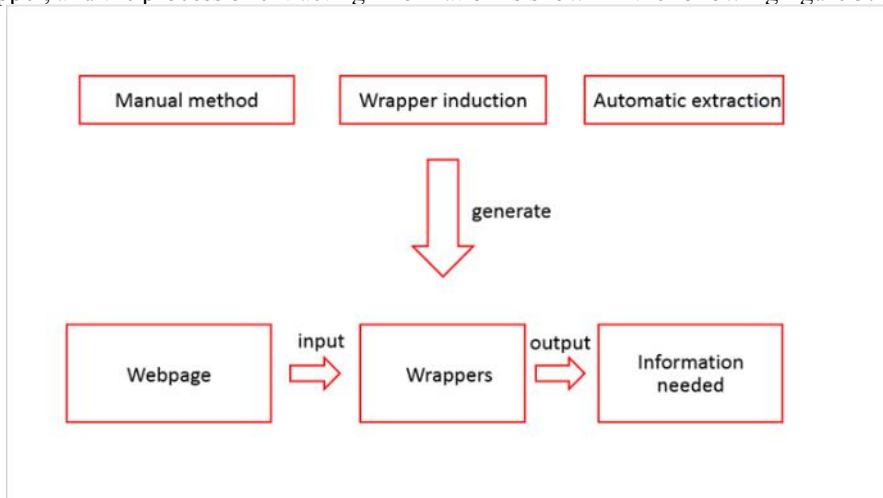


Figure 3. Information extraction process.

4.3. Keyword Extraction

Keyword extraction has important applications in dynamic data retrieval, automatic summarization, text clustering / classification and so on. Keyword extraction algorithms are generally divided into supervised and unsupervised.

Two unsupervised learning algorithms, TF-IDF algorithm and TextRank algorithm, are adopted in the advanced technology tracking and monitoring system.

① TF-IDF algorithm

TF-IDF is a numerical statistical method, which is used to reflect the importance of a technical dynamic hot word to a news, paper or patent. Its main algorithm is: if a word appears frequently in the same document, the TF of the word is high; if the word rarely appears in other documents, the IDF of the word is high. In this case, it is considered that the word has a good classification ability [3], and it is also proved to be a hot word.

② TextRank algorithm

An important feature of this algorithm is that it can analyze a single document and extract keywords without a corpus. The algorithm is as follows:

The first step is word segmentation and part of speech tagging. The effective data collected into the database are divided into sentence patterns, word segmentation and part of speech tagging are carried out sentence by sentence. At the same time, stop words are filtered, and only nouns, verbs, adjectives and other specified part of speech words are retained [4]. The second step is the construction of candidate keyword graph. The co-occurrence relationship is used to construct the edge between any two points in the keyword graph. There are edges between the two nodes only if their corresponding words appear together in the same technical dynamic [5,6,7]. In the third step, according to the formula of TextRank, the weight of each node is propagated iteratively until it converges. The node weights are sorted in reverse order to get the most important words as candidate

keywords [8]. Finally, these keywords are marked in the original text. If they form adjacent phrases, they are combined into multi word keywords [9,10].

4.4. Patent Feature Extraction

The method of artificial intelligence is used to extract the features of data and analyze patent features intelligently, such as rule-based clustering algorithm, text clustering algorithm. The patent keywords are extracted by automatic summarization method and rule-based method. According to the suggestions of domain experts, domain dictionaries or professional websites, the keywords are filtered and refined to extract patent features.

5. Conclusion

The strategic significance of big data technology is to be able to obtain useful information from massive data and apply it to practice after professional processing and mining analysis. The traditional methods and means of scientific and technological information analysis have been unable to meet the needs of the current world development situation. To complete the monitoring of new technologies from massive multi-source heterogeneous big data, it is necessary to build an accurate, efficient and stable intelligent monitoring system, and give full play to the decision-making and consulting ability of think tanks through big data analysis and artificial intelligence assistance. At the same time, the diversity of data sources puts forward higher requirements for the rapid integration and analysis ability of data. Therefore, this paper proposes the overall process structure and functional architecture of the new technology tracking and monitoring system based on multi-source heterogeneous data, and explores the establishment of a set of intelligent, automatic and more perceptive new technology tracking and monitoring system, so as to strengthen the support and leading role of new think tanks in scientific and technological innovation, and play a more obvious role.

References

- [1] Tan ZY, Wang Q, Cang HY, Rao YH, Yang NH, Nie L. Construction of monitoring and analysis platform for frontier information of scientific and technological development. *Sci. Res.*, 2010(02): 37-43
- [2] Zeng W, Li H, Fan YF, Liu GY, Li R, Xu Z. Research on the identification system of science and technology frontier in the open source information environment. *Information Studies: Theory & Application*, 2019(07): 34-38
- [3] Zhu J, Li HW, Peng X, Zhao WY. Research on the correlation between function and authority of Android application system. *Comp. Appl. Software*, 2014 (10): 33-39
- [4] Zhao Z. Research on the effect of text vectorization on text classification. China excellent master's thesis full text database, 2018 (01)
- [5] Chen Z, Zheng S. Research on chemical emergency information extraction based on multi algorithm fusion. *Comp. Digital Engin.*, 2018 (2): 6-6
- [6] Zhao MY. English short text measurement method based on part of speech and keywords. Chinese excellent master's thesis full text database, 2018, (01)
- [7] Xiang ZX. Design and implementation of cloud health information platform based on distributed crawler. China excellent master's dissertation full text database, 2018, (01)
- [8] Li MX. Design and implementation of an interactive mobile user interest discovery system. China excellent master's dissertation full text database, 2016, (11)
- [9] Wang Q. Analysis and research on microblog public opinion in emotional context. China excellent master's thesis full text database, 2018, (02)

- [10] Wang ZW, Qiu HP, Sun Y, Ke DL. Intelligent recommendation technology for military information service. *Command Contr. Simul*, 2019 (04): 120-125

An Octave Package to Perform Qualitative Analysis of Nonlinear Systems Immersed in R^4

Eder ESCOBAR ^a, Richard ABRAMONTE ^a Antenor ALIAGA ^a and Flabio GUTIERREZ ^{a,1}

^aScience Faculty, Universidad Nacional de Piura, Piura, Peru

Abstract. In this work, the AutonomousSystems4D package is presented, which allows the qualitative analysis of non-linear differential equation systems in four dimensions, as well as drawing the phase surfaces by immersing R^4 in R^3 . The package is programmed in the computational tool Octave. As a case study applied to the new Lorenz 4D System, sensitivity was found in the initial conditions, Lyapunov exponents, Kaplan Yorke dimension, a stable and unstable critical point, limit cycle, Hopf bifurcation, and hyperattractors. The package could be adapted to perform qualitative analysis and visualize phase surfaces to autonomous systems, e.g. Sprott 4D, Rossler 4D, etc. The package can be applied to problems such as: design, analysis, implementation of electronic circuits; to message encryption.

Keywords. Octave, Autonomous System, fourth dimension, qualitative analysis

1. Introduction

Several physical phenomena are modeled with systems of differential equations, but if they do in a fairly approximate way, the confidence in the results of the model is given by the qualitative analysis of the system.

In this work, a package in Octave 5.1.0 called **AutonomousSystems4D** is presented, which allows qualitative analysis, as well as visualizing phase surfaces of non-linear systems immersed in a four-dimensional (4D) system. As far as we know, there are no similar works.

Currently, to visualize the phase maps and the limit cycles of the systems that are in 4D, they are projected in three dimensions making zero any of the coordinated axes, which will have four surfaces to analyze (see for example [1–3]). In the developed package, no projections are used, the Velezmoreno model is used, which allows visualizing 4D objects in a three-dimensional (3D) system (see Section 2.1), so you will have a single surface, which can be rotated from different angles, this allows a better analysis of what happens in the system of 4D equations. The package could be adapted to perform qualitative analysis and visualize phase surfaces to autonomous systems, e.g. Lorenz 4D,

¹Corresponding Author: Flabio Gutierrez, Department of Mathematics, Universidad Nacional de Piura, Campus Universitario, Urb. Miraflores s/n, Castilla, Piura, Perú; E-mail: flabio@unp.edu.pe

Sprott 4D, Rossler 4D, etc. In this paper we apply it to perform qualitative analysis and visualization of a Lorenz hyper-caotic system immersed in R^4

Lorenz found the first strange attractor in the three-dimensional autonomous system of equations. In [4, 5] qualitative analyses are carried out for this type of system. Works on the Lorenz system in a four-dimensional space can be found in [6–8], these systems are hyper-caotic systems having two or more positive Lyapunov exponents.

The package can be applied to problems such as: design, analysis, implementation of electronic circuits [2, 3, 9]; to message encryption [10].

2. Preliminaries

Definition 1 . Hopf bifurcation [4].

Is the bifurcation corresponding to the presence $\lambda_{1,2} = \pm iw_0$, con $w_0 > 0$.

Definition 2 . Break-even point [5].

The point $X = a$ with $f(a) = 0$ is called the critical point of the equation $X' = f(X)$.

Theorema 1 Lyapunov [4].

Consider a dynamic system defined by $X' = F(X)$, where $X \in R^n$, F is continuous. Suppose you have an equilibrium X_0 , $A = F_X(X_0)$ (A es A is the Jacobian matrix of $F(X)$ evaluated in equilibrium). Then X_0 is stable if all the eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_n$ of A satisfy $Re\lambda < 0$.

Definition 3 . In order for the system of 4D equations to be hypercotic it must satisfy three conditions: [6].

- 1) A four-dimensional autonomous system is required.
- 2) Sensitivity in the initial conditions.
- 3) Two or more positive Lyapunov exponents and the sum of all Lyapunov exponents is less than 0.

Method for finding two roots in the imaginary axis. Consider the cubic equation.

$$\lambda^3 + L_1(c)\lambda^2 + L_2(c)\lambda + L_3(c) = 0 \tag{1}$$

Let $\lambda = iw_0$ an imaginary solution, then the cubic equation is equivalent to:

$$iw_0(L_2(c) - w_0^2) + (L_3(c) - L_1(c)w_0^2) = 0 \tag{2}$$

Where $L_2(c) - w_0^2 = 0$ y $L_3(c) - L_1(c)w_0^2 = 0$

that is to say, $L_3(c) = L_1(c)L_2(c)$

Let $c = c_0$ satisfy (3), and let $\lambda_{1,2} = \pm iw_0$ and $\lambda_3 = \lambda_0$ roots of Equation (1)

$$w_0 = \sqrt{L_2(c_0)} \tag{3}$$

In addition, it must be observed that $(\lambda - iw_0)(\lambda + iw_0)(\lambda - \lambda_0) = 0$

$$\lambda^3 - \lambda_0\lambda^2 + w_0^2\lambda + (-\lambda_0w_0^2) = 0 \tag{4}$$

From equations (1) and (4) the imaginary roots are obtained:

$$\lambda_{1,2} = \pm i\sqrt{L_2(c_0)} \text{ and } \lambda_3 = \lambda_0$$

2.1. Velezmoro Diving Model

Based on the model that allows three-dimensional (3D) objects to be displayed on a two-dimensional (2D) screen; Velezmoro and Ipanaque [11], propose a model that allows four-dimensional (4D) objects to be displayed in 3D.

Required: four non-collinear vectors in space

$$\hat{\mathcal{B}} = \{(e_{11}, e_{12}, e_{13}), (e_{21}, e_{22}, e_{23}), (e_{31}, e_{32}, e_{33}), (e_{41}, e_{42}, e_{43})\}$$

a dip that transforms a point $p = (p_1, p_2, p_3, p_4) \in 4D$ into a point $q = (q_1, q_2, q_3) \in 3D$.

For example $\varphi(p) = \frac{1}{\sqrt{3}}(p_1 + p_2 + p_3 + p_4, -p_1 + p_2 + p_3 - p_4, -p_1 - p_2 + p_3 + p_4)$

3. Octave Package for Qualitative Analysis and Visualization of 4D System Phase Surfaces

The **AutonomousSystems4D** package has been developed in octave 5.1.0, it allows you to visualize phase surfaces of ordinary differential equation systems, linear and non-linear, that are immersed in the fourth dimension. The Octave **ode45**, **lsode** libraries have been used to solve differential equations in numerical form. It also allows calculating the Lyapunov exponents and convergence or divergence of trajectories of the system's solutions. In this case it has been adapted to work with the Lorenz system in 4D.

3.1. Program for Graphing Phase Surfaces of 4D Systems

Subprogram 1 (System of Lorenz equations in 4D).

$x = 0;$

function $xdot = \text{func}(x, t)$

 % value of the parameters of the hypercotic system.

$a = 10; b = 8/3; c = 35; d = 0.5;$

$xdot(1, 1) = a * (x(2) - x(1)); xdot(2, 1) = c * x(1) - x(1) * x(3) - x(2) + x(4);$

$xdot(3, 1) = x(1) * x(2) - b * x(3); xdot(4, 1) = -d * x(1);$

end

$x0 = [0, 2, 0, 0];$ % Initial condition

$t = \text{linspace}(0, 200, 15000);$ % time interval;

tic;

$x = \text{lsode}('func', x0, t);$ % solves the system of 4D toc equations

toc;

$z = [x(:, 1)x(:, 2)x(:, 3)x(:, 4)];$ % solution matrix

% Velezmoro model

$p1 = (1/\text{sqrt}(3)). * [z(:, 1) + z(:, 2) + z(:, 3) + z(:, 4)]; p2 = (1/\text{sqrt}(3)). * [-z(:, 1) + z(:, 2) + z(:, 3) - z(:, 4)];$

$p3 = (1/\text{sqrt}(3)). * [-z(:, 1) - z(:, 2) + z(:, 3) + z(:, 4)];$

$\text{plot3}(p1, p2, p3, 'b', 'linewidth', 1.5); \text{xlabel}('X'); \text{ylabel}('Y'); \text{zlabel}('Z')$

3.2. Program to Calculate Lyapunov's Exponents

The method for calculating Lyapunov's exponents can be found in [4]

Subprogram 2 (Lyapunov's Algorithm).

```
function[Temp, Lexp] = lyapunov(n, rhs_ext_fcn, fcn_integrator, tstart, stept, tend, ystart);
    %n : number of non-linear OEDs. %n2 : total number of OEDs
    n1=n; n2=n1*(n1+1); % Number of iterations
    nit = round((tend-tstart) /stept); % Memory allocation
    y = zeros(n2,1); cum =; zeros(n1,1); y0 = y; gsc = cum; znorm = cum;
    y(1:n) = ystart(:); % Initial values
    for i = 1 : n1 y((n1+1)*i) = 1.0; end;
    t = tstart;
    % Main Loop
    for ITERLYAP = 1: nit
        % Extended ODE System Solution
        [T, Y] = feval(fcn_integrator, rhs_ext_fcn, [t, t + stept], y);
        t = t+stept; y = Y(size(Y,1),:);
        for i = 1:n1
            for j = 1:n1 y0(n1*i+j) = y(n1*j+i); end;
        end;
        % Construction of a new orthonormal base using gram-schmidt
        znorm(1) = 0.0;
        for j= 1:n1 znorm(1) = znorm(1)+ y0(n1*j+1)^2; end;
        znorm(1) = sqrt(znorm(1));
        for j=1:n1 y0(n1*j+1)= y0(n1*j+1)/znorm(1); end;
        for j = 2:n1
            for k = 1:(j-1)
                gsc(k) = 0.0;
                for l = 1:n1 gsc(k) = gsc(k)+y0(n1*l+j)*y0(n1*l+k); end;
            end;
            for k = 1:n1
                for l = 1:(j-1)
                    y0(n1*k+j) = y0(n1*k+j)-gsc(l)*y0(n1*k+l);
                end;
            end;
            znorm(j) = 0.0;
            for k= 1:n1 znorm(j) = znorm(j)+y0(n1*k+j)^2; end;
            znorm(j) = sqrt(znorm(j));
            for k = 1:n1 y0(n1*k+j)=y0(n1*k+j)/znorm(j); end
        end;
        for k = 1:n1 cum(k) = cum(k)+log(znorm(k)); end;
        % normalize exponents
        for k = 1:n1
            lp(k) = cum(k)/(t-tstart);
        end;
        % Output data
        if ITERLYAP == 1
```

```

    Lexp = lp; Texp = t;
  else
    Lexp = [Lexp; lp]; Texp = [Texp; t];
  end;
  for i = 1:n1
    for j = 1:n1
      y(n1*j+i) = y0(n1*i+j);
    end;
  end;
end;
end;

```

Subprogram 3 (Lorenz 4D system).

```

function f = exp_lyapunov4(t,X)
  a=10;b=8/3;c=35;d=0.5; x=X(1); y=X(2); z=X(3);w=X(4);
  Y= [X(5), X(9), X(13),X(17); X(6), X(10), X(14), X(18);...
  X(7), X(11), X(15), X(19); X(8), X(12), X(16), X(20)];
  f=zeros(16,1);
  f(1)=a*(y-x); f(2)=c*x-y-x*z+w; f(3)=x*y-b*z; f(4)=-d*x;
  Jac=[-a a 0 0;c-z -1 -x 1;y x -b 0;-d 0 0 0]; f(5:20)=Jac*Y;
end

```

Subprogram 4 (Calculation of Lyapunov's exponents).

```

[T, Res] =lyapunov(4, @exp_lyapunov4D, @ode45,0,0.5,200,[0 2 0 0]);
hold on
plot(T,Res(:,1),'r','linewidth',2); plot(T,Res(:,2),'b','linewidth',2);
plot(T,Res(:,3),'k','linewidth',2); plot(T,Res(:,4),'g','linewidth',2);
xlabel('Time'); ylabel('Lyapunov exponents'); Res(length(Res),:);

```

3.3. Program for Displaying Sensitivity under Initial Conditions

Subprogram 5

```

% calculates the solutions
x=0;
function xdot=func(x,t)
  a = 10; b = 8/3; c = 35; d = 0.5;
  xdot(1,1) = a * (x(2) - x(1)); xdot(2,1) = c * x(1) - x(1) * x(3) - x(2) + x(4);
  xdot(3,1) = x(1) * x(2) - b * x(3); xdot(4,1) = -d * x(1);
end
x0 = [0, 2, 0, 0]; X0=[0, 2, 0, 0.1]; t = linspace(0, 60, 200); tic;
x=lsode('func',x0,t);
z=[t',x(:,1),x(:,2),x(:,3),x(:,4)]; toc;
plot(z(:,1),z(:,5),'r','linewidth',1.5);
xlabel('T','FontSize',14); ylabel('w(t),W(t)','FontSize',14); hold on; tic;
X=lsode('func',X0,t); Z=[t',X(:,1),X(:,2),X(:,3),X(:,4)]; toc;
plot(Z(:,1),Z(:,5),'b','linewidth',1.5)

```


4. Qualitative Analysis of the Lorenz Hyperchaotic System in R⁴

In this section, the **AutonomousSystems4D** package (see Sección 3) is applied to perform the qualitative analysis of the Lorenz System (5), as well as to visualize the phase surfaces that are in 4D.

In ([6]), the four-dimensional (4D) Lorenz System was proposed, with the equations of state:

$$x' = a(y - x) \quad y' = cx - y - xz + w \quad z' = xy - bz \quad w' = -dx \quad (5)$$

Where x, y, z and w are state variables; a, b, c y d are positive control parameters of the new Lorenz 4D system. This system is hypercotic for $a = 10, b = 8/3, c = 38$ and $d = 0.5$ [6].

4.1. Sensitivity in Initial Conditions

To analyze the sensitivity in the initial conditions, perturbations are made in the initial condition of each variable. For $x = 0$ and $x = 0.1$. See Figure 1 (left); $y = 2, y = 2.1$. See Figure 1 (right); similarly for $z = 0$ and $z = 0.1$; $w = 0, w = 0.1$. It can be seen that the paths diverge in time.

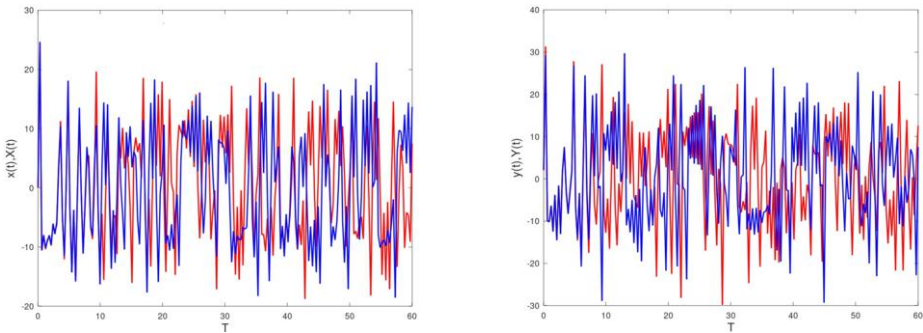


Figure 1. Disturbance to initial condition (0,2,0,0)

4.2. Lyapunov's Exponents

If in (5), the value of the control parameters are: $a=10, b=8/3, c=35, d=0.5$ and initial condition $(x,y,z,w)=(0,2,0,0)$; the Lorenz 4D System has three positive exponents and one negative Lyapunov exponent (See Figure 2 left)

$$\lambda_1 = 0.996745, \lambda_2 = 0.011837, \lambda_3 = 0.010623, \lambda_4 = -14.681099 \quad (6)$$

$$Suma = \lambda_1 + \lambda_2 + \lambda_3 + \lambda_4 = -13.662 < 0 \quad (7)$$

The Lyapunov dimension of the hyper-caotic attractor is fractional for parameter values, $a = 10, b = 8/3, c = 35, d = 0.5$ and initial conditions $(x,y,z,w) = (0,2,0,0)$.

$$D_{LY} = j + \frac{1}{|\lambda_{j+1}|} \sum_{i=1}^j \lambda_i = 3 + \frac{1}{|-14.681099|} (0.996745 + 0.011837 + 0.010623) = 3.0694$$

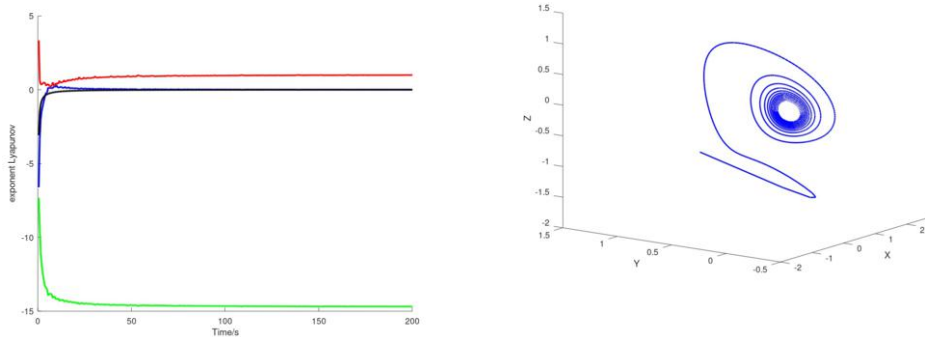


Figure 2. Lyapunov exponents (left), Hopf bifurcation (right)

4.3. Dissipation and Existence of the Attractor

The system divergence (5) is defined by: $\nabla V = \frac{\partial x'}{\partial x} + \frac{\partial y'}{\partial y} + \frac{\partial z'}{\partial z} + \frac{\partial w'}{\partial w}$

And it measures how quickly volumes change under system flow.

$$\nabla V = -a - 1 - b = -\frac{41}{3}; V(t) = V(0)e^{-\frac{41}{3}t}$$

When $t \rightarrow \infty$ the volume is reduced exponentially to zero. Any solution that enters the attractor will remain in time.

4.4. Vector Field Analysis

Applying the coordinate transformation to the original system: $(x, y, z, w) \rightarrow (-x, -y, z, -w)$, if the equations of state remain unchanged, it means that the phase surface of the system is symmetrical with respect to the z-axis.

$$\begin{aligned} x' &= -a((-y) - (-x)) = -x' \\ y' &= -c(-x + (-y) + (-x)z - (-w)) = -y' \\ z' &= (-x)(-y) - bz = z' \\ w' &= d(-x) = -w' \end{aligned}$$

There is a symmetry with respect to the z axes in the vector field. What happens when the flow of the system of equations crosses the coordinate axes, in the z axes, $x' = y' = 0$ and $z' = -bz$. This indicates that every orbit that passes through the z axes cannot leave it and therefore is invariant.

4.5. Calculation of the Equilibrium Point

To find the equilibrium point, each Lorenz 4D system equation is equated to zero.

$$a(y - x) = 0 \quad cx - y - xz + w = 0 \quad xy - bz = 0 \quad -dx = 0 \quad (8)$$

The equilibrium point obtained is $P_0 = (0, 0, 0, 0)$

4.6. Linearization of the Lorenz 4D Hyperchaotic System

Jacobian matrix of the Lorenz system of equations 4D:

$$J = \begin{pmatrix} -a & a & 0 & 0 \\ c-z & -1 & -x & 1 \\ y & x & -b & 0 \\ -d & 0 & 0 & 0 \end{pmatrix}$$

Jacobian matrix of the system evaluated at break-even (0,0,0,0):

$$A = \begin{pmatrix} -a & a & 0 & 0 \\ c & -1 & 0 & 1 \\ 0 & 0 & -b & 0 \\ -d & 0 & 0 & 0 \end{pmatrix}$$

4.7. Calculation of Eigenvalues

$$\det(A - \lambda I) = (\lambda + b)(\lambda^3 + (a + 1)\lambda^2 + (a - ac)\lambda + ad) = 0 \tag{9}$$

$$\text{Be } L_1 = a + 1, L_2 = a(1 - c), L_3 = ad$$

Cubic polynomial discriminant for values of parameters a = 10, b = 8/3 y d = 0.5.

- { Si $\Delta > 0$, three different real roots
- { Si $\Delta = 0$, multiple real roots
- { Si $\Delta < 0$, it has one real root and two complex roots

4.8. Stability Analysis at Origin

With the parameters a = 10, b = 8/3, d = 0.5 and c variation parameter.

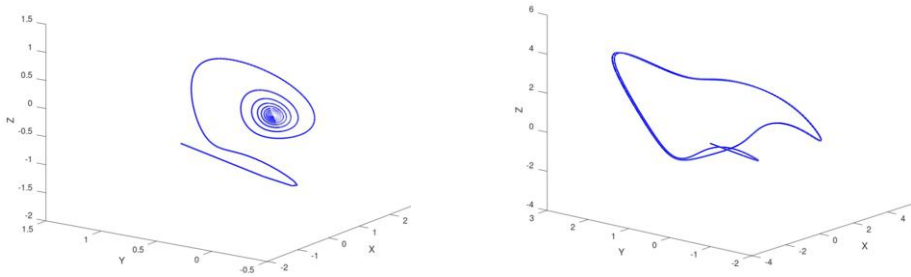


Figure 3. For c = 0.9 (left) the orbits tend to the origin, at c = 2.5 (right), the orbits have a stable limit cycle.

Case 1: When $0 < c < 0.95455$, the origin is asymptotically stable, there are two negative real eigenvalues and two imaginary eigenvalues that have a negative real part, therefore, all paths converge to the equilibrium point (0,0,0,0) (see Figure 3, left).

Case 2: When $0.95455 < c < 2.53$, the equilibrium point (0,0,0,0) is unstable, there are two negative real eigenvalues and two imaginary eigenvalues that have a positive real part, in this interval a limit cycle occurs (See Figure 3, right).

Case 3: When $c \geq 2.53$, the equilibrium point $(0,0,0,0)$ is saddle, there are two positive and two negative eigenvalues (See Figure 4). In this interval the origin becomes unstable, the trajectories move away and tend to limit attractor cycles.

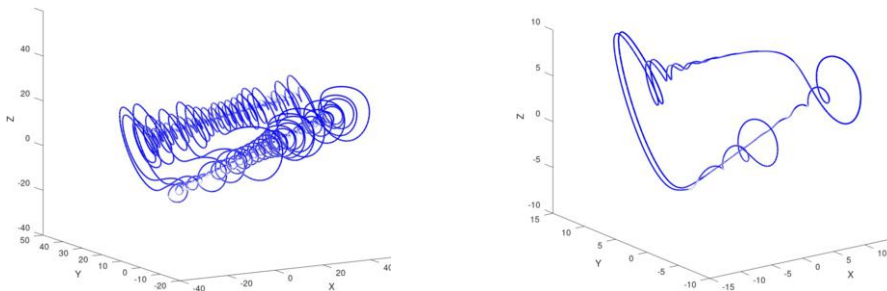


Figure 4. For $c = 20$ (left) and $c = 7$ (right), the orbits tend to two limit cycles of attraction.

4.9. Hopf Bifurcation in the Origin.

$$L_3 = L_1 * L_2; ad = (a - ac)(a + 1) \Rightarrow c = \frac{a - d + 1}{a + 1} = c^* \tag{10}$$

Hopf bifurcation occurs when $c = 0.95455$ and the eigenvalues are $\lambda_1 = -8/3$, $\lambda_2 = 0.67420 i$, $\lambda_3 = -0.67420 i$, $\lambda_4 = -11$ (See Figure 2).

4.10. Scenarios Presented by the Lorenz 4D System by Varying the Parameter c

The system (5) muestra diferentes hiper-atractores extraños de Lorenz, shows different strange Lorenz hypertractors, with fixed parameters $a = 10$, $b = 8/3$, $d = 0.5$ and the parameter c variable. For $c = 21, 22, 35, 40$ the origin is unstable, all the orbits move away from the origin and tend to two limiting attractor cycles (See Figures 5 and 6).

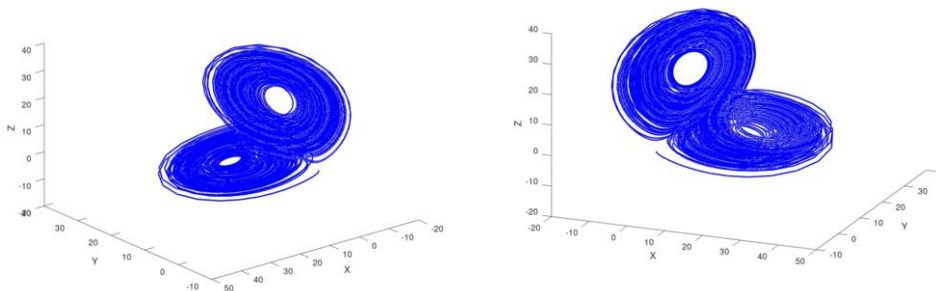


Figure 5. Display of the Lorenz Hyperattractor 4D: in $c = 21$ (left), in $c=22$ (right)

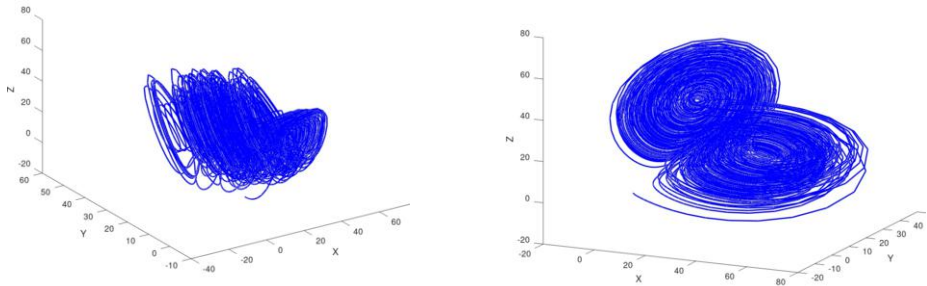


Figure 6. Display of the Lorenz Hyperattractor 4D: in $c = 35$ (left), in $c=40$ (right)

5. Conclusions

The **AutonomousSystems4D**, package allows qualitative analysis of ordinary non-linear differential equation systems in four dimensions, as well as visualization of phase surfaces. As a case study applied to the new Lorenz 4D System, sensitivity was found in the initial conditions, Lyapunov exponents, Kaplan Yorke dimension, a stable and unstable critical point, limit cycle, Hopf bifurcation, and hyperattractors.

References

- [1] Yu F, Gao L, Gu K, Yin B, Wan Q, & Zhou Z. A fully qualified four-wing four-dimensional autonomous chaotic system and its synchronization. *Optik*. 2017; 131:79-88.
- [2] Vaidyanathan S, Tielo-Cuautle E, Muñoz-Pacheco J, & Sambas A. A new four-dimensional chaotic system with hidden attractor and its circuit design. In 2018 IEEE 9th Latin American Symposium on Circuits & Systems (LASCAS). 2018; 1-4.
- [3] Benkouider K, Bouden T, & Halimi M. Analysis, circuit implementation and active control synchronization of a new 4D chaotic system with two quadratic nonlinearities. In 2019 IEEE 4th world conference on complex systems (WCCS). 2019; 1-6.
- [4] Kuznetsov YA. Elements of applied bifurcation theory. *Applied mathematical sciences*.1998; (112): 591.
- [5] Verhulst F. *Nonlinear Differential Equations and Dynamical Systems*. Springer. 1996.
- [6] Gang-Quan S, Hui C, Yan-Bin Z. A new four-dimensional hyperchaotic Lorenz system and its adaptive control. *Chinese Physics B*. 2012; 20(1): 010509.
- [7] Ma J, Yang Y. Hyperchaos numerical simulation and control in a 4D hyperchaotic system. *Discrete Dynamics in Nature and Society*. 2013; 980578.
- [8] Li C, Sprott JC. Coexisting hidden attractors in a 4-D simplified Lorenz system. *International Journal of Bifurcation and Chaos*.2014;24(03):1450034.
- [9] Panahi S, Pham V, Rajagopal K, Boubaker O, & Jafari S. A new four-dimensional chaotic system with no equilibrium point. In *Recent advances in chaotic systems and synchronization*. Academic Press. 2019; 63-76.
- [10] Tamba V, Kengne R, Kingni S, & Fotsin H. A Four-Dimensional Chaotic System With One or Without Equilibrium Points: Dynamical Analysis and Its Application to Text Encryption. In *Recent Advances in Chaotic Systems and Synchronization*. Academic Press. 2019; 277-300.
- [11] Velezmore R, Ipanaque R, & Mechato JA. A Mathematica Package for Visualizing Objects Immersed in R^4 . In *International Conference on Computational Science and Its Applications (ICCSA)*. Springer, Cham. 2019; 479-493.

Detecting Similar Versions of Software by Learning with Logistic Regression on Binary Opcode Information

Hyun-il LIM¹

Department of Computer Engineering, Kyungnam University, South Korea

Abstract. Logistic regression is widely used in decision problems to classify inputs through training from the previously known training data. In this paper, we propose an approach to detecting similar versions of software by learning with logistic regression on binary opcode information. Because the binary opcode information has detailed information for executing software on an individual machine, the learning from the binary opcode information can provide effective information in detecting similar versions of software. To evaluate the proposed approach, we experiment with two Java applications. The experimental results showed that the proposed logistic regression model can accurately detect similar versions of software after learning from training data. The proposed logistic regression model is expected to be applied in applications for comparing and detecting similar versions of software.

Keywords. Logistic regression, Software analysis, Similar version detection, Binary opcode analysis

1. Introduction

In recent computing environments, software plays an important role in various areas. To support efficient development and management of software, it is required to understand the characteristics of software. Software analysis is an approach to understand the specific characteristics of software. Detecting similar versions of software is one of the basic software analyses to figure out the similarity between versions of software and detect similar ones. The approach has various application areas, such as software similarity analysis [1, 2], code clone detection [3], or malware detection [4].

Machine learning [5] is an approach to generating a model for predicting solutions for given problems by learning previously known training data. Several related works on comparing the similarity of software by using machine learning have been studied, such as linear regression [6] and support vector machine [7]. Besides, deep neural networks have been applied through analyzing the n -grams of binary codes [8], common features of binary data [9], or images of binary codes [10].

Logistic regression is one of various machine learning approaches based on a statistical model to describe the relationship between independent and dependent

¹ Corresponding Author, Department of Computer Engineering, Kyungnam University, 7 Kyungnamdaehak-ro, Masanhappo-gu, Changwon, Gyeongsangnam-do 51767, South Korea; E-mail: hilim@kyungnam.ac.kr.

variables. This method is effectively used in binary decision problems for predicting the possibility of events. In this paper, we present an approach to applying binary code information to logistic regression to detect similar versions of software. Because it is not suitable to directly apply software as training data for analyzing and classifying software in logistic regression, it is essential to generate data for representing the features of software as training data. So, we present a method for generating training data of logistic regression for detecting similar versions of software. To evaluate the proposed approach, we experiment and show the experimental results for detecting similar versions in Java applications.

2. Logistic Regression for Software Analysis

Logistic regression [11] is a linear modeling approach to find the relation between the independent variable x and dependent variable y . In binary logistic regression, the dependent variable y has one of two values of 0 and 1 according to whether the predicting event occurred or not. The logistic regression model is generated from the learning of the relationship between the independent variable x and dependent variable y . The dependent variable y stands for the output data for presenting the occurrence of the event. The trained model can predict the possibility of the event for input data. So, in a binary prediction problem, the dependent variable has the value 0 or 1 according to the occurrence of the event. To formulate such a model, the representative logistic model is described as a function $g(x)$, that is continuously increasing between 0 and 1 according to the input variable x as follows:

$$g(x) = \frac{e^x}{1 + e^x}$$

This logistic model function can be used to model the possibility of an event on the input variable x to the result value between 0 and 1.

To design a logistic model for detecting similar versions of software, it is required to represent the features of software as input data of logistic regression. The output data are labeled 0 or 1, depending on whether the input data is for similar versions of software or not. To describe the features of software as input data, we consider the information of binary opcode of software. The binary opcode consists of instructions that are performed to accomplish a task in computing environments. So, the data represent how the task is performed to achieve the goal of the software. As the opcode information describes the characteristics of software at the instruction level, the information is an important criterion for distinguishing different versions of software. In this paper, we design a logistic regression model for detecting similar versions of software through data analyzed from binary opcode information that is derived by comparing the opcode distribution information of software. For example, when there are n types of instructions in software, let the opcode distribution information of two software A and B be (a_1, a_2, \dots, a_n) and (b_1, b_2, \dots, b_n) , respectively. Then, the distance data comparing the two software A and B is formulated as follows:

$$d(A, B) = (a_1 - b_1, a_2 - b_2, \dots, a_n - b_n)$$

So, the distance data is used as input data for logistic regression, and the model is trained with 1 or 0 according to the similarity of the two software.

Logistic regression is a method for finding a linear relationship between input and output to model the probability of a certain class. We design a logistic regression model for learning from the distance data of software for detecting similar versions. After the logistic regression model is trained with the distance data and label with 0 or 1 depending on the similarity of data, the model can be used as a classifier for detecting similar versions of software.

3. Designing Logistic Regression for Detecting Similar Versions of Software

In this section, we design the procedure for the logistic regression approach to detecting similar versions of software. Figure 1 shows the procedure for the designed approach. The binary code analyzer is a stage for analyzing the binary opcode from input software. It analyzes the structure of input software and generates the opcode distribution information. The binary opcode information needs to be compared with other data to distinguish similar versions of software. The training data generator compares the opcode information and makes training data set with labels 0 or 1 according to the similarity. After the logistic regression model is trained with the training data set, a logistic regression model is constructed for detecting similar versions of software.

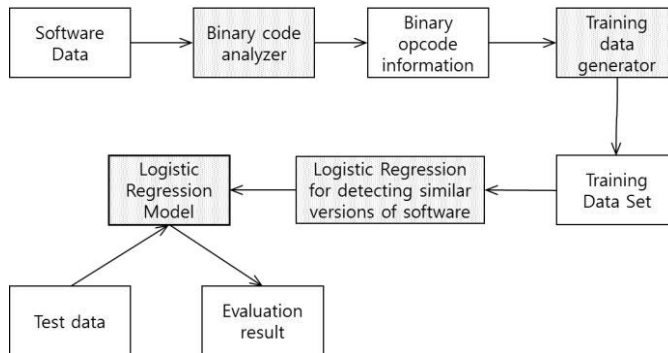


Figure 1. The procedure for applying logistic regression for detecting similar versions of software.

To evaluate the accuracy of the proposed approach, the generated model has experimented with a set of the other test data. The result can show the applicability of the proposed model in detecting similar versions by learning with logistic regression on binary opcode information.

4. Experimental Results

4.1. Experiments

In this section, we experiment on real-world Java applications to evaluate the accuracy of the proposed approach. As described in Figure 1, the binary code analyzer and the training data generator were implemented in Python to analyze binary opcode

information and generate the data set for logistic regression. The logistic regression for the data was implemented in Python and scikit-learn [12] to train the logistic regression model from the generated data and construct a model for detecting similar versions of software.

Table 1. The experimental environment for evaluating the logistic regression model for detecting similar versions of software.

Operating system	CPU	RAM	Binary opcode format
Microsoft Windows 10	Core i7-4790	32GB	Java bytecode

Table 1 shows the experimental environment for evaluating the proposed approach. The experiment was performed in Microsoft Windows 10 operating system with 32 GB of main memory. To evaluate the efficacy of the proposed method, we used Java bytecode as a binary opcode format. Java class files are executable files for Java Virtual Machine and they have Java bytecode as binary opcode to execute programs in the virtual machine. So, Java class files were analyzed to construct data sets of bytecode information. As benchmark software, we used Jakarta ORO and ANTLR for training and evaluation, respectively.

Table 2. The specification of the benchmark software used for training and testing in the experiment.

	Training Data	Test Data
Name of benchmark software	Jakarta ORO	ANTLR
Total # of class files	50	117
Total # of training or test data (similar versions)	1672 (38)	9672 (93)
Max # of bytecodes in a class file	923	1646
Average # of bytecodes in a class file	143.5	172.3
Max # of bytecodes in a similar version	1029	1705
Average # of bytecodes in a similar version	167.7	190.3

Table 2 shows the specification of the benchmark software for this experiment. The information on the numbers of Java bytecodes in benchmarking data is described. To ensure the reliability of the evaluation results, we used the class files that have more than 10 bytecodes as benchmark software. The data sets were generated by analyzing Java bytecode information and comparing the data from two versions of Java class files. The total numbers of data for training and testing were 1672 and 9672, respectively. Among the data, the numbers of data for similar versions were 38 and 93, respectively. To evaluate the detection accuracy for similar versions of software, we used the Smokescreen Java obfuscator to generate similar versions of the benchmark software. The Smokescreen obfuscator generates similar versions by modifying names, Java bytecode instructions, control flows to obfuscate internal structures of the original Java programs. So, the numbers of Java bytecodes in the similar versions were increased as compared to the original versions.

With the training data set from Jakarta ORO, we applied the logistic regression approach to generate a model for detecting similar versions. The generated detection model was evaluated with the test data set generated from ANTLR. From an analysis of

experimental results, we can confirm the applicability of the proposed approach in detecting similar versions of software.

Table 3. The evaluation results of the experiments for detecting similar versions of software with the test data described in Table 2.

	Evaluation Results
Total # of test data (similar versions)	9672 (93)
The # of detections for similar versions	93
The # of false detection for different versions	0
Detection accuracy	100%
Average training time (for 1672 training data)	12.0s

After generating a logistic regression model from the training data, the model was applied to the test data to evaluate how many similar versions of software can correctly be detected. Table 3 shows the evaluation results of the experiment. The number of total test data was 9672, and the number of data for similar versions was 93. The evaluation results showed that all the 93 similar versions in the test data were correctly detected. Besides, there was no false detection for the test data from different versions of software. The average training time was 12.0 seconds, and the detection accuracy was 100% for the test data. These results show that the model trained with the bytecode information of Java class files is highly effective in detecting similar versions of software.

4.2. Discussion and Future Work

From the evaluation results, we confirm that the logistic regression model on Java bytecode information can effectively detect similar versions of software. This is because the learning from the bytecode information can acquire the knowledge needed to distinguish different versions of software by adopting the features in the level of binary instructions. In the proposed approach, we have evaluated the accuracy for detecting similar versions of software with the model trained with the data of the same types of similar versions generated by the Smokescreen obfuscator. Because the training data with the previously known types of similar versions can reflect the exact information needed to detect similar versions of software, the accurate results can be achieved in the evaluation results. So, we confirm that well-prepared training data to reflect the actual environment is important for improving the accuracy of machine learning.

For the practical application of the proposed model for detecting similar versions of software, it needs to generalize the execution environment to adapt data with various types of similar versions. In the evaluation experiments, the model can be trained to fit well to the data of the same types of similarity. In the case where the exact type of similar versions is not clear, it will be difficult to construct accurate training data. In future work, we plan to evaluate the proposed model in environments without knowing the information on similar versions. To improve the ability to detect similar versions of software in real-world environments, we plan to analyze and reflect the additional information from software, such as control flows, opcode sequences, or function calls in the software, as the feature data for identifying the similarity. We also plan to experiment and compare the effectiveness in detecting similar versions of software with other machine learning algorithms.

5. Conclusion

Logistic regression is widely used in binary decision problems. In this paper, we designed an approach to detecting similar versions of software by learning with logistic regression on binary opcode information. Because the binary opcode information has detailed information for executing software, the learning from the information can be applied in modeling logistic regression for detecting similar versions of software. To evaluate the proposed approach, we implemented the bytecode analyzer and logistic regression model and experimented with two Java applications. The experimental results showed that the proposed logistic regression model can accurately detect similar versions of software. We confirm that the binary opcode information is effective data for reflecting the features of software for detecting similar versions. The proposed logistic regression model is expected to be applied as a reliable measure in detecting similar versions of software. Besides, a well-designed model for learning from data is expected to be applied to solve prediction problems in various areas.

Acknowledgements

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (Ministry of Education) (No. NRF-2017R1D1A1B03034769).

References

- [1] Niccolo M, Roberto G, Mila DP. A deep learning approach to program similarity. Proceedings of the 1st International Workshop on Machine Learning and Software Engineering in Symbiosis (MASES). September 2018. pp. 26-35.
- [2] Noam S, Nimrod P. Binary similarity detection using machine learning. Proceedings of the 13th Workshop on Programming Languages and Analysis for Security (PLAS). January 2018. pp. 42-47.
- [3] White M, Tufano M, Vendome C, Poshyvanyk D. Deep learning code fragments for code clone detection. Proceedings of the 31st IEEE/ACM International Conference on Automated Software Engineering (ASE), 2016. pp. 87-98. Singapore.
- [4] Danie G, Carles M, Jordi P. The rise of machine learning for detection and classification of malware: Research developments, trends and challenges. J Network Comp Appl. March 2020. 153(1).
- [5] Kevin P M. Machine Learning: A Probabilistic Perspective. The MIT Press. 2012.
- [6] Lim HI. A linear regression approach to modeling software characteristics for classifying similar software. Proceedings of COMPSAC, 2019. pp. 942-943.
- [7] Lim HI. Applying code vectors for presenting software features in machine learning. Proceedings of COMPSAC. 2018. pp. 803-804.
- [8] White M, Tufano M, Vendome C, Poshyvanyk D. Deep learning code fragments for code clone detection. 31st IEEE/ACM International Conference on Automated Software Engineering (ASE), 2016. pp. 87-98. Singapore.
- [9] Noam S, Nimrod P. Binary similarity detection using machine learning, Proceedings of the 13th Workshop on Programming Languages and Analysis for Security (PLAS). January 2018. pp. 42-47.
- [10] Niccolo M, Roberto G, Mila DP. A deep learning approach to program similarity. Proceedings of the 1st International Workshop on Machine Learning and Software Engineering in Symbiosis (MASES). September 2018. pp. 26-35.
- [11] Joseph MH. Practical Guide to Logistic Regression. Chapman and Hall. 2015.
- [12] Scikit-learn, Machine Learning in Python. <https://scikit-learn.org/stable/>.

Explainable AI: Using Shapley Value to Explain Complex Anomaly Detection ML-Based Systems

Jinying ZOU^a, Ovanes PETROSIAN^{b,1}

^a Faculty of Applied Mathematics and Control Processes, Saint-Petersburg State University, Saint-Petersburg, Russia

^b Faculty of Applied Mathematics and Control Processes, Saint-Petersburg State University, Saint-Petersburg, Russia

Abstract. Generally, Artificial Intelligence (AI) algorithms are unable to account for the logic of each decision they take during the course of arriving at a solution. This “black box” problem limits the usefulness of AI in military, medical, and financial security applications, among others, where the price for a mistake is great and the decision-maker must be able to monitor and understand each step along the process. In our research, we focus on the application of Explainable AI for log anomaly detection systems of a different kind. In particular, we use the Shapley value approach from cooperative game theory to explain the outcome or solution of two anomaly-detection algorithms: Decision tree and DeepLog. Both algorithms come from the machine learning-based log analysis toolkit for the automated anomaly detection “Loglizer”. The novelty of our research is that by using the Shapley value and special coding techniques we managed to evaluate or explain the contribution of both a single event and a grouped sequence of events of the Log for the purposes of anomaly detection. We explain how each event and sequence of events influences the solution, or the result, of an anomaly detection system.

Keywords. Anomaly detection, Log anomaly detection, Shapley value, DeepLog, Decision tree, Explainable AI

1. Introduction

Recently, deep learning has made great contributions toward the rapid development of AI. For machine learning, especially deep learning, explainable AI is a big challenge. Deep neural networks are a black box for us all. AI algorithms usually cannot explain the logic of each decision when providing a solution. Such opaque decisions are not adequately persuasive, especially in the fields of military, medical and financial security where stakes are high. Therefore explainable AI would be helpful:

- For users when AI technology is designed to offer solutions or help take decisions. System users should then understand why the system provides each specific solu-

¹Corresponding author is associate professor at the Department of Mathematical Modelling of Energetic Systems of Saint Petersburg State University. Email: petrosian.ovanes@yandex.ru

tion. For example, a doctor who makes a diagnosis needs to be able to understand why the medical diagnostic system makes such a recommendation [1].

- For developers in order to understand the black box of deep learning. They can improve their methods and models for building better machine learning models and improving system capabilities [2].

Anomaly detection is an important problem that has been well-studied within diverse research areas and domains of application. A common need when analyzing real-world data-sets is determining which instances stand out as being dissimilar to all others. Such instances are known as anomalies, and the goal of anomaly detection (also known as outlier detection) is to determine all such instances in a data-driven fashion [3]. Anomalies can be caused by errors in data but sometimes are indicative of a new, previously unknown, underlying process. Hawkins in [4] defines an outlier as an observation that deviates so significantly from other observations as to arouse suspicion that it was generated by a different mechanism. The most common causes of outliers or anomalies in a data set are *Data entry errors* (human errors), *Measurement errors* (instrument errors), *Experimental errors* (data extraction or experiment planning/executing errors), *Intentional* (dummy outliers made to test detection methods), *Data processing errors* (data manipulation or data set unintended mutations), *Sampling errors* (extracting or mixing data from faulty or disparate sources) and *Natural* (not an error but a novelty in the data). With respect to methods or algorithms, anomaly detection should be classified as *Supervised*, *Unsupervised*, *Hybrid Models* and *one-Class Neural Networks*. By application, anomaly detection can be classified by *Intrusion Detection*, *Fraud Detection*, *Malware Detection*, *Medical Anomaly Detection*, *Social Networks Anomaly Detection*, *Log Anomaly Detection*, *Internet of things (IoT) Big Data Anomaly Detection*, *Industrial Anomalies Detection*, *Anomaly Detection in Time Series*, *Video Surveillance*. More details can be found in a more comprehensive recent overview [5]. There are also some potential approaches, used to increase the performance and precision when obtaining the anomaly, such as successful geometric transformations model [6] combined with regression model [7] and Ito decomposition [8], to overcome the time limitations. In this paper we focus on *Log Anomaly Detection* and study two anomaly detection algorithms. One involves a very classic and widely-used decision tree algorithm and the second one a more modern and advanced deep learning DeepLog algorithm. However the aim of the paper is not to study at some point the best anomaly detection algorithms, but rather to advance Explainable AI techniques by applying them to anomaly detection systems of a different sort.

Anomaly detection algorithms are often thought to be limited because they cannot facilitate the process of validating results performed by domain experts. This is an urgent challenge for the industry. In 2019, Antwarg used the SHAP framework [9] to explain an anomaly detection system. They treat each feature as a player, and provide users with a more intuitive understanding by measuring each player's contribution to the solution. SHAP is based on the notion of optimal Shapley Value [10], which is a well-known solution concept from cooperative game theory [11]. Originally the Shapley value defines how to allocate profit, costs or, more generally, a utility among the players acting cooperatively. In the case of Explainable AI, the Shapley value can show the contribution each feature makes to the result of the anomaly detection system. It is important to note that the Shapley value does not only show the individual contribution of a feature to the result of the detection system, but also shows the contribution of a feature to all possible

combinations of the feature that constitutes the anomaly. The SHAP approach itself for XAI was proposed by Lundberg in [12]. Other related papers on SHAP are described below. In [13] authors present an improved SHAP using the Baseline Shapley (BShap) method which they further extend by using Integrated Gradients to the continuous domain. The paper [14] explores the dependence between SHAP values by extending the KernelSHAP method to handle dependent features. In the paper [15] authors described an extension of the SHAP method for trees under a framework called TreeExplainer to examine the global model structure using local explanations. Later the paper [16] describes a SHAP-based method to account for the predictions of time-series signals involving Long Short-Term Memory (LSTM) networks.

Besides SHAP, there are several other useful and applied algorithms for explaining black box (machine learning) algorithms, but in this paper our particular interest lies in the XAI approaches based on the use of the Shapley value:

- LIME is a method that interprets individual model predictions based on building a local approximation the model around a given prediction [17].
- DeepLIFT (Deep Learning Important Features) [18] is a method for decomposing the output prediction of a neural network on a specific input by back-propagating the contributions of all neurons in the network to every feature of the input.
- LRP (Layer-wise Relevance Propagation) [19] is a method that brings such explanative ability to potentially highly complex deep neural networks. It operates by propagating the prediction backward in the neural network using a set of purposely designed propagation rules.

For a more comprehensive and fundamental overview for Explainable AI approaches and models see [20].

In this paper we apply the Shapley approach to two anomaly detection systems with different structures. The first is the Decision tree [21], [22] in which we treat a single feature as a player (the feature-player approach). We differ from Antwarg's approach [9] in that we treat different events in the anomaly detection system as a player based on the data itself without considering the algorithm model. Nor do we use the SHAP framework [9], but develop our own framework based on the Shapley value. The second anomaly detection system we want to explain is DeepLog [23]. Current mainstream research for Shapley-related XAI treats a single feature as a player and then analyzes the contribution of each player to the result. But this approach has limitations for a class of anomaly detection systems where the anomaly can be the result not only of one feature, but by a sequence of features. Therefore, we consider the sequence of events as a player for modeling. The key challenge is a large number [24] of coalitions in the sequence-player approach in comparison to that of a feature-player approach. In order to avoid the problem of a big number of sequence players, we use the bi-level method for calculating the Shapley value. A related method for cooperative coalitional games can be found in [25]. On the first level we consider one event as a player, on the second level we consider a sequence of two events as a player.

In section 2, we briefly introduce algorithms for both Decision tree and DeepLog. In Section 3, we describe the Shapley value and how to apply it to explain the decision tree and DeepLog. In section 4, we attach our simulation result. In section 5 presents conclusions and discussions for future work.

All our research is based on the open source project "Loglizer", which is a machine learning-based log analysis toolkit for automated anomaly detection [26]. Our code for this research was uploaded to Github: <https://github.com/ZouJinying/XAIloglizer>.

2. Anomaly Detection Algorithms

2.1. Decision tree

In 2004 Mike Chen and others proposed using Decision tree in order to classify failed and successful requests for anomaly detection in large system logs [21]. Decision tree is a tree structure composed of nodes and branches. The nodes are divided into leaf nodes (representing a certain category) and internal nodes (representing a certain attribute or feature). Branches represent a test output. The basic idea of Decision tree is to use entropy as a measure to select the attributes of its nodes. Each selects the attributes with the largest gains, that is, the attribute with the smallest entropy value. When the entropy equals zero, all the instances in node are considered to be the same cluster. Below see a brief version of an anomaly detection Decision tree algorithm applied to system logs analysis:

1. Choose the event as a root. The best root will be chosen according to the information gain. The equation for calculating the information gain entropy is shown in Eq. (1) as below:

$$G(D, a) = H(D) - H(D | a) = H(D) + \sum_{i=1}^N \frac{|D^i|}{|D|} H(D^i), \quad (1)$$

where, $G(D, a)$ is information gain, $H(D)$ is a summary of information entropy for all features in the set D , $H(D | a)$ is the conditional information entropy under the condition of feature a .

2. Divide the samples into two sub-trees and find the maximum gain.
3. Continue the iterations (from step1 to step3) until there are no events or features left in input data set.
4. Each branch of tree is a prediction result that displays if the log is anomaly or normal.

2.2. DeepLog

DeepLog is a data-driven algorithm that uses large number of system logs for anomaly detection. The main intuition behind the design of DeepLog comes from natural language processing: treat log entries as sequence elements that follow certain patterns and grammatical rules [23]. As distinct from the log message counter-method, DeepLog is a deep neural network that uses long and short-term memory to model log sequences. Therefore, the importance of the sequence of events is greater than the message count of events. This study thus considers the sequence as a player to explain each specific solution. In this way we can analyze which sequences contribute more to the accuracy of the prediction and whether the sequences are meaningful.

Figure 1 shows the architecture of DeepLog which contains two parts: training and detection. The training data for DeepLog comes from the system's logs. The log is com-

bined by a log key and parameter value vector. At the training stage, the log needs to first be parsed. Then the obtained event sequence can be used as the training input for the detection model. After training is completed, the system can judge whether the log is normal according to the log key. If it is normal, DeepLog will further check the parameter value vector. If the parameter value vector is anomalous, it will be marked as an anomalous log.

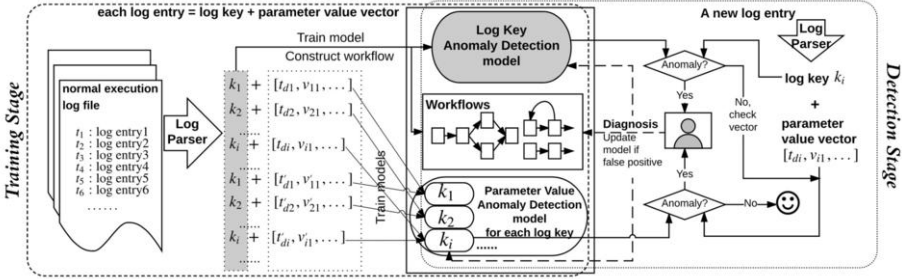


Figure 1. DeepLog architecture.

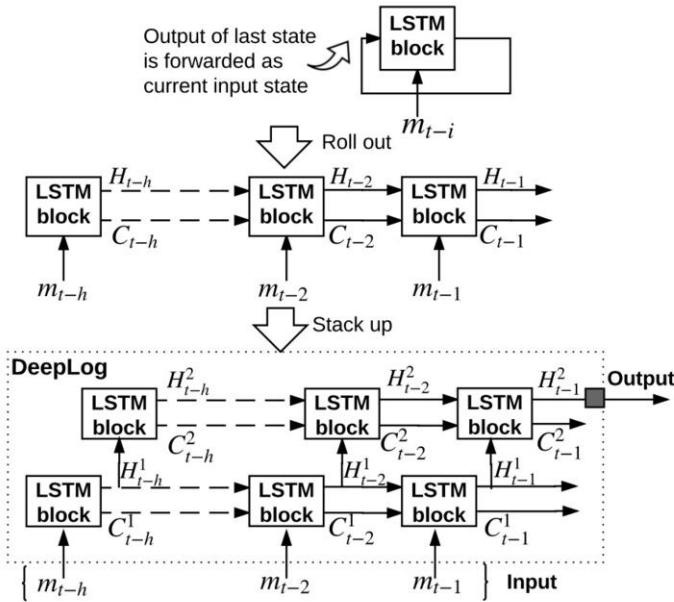


Figure 2. View of anomaly detection using stacked LSTM.

DeepLog uses Long Short-Term Memory (LSTM) in the recurrent neural network (RNN) as a framework to observe the long-term dependence of the sequence, figure 2. The top is single-nature LSTM block. Each LSTM block remembers the state for its input as a vector of fixed dimension. m_{t-i} is example of input. The center shows a series of LSTM blocks where a hidden vector H_{t-i} and state vector C_{t-i} are maintained in each cell. Both hidden vector and state vector will be passed to next as initial input to maintain

historical information. The bottom of figure 2 shows an example of DeepLog with two hidden layers of a deep neural network scheme. The input layer encodes n possible log keys from K as one-hot vectors. The output layer, which uses a standard multinomial logistic function, translates n layer of hidden state into a probability distribution function [23].

3. Shapley Value Approach for Explainable AI

3.1. Shapley Value in Game Theory

In the classical cooperative game theory it is assumed that players cooperate and together obtain a total reward and then distribute this reward between each other [27]. In order to share the total reward, the notion of imputation is introduced. In this paper as an imputation we use a classical cooperative solution Shapley value [10]. The explicit equation for the Shapley value is presented in Eq. (2) as below:

$$\varphi_i = \sum_{S|i \in S \subseteq N} \frac{(|S| - 1)!(|N| - |S|)!}{|N|!} [v(S) - v(S \setminus \{i\})], \quad i \in N, \quad (2)$$

where i is the number of players, N is the set of all players in the cooperative game, $S \subseteq N$ is a coalition of players, $|S|$ is the number of players in coalition S , $V(S)$ is a characteristic function of coalition $S \subseteq N$ that specifies the total payoff of coalition S .

In order to use the Shapley value in any domain it is necessary to calculate the values of characteristic function $V(S)$ for each coalition $S \subseteq N$. More details about how to calculate it for explaining anomaly detection in a log system will be presented in the following sections. In cooperative game theory, as can be seen from the Eq. (2), the Shapley value of player i shows the weighted sum of contributions of player i to the cooperation reward of each coalition S from N (term $[v(S) - v(S \setminus \{i\})]$). The left multiplier in the product (2) defines the probability that the coalition S itself will be formed, therefore the less the probability the less important is the individual contribution of player i to cooperation.

Using the Shapley value, the total reward is allocated among the players ($\sum_{i=1}^n \varphi_i = V(N)$). Generally speaking, if the contribution of the player to the cooperation is big, then his value of imputation is also big. In machine learning, the Shapley value approach can be applied to explain the contribution of each feature value to the overall solution.

3.2. Shapley Value for Decision Tree

The basic idea of Decision tree is to use a top-down recursive method to take information entropy as a measure to construct the fastest falling entropy value. The entropy value at the leaf node is zero. At this time, the instances in each leaf node belong to the same class. In other words, the essence of a Decision tree is a set of if-then rules. A Decision tree divides the feature space into disjoint units or regions.

In order to apply the approach based on Shapley value on the first step we need to calculate the values of the characteristic function for each coalition $S \subseteq N$, where N and S are the set and the subset of all features or unique events in the system log correspondingly. The meaning of characteristic function $V(S)$ of coalition S for an anomaly

detection system is the anomaly detection value or probability of anomaly based only on events from the set S . After calculating the characteristic function for each coalition S , it is possible to compute the Shapley value and, as a result, explain the result of Decision tree by explaining the contribution of each feature. The algorithm to calculate the Shapley value for the Decision tree anomaly detection is:

1. Define $n = |N|$ feature player from data set D .
2. Choose coalition $S \subseteq N$ of feature players. The total number of coalitions to consider is $\sum_{k=0}^n C_n^k$.
3. Run Decision tree algorithm for each coalition $S \subseteq N$ to obtain the value of characteristic function $V(S)$ or the accuracy of anomaly detection.
4. Repeat step 2 – 3 until all values of characteristic function $V(S)$ are calculated.
5. Using the Eq. (2) calculate the Shapley value for all players.

3.3. Shapley Value for DeepLog

The DeepLog approach allows a user not only to find a feature or event that is the anomaly, but also to define a sequence of events that can lead to an anomaly in the system log. Therefore, Explainable AI should also address this issue by introducing the explaining not only for a set of individual events, but also for a sequence of events.

Here we also define the accuracy of anomaly prediction for the set of features S using the value characteristic function $V(S)$. But we consider 2 events and 1 target event as a feature or player, figure 3. All sequences will be divided and sorted by the target event. We will calculate the value of characteristic function $V(S)$ for sets S of sequence events in system log instead of the set of single events to verify the contribution of each sequence. The workflow of DeepLog is presented below:

1. Log analysis: The goal of log analysis is to convert unstructured log messages into structured event mapping, based on which complex machine learning models can be applied.
2. Feature extraction: usually structured logs can be cut into short log sequences through interval windows, sliding windows or conversation windows. In DeepLog, we use sliding windows for feature extraction, and vectorize each extracted sequence.
3. Anomaly detection: After the model is obtained through the training data, anomaly prediction is performed on the extracted data.

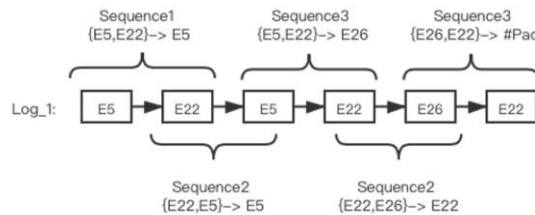


Figure 3. Example of sequence approach.

According to figure 3, short event sequences will be extracted from raw system long sequences to compose the set of features or players. This procedure is different from the

one in Decision tree algorithm. In DeepLog algorithm, the sequence approach is considered to measure the contribution of each sequence to the anomaly detection. The sequence approach leads to the problem of a large number of coalitions or a curse of dimensionality. To avoid this problem we use a bi-level approach for calculating the Shapley value or sequence estimation for anomaly detection. On the first level, as a feature or player we consider only the target event, while on the second level for each fixed target event we consider the sequence of two events as a player. Therefore the number of players or features on the first level corresponds to the number of unique target events $E \in N$ in the system log and is equal to $|N|$. On the second level the number of features or players is various depending on the related target event E , and is equal to the number of unique sequences of two events $\{E_i, E_j\} \in N_E$ before the target event E ($|N_E|$), $E \in N$. Later for the first level the values of characteristic function $V(S)$, $S \subseteq N$ are calculated using the same approach as for the Decision tree algorithm. But for the second level the values of characteristic function $V_E(S)$, $S_E \subseteq N_E$ are calculated for each fixed target event $E \in N$. Later Shapley values are calculated using the Eq. (2) both for first level φ^f and for the second level with sequences of events φ_s . The resulting Shapley value for each sequence of events is obtained by the multiplication of Shapley values on the first and second levels. The algorithm to calculate the Shapley value for DeepLog anomaly detection is:

1. Using sliding window = 3 extract features (target events and corresponding two sequence events).
2. Define G_n as sequence player group, where n is number of unique target players.
3. chose i -th coalition, i from 0 to $\sum_{k=0}^n C_n^k$.
4. Run DeepLog algorithm to obtain the accuracy of prediction.
5. Update the characteristic function for i -th coalition with accuracy prediction. $i++$, repeat step 3 to 5 till all characteristic function are obtained.
6. Use Eq. (2) to calculate the Shapley value for all players.

4. Simulation Results

In this section we present the simulation results of explaining the solution of pretrained Decision tree and DeepLog. For the Decision tree we use 2 data sets and for the DeepLog we use 1 data set of system logs. Simulations are run on the System: MacOS, CPU: 2.6 GHz, RAM: 8GB.

Link to data sets: <https://github.com/logpai/loglizer/tree/master/data/HDFS>

4.1. Simulation Results of Decision Tree

For the Decision tree we consider two input data sets. Both system logs consist of $n = 3900$ instances of the system log with 20 unique events. The initial system log is filtered to reduce the number of unique events to 14 because the other 6 appear rarely. Therefore, the total number of coalitions can be calculated in the following way: $\sum_{k=0}^{14} C_{14}^k = 16383$. The process of how to define players and coalitions using the initial system log is presented in figure 4. Using the Eq. (2), and values of characteristic function $V(S)$, we calculate the Shapley value. Below in figures 5 and 6 see the results for two test data sets. Here the contribution of each event of system log to the solution of Decision tree using the test data sets is defined.

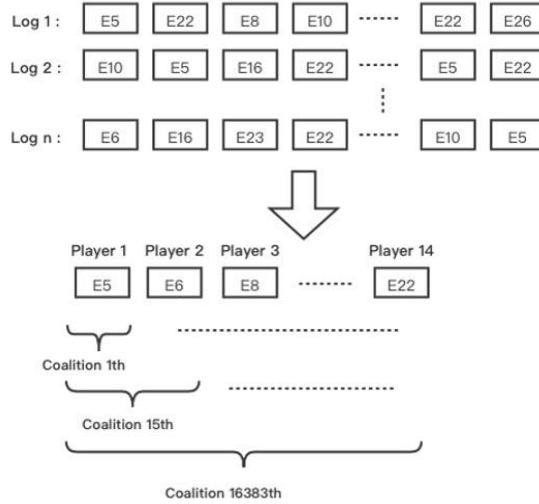


Figure 4. Connection between system log and components or cooperative game.

Figure 5 shows that the event *E11* plays the most important role in the anomaly detection for data set 1 for Decision tree. The important conclusion for the input data set 1 is that, even if we remove all events apart from event *E11* from the system log, then the anomaly detection would still be good because the Shapley value for event *E11* is equal to 0.9636. In figure 6 note the simulation results for the input data set 2. Here events *E9*, *E11* and *E26* play an important role. This means that each of these events provoked an anomaly in the system log, Shapley values for events *E9*, *E11* and *E26* are equal to 0.322, 0.313, 0.317 correspondingly.

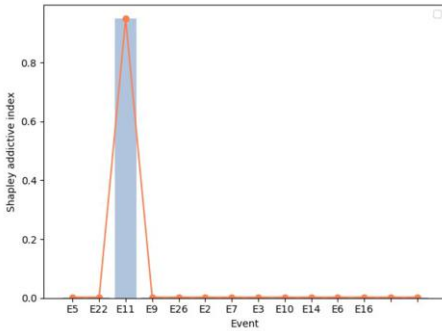


Figure 5. Shapley value with input data set 1.

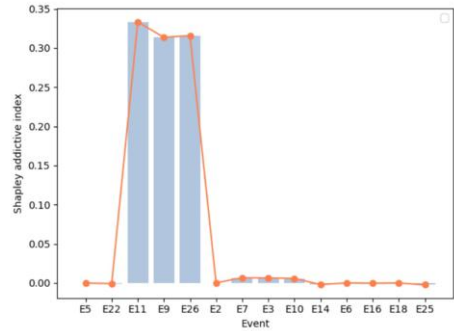


Figure 6. Shapley value with input data set 2.

4.2. Simulation Results of DeepLog

In figure 3 the process of bi-level method is shown. The window length for the second level of equal to 3.

After extracting features, we obtain a new data structure. The probability of an event after the sequence is the key to anomaly detection. Low probability means an anomaly

while a high probability means a normal sequence. Figure 7 also has element "#Pad", which means the end of the sequence. We can control whether players participate in the local level game by setting the event as Nan. On the second level all the sequences with the unique target event will be considered as players.

Test data set contains 1875 instances, that contain 15000 sequences of events. After cleaning the data, we obtain 7 unique events. In order to implement procedure described in section 3.3 we consider two cooperative game models: first level game model for groups of events, second game model for sequences of events in each group.

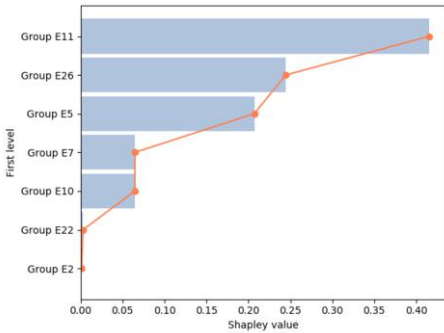


Figure 7. Shapley value for group players in the first level game.

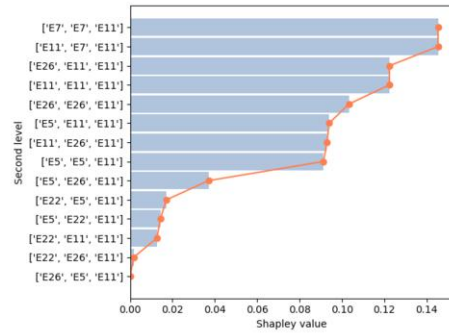


Figure 8. Shapley value for sequence players in the second level game for target event E11.

In the first level game model we define a player as a group of event sequences that correspond to a specific target event from the initial system log, see figure 3. Then we calculate the values of characteristic function and the Shapley value. It is important to notice that this is not yet the explaining result for the solution of DeepLog corresponding to the input data set. Figure 7 shows the contribution of each group of sequences corresponding to the target events: E11, E26, E5, E7, E10, E22, E2. It is easy to see that the contribution 0.415 to the result of DeepLog of the group E11 is the biggest, while the contribution of the group E2 is only 0.0011.

The second level game model is constructed separately for each group related to a target event. A player in the second level game is a sequence of events with the fixed target event. Using the Shapley value it is possible to define a contribution of each event sequence to the anomaly detection of event group. Figure 8 shows the contribution of each sequence of events to the anomaly detection for group of events related to the target event E11. Similar results are obtained for the target events E2, E5, E7, E10, E22, E26.

Finally when the Shapley values both for the first and the second level game models (figures 7 and 8) are obtained the Shapley value for each sequence of events is calculated using the approach described in section 3.3. The meaning of this Shapley value is the contribution of sequence of events to the anomaly detection of DeepLog using the test data set. Figure 9 shows the contribution of each sequence of events to the anomaly detection of DeepLog using the test data set. It is easy to see sequences {E7, E7, E11}, ..., {E5, E5, E11} have the biggest contribution to the anomaly detection. This means that the anomaly of the system log from test data set is concentrated in these sequences. The user of DeepLog system should check in detail what is the nature of these sequences to make a conclusion.

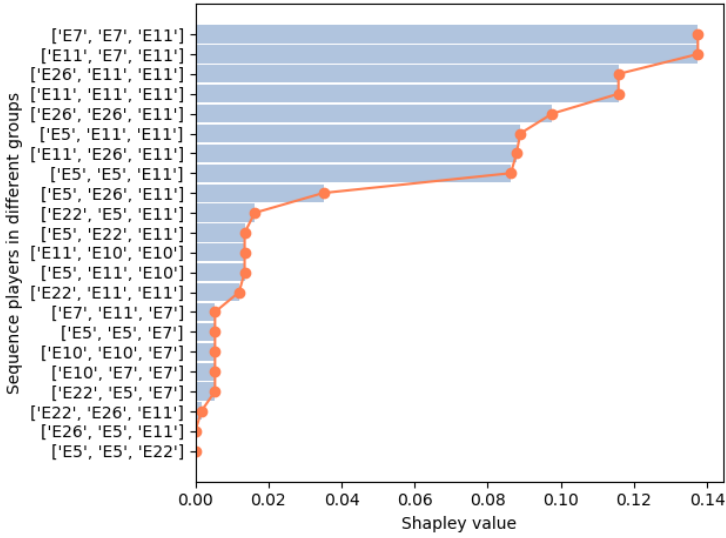


Figure 9. Shapley value for the result of anomaly detection using DeepLog with test data set.

5. Conclusion and Future work

At the current stage we are applying the Shapley value approach to static machine learning modeling, in particular to the Decision tree and DeepLog algorithms. As shown in the simulation section our approach can explain a specific solution using both single feature and sequence events. As it is indicated in figure 9, the user of anomaly detection system that has the explainable module could easily find the reason of the anomaly (specific event or sequence of events), not only the receive the anomaly alarm.

In the future, we will apply several sampling methods presented in the papers [28], [29], [30] to approximate the Shapley value. This would aid in evaluating the contribution for each unique sequence instead of using a bi-level method. And think about indicators to evaluate the precision and reliability of our approach. Besides, we plan to optimize the algorithm to decrease the time consumption and be comparable with well known framework SHAP [9].

6. Acknowledgement

The work of the second author is supported by Russian Foundation for Basic Research (RFBR) according to the research project No. 18-00-00727 (18-00-00725).

References

- [1] Lundberg SM et al. Explainable machine-learning predictions for the prevention of hypoxaemia during surgery. *Nature Biomed. Engin.*, 2.10 (2018): 749-760.
- [2] Holzinger A et al. What do we need to build explainable AI systems for the medical domain? arXiv preprint, arXiv:1712.09923. – 2017.
- [3] Chandola V, Banerjee A, Kumar V. Anomaly detection: A survey. *ACM Computing Surveys*, 2009, 41(3).
- [4] Hawkins D. Identification of outliers. Springer Netherlands, 1980, P. 188.
- [5] Chalapathy R, Chawla S. Deep learning for anomaly detection: a survey, arXiv: Learning, 2019.
- [6] Tkachenko R, Izonin I. Model and principles for the implementation of neural-like structures based on geometric data transformations. *Adv Intell Syst Comput* 754: 578–587.
- [7] Izonin I, Tkachenko R, Kryvinska N, Tkachenko P. Multiple linear regression based on coefficients identification using non-iterative SGTm Neural-Like Structure. In *International Work-Conference on Artificial Neural Networks*, Springer, Cham, 2019 June, pp: 467-497.
- [8] Tkachenko R, Izonin I, Vitynskyi P, Lotoshynska N, Pavlyuk O. Development of the noniterative supervised learning predictor based on the ITO decomposition and SGTm neural-like structure for managing medical insurance costs. *Data*, 2018, 3(4), 46.
- [9] Antwarg L, Shapira B. Explaining anomalies detected by autoencoders using SHAP. arXiv preprint, arXiv:1903.02407. – 2019.
- [10] Shapley LS. (August 21, 1951). Notes on the n-Person Game – II: The Value of an n-Person Game. Santa Monica, Calif.: RAND Corporation.
- [11] Leon A P, Nikolay A Z. *Game Theory* (2nd Edition), World Scientific, 2016.
- [12] Lundberg SM, Lee SI. A unified approach to interpreting model predictions. *Neural Inform. Processing Syst.*, 2017, pp. 4765–4774.
- [13] Sundararajan M, Najmi A. The many shapley values for model explanation. arXiv preprint, arXiv:1908.08474, 2019.
- [14] Aas K, Jullum M, Løland A. Explaining individual predictions when features are dependent: More accurate approximations to shapley values. arXiv preprint, arXiv:1903.10464, 2019.
- [15] Lundberg SM, Erion G, Chen H, DeGrave A, Prutkin JM, Nair B, Katz R, Himmelfarb J, Bansal N, Lee SI. From local explanations to global understanding with explainable AI for trees. *Nature machine intelligence*, 2020, 2(1): 2522–5839.
- [16] Vega Garcia M, Aznarte JL. Shapley additive explanations for NO2 forecasting. *Ecol. Inform.*, Mar 2020, 56: 101039.
- [17] Ribeiro MT, Singh S, Guestrin C. Why should I trust you? Explaining the predictions of any classifier. *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining*, – 2016. – . 1135-1144.
- [18] Shrikumar A, Greenside P, Kundaje A. Learning important features through propagating activation differences //arXiv preprint arXiv:1704.02685. – 2017.
- [19] Montavon G et al. Layer-wise relevance propagation: an overview. *Explainable AI: interpreting, explaining and visualizing deep learning*. – Springer, Cham, 2019. – . 193-209.
- [20] Arun D, Paul R. Opportunities and challenges in explainable artificial intelligence (XAI): a survey, arXiv, 2020.
- [21] Chen M, Zheng AX, Lloyd J, Jordan MI, Brewer E. Failure diagnosis using decision trees. *International Conference on Autonomic Computing*, 2004. *Proceedings.*, New York, NY, USA, 2004, pp. 36-43, doi: 10.1109/ICAC.2004.1301345.
- [22] Liang YL, Zhang YY, Xiong H, Sahoo R. Failure Prediction in IBM BlueGene/L Event Logs.
- [23] Anomaly detection and diagnosis from system logs through deep learning. *Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security*, 2017, 1285-1298.
- [24] Xu W, Huang L, Fox A, Patterson D, Jordan MI. Large-Scale System Problems Detection by Mining Console Logs.
- [25] Nikolay VK. Strategic stability of coalitions technological alliance parameters: a two-level cooperation. *Contributions to Game Theory and Management*, 2015, Volume 8, 111–136
- [26] He SL, Zhu JM, He PJ, Michael RL. Experience report: system log analysis for anomaly detection. *IEEE International Symposium on Software Reliability Engineering (ISSRE)*, 2016.
- [27] Shapley L. A value for n-person games. *Contributions to the Theory of Games*. 1953. 2(1): 307–317.

- [28] Mann I, Shapley LS. Values of large games 6: Evaluating the electoral college exactly. Tech. Rep., Rand Corp Santa Monica CA (1962).
- [29] Castro J, Gomez D, Tejada J. Polynomial calculation of the shapley value based on sampling. *Comput. Oper. Res.* 2009, 36: 1726–1730.
- [30] Maleki S, Tran-Thanh L, Hines G, Rahwan T, Rogers A. Bounding the estimation error of sampling based shapley value approximation. *arXiv preprint, arXiv*, 2013, 1306.4265.

A Simple Ensemble Learning Knowledge Distillation

Himel DAS GUPTA ^{a1}, Kun ZHANG ^b Victor S. SHENG ^c

^aDepartment of Computer Science, Texas Tech University, Lubbock, Texas, 79409, USA

^bDepartment of Computer Science, Bioinformatics facility of Xavier RCMI Center of Cancer Research, Xavier University of Louisiana, New Orleans, Louisiana, 70125, USA

^cDepartment of Computer Science, Texas Tech University, Lubbock, Texas, 79409, USA

Abstract. Deep neural network (DNN) has shown significant improvement in learning and generalizing different machine learning tasks over the years. But it comes with an expense of heavy computational power and memory requirements. We can see that machine learning applications are even running in portable devices like mobiles and embedded systems nowadays, which generally have limited resources regarding computational power and memory and thus can only run small machine learning models. However, smaller networks usually do not perform very well. In this paper, we have implemented a simple ensemble learning based knowledge distillation network to improve the accuracy of such small models. Our experimental results prove that the performance enhancement of smaller models can be achieved through distilling knowledge from a combination of small models rather than using a cumbersome model for the knowledge transfer. Besides, the ensemble knowledge distillation network is simpler, time-efficient, and easy to implement.

Keywords. ensemble, bagging, knowledge distillation

1. Introduction

The vast spectrum of artificial intelligence has imposed significant impacts over the years in different applications. We have witnessed the state of art results achieved in the field of computer vision [1], natural language processing [2], speech recognition [3], and autonomous vehicle [4]. The performance of machine learning models in such applications much depends on the architecture of the model even if adequate training data is available for training the networks. Integrating more layers and more parameters can certainly improve the accuracy of a machine learning model. However, it comes with the cost of requiring powerful computing systems. Small devices like mobile and embedded systems can barely provide such computational power. Besides, it takes a lot of time to complete these tasks. In order to deploy well in practical fields, such cumbersome architectures become hectic to run. The idea of knowledge distillation (KD) has brought

¹Corresponding Author: Himel Das Gupta, Department of Computer Science, Texas Tech University, Texas, USA; E-mail: Himel.Das@ttu.edu

prominent results in such situations where a small model is trained to mimic like a large model. The fundamental idea of knowledge distillation is to transfer the generalization ability of a complex teacher model to a simpler student model. A recent popular approach is to transfer the output class probabilities (from 0 to 1) of a big neural network to a student model. The student model tries to learn from these probabilities, instead of hard class labels. Thus, the student model can imitate the generalization of the teacher model on unseen data.

The classification accuracy of Neural networks can be improved with the help of ensemble learning [5] like bagging. With bagging, we can learn multiple independent models together and average their predictions to obtain a final prediction with a lower variance. There exist different ways to improve the performance of ensemble learning like in [6] where authors have demonstrated the relation in multistrategy ensemble learning between test error reduction and the generation of diversity in ensemble membership. However, in this paper, we propose a simple ensemble knowledge distillation method (called EKD) to improve the performance of a single model, instead of the ensemble model. Specifically, it first performs knowledge distillation from ensemble learned models, and then use the knowledge distilled to train a single model. Note that this approach is also different from traditional knowledge distillation (KD), since traditional knowledge distillation usually uses a comparatively bigger model (deep learning model) to distill the knowledge from the bigger model for training a simple single model. Our experimental comparisons show that knowledge transfer from a cluster of small models can be as effective as transferring knowledge from a cumbersome model.

Previous work like [7] has achieved a significant improvement in terms of accuracy using ensemble knowledge distillation. Now, there are some differences between their work and ours. First of all, their ensemble teacher model consists of multiple convolutional neural networks (CNN) with different architectures whereas we have used same CNN architecture for our ensemble teacher model. Secondly, their student model is a composition of multiple branches where each branch is represented as separate neural network architecture. On the other hand, our student model is a simple single neural network architecture. And most importantly, their framework they have connected these two networks (ensemble teacher network + compact student network) together whereas our teacher, student model are separate. To elaborate these, we have to focus on their distillation process where the main difference is lying. To start with, though their teacher model is trained at first separately, it gets fine-tuned simultaneously and collaboratively with the student model. And the distillation process is done by the objective function which is summation of "teacher model loss + student model loss + distillation loss" as the whole network is connected. In our model, training phase of teacher and student model is completely separate where we trained the ensemble teacher model first, calculate the softmax probabilities and then use these probabilities as constraint while training the student model. And the loss function for our student model is consists of only student model loss and distillation loss.

In a nutshell, We will discuss related work in Section 2. We are going to demonstrate the technical aspects of knowledge distillation and our EKD in Section 3. In Section 4, we will define our experiment setup and all the preprocessing steps. Analysis and empirical comparisons of our experiments will be discussed in Section 5. The paper ends with conclusions and future work opportunities in Section 6.

2. Related Works

The term "knowledge distillation" or "teacher-student model" has been first proposed by [8] which has been further improved by [9]. In [8] a novel method called "Model Compression" was first introduced where a complex, large network can be compressed into a smaller model. This paper has built the foundation of knowledge distillation. Additional improvement has been done by [10] where the authors have considered not only the output layers but also the intermediate layers for transferring class probabilities to the student model using L2 loss. Knowledge distillation has been used in reinforcement learning also [11]. In [12] authors have represented the distilled transfer knowledge as FSP Matrix which is generated by two layers. Instead of compressing model, distillation can also be achieved by training parameters of a student model identically to their teacher which is defined as Born-again networks in [13]. Adversarial based learning strategy has been used in [14] to distill the diverse knowledge from a compressed large trained ensemble networks. In [15] cluster of different architectural recurrent networks has been used as ensemble distillation learning to improving accuracy in speech recognition. Ensemble based knowledge distillation can have superior performance improvement over the traditional knowledge distillation shown in [16] where it uses data augmentation. What it does is, creates multiple copies of data with respect to the soft output targets from various teachers model. As mentioned in the introduction, in [7] ensemble learning has been improvised by using multiple branches of student model where the branches are trained by a teacher model using it's ground truth labels and information. Improvements in task like binary classification has been done using ensemble learning [17] where bagging technique is applied to an ensemble of pairs of neural networks. These networks have been used to predict degree of truth membership, indeterminacy membership, and false membership values in the interval neutrosophic sets. Ensemble learning can be very handy in improving classification performance for Deep Belief Network(DBN) too. In [18] a new mechanism called Pseudo Boost Deep Belief Network(PB-DBN) has been proposed in this regard where top layers are boosted while lower layers of the base classifiers share weights for feature extraction. A novel method called Generalized Regression Neural Network (GEFTS-GRNN) has been proposed in [19] where the authors combined a single GRNN from multiple base level GRNN to produce the final output. In our paper, we have shown a much easier way to implement a ensemble knowledge distillation network.

3. Background and Implemented Model

3.1. knowledge distillation (KD)

The architecture of a neural network is organized such way that we can get the probabilities of the classified classes by imposing "softmax" activation function on it. The general equation of such output layer is like this $y_i = \exp(x_i/T) / \sum_j \exp(x_j/T)$ in [8]. Where x_j is the logit, j is the number of classes, and y_i is the class probability. Here T is denoted as temperature value which is usually 1 [8]. The higher value in temperature signifies the softer probability distribution of the classes. The main idea of distillation is to transfer these probabilities as knowledge from the cumbersome model to the smaller model. It can be achieved by making the soften probabilities of teacher

model as target for the small model. For understanding, often the cumbersome model is defined as "teacher" model and the smaller model is defined as "student" to express the idea of student learning knowledge from the teacher.

In [8] the authors used weighted average of two functions to train the student model to produce correct labels in addition to the teacher's soft labels. The first objective function is the cross entropy with the correct labels whereas cross entropy with the soft labels is considered as second objective function. The distillation process is propagated by the custom loss function like in Eq. (1),

$$\begin{aligned}\mathcal{L}_{student} &= \alpha\mathcal{L}_{CL} + (1 - \alpha)\mathcal{L}_{KD} \\ \mathcal{L}_{KD} &= T^2KL(y_s, y_t)\end{aligned}\tag{1}$$

Here \mathcal{L}_{KL} is the built in KL DIVERGENCE loss, \mathcal{L}_{CL} is the normal cross-entropy loss, T is the temperature value and y_s, y_t are the targets softened for the student model. The hyperparameter α emphasizes between weighted average of the two loss functions.

3.2. Ensemble knowledge distillation (EKD)

To demonstrate our EKD framework, we have combined multiple small models as a single teacher model. That means, our teacher ensemble model is a composition of small models rather than one single cumbersome model like in a traditional knowledge distillation framework. **Figure 1** shows our combined network of ensemble knowledge distillation (EKD). Where $y_1 \dots y_n$ symbolizes the "softmax" predictions of the n number of ensemble models. In a traditional knowledge framework, the transferring of generalization ability of a cumbersome model to a small models is done by imposing the class probabilities of the cumbersome model as "soft targets" for the small model during the training phase. In our implemented method, instead of taking class probabilities of a single cumbersome model, we took the arithmetic mean of the class probabilities produced by each model of the ensemble model. So, $(y_1 \dots y_n)/n$ represents the probabilities that has been used as "soft targets" for the small student model in our implemented model.

4. Experiment Setup

4.1. Architecture of Used Networks

To conduct our experiments, we used four different neural networks which we are going to denote as **LargeCNNnet** (6 layer CNN), **SmallCIFAR10net** (2 layer CNN), **SmallCIFAR100net** (3 layer CNN), and **SmallMnistNet** (2 layer MLP). The application of these network in our experiments is summarized in **Table 1**. The motivation behind the architecture of these models was to show the implementation of the actual ensemble knowledge distillation method rather than getting the highest accuracy. We constructed the architecture of these models in such way that it doesn't take too much computation power and memory to run but significant enough to show the differences between a traditional knowledge distillation and ensemble knowledge distillation framework.

Table 1. Used neural network architectures.

Dataset	Teacher Model	Student Model
CIFAR10 (KD)	LargeCNNnet	SmallCIFAR10Net
CIFAR10 (EKD)	n*SmallCIFAR10Net	SmallCIFAR10Net
CIFAR100 (KD)	LargeCNNnet	SmallCIFAR100Net
CIFAR100 (EKD)	n*SmallCIFAR100Net	SmallCIFAR100Net
MNIST(KD)	LargeCNNnet	SmallMnistNet
MNIST(KD)	n*SmallMnistNet	SmallMnistNet

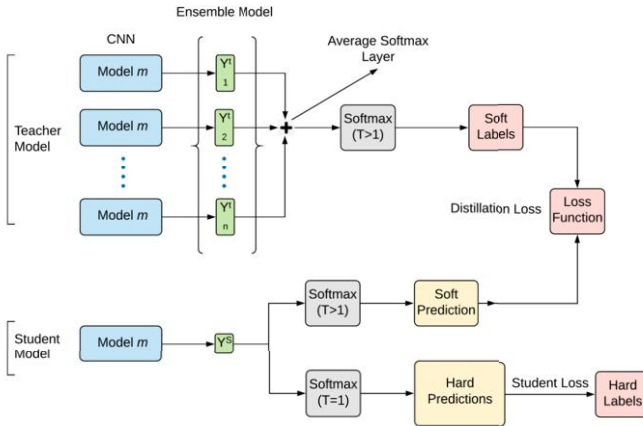


Figure 1. Combined Model: Ensemble Distillation Network

4.2. Datasets

To conduct our experiment, we have used three different datasets. The CIFAR10, CIFAR100 and MNIST. Both CIFAR10 and CIFAR100 datasets are consists of 60,000 images with pixel density of $32 \times 32 \times 3$. There are 10 classes in CIFAR10 and 100 classes in CIFAR100 with 6000 and 600 images of each class respectively. Out of 60,000 samples we used 50,000 for our training purposes and 10,000 for testing phase. The MNIST dataset consists of 60,000 handwritten digit images including 10,000 training samples.

4.3. Parameters, Hyperparameters and Loss Functions

Both CIFAR10 and CIFAR 100 experiment has been done using 25 epochs and 64 batch. MNIST experiment has been done using 4 epochs with 64 batch because 4 epochs was enough to provide accuracy around 80% in our EKD model. For bagging ensemble we used 5 combine models of SmallCIFAR10net as teacher model for CIFAR10 experiment and SmallCIFAR100net for CIFAR100 experiment. For distillation purpose, we used temperature value of 3. We used "adam" optimizer and "sparse categorical crossentropy" loss function for our teacher model.

5. Result Analysis and Discussion

We have created separate python script for all the experiments. Evaluating our experiments we see some significant results shown in **Table 2**. We are going to discuss on the result one by one.

Table 2. Empirical Comparison between normal knowledge distillation (KD) and ensemble knowledge distillation (EKD) in terms of accuracy

Model	CIFAR10		CIFAR100		MNIST	
	KD (%)	EKD (%)	KD (%)	EKD (%)	KD (%)	EKD (%)
Distilled Student Model	72.2	71.76	34.59	42.05	77.57	79.95
Student Model Alone	70.28	70.44	35.85	34.39	76.32	77.83
Improvement	1.92	1.32	-1.26	7.66	1.25	2.12

Table 3. Training time Comparison between normal knowledge distillation (KD) and ensemble knowledge distillation (EKD).

Datasets	Teacher Model			Student Model		
	KD (sec)	EKD (sec)	Improvement(sec)	KD (sec)	EKD (sec)	Improvement (sec)
CIFAR100	6625	3125	3500	850	750	100
CIFAR10	3825	2750	1075	600	650	-50
MNIST	575	87.5	487.5	25	25	0

5.1. CIFAR10 and CIFAR100

In the case of CIFAR10 dataset classification, although KD has shown better improvement than EKD, we can observe that EKD has obtained 1.32% performance gain on student model which emphasize the fact that teacher model can be a collection of small model and still can improve the standalone model accuracy. In the case of the CIFAR100 experiment, significant performance gain has been observed in the ERD mechanism whereas traditional distillation was not successful using the big model as a teacher model. But using EKD, performance has been enhanced by 7.66%. Additionally, we can observe that it took less time in case for EKD to perform the whole process compared to KD in **Table 3**. Also, to check the influence of hyperparameter like "number of ensemble network". We conducted three separate experiments with 5, 10 and 20 ensemble networks and for CIPAR10 we got improvement in our student model in this sequence 1.32%, 0.49%, and 2.83%. In our future research we will try to explore more on this. These results from the **Table 2** amplifies the fact that performance improvement in a classification task can be achieved through distillation by using a combination of small models like SmallCIFAR10net like the same way in using a cumbersome model like LargeCNNnet for distillation even better sometimes.

5.2. MNIST

In the case of MNIST dataset, our EKD network outperformed KD by providing 2.12% performance improvement. So, for both CIFAR100 and MNIST dataset, our model gained more accuracy than a traditional KD does.

These results from the **Table 2** approves the fact that performance improvement in a classification task which has been achieved by transferring distilled knowledge from a cluster of small models can be as powerful as transferring from a cumbersome model, and sometimes even better. In future, we also want to provide more statistical evaluation with additional experiments of our proposed mechanism to signifies the impact of our simple ensemble knowledge distillation framework.

6. Conclusions and Future Work

In this paper, we proposed a simple ensemble knowledge distillation approach called EKD. Our experimental results showed that EKD can improve the accuracy of a single learning model through transferring the distilled knowledge from an ensemble models. It performs better than the traditional knowledge distillation using a cumbersome model as the teach model, especially on the CIFAR100 and the MNIST dataset. Specifically, on the CIFAR100 dataset, the experimental result shows that EKD achieved the highest accuracy (around 42.05%), which is much higher than 34.59% (achieved by KD).

Although EKD performs significantly better than KD on CIFAR 100, it loses to KD on CIFAR 10. This motivates us to conduct further study to investigate potential reasons. We will also investigate the performance of hierarchical knowledge distillation, including traditional knowledge distillation and ensemble distillation.

References

- [1] Forsyth DA, Ponce J. Computer vision: a modern approach. Prentice Hall Professional Technical Reference; 2002 Mar 1.
- [2] Manning C, Schütze H. Foundations of statistical natural language processing. MIT press; 1999 May 28.
- [3] Graves A, Mohamed AR, Hinton G. Speech recognition with deep recurrent neural networks. In 2013 IEEE international conference on acoustics, speech and signal processing 2013 May 26 (pp. 6645-6649). IEEE.
- [4] Schwarting W, Alonso-Mora J, Rus D. Planning and decision-making for autonomous vehicles. Annual Review of Control, Robotics, and Autonomous Systems. 2018 May 29.
- [5] Dietterich TG. Ensemble learning. The handbook of brain theory and neural networks. 2002 Mar;2:110-25.
- [6] Webb GI, Zheng Z. Multistrategy ensemble learning: Reducing error by combining ensemble learning techniques. IEEE Transactions on Knowledge and Data Engineering. 2004 Aug 2;16(8):980-91.
- [7] Asif U, Tang J, Harrer S. Ensemble knowledge distillation for learning improved and efficient networks. arXiv preprint arXiv:1909.08097. 2019 Sep 17.
- [8] Bucilua C, Caruana R, Niculescu-Mizil A. Model compression. In Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining 2006 Aug 20 (pp. 535-541).
- [9] Hinton G, Vinyals O, Dean J. Distilling the knowledge in a neural network. arXiv preprint arXiv:1503.02531. 2015 Mar 9.
- [10] Romero A, Ballas N, Kahou SE, Chassang A, Gatta C, Bengio Y. Fitnets: Hints for thin deep nets. arXiv preprint arXiv:1412.6550. 2014 Dec 19.
- [11] Schmitt S, Hudson JJ, Zidek A, Osindero S, Doersch C, Czarnecki WM, Leibo JZ, Kuttler H, Zisserman A, Simonyan K, Eslami SM. Kickstarting deep reinforcement learning. arXiv preprint arXiv:1803.03835. 2018 Mar 10.
- [12] Yim J, Joo D, Bae J, Kim J. A gift from knowledge distillation: Fast optimization, network minimization and transfer learning. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition 2017 (pp. 4133-4141).
- [13] Furlanello T, Lipton ZC, Tschannen M, Itti L, Anandkumar A. Born again neural networks. arXiv preprint arXiv:1805.04770. 2018 May 12.
- [14] Shen Z, He Z, Cui W, Yu J, Zheng Y, Zhu C, Savvides M. Adversarial-based knowledge distillation for multi-model ensemble and noisy data refinement. arXiv preprint arXiv:1908.08520. 2019 Aug 22.
- [15] Chebotar Y, Waters A. Distilling Knowledge from Ensembles of Neural Networks for Speech Recognition. In Interspeech 2016 Sep (pp. 3439-3443).
- [16] Fukuda T, Suzuki M, Kurata G, Thomas S, Cui J, Ramabhadran B. Efficient Knowledge Distillation from an Ensemble of Teachers. In Interspeech 2017 Aug 20 (pp. 3697-3701).
- [17] Kraipeerapun P, Fung CC. Binary classification using ensemble neural networks and interval neutrosophic sets. Neurocomputing. 2009 Aug 1;72(13-15):2845-56.
- [18] Duan T, Srihari SN. Pseudo boosted deep belief network. In International Conference on Artificial Neural Networks 2016 Sep 6 (pp. 105-112). Springer, Cham.
- [19] Gheyas IA, Smith LS. A novel neural network ensemble architecture for time series forecasting. Neurocomputing. 2011 Nov 1;74(18):3855-64.

Using Conversation Analysis for Examining Social Media Interactions

Matteo Farina^{a,1}

^a*The University of Adelaide and Flinders University*

Abstract. Social media research has grown exponentially in recent years. However, it seems that to date only a few studies have applied Conversation Analysis (CA) to study social media interactions. The aim of this paper is to show the benefits of using CA for the analysis of this type of social interactions.

Keywords. Conversation Analysis (CA), social media research, Facebook.

1. Introduction

In recent years many statisticians, mathematicians and computational scientists have turned their attention to social media research [1]. This has resulted in a huge amount of literature [2-6] that applied scientific approaches to analyze social media interactions. Although these studies identified the general patterns that characterize these interactions, they did not analyze them from an interactional perspective. This paper tries to fill this gap in social media research. It proposes to use Conversation Analysis (CA) for the study of social media interactions. After a brief summary of the origins of CA and a discussion on why this methodological approach was used in the past for analyzing technology-mediated social interactions, this article discusses the case of @-formulations in Facebook comment threads to show the benefits of using CA for social media research.

2. Conversation Analysis

2.1. Origins

Conversation Analysis (CA) is a methodological approach that studies social interactions [7-9]. It emerged in the late 1960s from the field of ethnomethodology through the work of Harvey Sacks, Emmanuel A. Schegloff and Gail Jefferson [10]. According to Sacks [7], people perform actions through talk. They can greet friends, request information, accept or decline invitations, ask questions and so on. Furthermore, actions performed through talk are interactionally achieved; they are the result of the collaborative work performed by participants in an interaction [7, 9]. Thus, if a speaker wants to successfully perform a greeting, s/he may say “hello” to an intended recipient

¹ Corresponding Author, Matteo Farina, College of Humanities, Arts and Social Sciences, Flinders University, Bedford Park South Australia 5042; E-mail: alicematteo@gmail.com.

and wait for a response. If the recipient fails to respond, the greeting is incomplete; only the First Pair Part (FPP) of an adjacency pair sequence has been performed, while a Second Pair Part (SPP) is missing [8]. This shows that actions accomplished through talk are organized in pairs, they are made up of two components, FPPs and SPPs; FPPs are turns at talk that launch sequences whereas SPPs are pieces of talk that complete sequences [8]. If one part is missing, the action is incomplete. Moreover, if like in the example above, a speaker utters a FPP and a recipient fails to produce a SPP, the action is not mutually achieved. Only one participant contributed to the interaction.

Although many scholars claim that CA is mainly suitable for investigating spoken conversation, this is not the case. Since its development, this methodological approach has been used for the examination of technology-mediated social interactions. For example, Schegloff used CA to study phone conversations [11]. This study describes the impact of the technology on the sequential organization of these interactions. It shows how the affordances and constraints of the telephone affect communications. This research also demonstrates that phone conversations are opened by identification and recognition sequences. These sequences are typical of phone conversations; they are used by conversationalists to identify their interlocutors. In telephone conversations people do not see each other; thus, participants use talk to overcome a limitation of the technology. Therefore, previous studies have pointed out that CA is a suitable approach for the examination of technology-mediated social interactions.

2.2. CA and Computer-Mediated Communication

With the introduction of the Internet in the 1980s social interactions became computerized. People started communicating with each other using Internet-Relay Chats (IRC)s, emails and online discussion forums [12]. In other words, everyday social interactions began to shift to digital environments. As a result, conversation analysts turned their attention to Computer-Mediated Communication (CMC). They studied synchronous as well as asynchronous interactions such as IRCs, discussion forums, weblogs and email exchanges [13-19]. They used CA to analyze the moment-by-moment unfolding of computer-mediated interactions. Conversation analysts also examined the role of technology in CMC. They noticed that despite being either synchronous or asynchronous, computer-mediated communications are *conversational-like*. They are characterized by turn-taking [13, 14], repair [15] and sequential organization [19]. However, like phone conversations, the affordances and constraints of the technology have an impact on the interaction. For example, although turn-taking is identifiable in online chats, it is affected by the constraints of the platform used for communicating. Turn-taking was introduced by conversation analysts to describe what normally happens in spoken conversation, that is, that conversationalists speak one at a time. In other words, speakers do not talk simultaneously, when a speaker talks another listens and speaker changes occur with limited gaps and overlaps [7-9]. However, this is not the case in IRCs. In many IRCs systems, there is a delay between the processes of message production and message availability. Although messages that are published are immediately available to participants for inspection, messages that are being constructed are not available to participants. Therefore, there is a delay between the time in which a message is typed and when it is made available to recipients on screen. As a result, Garcia and Jacobs classified IRCs as quasi-synchronous [14]. More importantly, the delay between message construction and message availability affects the turn-taking system. That is, in online chats a participant might post a FPP that

initiates a sequence, and before a recipient has posted a SPP that completes the pair, s/he might publish another FPP that launches another sequence. Thus, disrupting the linearity of the turn-taking system that establishes that FPPs and SPPs should be adjacently positioned one after another within adjacency pairs [7, 9]. This disruption of the sequential organization of chat interactions is possible because of the affordances of the technology [20, 21]. In other words, chat platforms allow adjacency pairs to be disrupted, which means that participants can post messages that start new sequences even before other participants have responded to them.

Therefore, CA is a powerful tool for the micro analysis of the moment-by-moment unfolding of CMC interactions. This approach enables researchers to reveal how sequentiality is achieved in computer-mediated communications. Moreover, it allows scholars to examine the role of the features of the technology in the interaction.

2.3. Social Media as Loci for Social Interaction

In recent years social media have become loci for social interaction [22]. 3.6 billion people use social media websites worldwide [23]; 2.5 billion are active on Facebook, 2.0 billion watch videos on YouTube and use WhatsApp, and 1.2 communicate on WeChat [24]. Individuals use social networking sites to communicate and maintain social relationships [25, 26]. Moreover, as per phone conversations and other types of computer-mediated communications, interactions occurring on social media are *conversational-like*. Therefore, in the last few years, a growing number of studies has started applying CA to examine these interactions [27-34]. These studies have shown that interactions occurring on social media are multimodal, they consist of textual and visual elements such as texts, images, videos, hyperlinks as well as messages that are made up of the combinations of both these elements such as texts with images, texts with videos and texts with hyperlinks [27]. Moreover, social media interactions are sequentially organized. Facebook comment threads [28, 29] for instance, mainly consist of contributions that respond and orient to messages that launch comment threads. In other words, comments that follow an initial post generally respond to it, even when they are not adjacently positioned after this contribution. Thus, as per online chats, linearity does not seem to affect the sequential organization of Facebook comment threads. Nevertheless, Facebook users seem to successfully interpret and respond to comments even when they are not adjacently positioned one after the other in a comment thread. This depends on visual saliency [35], an affordance of many text-based communication systems. On Facebook, visual saliency enables users to access comments and create sequentiality. In other words, Facebook users can scroll comments up or down on screen to put FPPs and SPPs together and create sequences, regardless of their position in a comment thread.

The studies discussed in this and the previous paragraphs demonstrated that CA is a useful methodological approach for analyzing technology-mediated social interactions. They have shown that CA is fruitful to examine how “conversationalists” adapted their interactional practices to suit the affordances and constraints of the technology. The following section, on the other hand, will show that CA can also be used to prove that sometimes technologies have been modified to accommodate the interactional practices developed by their users.

3. @-formulations in Facebook interactions

Facebook introduced the *Reply* button in 2013 [36]. This button is positioned under every comment posted in a comment thread (with the exception of the first comment in a thread) and allows users to respond to comments from Friends. In other words, this button enables Facebook users to post comments one after the other in a comment thread and achieve sequentiality. It appears that the *Reply* button was introduced to overcome a limitation of Facebook, that was that users did not have control over the positioning of comments within comment threads. Before the introduction of the *Reply* button, comments were organized in a chronological order; the last comment presented on screen was the most recent contribution published in a comment thread. So, when a user published a comment in a thread, s/he did not have control over its positioning in the interaction, this comment was always positioned as the last comment in the thread. This disrupted the sequentiality of interactions creating an interactional issue for Facebook users; they could not post a comment directly after the comment it was responding to. To overcome this limitation of the technology Facebook users started using @-formulations such as the @ symbol followed by a user's name (e.g. @Matteo Farina) to indicate to which comment they were responding to as clearly shown by extract 1 below:

Extract 1 [@-formulations in Facebook comment threads]²

1 Mauro: ho Deciso . . . A CARNEVALE MI VESTO DA METROMAN

Mauro: I have decided . . . AT THE NEXT CARNEVALE PARTY I'LL DRESS UP AS METROMAN

[. . .]

9 Mina: Lo so che è ovvio . . . purtroppo ti conosco :D

Mina: I know it's obvious . . . unfortunately I know you :D

10 Richard: sexy

Richard: sexy

11 Daniele: @Mina: purtroppo ti conosco è un insulto gravissimo . . . io non ti parlerei + a vita ;) @Mauro: io mi vesto da palo della metro, così puoi attaccarti al palo e girare!!!

Daniele: @Mina: unfortunately "I know you" is a terrible offence . . . I would feel like not talking to you anymore;) @Mauro: I will dress up as an underground pole, so you can attach yourself to the pole and swing around me!!!

Extract 1 is the transcript of a comment thread about Metroman, a busker who performs on the underground in Milan, Italy. After a series of comments [1-10] where Friends mock Mauro for humorously announcing his plan for the next Carnevale³ party, which is to dress up as Metroman, at comment 11 Daniele uses two @-formulations (@Mina; @Mauro) to select the recipients of his contribution. These @-formulations not only explicitly indicate the addressees of Daniele's comment but they also preserve the sequentiality of this comment thread. In other words, Daniele uses them to indicate to which comments he is responding to. The initial component of comment 11 clearly

² Please note that data in this extract have been transcribed and translated from Italian.

³ Carnevale is an Italian celebration where people go to parties wearing masks and dressing up as fantasy characters, such as Pinocchio, Mickey Mouse, and so on.

orients to the comment at Post 9, where Mina uses the expression “ti conosco” (*I know you*) to mock Mauro’s plan of dressing up as Metroman. In this part of the comment, Daniele jokingly tells Mina that he considers “ti conosco” to be a terrible offence (@Mina: *purtroppo ti conosco è un insulto gravissimo . . . io non ti parlerei + a vita ;*)). On the other hand, the final component of the comment at Post 11 orients to Post 1. Daniele humorously says that if Mauro dresses up as Metroman, he will dress up as an underground pole, so Mauro can swing around him (@ Mauro: *io mi vesto da palo della metro, così puoi attaccarti al palo e girare!!!*). This comment refers to a video posted in another comment thread where Metroman swings around a pole while performing on the underground. Although occurring 10 comments after the contribution at post 1, this message responds to it.

This analysis of Extract 1 shows that CA is a useful approach to demonstrate how technologies have sometimes been modified to suit the needs of their users.

4. Conclusion

This paper has demonstrated that CA is a useful approach for the investigation of social media interactions. This methodology enables researchers to make a fine-grained micro analysis of the moment-by-moment unfolding of these interactions. Furthermore, CA can be used to examine the role of the technology in social media interactions. This methodological approach can also reveal the interactional practices developed by social media users to overcome the limitations of the technology. In other words, CA provides researchers with an opportunity to study social media interactions from an interactional perspective. Therefore, this paper argues for the integration of CA into scientific research on social media. On one hand, this kind of interdisciplinary research might result in the development of more effective computational tools for social media interaction. On the other hand, it might be used for managing social problems; for instance, by creating softwares that detect posts that promote cyberhate and extremism in social media.

References

- [1] Resnyansky L. Conceptual frameworks for social and cultural Big Data analytics: Answering the epistemological challenge. *Big Data Society*. 2019;6(1):1-12.
- [2] Garton L, Haythornthwaite C, Wellman B. Studying online social networks. *J Computer-Mediated Comm*. 1997;3(1).
- [3] O’Callaghan D, Greene D, Conway M, Carthy J, Cunningham P. An Analysis of interactions within and between extreme right communities in social media. *Ubiquitous Social Media Analysis*. 2012:88-107.
- [4] Honey C, Herring CS. Beyond microblogging: Conversation and collaboration via Twitter. 2009 42nd Hawaii International Conference on System Sciences Ieee; 2009. p. 1-10.
- [5] Ritter A, Cherry C, Dolan B. Unsupervised modeling of twitter conversations. *Human Language Technologies: The 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics* 2010:172-80.
- [6] Weng J, Lim E-P, Jiang J, He Q, editors. *Twitterrank: finding topic-sensitive influential twitterers*. 3rd ACM international conference on Web search and data mining; 2010.
- [7] Sacks H. *Lectures on conversation*. Jefferson G, editor. Cambridge: Blackwell Publishers; 1992.
- [8] Sacks H, Schegloff E, Jefferson G. A simplest systematics for the organization of turn-taking for conversation. *Language*. 1974; 50(4): 696-735.
- [9] Schegloff E. *Sequence Organization in Interaction*. Cambridge: Cambridge University Press; 2007.

- [10] Liddicoat AJ. An introduction to conversation analysis. London: Continuum International Publishing Group; 2011.
- [11] Schegloff E. Identification and recognition in telephone conversation openings. In: Psathas G, editor. *Everyday Language: Studies in ethnomethodology*. New York: Irvington; 1979. p. 23-78.
- [12] Giles D, Stommel W, Paulus T, Lester J, Reed D. Microanalysis of online data: the methodological development of "digital CA". *Discourse, Context & Media*. 2015(7):45-51.
- [13] Garcia CA, Jacobs BJ. The interactional organization of computer mediated communication in the college classroom. *Qualitative Sociology*. 1998;21(3):299-317.
- [14] Garcia CA, Jacobs BJ. The eyes of the beholder: understanding the turn-taking system in quasi-synchronous computer-mediated communication. *Res. Language Social Inter*. 1999; 33(4):337-67.
- [15] Schönfeldt J, Golato A. Repair in chats: a conversation analytic approach. *Res. Language Social Interaction*. 2003; 36(3):241-84.
- [16] Herring CS. A faceted classification scheme for computer-mediated discourse. *Language@Internet*. 2007;1(1):1-37.
- [17] Herring CS. An approach to researching online behavior. In: Barab SA, Kling, Kling R, Gray JH, editors. *Designing for Virtual Communities in the Service of Learning*. New York: Cambridge University Press; 2004. p. 338-76.
- [18] Antaki C, Ardévol E, Núñez F, Vayreda A. "For she who knows who she is:" Managing accountability in online forum messages. *J Computer-Mediated Commun*. 2005; 11(1):114,32.
- [19] Reed D. 'Making Conversation': sequential integrity and the local management of interaction on internet newsgroups. *Proceedings of the 34th Hawaii International Conference on System Sciences*; Maui, Hawaii2001.
- [20] Herring CS. Interactional Coherence in CMC. *J Computer-Mediated Comm*. 1999; 4(4).
- [21] Berglund TÖ. Disrupted turn adjacency and coherence maintenance in instant messaging conversations. *Language & Internet*. 2009;6(2):1-25.
- [22] Golder S, Wilkinson D, Huberman B. Rhythms of social interaction: messaging within a massive online network. In: Steinfield C, Pentland BT, Ackerman M, Contractor N, editors. *Communities and Technologies*. London: Springer; 2007. p. 41-66.
- [23] Clement J. Number of social network users worldwide from 2017 to 2025 2020 [cited 2020 22 July]. Available from: <https://www.statista.com/statistics/278414/number-of-worldwide-social-network-users/>.
- [24] Clement J. Most popular social networks worldwide as of April 2020, ranked by number of active users: Statista.com; 2020 [cited 2020 22 July]. Available from: <https://www.statista.com/statistics/272014/global-social-networks-ranked-by-number-of-users/>.
- [25] Kabilan MK, Ahmad N, Abidin MJZ. Facebook: An online environment for learning of English in institutions of higher education? *The Internet and Higher Education*. 2010; 13: 179-87.
- [26] Leonardi M. *Narrative as self performance: the rhetorical construction of identities on Facebook profiles*. Albuquerque: The University of Nex Mexico; 2009.
- [27] Farina M. Facebook first post telling. *Journal of pragmatics*. 2015; 90:1-11.
- [28] Farina M. *Facebook and Conversation Analysis*. London: Bloomsbury; 2018.
- [29] Frobenius M, Harper R. Tying in comment sections: The production of meaning and sense on Facebook. The self-explicative organization of communication acts on and through Facebook. *Semiotica*. 2015(204):121-43.
- [30] Meredith J. Analysing technological affordances of online interactions using conversation analysis. *Journal of pragmatics*. 2017(115):42-55.
- [31] Meredith J, Stokoe E. Repair: Comparing Facebook 'chat' with spoken interaction. *Discourse Comm*. 2014;8(2):181-207.
- [32] Housley W, Webb H, Edwards A, Procter R, Jirotko M. Digitizing Sacks? Approaching social media as data. *Qualitative Res*. 2017;17(6):627-44.
- [33] Tolmie P, Procter R, Rouncefield M, Liakata M, Zubiaga A. Microblog analysis as a program of work. *ACM Transactions on Social Computing*. 2018;1(1):1-40.
- [34] Zubiaga A, Liakata M, Procter R, Wong Sak Hoi G, Tolmie P, Masuda N. Analysing how people orient to and spread rumours in social media by looking at conversational threads. *PloS One*. 2016; 11(3).
- [35] Pelletier J. Negotiation in cyberspace: The role of chatting in the development of grammatical competence. *Network-based language teaching: Concepts and Practice*. Cambridge, New York: Cambridge University Press; 2000. p. 9-86.
- [36] Murphy S. Facebook Finally Adds 'Reply' Option to Comments 2013 [cited 2020 27 July]. Available from: <https://mashable.com/2013/03/25/facebook-replies/>.

A Study on the Effect of DropConnect to Control Overfitting in Designing Neural Networks

Hyun-il LIM¹

Department of Computer Engineering, Kyungnam University, South Korea

Abstract. The neural network is an approach of machine learning by training the connected nodes of a model to predict the results of specific problems. The prediction model is trained by using previously collected training data. In training neural network models, overfitting problems can occur from the excessively dependent training of data and the structural problems of the models. In this paper, we analyze the effect of DropConnect for controlling overfitting in neural networks. It is analyzed according to the DropConnect rates and the number of nodes in designing neural networks. The analysis results of this study help to understand the effect of DropConnect in neural networks. To design an effective neural network model, the DropConnect can be applied with appropriate parameters from the understanding of the effect of the DropConnect in neural network models.

Keywords. DropConnect, neural network, overfitting, machine learning

1. Introduction

As information technology advances, the software is becoming an important factor in solving various types of problems. Machine learning [1-3] is a way to solve problems through training models with data using software technology. This is a way to solve problems using data models trained with previous data. As machine learning has been used in various areas, the design of the machine learning model for solving problems is important in improving the accuracy of the model. Therefore, it is necessary to understand the features of the machine learning model to design it well.

In training neural network models, the accuracy of results should be increased from the information on the training data. The overfitting problem [4-6] may occur when the trained models are excessively customized for the training data, which can lead to poor performance in general data. Therefore, it is needed to reduce such overfitting in training neural network models [5]. The DropConnect [7, 8] is an approach to reducing overfitting by temporarily excluding connected links in the process of training. Such simplification of the link structures of neural networks can reduce the model's over-customization of training data.

In this paper, the effects of DropConnect according to various factors in a deep neural network are analyzed so that the effects of DropConnect can be understood to be

¹ Corresponding Author: Hyun-il LIM, Department of Computer Engineering, Kyungnam University, 7 Kyungnamdaehak-ro, Masanhappo-gu, Changwon, Gyeongsangnam-do 51767, South Korea; E-mail: hilim@kyungnam.ac.kr.

effectively applied in designing neural networks. The experimental results presented in this paper can be considered to apply DropConnect with several factors to reduce overfitting in designing deep neural networks.

2. The Concept of DropConnect in Deep Neural Networks

To create a neural network model, a training process is carried out on training data. At this training process, too dependent customization on the training data results in poor performance in applying in real data. This problem is mainly caused by the dependence of the model on noise data that may be included in the training data. This excessive dependence of the neural network model on the training data is called overfitting. To overcome this problem, it is needed to consider an effective method to reduce overfitting in designing neural networks.

The DropConnect is a way of temporarily excluding several links connected between nodes in the original neural network based on the DropConnect rate at a training phase. The temporary removal of links in training can reduce excessive dependence of the model on the training data. So, an active application of DropConnect can reduce the overfitting in training neural network models. In applying DropConnect to neural network models, several factors should be considered to achieve good performance, such as DropConnect rate and the number of nodes in neural network models.

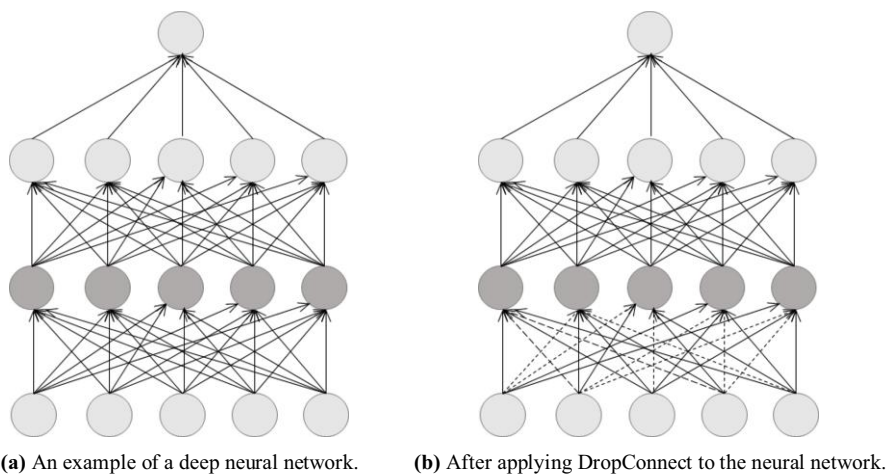


Figure 1. The structure of a deep neural network after applying DropConnect.

Figure 1 shows the structure of applying DropConnect to the hidden layer of a deep neural network. Figure 1(a) shows an example of a general deep neural network model that is fully connected to the nodes on each layer. By applying DropConnect to the neural network, several links are temporarily removed from the training of the network according to the DropConnect rate at the training phase of the model. Therefore, several dashed links in Figure 1(b) are temporarily excluded according to the DropConnect rate, and only connected links participate in the training of the neural network model. Such simplification of the structure of a fully connected neural network model can reduce the overfitting of the network by reducing excessive customization of the model to the

training data. In applying DropConnect to neural network models, the effects of applying DropConnect should be reflected to determine several factors for designing neural networks, such as the DropConnect rate, and the structure of each layer. For example, how many nodes should be placed in the original neural network models or how many links to be excluded in training the neural network models by DropConnect to achieve good performance.

3. The Effect of Applying DropConnect according to the DropConnect Rate

In this experiment, the effect of applying DropConnect according to the DropConnect rate was analyzed in a neural network model. This experiment was performed in a neural network for recognizing the handwritten digits of the MNIST [9]. The MNIST database is popularly used in designing and evaluating machine learning models for image recognition [10-12]. The neural network was designed with one input layer for input data and two hidden layers of 128 and 512 nodes, respectively. The output layer was designed with the softmax function to classify handwritten digits into the corresponding number. The neural network model was implemented in Python with Keras. To analyze the effect of DropConnect in the neural network model, the accuracy and loss values were measured with the varying DropConnect rates between 0% and 80% for the second hidden layer.

Figure 2 shows the experimental results on the effect of applying DropConnect to the neural network model according to DropConnect rates. The DropConnect rate means the ratio of temporarily excluded links at each training epoch, and it should be determined in the design of neural network models. The graphs show how the loss and accuracy change as the training epoch progresses. In the graphs, as the training epoch progressed, training losses decreased and accuracies increased in common. In the training loss and accuracy graphs, the results show that the higher the DropConnect rate, the lower the training loss and the higher the accuracy. The high DropConnect rate means that more connected links will be excluded at each training phase. Therefore, the remaining links can be highly adjusted to the training data because only the weights of links participating in the training are adjusted. Thus, the loss and accuracy improve as the training epoch progresses, and the overall training accuracy of the model is expected to be improved.

Validation loss and accuracy were somewhat different from the training loss and accuracy. The validation loss and accuracy no longer improve at a certain point, even if the training epoch progresses. The high DropConnect rate results in lower loss and higher accuracy, but repeated training does not improve the loss and accuracy. In this experimental result, it can be confirmed that the validation loss and accuracy of the model do not improve after the fifth epoch of training. These results indicate that repeated training no longer affects the improvement of results for the validation data because the model is highly overfitted to the training data. The results show the lowest loss and highest accuracy when the DropConnect rate is around 60%. If the DropConnect rate is determined above the appropriate threshold, it can be confirmed that too many links are excluded to train the neural network model, so the training of the model cannot be performed sufficiently.

In summary, losses show better results when DropConnect is applied to the neural network. It is because the loss reduces the overfitting of the model. However, the excessive DropConnect rate can decrease accuracy because too many links between nodes are excluded in training the model.

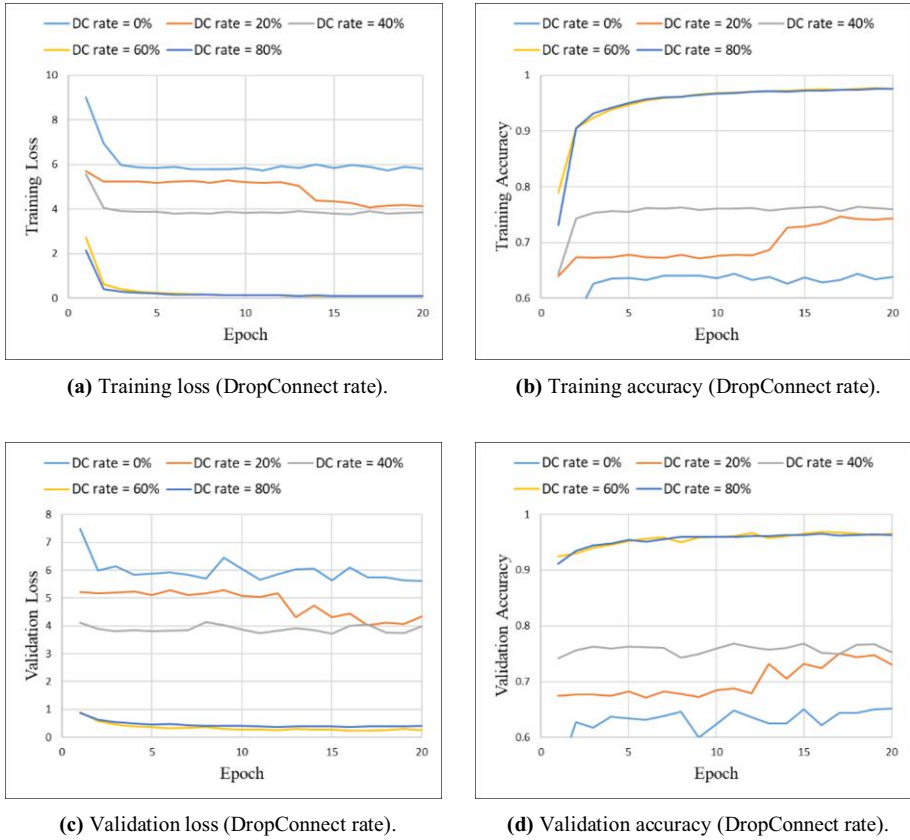


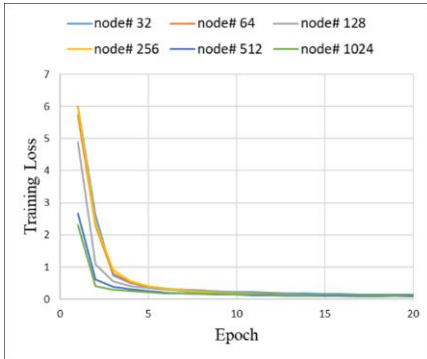
Figure 2. The experiments on the effect of applying DropConnect to a neural network (DropConnect rates).

4. The Effect of the Number of Nodes in Applying DropConnect

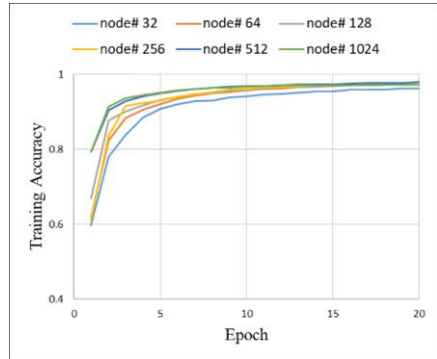
In this experiment, the effect of applying DropConnect was analyzed in a neural network model according to the number of nodes. This experiment was performed with a model for recognizing the handwritten digits of the MNIST as in the previous section. The number of nodes of neural network models is an essential design factor to achieve a good training model. So, it should be considered together with the effect of DropConnect to complete the structure of neural network models. In this experiment, the DropConnect was applied with the rate of 60% and the number of nodes was experimented with from 32 to 1024 at the second hidden layer of the neural network.

Figure 3 shows the experimental results of analyzing the effects of the number of nodes of hidden layers in applying DropConnect. In the early epoch of training, the training loss for the neural network model with many nodes was lower than that with the fewer nodes, so the larger number of nodes could reduce training loss. However, as the training epoch progressed, the training loss remained almost alike, regardless of the number of nodes. The training accuracy showed slightly high for the neural network models with large numbers of nodes.

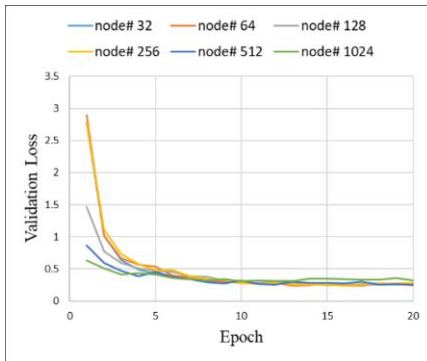
The validation loss for the model with the largest number of nodes (1024) showed the smallest value in the early epochs of training. As training progressed, the validation loss increased rather than the other models with less number of nodes. Unlike the training loss, high validation loss implies the possibility of overfitting in the trained model. Therefore, a complex model with a larger number of nodes than necessary can be more likely to be overfitted. In the early epoch of the training, the training accuracy was different in terms of the number of nodes. However, as the training epochs progressed, validation accuracy maintained similar regardless of the number of nodes. Therefore, it can be seen that the number of nodes does not have a significant impact on models with well-applied DropConnect rates.



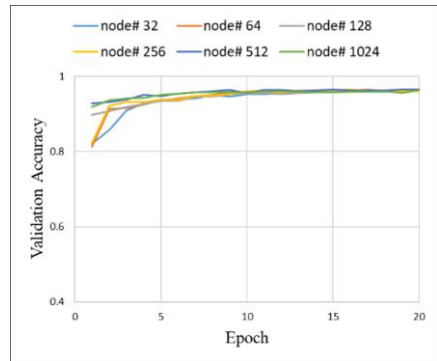
(a) Training loss (number of nodes).



(b) Training accuracy (number of nodes).



(c) Validation loss (number of nodes).



(d) Validation accuracy (number of nodes).

Figure 3. The experiments on the effect of applying DropConnect (the number of nodes).

In the experimental results, in the early epochs of training, the training loss of the model with the largest number of nodes showed the best result. However, as training progressed, the training loss maintained similarly regardless of the number of nodes. In the result of validation loss, the model with many nodes was overfitted, and its validation loss was increased as compared to the other models with less number of nodes. So, it can be seen that too many nodes do not help to train the neural network model.

In applying the DropConnect in designing neural networks, the effect of applying DropConnect should be understood, and the proper application of DropConnect can help

to achieve stable results of neural network models. In future work, we plan to analyze the effect of DropConnect in reducing overfitting for applications in various areas together with the other design factors of neural network models, such as the structure of nodes, links, and layers of neural networks.

5. Conclusion

In this paper, the effect of applying DropConnect according to the DropConnect rates and the number of nodes in a neural network model has been analyzed. In the experimental results, it is confirmed that applying DropConnect is helpful to improve the accuracy of the trained model by temporarily excluding several links of the model to reduce overfitting. To apply the DropConnect to neural network models, the effect should be understood and the DropConnect rate should be applied properly to achieve stable results because the inconsiderate application of DropConnect may reduce the accuracy. The proper application of DropConnect is expected to help to design neural network models for controlling overfitting and improving the accuracy of the models.

Acknowledgements

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (Ministry of Education) (No. NRF-2017R1D1A1B03034769).

References

- [1] Kevin PM. Machine learning: a probabilistic perspective. The MIT Press. 2012.
- [2] Shai SS, Shai BD. Understanding machine learning: from theory to algorithms. Cambridge University Press. 2014.
- [3] Pedro D. A few useful things to know about machine learning. *Communications of the ACM*. 2012; 55(10): 78-87.
- [4] Salman S, Liu XW. Overfitting Mechanism and Avoidance in Deep Neural Networks. 2019.
- [5] Xue Y. An overview of overfitting and its solutions. *J. Physics: Conference Series*. 2019. 1168(022022).
- [6] Paul C, David J. Overfitting explained. In: *Proceedings of the Sixth International Workshop on Artificial Intelligence and Statistics*. 2000.
- [7] Li W, Zeiler M, Zhang SX, Cun YL, Fergus R. Regularization of neural networks using dropconnect. In: *Proceedings of the 30th International Conference on International Conference on Machine Learning*; 2013 June; vol. 28, p. 1058-1066.
- [8] Jensun R, Sascha S, Thomas V. DropConnect for evaluation of classification stability in learning vector quantization. In: *Proceedings of European Symposium on Artificial Neural Networks (ESANN 2019)*; 2019 April 24-26; Belgium; p. 19-24.
- [9] The Mnist Database of handwritten digits, <http://yann.lecun.com/exdb/mnist/>.
- [10] Alejandro B, Yago S, Pedro I. A survey of handwritten character recognition with MNIST and EMNIST. *Applied Sciences*. 2019 August; 9(15):3169.
- [11] Wu M, Zhang Z. Handwritten Digit Classification using the MNIST Data Set. 2010.
- [12] F. Siddique, S. Sakib and M. A. B. Siddique. Recognition of handwritten digit using convolutional neural network in python with tensorflow and comparison of performance for various hidden layers. In: *Proceedings of the 5th International Conference on Advances in Electrical Engineering (ICAEE)*, 2019; Dhaka, Bangladesh; p. 541-546.

The Evolution from Digital Mock-Up to Digital Twin

Sheng DAI^a, Gang ZHAO^a, Yong YU^{a,b,1} and Qiangwei BAO^a

^a*School of Mechanical Engineering and Automation, Beihang University, Beijing 100191, PR China*

^b*Beijing Engineering Technological Research Center of High-Efficient & Green CNC Machining Process, Beijing 100191, PR China*

Abstract. This paper summarizes the development and evolution from digital mock-up (DMU) to digital twin (DT) by clarifying the connotation of DT from prospective of digital product definition (DPD). Firstly, taking Airbus as example, the evolution of DMU is introduced, with the detailed analysis of configured DMU, functional DMU, and industrial DMU. Secondly, based on the literature review of DT, the definition, purpose and several applications of DT concept are clearly expounded. Finally, the augmentation for DT and DPD's relationship are deduced.

Keywords. digital product definition, digital mock-up, digital twin

1. Introduction

Since 1950s, the digital manufacturing concept has emerged with the rapid development of various supporting technologies. Digital manufacturing is an integrated computer system supported by simulation tools, three-dimensional visualization tools, and various collaborative tools, which realizes the DPD and process definition simultaneously. It can be seen that DPD provides the data foundation for digital manufacturing. To achieve full-element and full-process digital manufacturing, the DPD have to be done accurately and precisely.

Based on the motivation above, the digital manufacturing concept has experienced a development from two-dimensional to three-dimensional. The industry has proposed DMU, virtual prototypes (VP) and other concepts that focus on product design, leading to the improvement of the definition and expression of product information.

Recently, the experience of domestic and foreign manufacturing industries has shown that the product model defined by three-dimensional digitalization has matured and its benefits have been verified repeatedly. However, there are still some challenges in digital manufacturing:

- From the point of view of its connotation, DPD mainly focuses on the designing information and always ignores the data of manufacturing, operation, and main-

¹Corresponding Author: Yong Yu, School of Mechanical Engineering & Automation, Beihang University, Beijing, PR China; E-mail: yuyong@buaa.edu.cn.

tenance stages. As a result, there is a gap in terms of representing middle- and end-of-product lifecycle-related semantics [1].

- There is a lack of connection between product definition and process definition. For example, DMU defines the assembly relationship of parts/components, but it cannot accurately describe the assembly process.
- The definition of product information is the “idealized description” given by the CAD (computer-aided design) software. Thus, the simulation based on the “idealized description” has limited guidance for real physical objects.

All in all, to meet the needs of intelligent manufacturing and mirror the actual status of the physical objects, the digital twin concept was introduced by Grieves in 2011 [2]. It is suggested that DT can reflect the dynamic mapping of physical product and its digital counterpart and it is the technical core of cyber-physical system (CPS) [3].

The remainder of this work is organized as follows: Section 2 reviews the literature in the evolution of DMU concept. The connotation of DT concept is presented in Section 3 with detailed discussions. The relationship between DT and DPD is discussed in Section 4. Finally, discussion is provided in Sections 5.

2. The evolution of DMU in Airbus

With the development of the CAD software, the connotation of DMU has undergone great changes. Take Airbus company as an example, DMU has gradually evolved into CDMU (configured DMU), FDMU (functional DMU) and iDMU (industrial DMU), as shown in Figure 1.

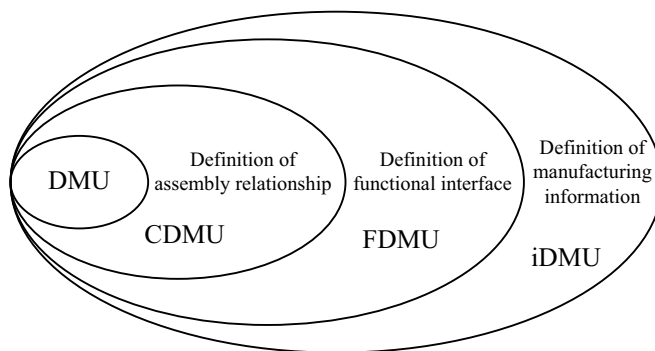


Figure 1. The evolution of DMU in Airbus.

It is believed that a complete DMU should be composed of a three-dimensional model, product structure and product attributes in Airbus. The product structure describes the hierarchical dependence and organizational relationship of the digital model, and the attributes describe the state. This data organization is called configured DMU (configurable DMU) [4], as illustrated in Figure 2.

On the other hand, in order to enrich the functions and behaviors of the DMU in the designing stage, Airbus proposed the functional DMU [5]. FDMU uses the DMU as a carrier to extract components and support requirements and functional analysis, and com-

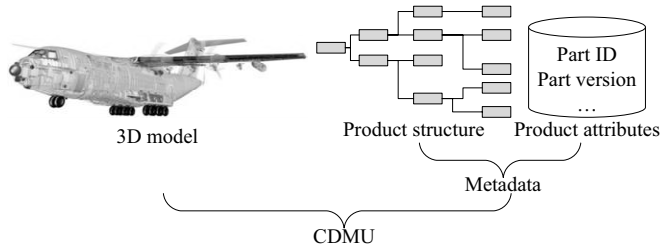


Figure 2. The CDMU concept.

prehensive simulation. The emergence of the FMI standard has profoundly affected the research of FDMU. The FMI standard was first proposed by Daimler AG. The ITI company developed the SimulationX software based on the FMI standard. Dassault Systèmes acquired Dymola software that supports FMI standard.

In order to explore the usage of DMU and optimize the manufacturing process, Airbus implemented the collaborative engineering and iDMU concept in the development of A320neo [6]. In the engineering phase, Airbus reintegrated the functional designing and process manufacturing, and focuses on the definition of products, processes and manufacturing resources from the conceptual stage to mass production. The delivery at this stage were called iDMU. In the manufacturing phase, real data such as the deviation were integrated into iDMU. As shown in Figure 3, when the manufacturing works ends, the output is called as-built iDMU, it is a complete definition of the as-built product.

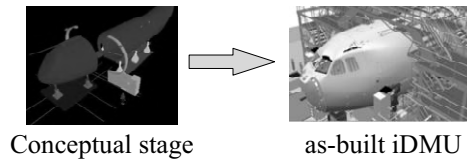


Figure 3. The iDMU concept.

As a result, An iDMU gathers all the product, processes and resources information to model a virtual assembly line and provides a single platform to define and validate the assembly line industrial design [7]. From the perspective of reflecting the physical objects, the as-built iDMU concept is similar with the DT concept. Unfortunately, since it appeared in the research report in 2013, the detailed method of iDMU is still unclear as a commercial secret.

3. From DMU to DT

As mentioned above, although the DMU concept in digital manufacturing has been broadened to express more and more information, it still ignores the data of manufacturing, operation, and maintenance stages. But DT technology is considered to be able to describe physical products and effectively manage the life-cycle information. Thus, the DT concept has gradually attracted the attention of academia and industry.

3.1. The proposal of DT concept

As illustrated in Figure 4, the original assumption of DT was introduced by Dr. Michael Grieves in 2002, which was named “conceptual ideal for PLM (product life management) information mirroring”. Subsequently, “Mirrored Space Model” and “Information Mirroring Model” had been used to redefine the concept. Finally, the “Digital Twin” was introduced by Grieves and his partner in 2011.

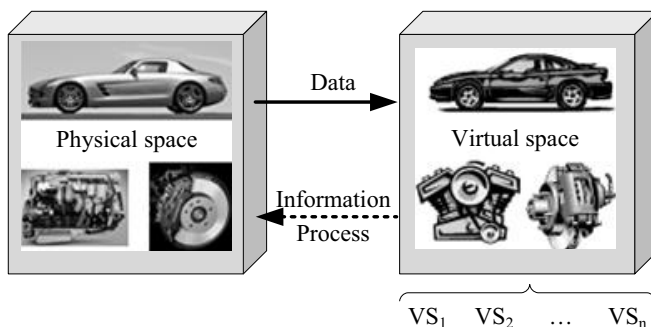


Figure 4. The conceptual ideal for PLM.

Later in 2011, United States Air Force research laboratory adopted the DT concept as a new way to predict the aircraft structural life. In 2012, National Aeronautics and Space Administration defined DT as “an integrated multi-physics, multi-scale, probabilistic simulation of a vehicle or system that uses the best available physical models, sensor updates, fleet history, and so forth, to mirror the life of its flying twin” [8].

3.2. The definition of DT concept

After several years, different definitions and explanation of DT came out. For example, DT was seen as the next generation of simulation [9]. Grieves redefined DT as “a set of virtual information constructs that fully describes a potential or actual physical manufactured product from the micro atomic level to the macro geometrical level and any information that could be obtained from inspecting a physical manufactured product can be obtained from its digital twin” [10].

Recently, DT refers to a virtual representation of manufacturing elements, a living model that continuously updates and changes as the physical counterpart changes in a synchronous manner [11]. The role of DT in digital manufacturing is to allow a physical “thing” to be understood for computers and machines by reflecting its physical status.

3.3. Applications of DT in digital manufacturing

Since the DT concept was proposed, it has been applied in many industrial fields. The possibility of DT-driven product manufacturing is proposed with an example about the drive shaft machining process [12]. A DT-based manufacturing execution system is developed and validated, combining MTConnect data with production data collected from operators [13]. An approach to modeling product DTs, process DTs, and operation DTs using Automation Markup Language is proposed [14]. A DT-based process evaluation

method is developed, including the real-time mapping mechanism, the construction of DT-based machining process evaluation framework, and the process evaluation [15]. A generic architecture for cloud-based manufacturing equipment based on DT and big data analysis is illustrated [16]. A novel DT-driven approach for rapid reconfiguration of automated manufacturing systems was proposed and key enabling techniques, including how to twin cyber and physical system and how to quickly bilevel program the production capacity and functionality of manufacturing systems to adapt rapid changes of products were detailed [17]. Zheng et al. introduced a generic CPS system architecture for DT establishment in smart manufacturing with a novel tri-model-based approach for product-level DT development and then conducted a case study of an open source 3D printer DT establishment [18]. Based on the discrete event system modeling theory, a virtual modeling method at a conceptual level and the implementation mechanisms of the virtual-physical connections in practice for establishing DT shop floor was given [19]. Zhuang et al. proposed a DT-based assembly data management and process traceability approach for complex products and the DT-based Assembly Process Management and Control System (DT-APMCS) was designed to verify the efficiency of the proposed approach [20].

In summary, DT concept can be applied to many fields of digital manufacturing, reflecting real conditions, accumulating actual data, paying attention to previously ignored information, mining hidden new knowledge, and optimizing the entire stage of digital manufacturing.

4. DT and DPD

Normally, DPD is a process that allows the design team to input all their information into the 3D model, thus eliminating the need to create a 2D drawing. For example, DMU involves making the 3D CAD product model the authority dataset, only including the 3D Model and additional information such as parts lists, part coordination documents, etc, within the 3D viewing area of the model.

On the other hand, with the development of PLM technology, DPD focuses on the research on the unified and consistent definition of product information in different stages. PLM is proposed to manage all the product life-cycle data, and it contains all the contents of PDM (product data management). It supports related activities including designing, manufacturing, sales, and maintenance and effectively manages the data generated in these activities. These understandings are not different from the conceptual goals of DT, so this section discusses the relationship between DPD and DT based on the early research.

4.1. DMU is the output of design rationale

The original intention of DPD is to express real physical products. However, taking DMU as an example, the final result becomes “the ideal definition of the product”. Due to actual factors such as product deviation, it is impossible to fabricate an ideal product in the real world. Thus, DMU is the output of continuous iteration process of design activities and reflects human design rationale.

4.2. DT is the representation of physical product

The realism holds that the world does not depend on our representation of it. no matter how we describe the world or the description does not exist, the world still exists independently. On the other hand, “model-dependent realism” asserts that all we can know about “reality” consists of networks of world pictures that explain observations by connecting them with rules to concepts defined in models [21]. The realism also suggests that we cannot know “reality-as-it-is-in-itself”, but only an “approximation” of it provided by the intermediary of models.

Therefore, in any case, DT cannot be completely reflect the physical product in the real world, but a high fidelity approximation. The DT concept is an available form of representation of the physical product, and it reflects the real world in the virtual space.

4.3. DT expands the meaning of DPD

Based on the discussion above, it is suggested that DPD should not be limited to designing phase, but the digital description of all the stages of product life cycle. In fact, the connotation of DPD has been expanding with the development of new technology. When the DMU cannot describe the function and performance of the product, the researchers introduced “FDMU” to describe the functional interface. When the working mode changes from parallel engineering to collaborative engineering, researchers introduced “iDMU” to describe the digital definition of the entire development phase.

In summary, DMU and DT, as two aspects that define ideal products and physical products, can be organically unified, as shown in Figure 5. The merge of DMU and DT reflects the high integration of cyberspace, physical space, and consciousness space. Therefore, DTs will become the “middleware” for in-depth communication between human and machine.

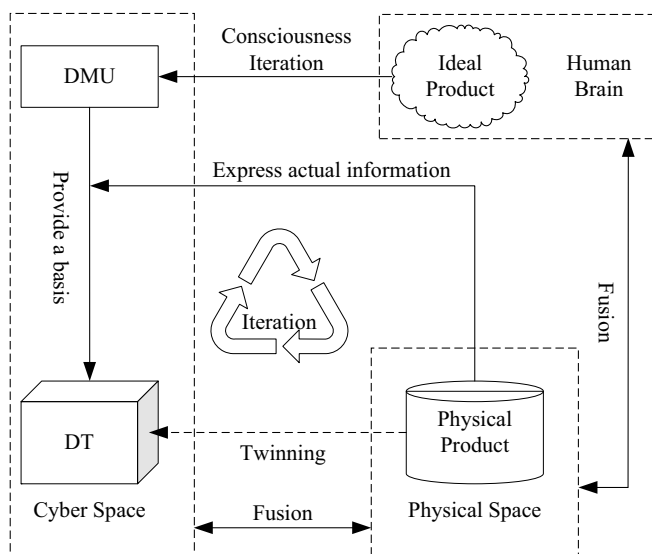


Figure 5. The relationship between DMU & DT.

5. Discussions

By combing the development trend of DPD, this article summarized the development from DMU to DT, expanded the DPD concept, incorporated DT into DPD, and improved the connotation and connection of DPD in the virtual world and physical world.

With the booming growth of information technologies in the manufacturing industry, remarkable efforts about DT-driven applications are going to challenge the fundamentals of manufacturing systems and operations. Many manufacturing companies nowadays believe that the DT concept could improve the utilization of process knowledge for machining planning.

DTs are the foundation of intelligent manufacturing systems and modeling of DT is an effective way to realize CPS. At present, the research on DT is still at early stage. How to build DTs of physical products requires further research. How to realize the product operation and maintenance based on DTs, and improve the perception feedback, precise control and precise execution of the product life cycle process is the main problem.

References

- [1] Kadiri SE, Kiritsis D. Ontologies in the context of product lifecycle management: state of the art literature review. *Int J Prod Res.* 2015 Jun;53(18):5657–5668.
- [2] Grieves M. *Virtually perfect: driving innovative and lean products through product lifecycle management.* Florida: Space Coast Press; 2011.
- [3] Rosen R, von Wichert G, Lo G, Bettenhausen KD. About the importance of autonomy and digital twins for the future of manufacturing. *IFAC-PapersOnLine.* 2015 Aug;48(3):567–572.
- [4] Garbade R, Dolezal WR. DMU@Airbus-evolution of the digital mock-up (DMU) at Airbus to the centre of aircraft development. In: Krause FL, editor. *The future of product development.* Berlin, Heidelberg: Springer Berlin Heidelberg; 2007. p. 3–12.
- [5] Shuichi F, Zoran L, Josip S. FDMU-functional spatial experience beyond DMU? In: *Proceedings of the 20th ISPE International Conference on Concurrent Engineering; 2013 Sep 2-6; Melbourne, Australia.* Amsterdam: IOS Press; 2013. p. 431–440.
- [6] Mas F, Menéndez JL, Oliva M, Ríos J. Collaborative engineering: an Airbus case study. *Procedia Eng.* 2013 Sep;63:336–345.
- [7] Menéndez JL, Mas F, Servan J, Arista R, Ríos J. Implementation of the iDMU for an aerospace industrialization in Airbus. *Procedia Eng.* 2013 Sep;63:327–335.
- [8] Glaessgen E, Stargel D. The digital twin paradigm for future NASA and U.S. air force vehicles. In: *Proceedings of the 53rd AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference; 2012 Apr 23-26; Honolulu, Hawaii.* Reston(VA): AIAA Press; 2012. p. 1818.
- [9] Weyer S, Meyer T, Ohmer M, Gorecky D, Zühlke D. Future modeling and simulation of CPS-based factories: an example from the automotive industry. *IFAC-PapersOnLine.* 2016 Feb;49(31):97–102.
- [10] Grieves M, Vickers J. Digital twin: mitigating unpredictable, undesirable emergent behavior in complex systems. In: Kahlen FJ, Shannon F, Anabela A, editors. *Transdisciplinary perspectives on complex systems.* Cham: Springer International Publishing; 2017. p. 85–113.
- [11] Lu Y, Liu C, Wang KIK, Huang H, Xu X. Digital twin-driven smart manufacturing: connotation, reference model, applications and research issues. *Robot Cim-Int Manuf.* 2020 Feb;61:101837.
- [12] Tao F, Cheng J, Qi Q, Zhang M, Zhang H, Sui F. Digital twin-driven product design, manufacturing and service with big data. *Int J Adv Manuf Tech.* 2017 Mar;94(9-12):3563–3576.
- [13] Coronado PDU, Lynn R, Louhichi W, Parto M, Wescoat E, Kurfess T. Part data integration in the Shop Floor Digital Twin: Mobile and cloud technologies to enable a manufacturing execution system. *J Manuf Syst.* 2018 Jul;48:25–33.
- [14] Bao J, Guo D, Li J, Zhang J. The modelling and operations for the digital twin in the context of manufacturing. *Enterp Inf Syst.* 2018 Oct;13(4):534–556.
- [15] Liu J, Zhou H, Liu X, Tian G, Wu M, Cao L, et al. Dynamic evaluation method of machining process planning based on digital twin. *IEEE Access.* 2019;7:19312–19323.

- [16] Lu Y, Xu X. Cloud-based manufacturing equipment and big data analytics to enable on-demand manufacturing services. *Robot Cim-Int Manuf.* 2019 Jun;57:92–102.
- [17] Leng J, Liu Q, Ye S, Jing J, Wang Y, Zhang C, et al. Digital twin-driven rapid reconfiguration of the automated manufacturing system via an open architecture model. *Robot Cim-Int Manuf.* 2020 Jun;63:101895.
- [18] Zheng P, Sivabalan AS. A generic tri-model-based approach for product-level digital twin development in a smart manufacturing environment. *Robot Cim-Int Manuf.* 2020 Aug;64:101958.
- [19] Jiang H, Qin S, Fu J, Zhang J, Ding G. How to model and implement connections between physical and virtual models for digital twin application. *J Manuf Syst.* 2020 Jun.
- [20] Zhuang C, Gong J, Liu J. Digital twin-based assembly data management and process traceability for complex products. *J Manuf Syst.* 2020 Jun.
- [21] Koonin EV. *The logic of chance: the nature and origin of biological evolution.* New York: Pearson Education; 2011.

Adapted NRC Based Sentiment Analysis in Event Distraction

Zeling WANG¹, Bing WEI and Yibing LI

School of Computer Science and Technology, Hefei Normal University, Hefei, 230601, China

Abstract. The paper was trying to extract entities from related tweets collected from twitter. This project first collected real-time tweets from twitter searching API with related topic-based hashtags during the death of American black man George Floyd. We then used two approaches to identify the polarities or emotions of each tweets and generated over-time sentiment flow chart in detecting entities. We found that some extreme sentiment score was correlated with some key entities over time. And our adapted NRC-lexicon based approach obtained better results. This paper revealed that public's sentiment displayed on tweets was generally consistent with the correlated events previously. It might help researchers in predicting or preventing public events in the future.

Keywords. Sentiment analysis, event distraction, social networking mining, NLP

1. Introduction

With the rapid development of online social networking, the analyzing of online social networking seems more popular for these decades. Many social networking users prefer to share their emotions such as joy, happiness, sadness, or sorrow on various social networking sites.

With more sharing information and opinions from users as well as the easy access API from various social networks, researchers nowadays can easily gather information to do some enterprise development, public sentiment surveys and customer satisfaction surveys. Sentiment analysis as a method using in social network analyzing, can be applied in different situations. For example, customers can judge a product before purchasing by the sentiment analysis of existing comments. Also, companies could use this sentiment research to find and analyze their brand influence among all their customers. In addition, the government or organizations can utilize this information to detect or prevent illegal issues before happening [1].

The experiment in this article is mainly done with Twitter. Users on twitter can create a short message called tweets, which can always represent different opinions or attitudes towards to various topics.

These years, sentiment analysis is widely applied in many areas in order to promote business sales or as a tool to collect customer's attitudes toward for later improvements. We are trying to calculate and analyze the sentiment change towards to one topic over time, to find some events or issues happening.

¹ Corresponding Author, Zeling Wang, School of Computer Science and Technology, Hefei Normal University, Hefei, 230601, China; E-mail: zw6e12@soton.ac.uk.

2. Background

2.1. *Online Social Networking and Analysis*

With the development of World Wide Web, the social networking seems to be the key part in current internet. However, it is not coming directly from the born of World Wide Web. In Web 2.0, the Internet users become more important in contribution of the Internet. Everybody can be the content generator online, which gives the users brilliant ways to communicate and interact with each other [2, 3].

With the analyzing of social networking graph, researchers can both improve the online system right now and try to find the deeper meaning from the large scale of information provided by online users [4].

2.1.1. *Micro Blog and Twitter*

Microblogging is an online phenomenon which gains popularity in recent years. It is a short blog or text (usually less than 200 words) to represent the users' status or lives. The text can be messaged to or interacted with their friends via text message, mobile apps or computer browsers. This type of microblogging service is provided by many companies (e.g. Twitter, Facebook). The most popular microblogging site is Twitter, which has many active users and tweets interacted online everyday [5]. On twitter, users form relationship with following and followed. Following a user means subscribing all the tweets they post. [6].

2.1.2. *Sentiment Analysis*

Sentiment analysis is a textual information mining based on a given texts. From Liu's research, the textual information can be categorized into 2 main types: facts and opinions. Facts are always the objective representation from the text, such as the events, activities or entities. Opinions mean the subjective representation from people, such as their sentiment, their mood, even their thought about some topics or entities. From the popularity of search engine, a lot of work about textual information has been done, such as text classification, text mining and natural language processing. Less focus is on the opinion analyzing; however, opinion is one of the most important element when people are judging or deciding. It may not only be used by individuals, but also the government or organizations [3,7].

Due to the huge changes resulted from internet, the way how people represent their opinion changes a lot. People can easily express their feelings about some topics or events on forums, blogs or social networking sites. Those large amounts of information provide users and researchers a great chance to do the sentiment mining and analyzing work. For instance, if one prefers to buy some products, they can easily check the review or comments from other customers, instead of finding and asking their friends who have bought it before. In additional, with the collection of a large amount of sentiment data, the analyzing work not only give users or researchers great feedback, but also help us to identify or locating specific events about a topic, even it has been proved to complete some prediction work [8,9,10].

3. Related work

With the popularity of micro blog, many different approaches have been applied on the analysis of sentiment mining. However, the sentiment analysis on twitter is different from the sentiment analysis based on a given text, due to the limited length of a given tweet (140 characters), the irregular words in tweets and the creation of internet slang. Then, the data collected should always be cleaned up or processed in advance in order to generate high-quality result [5].

The study conducted by Go et al [11] applying emotion such as “:)” for positive “:(” for negative to classify the data. They first pre-processing the data and then conducted with three standard classifiers: multinomial Naïve Bayes, Maximum Entropy and Support Vector Machine (SVM). The best result is coming from the Maximus Entropy with approximately 83% accuracy with unigram and bigram together.

The approach used above did not identify and manage the neutral sentiment. The experiment did by Pak and Paroubek [12] is trying to improve. They collected the neutral tweets data from different newspaper and magazines, then with the three class NB classifier (positive, negative and neutral), which will be able to detect the neutral words from tweets data. However, this did not achieve an expected result. The accuracy of this experiment only comes to 40%, however, the consideration of neutral would be useful and helpful in future opinion minding work.

From the idea coming from Speriosu et al [13], the sentiment analysis on twitter can be improved with the help of graph theories. They proposed that the combined sentiment label from propagation work can achieve a better accuracy in sentiment analysis [11].

Another research using a two-phased approach has been done by Barbosa and Feng [14]. The two-phases are separated the analyzing process into two steps. First step is to classify the tweets into objective or subjective class. Instead of the n-grams feature, they decided to use two different features “Meta-features” and “Syntax features”.

additionally, the study from Agarwal et al. [15] shows a comparison of two different models based on their features and kernel trees. From their conclusion, they explored 50 different types of features in the experiment and found that both of their two different models from feature based and kernel tree based are better than the unigram baseline. Also, Kouloumpis [16] has tried to use variety of features such as n-gram features, lexicon features, part-of-speech features and micro-blogging features based on different corpus. The result comes from this experiment demonstrate that the micro-blogging features are clearly useful in sentiment classification of online tweets. However, other features are still required research or experiment.

Araque et al. [17] demonstrated that deep learning approach could be involved into sentiment analysis with the classification of different tweets. Apart from those traditional methods, Xiong [18] revealed a multi-layers sentiment-enriched word embedding methods, which involved asymmetric neutral layers can be greatly efficient in learning process in sentiment analysis. Reis [19] presented a random forest-based approach with the ROC curve and F1 score in detecting fake news online.

3.1. Methodology

The aim of this paper is to analyze the sentiment changes of a specific topic with hashtag in a given period (two weeks after George’s death). With the death of black people George Floyd, there is a huge burst of tweets generated on twitter. The progress is shown

in Figure 1. We first tried to collect the related tweets from different hashtags such as Floyd, WalkWithUs, BlackLivesMatter etc. in a consecutive period. After Analyzing the tweets with our two lexicons (The AFINN-111 and adapted NRC), we then tried to locate the specific events or sentiment burst from the sentiment changes overtime. The dataset used in the experiment is collected through Twitter's API with specific hashtag related to the topic.

The first step is to obtain the authentication from twitter. From the API correlated application, we can achieve some parameters in order to acquire the permission to collect data. Before the data is then stored in our csv files, we set up the hash tag (eg. WalkWithUs) to search and mine the tweets we expected with the given token.[20]

Data pre-process is done with some cleaning work since there are many noises affecting our result. Data cleaning [21] involves the removal of unnecessary data, such as Html tags, emojis, numbers etc.

After the preparation of raw data, the next step is to analyze and model the data. We first used the polarity-based corpus [13] to identify the sentiment of each tweet overtime. Then generated the charts to show the sentiment flow overtime to identify the burst or events. In this paper, the sentiment extraction was done with two approaches before comparison. One is using AFINN, which is a tweets-based lexicon with sentiment score ranging from -5 to 5. The other is adapted NRC-lexicon, which included 8 different emotional categories.

When the sentiment detection was done, the next step is to generate the sentiment flow over time in the consecutive time-slots. The generated plots of sentiment fluctuation will be generated for us to identify the events.

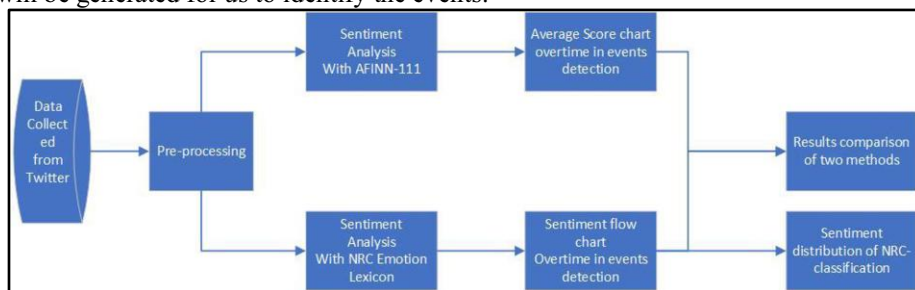


Figure 1. The Flow chart of the project.

4. Implementation and Results

The python was used to create the project. After first collecting data from the twitter API, we obtained the datasets for further sentiment analysis. We chose the Streaming API rather than the REST API, due to the REST one provides short period connection which is difficult to generate the continuous dataset. Then we used lexicon-based approach to analyze the sentiment for each time-slot (every minute). To fairly evaluate our approach, we also included the NRC dictionary to analyze and compare each result, the comparison would help us better insight the efficiency in sentiment output and event extraction [22].

4.1. Lexicon Based Approach for Sentiment Classification

The lexicon provides a series of labeled words or phrases for users to identify and classify the datasets. We conduct the sentiment analysis with the lexicon described above- the AFINN-111 [23]. Due to the flexibility and word-limit of online tweets, it is more suitable to choose a lexicon that constructed based on the tweets before. The AFINN-111 is built by Nielsen with the tweets before, so that it is related to the language and structure of online tweets. AFINN-111 contains 2477 words with labels of sentiment strength from very negative (-5) to very positive (+5). In the experiment process, we apply the lexicon to analyze each tweet over time, and then we can generate the result with sentiment score of each tweet. This score intended to show the polarities of the tweets which can be utilized to monitor the sentiment change or flip over time [25]. The score is obtained from Eq. (1), in which the ps represents the positive score and pn means the number of the tweets, whereas the ns means negative score and nn stands for the number of related tweets.

$$\text{Sentiment scores} = \frac{\sum ps * pn + ns * nn}{n} \quad (1)$$

The key-word based sentiment classification approach is working in this way: in each tweet, the positive words and negative words found are count. Then by checking these keywords score from the lexicon, we can count the final score of each tweet. If the final score is larger than 0, it means that the positive words are stronger than negative words in the tweet, so that the tweet is positive. By contrast, if the final score is less than 0, which means that the negative words are stronger than the positive words in the tweet, the tweet will be classified as negative. However, if the final score equals 0, it means that this is a neutral tweet or a tie.

5. NRC Word-Emotion Association Lexicon

The NRC Emotion Lexicon is a list of English words classified into 8 different emotions with simply 2 polarized sentiment positive and negative. The positive sentiment contains emotions of anticipation, trust, surprise and joy however the negative sentiment includes anger, fear, sadness and disgust.

The previous lexicon would count the words in the range of different sentiment scores, but due to the limitation of the size, the sentiment extraction would be restrained. However, the NRC Emotion Lexicon [24] is much more widely classified. With 8 different emotions, words and phrases can be better categorize. But it might receive better efficiency by add the factors in determine the polarity of each emotion.

From the related work done by Bandhakavi and Wiratunga [25], we realized that each emotion appeared in varied rates but represent different intensity. We then try to put coefficient for different emotion to receive better processing. We treat anger and disgust as extreme negative emotion by timing the count with 1.2. but sadness and fear as less-negative emotion, then times 0.8. For the positive sentiments, we count joy and anticipation as extreme positive with timing 1.2, but surprise and joy as less-positive with the coefficient 0.8.

$$\text{Sentiment score} = \frac{\sum_{n=1}^8 \text{coefficient} * \text{count} * \text{score}}{\# \text{tweets number}} \tag{2}$$

6. Sentiment Analysis based on the events related to the death of George Floyd

Tweets can always provide sentiment information while in the period or around a specific event. In this paper, we applied our discussed sentiment approaches to analyze the collected dataset and then tried to identify what these polarity information means and what they are associated with different entities [26]. On May 25th 2020, a 46-year-old black man, George Floyd died after officer kneels on his neck in custody. This event then trigger huge discussion on Twitter. We began to collect the tweets associated with this event during the next 2 weeks, and successfully identified the association between the sentiment changes correlated with the follow up events. The sentiment classification in 2 methods are showed in the figure 2. From the NRC-lexicon based approach [24], we identified 3 abnormal peaks in May 29th, June 1st and June 4th respectively. At the first week after George’s death, the general emotional polarity of public is negative, but after June 1st, especially the polarity began to flip to positive on 2nd June mainly because people began to protest peacefully in downtown Houston and government began to take the responsibilities of some issues. The 3 peak polarities are corelated with the events followed:

- May 29th: Officers arrested, charged in Floyd’s death
- June 1st: Family autopsy revealed Floyd was asphyxiated by sustained pressure
- June 4th: Minneapolis memorial service for George Floyd honored his life, called Americans to action

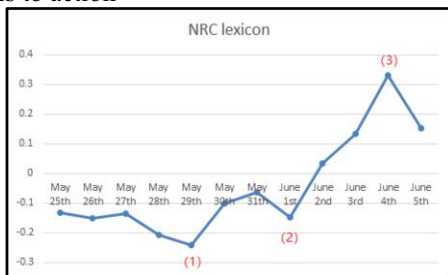


Figure 2. Sentiment score distribution from NRC approach.

Also, with the specific classification of 8 different emotions (Figure 3), we can then easily identify the tweets score of different categories, and then help us to understand the general public emotions towards the event.

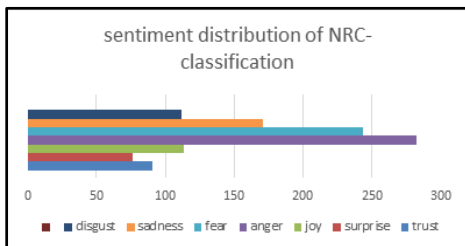


Figure 3. Categorized tweets in NRC approach.

However, for the traditional lexicon AFINN-111 (Figure 4), we still can identify the general trend that from negative gradually flip to positive. But the peak sentiment polarities are not as obvious as the NRC approach. It might due to the limitation amount of sentimental words included or because the words are only classified into 3 categories compared with 8 categories in NRC approach.

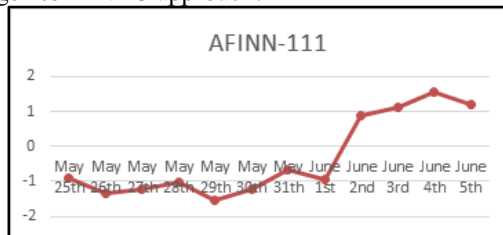


Figure 4. Sentiment score distribution from AFINN-111 approach.

7. Future Work

Future work will mainly try to deal with scalable amount of data for a specific event instead of the hashtag key words in twitter. While the dimension of data hugely increased to an event, the traditional lexicon-based analysis may be possible to combine with neural network. Apart from that, the coefficient we applied to modify the strength of positive and negative can be adjusted due to the general trends towards the event and it is important to find out the reliable coefficient in handling different issues.

8. Conclusion

We try to use the sentiment polarities in measuring the events happening from online social networking. From the approaches people talk about the events either positively or negatively, we figure out some correlations between the events happened at the same time. Analyzing tweets allowed us to extract related events and sentiment trends around the hot issues such as the public events followed with the death of George Floyd. In this paper, 2 lexicon-based approaches are used to classify the sentiment of our corpus. Moreover, we try to associate the polarities of tweets variation over time with the specific events happened incidentally. This approach of sentiment analysis helps us to extract events based on polarized-opinioned tweets. This may help the government or public organization in preventing the terrorism or avoiding the illegal events or protests in advance.

9. Acknowledgement

This research work is supported and funded by The Youth projects of Natural Science Foundation in Anhui Province(1908085QF287), and is also supported and funded by the Key projects of the National Natural Science Foundation of universities in Anhui Province (NO. KJ2020A0112, NO. KJ2018A0499)

References

- [1] Salloum SA, Al-Emran M, Monem AA, Shaalan K. A survey of text mining in social media: facebook and twitter perspectives. *Adv. Sci. Technol. Eng. Syst. J.* 2017 Jan; 2(1):127-33.
- [2] O'reilly T. What is Web 2.0: Design patterns and business models for the next generation of software. *Communications & strategies.* 2007 Mar 1(1):17.
- [3] Mislove A, Marcon M, Gummadi KP, Druschel P, Bhattacharjee B. Measurement and analysis of online social networks. In *Proceedings of the 7th ACM SIGCOMM conference on Internet measurement 2007 Oct 24* (pp. 29-42).
- [4] Peng S, Wang G, Xie D. Social influence analysis in social networking big data: Opportunities and challenges. *IEEE network.* 2016 Nov 3; 31(1):11-7.
- [5] Java A, Song X, Finin T, Tseng B. Why we twitter: understanding microblogging. In *Proceedings of the 9th WebKDD and 1st SNA-KDD 2007 Workshop on Web Mining and Social Network Analysis 2018* (pp. 56-65).
- [6] Kwak H, Lee C, Park H, Moon S. What is Twitter, a social network or a news media? In *Proceedings of the 19th international conference on World wide web 2010 Apr 26* (pp. 591-600).
- [7] Liu B. Sentiment analysis and subjectivity. *Handbook of natural language processing.* 2010 Feb; 2(2010): 627-66.
- [8] Bakshi RK, Kaur N, Kaur R, Kaur G. Opinion mining and sentiment analysis. In *2016 3rd International Conference on Computing for Sustainable Global Development (INDIACom) 2016 Mar 16* (pp. 452-455). IEEE.
- [9] Li N, Wu DD. Using text mining and sentiment analysis for online forums hotspot detection and forecast. *Decision support systems.* 2010 Jan 1;48(2):354-68.
- [10] Lee D, Jeong OR, Lee SG. Opinion mining of customer feedback data on the web. In *Proceedings of the 2nd international conference on Ubiquitous information management and communication 2008 Jan 31* (pp. 230-235).
- [11] Go A, Bhayani R, Huang L. Twitter sentiment classification using distant supervision. *CS224N project report, Stanford.* 2009 Dec; 1(12):2009.
- [12] Pak A, Paroubek P. Twitter as a corpus for sentiment analysis and opinion mining. In *LREc 2010 May 17* 10(2010): 1320-1326.
- [13] Speriosu M, Sudan N, Upadhyay S, Baldridge J. Twitter polarity classification with label propagation over lexical links and the follower graph. In *Proceedings of the First workshop on Unsupervised Learning in NLP 2011 Jul* (pp. 53-63).
- [14] Barbosa L, Feng J. Robust sentiment detection on twitter from biased and noisy data. In *Coling 2010: Posters 2010 Aug* (pp. 36-44).
- [15] Agarwal A, Xie B, Vovsha I, Rambow O, Passonneau RJ. Sentiment analysis of twitter data. In *Proceedings of the workshop on language in social media (LSM 2011) 2011 Jun* (pp. 30-38).
- [16] Kouloumpis E, Wilson T, Moore J. Twitter sentiment analysis: The good the bad and the omg! In *Fifth International AAAI conference on weblogs and social media 2011 Jul 5*.
- [17] Symeonidis S, Efrosynidis D, Arampatzis A. A comparative evaluation of pre-processing techniques and their interactions for twitter sentiment analysis. *Expert Systems with Applications.* 2018 Nov 15; 110:298-310.
- [18] Xiong S, Lv H, Zhao W, Ji D. Towards Twitter sentiment classification by multi-level sentiment-enriched word embeddings. *Neurocomputing.* 2018 Jan 31; 275: 2459-66.
- [19] Reis JC, Correia A, Murai F, Veloso A, Benevenuto F. Supervised learning for fake news detection. *IEEE Intelligent Systems.* 2019 May 8;34(2):76-81.
- [20] Trupthi M, Pabboju S, Narasimha G. Sentiment analysis on twitter using streaming API. In *2017 IEEE 7th International Advance Computing Conference (IACC) 2017 Jan 5* (pp. 915-919). IEEE.
- [21] Saif H, Fernández M, He Y, Alani H. On stopwords, filtering and data sparsity for sentiment analysis of twitter.
- [22] Nausheen F, Begum SH. Sentiment analysis to predict election results using Python. In *2018 2nd international conference on inventive systems and control (ICISC) 2018 Jan 19* (pp. 1259-1262). IEEE.
- [23] Miura Y, Sakaki S, Hattori K, Ohkuma T, Team X. A sentiment analyzer with enhanced lexicon mapping and weighting scheme for unbalanced data. In *Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014) 2014 Aug* (pp. 628-632).
- [24] Mohammad SM, Turney PD. NRC emotion lexicon. National Research Council, Canada. 2013 Nov 15;2.
- [25] Bandhakavi A, Wiratunga N, Massie S, Padmanabhan D. Lexicon generation for emotion detection from text. *IEEE intelligent systems.* 2017 Feb 13;32(1):102-8.
- [26] Yu Y, Wang X. World Cup 2014 in the Twitter World: A big data analysis of sentiments in US sports fans' tweets. *Computers in Human Behavior.* 2015 Jul 1; 48:392-400.

Intelligence Accounting Information Fusion System: Theory, Model and Framework

Xinsheng DUAN¹

*School of Accounting, Capital University of Economics and Business, Beijing 100070
P. R. China*

Abstract. It is defined in this paper that the intelligence accounting information fusion system is an intelligence accounting information system based on accounting information fusion (briefly AIF). This is a new accounting information system that goes beyond the traditional accounting information system, based on AIF and has the function of intelligent decision-making support. Intelligence accounting information fusion system must have four conditions or functions. First, the system should have basic accounting information generation functions. Second, the system should have the ability to obtain other relevant information from various channels, called information acquisition intelligence. Third, the information obtained should be expressed and stored intelligently, called information expression and information storage intelligence. Fourth, the system should have intelligent decision support function. This paper will research firstly on AIF, establish theory for AIF to support the stakeholders to make decision using both accounting information and non-accounting information. Then, the paper will study the intelligence accounting information system based on AIF. The theory and the realization of intelligence accounting information fusion system will be studied based on methods and technologies in the field of Artificial Intelligence and Expert Systems.

Keywords. Big data analysis, accounting information fusion, decision support, artificial intelligence, intelligence accounting information fusion system

1. Introduction

In recent years, the development of artificial intelligence has made unprecedented achievements, and the application of artificial intelligence in accounting, auditing and financial management has been carried out in a large number of studies. The big four accounting firms have launched their own financial robots. The use of electronic invoices, the establishment of financial sharing centers and automatic accounting are all important applications of artificial intelligence in accounting and auditing. But these financial intelligences are just automation in a particular area of financial accounting, not real artificial intelligence. What is artificial intelligence in the true sense? In this paper, decision-making is the true intelligence of human beings, especially, the ability to apply unstructured information to make decisions is the highest embodiment of human intelligence. This intelligence is currently difficult for machines to achieve. This paper

¹ Corresponding Author: Xinsheng DUAN, School of Accounting, Capital University of Economics and Business, Beijing 100070 P. R. China; E-mail: xsduan@163.com.

will study the accounting information system with decision support function, which should be the higher level of intelligence accounting information system.

This paper first discusses the concept and theory of AIF, and discusses theory and model of using accounting information and non-accounting information to make decisions jointly. Then, from the perspective of decision support, the intelligence accounting information system is studied, and the theoretical system and design concept of the intelligence accounting information system based on AIF are put forward. Third, the computer implementation of intelligence accounting information fusion system is discussed by using the concept and technologies in the field of Artificial Intelligence and Expert Systems.

2. Research Background

This paper will apply artificial intelligence, big data analysis, cloud computing technology, information fusion, XBRL financial reporting and Dempster-Shafer evidence theory to study intelligence accounting information system from the perspective of decision support, and put forward the research ideas and design concepts of intelligence accounting information system based on AIF.

Big data and cloud computing are the hot topics in the information technology industry in recent years. In May 2011, the McKinsey Global Institute released a report entitled "Big Data: The Next Frontier of Innovation, Competition, and Productivity"[1], formally introducing the concept of big data. Since then, big data technology has been widely used in all of the world, creating tremendous economic value. This paper will apply big data analysis and cloud computing technology to study the theory and practice of AIF and intelligence accounting information system, which is an attempt to apply big data in the field of financial accounting.

The research of this paper will start with the concept of information fusion, put forward the concept, theory and method of accounting information fusion, and then use it for the research of intelligence accounting information system. Information fusion, also known as data fusion, multi-sensor data fusion, or multi-source information fusion, originated in military applications. In 1991, JDL: Joint Directors of Laboratories in USA defined information fusion as a multi-layered, multi-faceted process to detect, correlate, estimate, and combine of multi-source data for accurately estimate the status, risks and threats of the battlefield.

Since the 1990s, the rapid development of computer technology has greatly promoted the research of information fusion theory, and the application field of information fusion has been rapidly extended from military to other fields. Such as: robot and intelligent instrument systems, intelligent manufacturing systems, aerospace applications, image analysis and understanding, inertial navigation, pattern recognition and so on.

Unfortunately, however, there is little research on the application of information fusion in enterprise management. Our research makes an effort to use information fusion in the area of management, especially accounting, financial management, and auditing.

The important problem in the theoretical research of information fusion is the expression and combination of information. Among the many combination methods, Dempster-Shafer evidence theory [2][3] provides a natural and powerful method for the expression and synthesis of uncertain information, so Dempster-Shafer evidence theory

has become a widely used and more researched field in information fusion theory [4-8]. Therefore, this paper will also take Dempster-Shafer evidence theory and its expansions as the main method of AIF, and establish the theoretical framework of AIF based on Dempster-Shafer evidence theory.

The author of the paper has also done some work in the research of Dempster-Shafer evidence theory. In 1993, he published a book, called "Evidence Theory, Decision Making, and Artificial Intelligence" [9]. Once published, the book has been well received by researchers and users of evidence theory and has been cited many times by various magazines and many researchers. In 1994, the project "Research on the belief function model of decision-making" (approval number: 79300020) is applied successfully from the National Natural Science Foundation of China, and after several years of research, another book "Evidence Decision-making" was published in 1996 [10].

In recent years, I have applied the theory of evidence to the solution of the problem of comprehensive evaluation, and put forward a model of "Evidential Synthetic Evaluation Model" [11].

In the paper "Using evidential synthetic model to evaluate the performance of RMB PE funds", the application of evidential synthetic model in private equity fund performance evaluation is discussed [12]. In the paper "Belief Function Model of Private Equity Fund Performance Measurement", the fund performance measurement is discussed using belief function model [13].

Regarding the concept and theory of AIF, Duan Xinsheng established the theoretical framework of AIF in two papers in 2014 and 2016, "Research on the Limitations of Accounting Information and the Integration of Accounting Information" [14] and "Accounting Information Fusion for Decision Making" [15]. In these two papers, the theory and thought of information fusion are applied to the comprehensive processing and use of accounting information, put forward the concept of AIF and established the theoretical framework of AIF based on Dempster-Shafer evidence theory and its expansions.

3. Intelligence Accounting Information System

The main function of accounting information system is to provide accounting information, and mainly financial accounting information, which does not include the use of accounting information and decision support. This paper holds that the function of accounting information system should be expanded to provide not only information, but also decision support. In such a system, the system can automatically obtain the required information, including not only accounting information but also non-accounting information, according to different decision objectives, and can make the required decisions on the basis of this information. Among them, accounting information, including financial accounting information and management accounting information, can be produced by the traditional accounting information system and provided to decision makers or the decision engine of the system in a reasonable manner. Non-accounting information, including text, image, and video information, can be automatically crawled from the network or other sources using big data analysis. The decision engine of the system can make decision suggestions based on the unified expression, unified storage and unified use of the obtained accounting information and non-accounting information. If new information is found, including accounting and non-

accounting information, the system can automatically add new information to correct the original decision. Information can be added and decisions can be corrected one by one. Of course, the system can automatically identify the new information is financial accounting information, management accounting information, or non-financial information, so as to use different decision-making models to deal with. Such an accounting information system is called an intelligence accounting information system based on decision support perspective.

The idea of intelligence accounting information system can be realized by using the theory of AIF established in the earlier paper and here after.

4. Intelligence Accounting Information Fusion System: A Framework

This paper holds that the intelligence accounting information system based on AIF, referred to as the intelligence accounting information fusion system, is an accounting information system based on AIF and has ability providing decision-making support intelligently. To this end, intelligence accounting information fusion system must have three conditions or functions.

First, the system should have basic accounting functions, should be in accordance with the current accounting standards to complete the usual financial accounting, generating the usual financial reports and financial statements. It also has the ability to complete the management accounting and generate management accounting information.

Second, in addition to its own ability to generate financial accounting and management accounting information, it should also have the ability to obtain other relevant information from different sources, called information acquisition intelligence. It cannot refuse the entry and extraction of any information, can automatically identify the obtained information, financial accounting information or management accounting information, accounting information or non-accounting information, numerical information or text information, etc. The obtained information should be intelligently expressed and stored, called information expression and information storage intelligence. Therefore, intelligence accounting information fusion system is able to express and store any information, so there must be a strong database, knowledge base and rule base to support.

Third, intelligence accounting information fusion system should have intelligent decision-making mechanism. The information stored in different sources should be intelligently and uniformly processed. According to different decision-making purposes, the information stored in different sources can be identified and the relevant information can be extracted. The extracted information can be combined to give a suggestion that are useful for decision-making.

The realization of intelligence accounting information fusion system can be carried out in a layered way.

- It is the first step to discuss the information fusion of financial accounting and management accounting, and establish a unified financial accounting and management accounting processing engine. The theory and method of information extraction and data processing of XBRL financial report can be studied at the beginning. Especially, the XBRL representing and storing of management accounting information should be studied in depth. The technical background and theoretical support of the fusion algorithm of XBRL financial

accounting information and management accounting information should be explored in this stage.

- Second step discusses the fusion mechanism of accounting information and non-accounting information, and establishes a data processing engine that can handle accounting information (including financial accounting information and management accounting information) and non-accounting information based on the theory of AIF. The theories and methods of using big data and cloud computing technology to retrieve, extract, store and process information should be explored. In particular, the crawling and processing of information from internet, such as text information, image information and video information are studied in depth to lay the foundation for the realization of the subsequent intelligent system.
- Next discusses the technical realization of the integration of accounting information (including financial accounting information and management accounting information) and non-accounting information. To study the theoretical and practical problems of database, knowledge base and rule base. To study the identification and evaluation of the usefulness of information and establish a mechanism for real-time clean-up and updating of outdated information. The fusion algorithm and the establishment of self-learning algorithm are studied to realize the goal of automatic update, self-improvement and automatic learning of the system.

5. Conclusions and Future Perspectives

This paper is to study the acquisition, storage, integration and utilization of accounting information and non-accounting information in big data environment, to establish a theoretical model for the integration of financial accounting information with management accounting information, accounting information and non-accounting information, and to explore the theory and realization of intelligence accounting information system based on AIF.

Traditional accounting information systems either provide information to shareholders out of a duty of responsibility, or provide information for stakeholders to make decisions. In short, the goal of the traditional accounting information system is to provide information, the differences for two different views only lay on the place that the view of duty responsibility is to provide information to shareholders, while the view of decision support is to provide information to stakeholders. The decision support here is that the accounting system provides information to support the decision-making of stakeholders, not the accounting information system itself can make decisions or the accounting information system itself has the function of decision support. The intelligence accounting information fusion system discussed in this paper does not provide information for decision support alone, but the system itself has the function of decision support and even making decisions like human beings. Therefore, the decision support in this paper and the decision support in the traditional accounting information system are two different meanings.

The decision-making information generation mechanism in the intelligence accounting information fusion system discussed in this paper is fundamentally different from the traditional accounting theory. According to the three-tiered model established in paper "Accounting Information Fusion for Decision Making" (Xinsheng Duan, 2016),

the traditional accounting theory can only provide the first level of information, while the information provided by the intelligence accounting information fusion system is the third level of information. Therefore, the intelligence accounting information fusion system may be the innovation and subversion of the traditional accounting information system, which will also have a great impact on today's accounting theory, and may even lead to a major revolution in accounting theory.

The one of the roles or important functions of intelligent accounting information fusion system is to make decisions using accounting information and non-accounting information. We know that making decisions is a human-specific ability and an advanced intelligence. Therefore, compared with the traditional accounting information system and even the intelligent system mentioned in many literatures, such as automatic accounting system, it has a higher level of intelligence. With the improvement and development of intelligent accounting information fusion system, the level of intelligence will be higher and higher, and it may not be impossible to reach or approach the level of human intelligence.

References

- [1] MacKinsey Global Institute. Big data: The next frontier for innovation, competition, and productivity. May 2011.
- [2] Dempster AP. Upper and lower probabilities induced by a multivalued mapping. *Annals of Mathematical Statistics*. 1967, 38: 325-339.
- [3] Shafer G. A mathematical theory of evidence. Princeton University Press, Princeton, 1976.
- [4] Dezert J. Foundations for a new theory of plausible and paradoxical reasoning. *Inform. Security*. 2002, NSFC 2018, 9: 90-95.
- [5] Dubis D, Prade H. On several representation of an uncertain body of evidence. In *Fuzzy Information and Decision Processes*, North-Holland, 1982.
- [6] Dubis D, Prade H. Representation and combination of uncertainty with belief functions and possibility measures. *Comput. Intellig.* 1988, 4: 244-264.
- [7] Murphy C K. Combination belief functions when evidence conflicts. *Decision Support Systems*. Elsevier Publisher, 2000, Vol. 29, 1-9.
- [8] Smets P. The combination of evidence in the transferable belief model. *IEEE Trans. On Pattern Analysis and Machine Intelligence*. 1990, 12(5): 447-458.
- [9] Duan XS. Evidence Theory and Its Applications to Decision Making and Artificial Intelligence. People University of China Press, Beijing, 1993 (in Chinese).
- [10] Duan XS. Evidence Decision. Economic Science Press, Beijing, 1996 (in Chinese).
- [11] Duan XS. Evidential synthetic evaluation model. *Proceedings of 2nd International Conference on Artificial Intelligence, Management Science and Electronic Commerce*. 2011, 4(I): 3030-3033.
- [12] Duan XS. Using evidential synthetic evaluation model to evaluate the performance of RMB PE funds. *Int J Wireless Mobile Computing*. 2011, 5(1): 60-69.
- [13] Duan XS. Belief function model of private equity fund performance measurement. *Proceedings of 2011 International Conference on E-Business and E-Government*. 2011, 2114-2117.
- [14] Duan XS. Research on the Limitations of Accounting Information and the Integration of Accounting Information. *Journal of Accountant Friendship*. 2014, 18: 6-8 (in Chinese).
- [15] Duan XS. Accounting information fusion for decision making. Diversity of Managerial Perspectives from Inside China. Check Teck Foo, Springer, pp 67-81, 2016.

Modeling Multi-Order Adaptive Processes by Self-Modeling Networks

Jan TREUR¹

*Social AI Group, Vrije Universiteit Amsterdam, De Boelelaan 1111, 1081HV
Amsterdam, the Netherlands*

Abstract. A self-modeling network for some base network is a network extension that represents part of the base network structure by a self-model in terms of added network nodes and connections for them. By iterating this construction, multi-order network adaptation is easily obtained. A dedicated software environment for self-modeling networks that has been developed supports the modeling and simulation processes. This will be illustrated for a number of adaptation principles from a number of application domains.

Keywords. Adaptive network, self-modeling network, multi-order adaptive

1. Introduction

A self-modeling network is a network that represents part of its own network structure by a self-model in terms of dedicated network nodes and connections for them. A network structure can be described by network characteristics for connectivity for connections between nodes, aggregation for combining multiple incoming impacts on a node, and timing for the speed of node state dynamics; e.g., [1, 2, 3]. Any base network can be extended to a self-modeling network for it, by adding a self-model for part of the base network's structure. In this case, the added self-model consists of a number of added nodes representing specific characteristics of the base network structure, such as connection weights and excitability thresholds, plus connections for these added self-model nodes. For the approach considered here, in general nodes in a network are assumed to have activation levels that can change over time due to impact from other nodes from which they have incoming connections. If in particular the nodes from a self-model representing some of the network characteristics of a base network are dynamic, these base network characteristics become adaptive, thus an adaptive base network is obtained, in the sense that adaptation of the base network is modeled by the dynamics *within* the self-modeling network extending the base network.

Moreover, multi-order network adaptation can be obtained by iterating this self-modeling construction. If multi-order self-models are included in a self-modeling network, any included self-model (of some order) can have its own (next-order) self-model within the overall network where the latter self-model represents some of the network characteristics of the former self-model. For example, this allows to control the dynamics of self-models, so that self-controlled adaptive networks are obtained.

¹ Corresponding Author: Jan Treur, Social AI Group, Vrije Universiteit Amsterdam, De Boelelaan 1111, 1081HV Amsterdam, the Netherlands; Email: j.treur@vu.nl.

A dedicated software environment for self-modeling networks that has been developed supports these modeling and simulation processes; see [3], Ch. 9. In this paper, for a number of adaptation principles from different application domains, it will be illustrated how they can be modeled by proper pre-specified self-models that can be used as building blocks to extend any base network to make it adaptive.

In the paper, first in Section 2 the modeling approach from [3] based on self-modeling networks is briefly described. In Section 3 nine different adaptation principles from the Cognitive Neuroscience and Social Science literature are described. Next, in Section 4, for the adaptation principles described in Section 3 it is shown in more detail how they can be modeled by self-models. Section 5 is a discussion.

2. Networks Using Self-Models: Self-Modeling Networks

In this section, the network-oriented modelling approach used from [3] is introduced. Following [3, 4], a temporal-causal network model is characterized by (here X and Y denote nodes of the network, also called *states*):

- *Connectivity characteristics* Connections from a state X to a state Y and their weights $\omega_{X,Y}$
- *Aggregation characteristics* For any state Y , some combination function $c_Y(\cdot)$ (usually with some parameters) defines the aggregation that is applied to the impacts $\omega_{X,Y}X(t)$ on Y from its incoming connections from states X
- *Timing characteristics* Each state Y has a speed factor η_Y defining how fast it changes for given impact.

The following difference (or differential) equations that are used for simulation purposes and also for analysis of temporal-causal networks incorporate these network characteristics $\omega_{X,Y}$, $c_Y(\cdot)$, η_Y in a standard numerical format:

$$Y(t + \Delta t) = Y(t) + \eta_Y [c_Y(\omega_{X_1,Y}X_1(t) \cdots \omega_{X_k,Y}X_k(t)) - Y(t)] \Delta t \tag{1}$$

for any state Y and where X_1 to X_k are the states from which Y gets its incoming connections. Here the overall combination function $c_Y(\cdot)$ for state Y is the weighted average of available basic combination functions $c_j(\cdot)$ by specified weights $\gamma_{j,Y}$ (and parameters $\pi_{1,j,Y}$, $\pi_{2,j,Y}$ of $c_j(\cdot)$) for Y :

$$c_Y(V_1, \dots, V_k) = \frac{\gamma_{1,Y} c_1(V_1, \dots, V_k) + \dots + \gamma_{m,Y} c_m(V_1, \dots, V_k)}{\gamma_{1,Y} + \dots + \gamma_{m,Y}} \tag{2}$$

Such Eq. (1) and (2) are hidden in the dedicated software environment; see [3], Ch 9. Within this software environment, currently around 40 useful basic combination functions are included in a combination function library; see Table 1 for some of them. The above concepts enable to design network models and their dynamics in a declarative manner, based on mathematically defined functions and relations.

Table 1. Examples of basic combination functions from the library.

	Notation	Formula	Parameters
Euclidean	$\mathbf{eucl}_{n,\lambda}(V_1, \dots, V_k)$	$\sqrt[n]{\frac{V_1^n + \dots + V_k^n}{\lambda}}$	Order $n > 0$ Scaling factor $\lambda > 0$
Advanced logistic sum	$\mathbf{alogistic}_{\sigma,\tau}(V_1, \dots, V_k)$	$\left[\frac{1}{1 + e^{-\sigma(V_1 + \dots + V_k - \tau)}} - \frac{1}{1 + e^{\sigma\tau}} \right] (1 + e^{-\sigma\tau})$	Steepness $\sigma > 0$ Excitability threshold τ
Scaled maximum	$\mathbf{smax}_{\lambda}(V_1, \dots, V_k)$	$\mathbf{max}(V_1, \dots, V_k) / \lambda$	Scaling factor $\lambda > 0$
Scaled minimum	$\mathbf{smin}_{\lambda}(V_1, \dots, V_k)$	$\mathbf{min}(V_1, \dots, V_k) / \lambda$	Scaling factor $\lambda > 0$

Note that there is a crucial distinction for network models between network characteristics and network states. Network *states* have values (their activation levels) and are explicit representations that may be accessible for network states by connections to and from them and can be handled or manipulated in that way. They can be considered to provide an informational view on the network; usually the states are assumed to have a certain informational content. In contrast, network *characteristics* (such as connection weights and excitability thresholds) have values (their strengths) and determine (e.g., cognitive) processes and behavior in an implicit, automatic manner. They can be considered to provide an embodiment view on the network. In principle, these characteristics by themselves are not directly accessible nor observable for network states; in principle you can make connections between states but you cannot make connections between network characteristics or between states and characteristics.

As indicated above, ‘network characteristics’ and ‘network states’ are two distinct concepts for a network. Self-modeling is a way to relate these distinct concepts to each other in an interesting and useful way. A *self-model* is making the network characteristics (such as connection weights and excitability thresholds) explicit in the form of adding states (called *self-model states*) for these characteristics and also connections for these additional states. Thus, the network gets an internal self-model of part of its network structure: it explicitly represents information about its own network structure. In this way, by iteration different self-modeling levels can be created where network characteristics from one level relate to network states at a next level. Thus, an arbitrary number of self-modeling levels can be modeled, covering *second-order* or *higher-order* effects.

More specifically, adding a self-model for a temporal-causal base network is done in the way that for some of the states Y of the base network and some of the network structure characteristics for connectivity, aggregation and timing (i.e., some from $\omega_{X,Y}$, $\gamma_{j,Y}$, $\pi_{i,j,Y}$, η_Y), additional network states $\mathbf{W}_{X,Y}$, $\mathbf{C}_{j,Y}$, $\mathbf{P}_{i,j,Y}$, \mathbf{H}_Y (*self-model states* or *reification states*) are introduced and connected to other states:

a) Connectivity self-model

- Self-model states $\mathbf{W}_{X,Y}$ are added representing connectivity characteristics, in particular connection weights $\omega_{X,Y}$

b) Aggregation self-model

- Self-model states $\mathbf{C}_{j,Y}$ are added representing aggregation characteristics, in particular combination function weights $\gamma_{j,Y}$
- Self-model states $\mathbf{P}_{i,j,Y}$ are added representing aggregation characteristics, in particular combination function parameters $\pi_{i,j,Y}$

c) Timing self-model

- Self-model states \mathbf{H}_Y are added representing timing characteristics, in particular speed factors η_Y

The notations $\mathbf{W}_{X,Y}$, $\mathbf{C}_{j,Y}$, $\mathbf{P}_{i,j,Y}$, \mathbf{H}_Y for the self-model states indicate the referencing relation with respect to the characteristics $\omega_{X,Y}$, $\gamma_{j,Y}$, $\pi_{i,j,Y}$, η_Y : here \mathbf{W} refers to ω , \mathbf{C} refers to γ , \mathbf{P} refers to π , and \mathbf{H} refers to η , respectively. For the processing, these self-model states define the dynamics of any state Y in a canonical manner according to Eq. (1) and (2) whereby the values of $\omega_{X,Y}$, $\gamma_{j,Y}$, $\pi_{i,j,Y}$, η_Y are replaced by the state values of $\mathbf{W}_{X,Y}$, $\mathbf{C}_{j,Y}$, $\mathbf{P}_{i,j,Y}$, \mathbf{H}_Y at time t , respectively.

Note that concerning the terminology used, only the states that represent some network characteristics are called *self-model states*. The states to which these self-model states are connected still belong to the self-model (e.g., as depicted in Figure 1 and

further) but they can either be other self-model states or other states that are not self-model states, such as the states X and Y . An example of an aggregation self-model state $\mathbf{P}_{i,j,Y}$ for a combination function parameter $\pi_{i,j,Y}$ is for the excitability threshold τ_Y of state Y , which is the second parameter of a logistic sum combination function; then $\mathbf{P}_{i,j,Y}$ is usually indicated by \mathbf{T}_Y , where \mathbf{T} refers to τ . The network constructed by the addition of a self-model to a base network is called a *self-modeling network* or a *reified network* for this base network. This constructed network is also a temporal-causal network model itself, as has been shown in [3], Ch. 10; for this reason, this construction can easily be applied iteratively to obtain multiple levels or orders of self-models, in which case the resulting network is called a *multi-order* or *higher-order self-modeling network* or *reified network*.

3. Adaptation Principles from Different Domains

In this section, a number of adaptation principles of different orders are described as can be found in the literature on Cognitive Neuroscience and Social Sciences.

3.1. First-order Adaptation Principles

First-order adaptation principles for some base network address adaptation of some of the base network's characteristics concerning its connectivity, aggregation of multiple connections and timing of node state dynamics. Much research has focused in particular on learning of connectivity characteristics based on adaptive connections, but also other characteristics can be made adaptive, as will be discussed.

3.1.1. The Hebbian Learning Adaptation Principle

As a first example, for mental or neural networks, the *Hebbian learning* adaptation principle [5] can be formulated by:

‘When an axon of cell A is near enough to excite B and repeatedly or persistently takes part in firing it, some growth process or metabolic change takes place in one or both cells such that A’s efficiency, as one of the cells firing B, is increased.’ (3)
[5], p. 62

This is sometimes simplified (neglecting the phrase ‘one of the cells firing B’) to:

‘What fires together, wires together’ [6, 7]

This can easily be modeled by using a *connectivity self-model* based on self-model states $\mathbf{W}_{X,Y}$ representing connection weights $\omega_{X,Y}$.

3.1.2. The Bonding by Homophily Adaptation Principle

An example of the use of a network's self-model for the social domain is the bonding by homophily adaptation principle

‘Birds of a feather flock together’ (4)

This expresses how being ‘birds of a feather’ or ‘being alike’ strengthens the connection between two persons [8-13]. Similar to the Hebbian learning case, this can be modeled by a social network's *connectivity self-model* based on self-model states $\mathbf{W}_{X,Y}$ representing connection weights $\omega_{X,Y}$.

3.1.3. The More Becomes More Adaptation Principle

Another first-order adaptation principle for social networks is the ‘more becomes more’ principle expressing that more popular people attract more connections:

‘Persons with more connections attract more connections’ [4], p. 311 (5)

In a wider context this more becomes more principle relates to what sometimes is called ‘the rich get richer’ [14, 15], ‘cumulative advantage’ [16], ‘the Matthew effect’ [17] or ‘preferential attachment’ [18]. Similar to the Hebbian learning and bonding by homophily cases, this can be modeled by a social network’s *connectivity self-model* based on self-model states $\mathbf{W}_{X,Y}$ representing connection weights $\omega_{X,Y}$.

3.1.4. The Interaction Connects Adaptation Principle

The idea behind the Interaction Connects adaptation principle from Social Science is that ‘The more interaction you have with somebody, the stronger you will become connected’ (6)

See, for example, [19-23]. Similar to the Hebbian learning and bonding by homophily cases, this can be modeled by a social network’s *connectivity self-model* based on self-model states $\mathbf{W}_{X,Y}$ representing connection weights $\omega_{X,Y}$.

3.1.5. The Enhanced Excitability Adaptation Principle

Although connectivity adaptation has some popularity in the literature, also other characteristics can be made adaptive. Instead of a connectivity self-model to model adaptive connection weights, also an *aggregation self-model* can be used, for example, to model intrinsic neuronal excitability, as described in [24]:

‘Long-lasting modifications in intrinsic excitability are manifested in changes (7) in the neuron’s response to a given extrinsic current (generated by synaptic activity or applied via the recording electrode).’ [24], p. 30

This form of adaptation can be modeled by an *aggregation self-model* based on self-model states \mathbf{T}_Y for adaptive excitability thresholds. For example, this type of self-model has been used to model adaptation (desensitization) to spicy food; see [25].

3.2. Second-Order Adaptation Principles

The examples of adaptation principles in Section 3.1 refer to forms of *plasticity*, which can be described by a first-order adaptive network that is modelled using a dynamic first-order self-model for connectivity or aggregation characteristics of the base network, in particular for the connection weights and/or the excitability thresholds used in aggregation. Whether or not and to which extent such plasticity as described above actually takes place is controlled by a form of *metaplasticity*; e.g. [26-31].

3.2.1. The Exposure Accelerates Adaptation Speed Adaptation Principle

For example, in [29] the following compact quote is found indicating that due to stimulus exposure, the adaptation speed will increase:

‘Adaptation accelerates with increasing stimulus exposure’ [29], p. 2. (8)

This indeed refers to a form of metaplasticity, which can be described by a second-order adaptive network that is modeled using a dynamic second-order *timing self-model*, for timing characteristics of a first-order self-model for the first-order adaptation, based on self-model states $\mathbf{H}\mathbf{w}_{X,Y}$ for adaptive learning speed.

3.2.2. The Exposure Modulates Persistence Adaptation Principle

A similar perspective can be applied to obtain a principle for modulation of persistence.

‘Stimulus exposure modulates persistence of adaptation’ (9)

Depending on further context factors, this can be applied in different ways. Reduced persistence can be used in order to be able to get rid of earlier learnt connections that do not apply. However, enhanced persistence can be used to keep what has been learnt. This also refers to a form of metaplasticity, which can be described by a second-order adaptive network that is modeled using a dynamic second-order *aggregation self-model*, for persistence characteristics of a first-order self-model for the first-order adaptation, based on self-model states $\mathbf{M}\mathbf{w}_{XY}$ for an adaptive persistence factor.

3.2.3. The Plasticity Versus Stability adaptation principle

In a similar direction [31] it is more generally discussed how it depends on the circumstances when the extent of plasticity is or should be high and when it is or should be low in favor of stability:

‘The Plasticity Versus Stability Conundrum’ [31], p. 773. (10)

This principle relates to the previous two and can use these second-order self-models.

3.2.4. The Stress Blocks Adaptation Principle

Yet another principle that is indicated in the literature refers to the effect of high stress levels on the extent of plasticity:

‘High stress levels slow down or block adaptation’ (11)

See, for example, the following quote from [27], where such slowing down or blocking of adaptation is called negative metaplasticity:

‘Numerous electrophysiological studies have shown that ‘negative’ metaplasticity develops in brain areas such as the hippocampus and its related structures (e.g., the lateral septum and the nucleus accumbens) following stress.’ [27], p. 631

This can be described by a second-order adaptive network modeled using a dynamic second-order *timing self-model*, for timing characteristics of a first-order self-model for the first-order adaptation, based on self-model states $\mathbf{H}\mathbf{w}_{XY}$ for adaptive learning speed.

The first- and second-order adaptation principles such as the one summarized in (3) to (11) above have been formalized in the form of self-models used in first- and second-order adaptive network models that have been designed, as discussed in Section 4.

4. Using Self-Models to Formalize Adaptation Principles

In this section, it will be shown how the modeling approach for self-modeling network models described in Section 2 can be used to model the adaptation principles of different orders discussed in Section 3. In particular the connectivity and aggregation characteristics of the addressed self-models are discussed. Timing characteristics for these self-models are just values (speed factors for each of the states) that will usually be set depending on a specific application. When self-models are changing over time in a proper manner, this offers a useful method to model adaptive networks based on any adaptation principles. This does not only apply to first-order adaptive networks, but also to higher-order adaptive networks, by using higher-order self-models.

4.1. First-Order Self-Models for First-Order Adaptation Principles

First the adaptation principle for Hebbian Learning will be addressed, as described in Section 3.1.1. To incorporate the ‘firing together’ part, for the self-model’s *connectivity characteristics*, upward causal connections to connectivity self-model state $\mathbf{W}_{X,Y}$ from X and Y are used to formulate a Hebbian learning adaptation principle; see Figure 1. The upward connections have weight 1 here. Also a connection from $\mathbf{W}_{X,Y}$ to itself with weight 1 is used; in pictures they are usually left out.

So, the *connectivity characteristics* of the self-model here consist of the three nodes $\mathbf{W}_{X,Y}$, X , and Y , together with the two incoming upward connections (the blue arrows) from X and Y to $\mathbf{W}_{X,Y}$, one outgoing connection from $\mathbf{W}_{X,Y}$ to Y (the pink downward arrow), and the leveled connection (black arrow) from X to Y . Note that as mentioned in the last paragraph of Section 3.2, only the states that represent a network characteristic are called self-model states, in this case $\mathbf{W}_{X,Y}$. In connectivity pictures such as Figure 1 and further, the self-model states are the states with an outgoing (pink) downward connection. Some other states to which they are connected such as in this case X and Y are still part of the self-model, but will not be called self-model states; they do not have an outgoing downward connection. The downward connection takes care that the value of $\mathbf{W}_{X,Y}$ is actually used for the connection weight of the connection from X to Y . For the *aggregation characteristics* of the self-model, one of the options for a learning rule is defined by the combination function $\mathbf{hebb}_{\mu}(V_1, V_2, W)$ from Table 2, where V_1, V_2 refer to the activation levels of the connected states X to Y , and W to the value of $\mathbf{W}_{X,Y}$ representing the connection weight. For more options of Hebbian learning combination functions and further mathematical analysis of them, see, for example [3], Ch. 14.

Table 2 Combination functions for self-models modeling the first- and second-order adaptation principles. The first five rows cover the first-order adaptation principles from Section 3.1 and the last four rows the second-order adaptation principles from Section 3.2.

Adaptation principle and self-model state	Combination function options	Variables and Parameters
Hebbian Learning $\mathbf{W}_{X,Y}$	3.1.1 $\mathbf{hebb}_{\mu}(V_1, V_2, W) = V_1 V_2 (1 - W) + \mu W$	V_1, V_2 activation levels of connected states W activation level of self-model state for connection weight μ persistence factor
Bonding by Homophily $\mathbf{W}_{X,Y}$	3.1.2 $\mathbf{shomo}_{\alpha, \tau}(V_1, V_2, W) = W + \alpha W (1 - W) (\tau - V_1 - V_2)$	V_1, V_2 activation levels of connected persons W connection weight α modulation factor τ tipping point
More Becomes More $\mathbf{W}_{X,Y}$	3.1.3 $\mathbf{eucl}_{\lambda}(W_1, \dots, W_k)$ $\mathbf{alogistic}_{\sigma, \tau}(W_1, \dots, W_k)$	W_1, \dots, W_k activation levels of self-model states for connection weights of persons connected to B
Interaction Connects $\mathbf{W}_{X,Y}$	3.1.4 $\mathbf{eucl}_{\lambda}(V_1, \dots, V_k)$ $\mathbf{alogistic}_{\sigma, \tau}(V_1, \dots, V_k)$	V_1, \dots, V_k impacts from interaction states for the connected person
Enhanced Excitability \mathbf{T}_Y	3.1.5 $\mathbf{eucl}_{\lambda}(V_1, \dots, V_k)$ $\mathbf{alogistic}_{\sigma, \tau}(V_1, \dots, V_k)$	V_1, \dots, V_k impacts from base states
Exposure Accelerates Adaptation Speed $\mathbf{H}_{W_{X,Y}}$	3.2.1 $\mathbf{eucl}_{\lambda}(V_1, \dots, V_k)$ $\mathbf{alogistic}_{\sigma, \tau}(V_1, \dots, V_k)$	V_1, \dots, V_k impacts from base states and first-order self-model states
Exposure Modulates Persistence $\mathbf{M}_{W_{X,Y}}$	3.2.2 $\mathbf{eucl}_{\lambda}(V_1, \dots, V_k)$ $\mathbf{alogistic}_{\sigma, \tau}(V_1, \dots, V_k)$	V_1, \dots, V_k impacts from base states and first-order self-model states
Plasticity Versus Stability $\mathbf{H}_{W_{X,Y}}, \mathbf{M}_{W_{X,Y}}$	3.2.3 $\mathbf{eucl}_{\lambda}(V_1, \dots, V_k)$ $\mathbf{alogistic}_{\sigma, \tau}(V_1, \dots, V_k)$	V_1, \dots, V_k impacts from base states and first-order self-model states
Stress Blocks Adaptation $\mathbf{H}_{W_{X,Y}}$	3.2.4 $\mathbf{eucl}_{\lambda}(V_1, \dots, V_k)$ $\mathbf{alogistic}_{\sigma, \tau}(V_1, \dots, V_k)$	V_1, \dots, V_k impacts from base states for stress level and first-order self-model states

Next, the adaptation principle for Bonding by homophily will be addressed, as described in Section 3.1.2. It happens that for this connectivity self-model exactly the same *connectivity characteristics* apply as for Hebbian learning, as depicted in Figure 1.

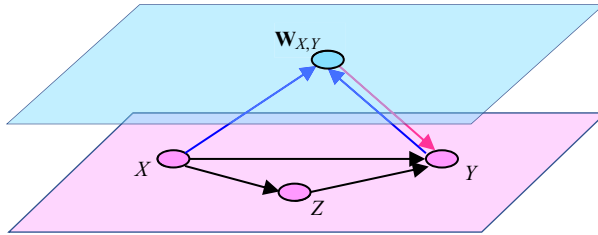


Figure 1. Connectivity characteristics of the self-model for the Hebbian Learning adaptation principle for Mental Networks or the Bonding by Homophily adaptation principle for Social Networks

For *aggregation characteristics* of this self-model, an option for an adaptation rule is defined by the combination function $\mathbf{slhomo}_{\alpha,\tau}(V_1, V_2, W)$ from Table 2, where V_1, V_2 refer to the activation levels (for example, for some opinion) of the connected persons and W to the value of $\mathbf{W}_{X,Y}$ representing the connection weight. For more options and further mathematical analysis, see, for example [3], Ch. 13, or [13].

The More Becomes More adaptation principle as described in Section 3.1.3 has *connectivity characteristics* as shown in Figure 2. Here, the connectivity self-model states for different connections affect each other, as a connection of a person X_3 to a given person Y depends on the existence and strengths of connections from other persons X_i to the same person Y ; see the black leveled arrows in the upper plane.

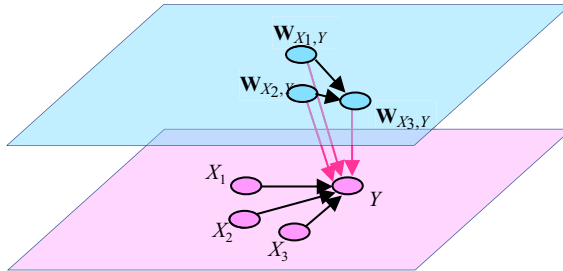


Figure 2. Connectivity characteristics of a self-model for the More Becomes More adaptation principle for person X_3 with respect to person Y

So, in this case the *connectivity characteristics* of the self-model are the nodes $\mathbf{W}_{X_1,Y}$, $\mathbf{W}_{X_2,Y}$, $\mathbf{W}_{X_3,Y}$, and Y , together with leveled connections (black arrows) from each $\mathbf{W}_{X_j,Y}$ to $\mathbf{W}_{X_3,Y}$ and downward connections (pink arrows) from each $\mathbf{W}_{X_j,Y}$ to Y . Again, these (pink) downward connections takes care that the value of $\mathbf{W}_{X_j,Y}$ is actually used for the connection weight of the connection from X_j to Y . For the *aggregation characteristics* of this self-model, some form of aggregation of the weights of these other connections represented by the $\mathbf{W}_{X_j,Y}$ can be used, such as by using a Euclidean or logistic sum combination function; see Table 2. For example, in [32] a logistic sum function was used, and in [33] a scaled sum (with scaling factor the number of existing connections for Y resulting in an average weight), which is a first-order Euclidean combination function.

For the Interaction Connects adaptation principle described in Section 3.1.4, the connectivity self-model states for the connection weights are affected by certain states

$I_{X_i,Y}$ representing the strength of (actual) interaction. Therefore, the *connectivity characteristics* of a self-model for this adaptation principle are as shown in Figure 3, with (blue) upward connections from interaction states $I_{X_i,Y}$ to the self-model states $\mathbf{W}_{X_i,Y}$ and (pink) downward connections from $\mathbf{W}_{X_i,Y}$ to Y . Note that there also multiple interaction states can be used for one connection, for example, for different interaction channels. The *aggregation characteristics* of the self-model states $\mathbf{W}_{X_i,Y}$ can be specified, for example, by a Euclidean or logistic sum function, as shown in Table 2.

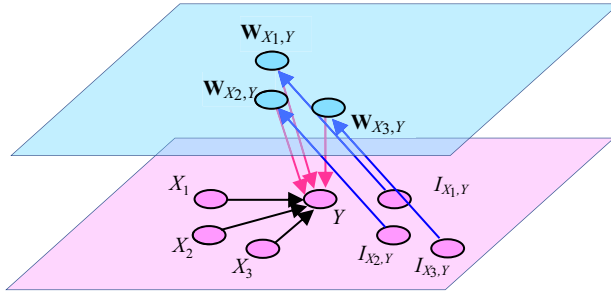


Figure 3. Connectivity characteristics of a self-model for the Interaction Connects adaptation principle for persons X_1 , X_2 and X_3 with respect to person Y .

For the Enhanced Excitability adaptation principle described in Section 3.1.5, an aggregation self-model with *connectivity characteristics* depicted in Figure 4 can be used.

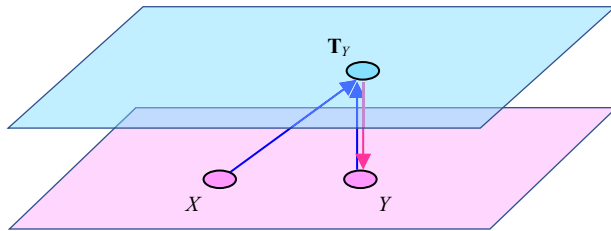


Figure 4. Connectivity characteristics of a self-model for the Enhanced Excitability adaptation principle for persons X_1 , X_2 and X_3 with respect to person Y .

In this case state Y is assumed to use a logistic sum combination function, which has an excitability threshold parameter τ (or any other function with such a parameter). Here this excitability threshold is represented by aggregation self-model state \mathbf{T}_Y which is affected by exposure from activation of the involved states. Note that to enhance excitability, the value of self-model state \mathbf{T}_Y representing the excitability threshold has to decrease. Therefore, these upward connections need to get negative connection weights, whereas a positive connection weight from \mathbf{T}_Y itself can be used. In this case, the (pink) downward connection from \mathbf{T}_Y to Y takes care that the value of \mathbf{T}_Y is actually used for the threshold value of the logistic sum function of Y . Also a connection from a related connectivity self-model state $\mathbf{W}_{X,Y}$ to \mathbf{T}_Y with positive connection weight might be added in this self-model to obtain some balancing effect. For the *aggregation characteristics*, for example, a Euclidean (with odd order n to keep the negative impacts negative) or logistic sum function can be used for \mathbf{T}_Y , as shown in Table 2.

4.2. Second-Order Self-Models for Second-Order Adaptation Principles

The first second-order adaptation principle discussed is the Exposure Accelerates Adaptation Speed principle described in Section 3.2.1. This is modeled by a second-order timing self-model. As it is a second-order adaptation principle for some first-order adaptation principle, for the sake of clarity it is described here with respect to the first-order adaptation principle for Hebbian Learning; although it might be applied to other first-order adaptation principle as well, but then it will have a similar structure to what is shown here. The *connectivity characteristics* of this timing self-model are shown in Figure 5; they consist of the states $\mathbf{Hw}_{X,Y}$, $\mathbf{W}_{X,Y}$, X , and Y , together with the (positive, blue) upward connections from the two base states X and Y to the self-model state $\mathbf{Hw}_{X,Y}$ expressing the part of the principle referring to ‘exposure’, the (negative, blue) upward connection from $\mathbf{W}_{X,Y}$ to the self-model state $\mathbf{Hw}_{X,Y}$ and the downward (pink) connection from $\mathbf{Hw}_{X,Y}$ to $\mathbf{W}_{X,Y}$ that takes care that the value of $\mathbf{Hw}_{X,Y}$ is actually used as speed factor for $\mathbf{W}_{X,Y}$. By the upward connections, stronger activation of the base states X and Y will lead to an increased value of $\mathbf{Hw}_{X,Y}$, as indicated by the part of the principle referring to ‘accelerates’. The (negative) upward connection from the considered state $\mathbf{W}_{X,Y}$ to $\mathbf{Hw}_{X,Y}$ can be used for (counter)balancing. For the *aggregation characteristics*, for example a Euclidean (with odd order n to keep the negative impacts negative) or logistic sum function can be used for $\mathbf{Hw}_{X,Y}$, as shown in Table 2.

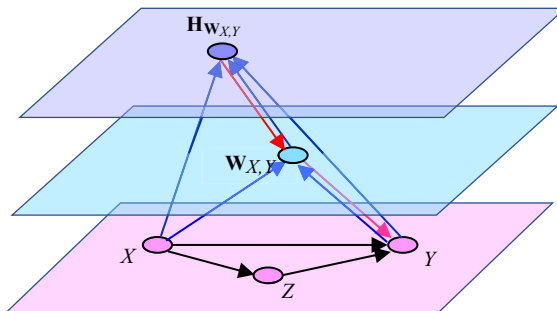


Figure 5. Connectivity of a second-order self-model for the Exposure Accelerates Adaptation Speed adaptation principle with a first-order self-model for Hebbian learning.

Next, the second-order Exposure Modulates Persistence adaptation principle (for the first-order Hebbian Learning principle) described in Section 3.2.2 is addressed, based on second-order aggregation self-model state $\mathbf{Mw}_{X,Y}$ representing persistence of the first-order adaptation. For the *connectivity characteristics* of this self-model, see Figure 6.

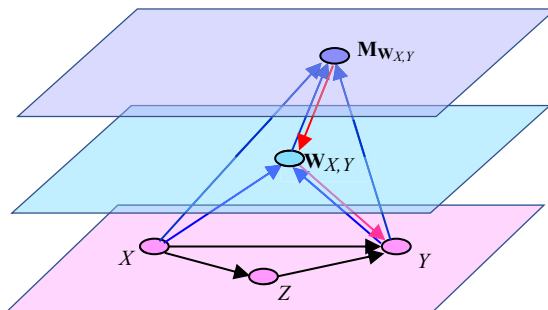


Figure 6. Connectivity of a second-order self-model for the Exposure Modulates Persistence adaptation principle with a first-order self-model for Hebbian learning

The upward connections from base states X and Y to $\mathbf{M}\mathbf{W}_{X,Y}$ may suppress the persistence (when they are negative). This paves the road to get rid of the learnt effects from the past in case they are no longer applicable. The positive upward connection from first-order state $\mathbf{W}_{X,Y}$ to $\mathbf{H}\mathbf{W}_{X,Y}$ can be used for counterbalancing. However, the upward connections from base states X and Y to $\mathbf{M}\mathbf{W}_{X,Y}$ can also be made positive in which case they increase persistence during a learning process to keep the learnt effect well. This also illustrates the Plasticity Versus Stability Conundrum adaptation principle described in Section 3.2.3. The (pink) downward connection from $\mathbf{M}\mathbf{W}_{X,Y}$ to $\mathbf{W}_{X,Y}$ takes care that the value of $\mathbf{W}_{X,Y}$ is actually used for the connection weight of the connection from X to Y . For the *aggregation characteristics*, for example a Euclidean (with odd order n) or logistic sum function can be used for $\mathbf{M}\mathbf{W}_{X,Y}$, as shown in Table 2.

Finally, a second-order self-model for the Stress Blocks Adaptation principle described in Section 3.2.4 can be obtained in a similar way as the one for Exposure Accelerates Adaptation Speed principle (see connectivity in Figure 5) but this time with *connectivity characteristics* based on a negative upward connection from a base state representing the stress level, which brings the timing characteristic self-model state $\mathbf{H}\mathbf{W}_{X,Y}$ to low values or even 0. For the *aggregation characteristics*, again for example a Euclidean or logistic sum function can be used for $\mathbf{H}\mathbf{W}_{X,Y}$; see Table 2.

5. Discussion

In this paper the use of self-modeling networks to model adaptive biological, mental and social processes of any order of adaptation was addressed. Following the network-oriented modeling approach described in [3], it was shown how self-models for networks provide useful pre-specified building blocks to design complex multi-order adaptive network models in the form of self-modeling networks. This was illustrated for a number of adaptation principles from different application domains. A dedicated software environment for self-modeling networks that has been developed supports the modeling and simulation: <https://www.researchgate.net/project/Network-Oriented-Modeling-Software>.

As an illustration, in [3], Ch. 4, four of the adaptation principles known from the literature and specified in Section 4 were applied to obtain a network model involving both plasticity and metaplasticity. In particular, two first-order adaptation principles (for Hebbian Learning and for Enhanced Excitability) and two second-order adaptation principles (for Exposure Accelerates Adaptation Speed and for Exposure Modulates Persistence) are covered in this network model.

References

- [1] Treur J. Multilevel network reification: representing higher order adaptivity in a network. In: Aiello L, Cherifi C, Cherifi H, Lambiotte R, Lió P, Rocha L. editors. Proc. of the 7th Int. Conf. on Complex Networks and their Applications, ComplexNetworks'18, vol. 1. Studies in Computational Intelligence, vol. 812, Springer Nature, 2018, p. 635-51.
- [2] Treur J. Modeling higher-order adaptivity of a network by multilevel network reification. Network Science 2020;8: S110-44.
- [3] Treur J. Network-oriented modeling for adaptive networks: designing higher-order adaptive biological, mental, and social network models. Cham, Switzerland: Springer Nature Publishing; 2020. 412 p.

- [4] Treur J. Network-Oriented Modeling: Addressing Complexity of Cognitive, Affective and Social Interactions. Cham, Switzerland: Springer Publishers; 2016. 499 p.
- [5] Hebb DO. The organization of behavior: A neuropsychological theory. New York: John Wiley and Sons; 1949. 335 p.
- [6] Shatz CJ. The developing brain. *Sci. Am.* 1992; 267:60–67. (10.1038/scientificamerican0992-60)
- [7] Keysers C, Gazzola V. Hebbian learning and predictive mirror neurons for actions, sensations and emotions. *Philos Trans R Soc Lond B Biol Sci* 2014;369: 20130175.
- [8] Pearson M, Steglich C, Snijders T. Homophily and assimilation among sport-active adolescent substance users. *Connections* 2006;27(1):47–63.
- [9] McPherson M, Smith-Lovin L, Cook JM. Birds of a feather: homophily in social networks. *Annu. Rev. Sociol.* 2001; 27:415–44.
- [10] Levy DA, Nail PR. Contagion: A theoretical and empirical review and reconceptualization. *Genetic, social, and general psychology monographs* 1993;119(2):233-284.
- [11] Holme P, Newman MEJ. Nonequilibrium phase transition in the coevolution of networks and opinions *Phys. Rev. E* 2006;74(5):056108.
- [12] Sharpanskykh A, Treur J. Modelling and analysis of social contagion in dynamic networks. *Neurocomputing* 2014; 146:140–50.
- [13] Treur J. Mathematical analysis of the emergence of communities based on coevolution of social contagion and bonding by homophily. *Applied Network Science* 2019;4: article 1.
- [14] Simon HA. On a class of skew distribution functions *Biometrika* 1955; 42: 425–40.
- [15] Bornholdt S, Ebel H. World wide webscaling exponent from Simon’s 1955 model *Phys. Rev. E* 2001;64: article 035104.
- [16] Price DJ. de S. A general theory of bibliometric and other cumulative advantage processes *J. Amer. Soc. Inform. Sci.* 1976; 27: 292–306
- [17] Merton RK. The Matthew effect in science. *Science* 1968;159: 56–63.
- [18] Barabási AL, Albert R. Emergence of scaling in random networks. *Science* 1999; 286: 509-512.
- [19] Hove MJ, Risen JL. It’s all in the timing: interpersonal synchrony increases affiliation. *Soc. Cognit.* 2009; 27: 949–60. (doi:10.1521/soco.2009.27.6.949)
- [20] Pearce E, Launay J, Dunbar RIM. (). The Ice-breaker Effect: singing together mediates fast social bonding. *Royal Society Open Science* 2015;2: article 150221 <http://dx.doi.org/10.1098/rsos.150221>.
- [21] Weinstein D, Launay J, Pearce, E, Dunbar RIM, Stewart L. Singing and social bonding: Changes in connectivity and pain threshold as a function of group size. *Evolution & Human Behaviour* 2016;37(2):152-58. doi: 10.1016/j.evolhumbehav.2015.10.002
- [22] Gilbert E, Karahalios K. Predicting tie strength with social media. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems CHI’09, 2009*, p. 211-20.
- [23] Morris MR, Teevan J, Panovich K. What do people ask their social networks, and why? a survey study of status message Q&A behavior. *CHI 2010*. 2010.
- [24] Chandra N, Barkai E. A non-synaptic mechanism of complex learning: modulation of intrinsic neuronal excitability. *Neurobiology of Learning and Memory* 2018; 154: 30-36.
- [25] Choy M, El Fassi S, Treur J. An adaptive network model for pain and pleasure through spicy food and its desensitization. *Cognitive Systems Research* 2020: in press
- [26] Abraham WC, Bear MF. Metaplasticity: the plasticity of synaptic plasticity. *Trends in Neuroscience* 1996;19(4):126-130.
- [27] Garcia R. Stress, metaplasticity, and antidepressants. *Current Molecular Medicine* 2002; 2: 629-38.
- [28] Magerl W, Hansen N, Treede RD, Klein T. The human pain system exhibits higher-order plasticity (metaplasticity). *Neurobiology of Learning and Memory* 2018; 154:112-20.
- [29] Robinson BL, Harper NS, McAlpine D. Meta-adaptation in the auditory midbrain under cortical influence. *Nat. Commun.* 2016; 7: article 13442.
- [30] Sehgal M, Song C, Ehlers VL, Moyer Jr JR. Learning to learn – intrinsic plasticity as a metaplasticity mechanism for memory formation. *Neurobiology of Learning and Memory* 2013; 105: 186-99.
- [31] Sjöström PJ, Rancz EA, Roth A, Hausser M. Dendritic excitability and synaptic plasticity. *Physiol Rev* 2008; 88: 769–840.
- [32] Beukel S van den, Goos SH, Treur J. An adaptive temporal-causal network model for social networks based on the homophily and more-becomes-more principle. *Neurocomputing* 2019; 338: 361-71
- [33] Blankendaal R, Parinussa S, Treur J. A temporal-causal modelling approach to integrated contagion and network change in social networks. *Proceedings of the Twenty-second European Conference on Artificial Intelligence, ECAI’16, 2016*, p. 1388–96

This page intentionally left blank

Part II
Communication and Networks

This page intentionally left blank

Application of Twin Objective Function SVM in Sentiment Analysis

Qiaoman YANG^{a1}, Chunyu LIU^b

^aShaanxi College of Communication Technology, China

^bBeijing Normal University, China

Abstract. Classification modeling is one of the key issues in sentiment analysis. Support vector machine (SVM) has been widely used in classification as an effective machine learning method. Generally, a common SVM is only for decision-making that sacrifices the distribution of data. In practice, sentiment data are big and mazy, which results in the deficiency of accuracy and stability when common SVM is used. The study investigates sentiment analysis by applying the twin objective function SVM, including nonparallel SVM(NPSVM) and twin SVM (TWSVM). From the experiments, we concluded that twin objective function SVMs are superior to NB and single objective function SVM in accuracy and stability.

Keywords. sentiment analysis, twin objective functions, SVM, NB

1. Introduction

Because of the flourishing development of the internet and Web2.0, an increasing number of people are depending on the network. They express their feelings or evaluate on things via discussion forums, blogs, twitter, etc., which results in massive and big data explosion. However, it is difficult to collect and process vast amounts of information online by artificial methods. Sentiment analysis technology came into being in this environment. Sentiment analysis technology is an emerging field that mine unstructured information. Its aim is to take advantage of automation and intelligent technology analyzing mining, reasoning and learning the subjectivity text messages that have some emotional colors, obtaining the potential and valuable hidden information. Currently, sentiment analysis technology has many widely used applications in fields of business intelligence application, recommender systems, message filtering and so on [1][2]. Therefore sentiment analysis technology is a hot research field. Sentiment information classification, which classifies the binary appraisal data, that have subjective information, is one of the main tasks in the sentiments analysis. Sentiment analysis, which is based on supervised learning, emotional research is a hot research area based on machine learning. The machine learning technique of emotional classification is of because it can model many functions and capture context [3] in the process. And they are relatively easy to adapt to

¹ Corresponding Author. Assistant, Shaanxi College of Communication Technology, 19 Wenjing Road, Xi'an, Shaanxi China; E-mail: yangqman@163.com.

changing inputs and the possibility of measuring the uncertainty of the classification. The method of monitoring training from the manual classification is the most popular.[4],[5],[6],[7].

In addition to the Non-negative Matrix Tri-factorization[8],Genetic Algorithms[9], the most supervised learning algorithm which used on the sentiment analysis are still Naïve Bayes (NB),Maximum Entropy (ME)and Support Vector Machine (SVM) [10],[11]. A SVM [12] Operate by constructing a hyper plan with the closest training distance with the largest Euclidean distance. This can regard as the range between the separating hyper plane and the two parallel hyper planes on each side, indicate the border of an example of a class in the feature space. It is known that SVM is robust in the case of many features, not destined by the curse of dimensions, and produces the highest accuracy in sentiment classification [13]. However, a single objective classification, that aims at structuring the decision-surface of data, does not fit the data well, and it has some disadvantages, such as, running slowly, and being applicable only to small data.

Twin objective function SVM are now a hot in the field of studying support vector machine. The method has two objective functions, aiming to produce two separating hyper planes, thus, having more advantages than single objective function SVM. Such as, considering more data characteristics, more amenable to parallel computing, making the classification more accurate, faster, broadening the scope of data, and so on. This study applies two general and representative twin objective function SVM to the sentiment analysis. These two methods are Twin Support Vector Machine (TWSVM) introduced by R.K.jayaadeva. et. al. and Non-parallel Support Vector Machine introduced by Tian [14]. In our experiment, we compare Naïve Bayes (NB), the classical SVM algorithm Lib SVM and SMO with the two methods that are used in this study. The results show that twin objective function SVM are prior to other machine learning methods in accuracy and stability on the sentiment analysis data sets.

This paper is organized as follows. Section 2 briefly dwells on Naïve Bayes, the single objective function support vector machine. Section 3 describes two twin objective function support vector machines (TWSVM and NPSVM). Section 4 presents the experimental results and section 5 contains the conclusion.

2. Methods

2.1 Naïve Bayes

One approach to text classification is to assign to a given document a class $b^* = \arg \max_b p(b|a)$. We derive the NB classifier by first observing that by Bayes' rule:

$$p(b|a) = \frac{p(b)p(a|b)}{p(a)},$$

where $p(a)$ plays no role in selecting b^* . To estimate the term $p(a|b)$, NB decomposes it by assuming the f_i 's are conditionally independent given a 's class:

$$p_{NB}(b|a) = \frac{p(b)(\prod_{i=1}^m p(f_i|b)^{n_i(a)})}{p(a)}$$

That is to say, assuming that each independent function is an indication of the allocation class, independent of each other, we chose this classifier because of its

simple implementation, high computational efficiency, and straightforward features in learning. [15], [16]. We used the NB classifier of WEKA.

2.2 SVM

SVM, which was introduced by Vapnik and his co-workers in the early 1990s, is a classifier for the two types of the optimal classification that are linear inseparable. The goal of the SVM is to find an optimal separating hyper plane by constructing a hyper plane separating two classes with no error and maximizing the margin of the two types.

For classification about the training data

$$T = \{(x_i, y_i) | i = 1, \dots, l\}, \tag{1}$$

where $x_i \in \mathfrak{R}^n, y_i \in Y = \{+1, -1\}, i = 1, \dots, l$. Linear SVM is applied to solve the following primal quadratic programming problem (QPP):

$$\min_{w, b, \xi} \frac{1}{2} \|w\|_2^2 + C \sum_{i=1}^l \xi_i, \tag{2}$$

$$s.t. y_i ((w \cdot x_i) + b) \geq 1 - \xi_i, \quad \xi_i \geq 0, i = 1, 2, \dots, l,$$

where C is a penalty parameter and ξ_i are the slack variables. The Wolf Dual of

(2) can be expressed as:

$$\max_{\alpha} \sum_{j=1}^l \alpha_j - \frac{1}{2} \sum_{j=1}^l \sum_{j=1}^l y_j y_j (x_j \cdot x_j) \alpha_j \alpha_j \tag{3}$$

$$s.t. \sum_{j=1}^l y_j \alpha_j = 0, \quad 0 \leq \alpha_i \leq C, i = 1, l, l,$$

where $\alpha \in \mathfrak{R}^l$ are Lagrangian multipliers. The optimal separating hyper plane of (3) is expressed as:

$$w = \sum_{i=1}^l \alpha_i^* y_i x_i, \quad b = \frac{1}{N_{sv}} (y_j - \sum_{i=1}^{N_{sv}} \alpha_i^* y_i (x_i \cdot x_j)) \tag{4}$$

A new sample is classified as +1 or -1 according to the final decision function $f(x) = \text{sgn}((w \cdot x) + b)$.

Lib SVM and SMO are the effective algorithms to solve SVM. Lib SVM is developed and designed by Zhiren Lin and others of the Taiwan University. It is a simple, easy to use fast, and efficient packages to solve the SVM recognition and regression. It involves few adjustment parameters, and provides cross-validation functions. SMO is a fast and efficient algorithm for the SVM. It transfers the large convex programming problem into a small convex programming problem, and then it can train the larger data sets. The study described in this paper mainly uses the following two methods to solve the problem of the SVM.

2.3 Twin Objective Function SVM

Single objective SVM has some of the following problems when it is applied in the sentiment analysis. A single objective function is aimed at constructing the decision-surface of the data, but it is not fit well to the distribution of data with the different emotional categories. Twin objective function SVM can achieve optimal data classification data while approaching data effectively.

2.3.1 TWSVM

TWSVM is proposed by Jayadeva based on SVM. Its goal is to find two hyper planes for binary classification, with each super plane approaching one of the categories and being far away from other categories as soon as possible.

Consider the binary classification problem with the training set (5)

$$T = \{(x_1, +1), \dots, (x_p, +1), (x_{p+1}, -1), \dots, (x_{p+q}, -1)\},$$

TWSVM seeks two hyper planes (6)

$$(w_+ \cdot x) + b_+ = 0 \text{ and } (w_- \cdot x) + b_- = 0$$

pass through seek two small QPPS (7)

$$\begin{aligned} \min_{w_+, b_+, \xi_+} & \frac{1}{2} \sum_{i=1}^p ((w_+ \cdot x_i) + b_+)^2 + d_1 \sum_{j=p+1}^{p+q} \xi_j \\ \text{s.t.} & (w_+ \cdot x_j) + b_+ \leq -1 + \xi_j \\ & \xi_j \geq 0, j = p+1, \dots, p+q \end{aligned}$$

and (8)

$$\begin{aligned} \min_{w_-, b_-, \xi_-} & \frac{1}{2} \sum_{i=p+1}^{p+q} ((w_- \cdot x_i) + b_-)^2 + d_2 \sum_{j=1}^p \xi_j \\ \text{s.t.} & (w_- \cdot x_j) + b_- \leq 1 - \xi_j \\ & \xi_j \geq 0, j = 1, \dots, p \end{aligned}$$

where $d_i \geq 0, i=1,2$ are the penalty parameters .First, we obtain the dual problems of (7) and (8) by using the Lagrangian equation and KKT conditions:

$$\begin{aligned} \max_{\alpha} & e_+^T \alpha - \frac{1}{2} \alpha^T G (H^T H)^{-1} G^T \alpha \\ \text{s.t.} & 0 \leq \alpha \leq d_1 e_- \end{aligned} \tag{9}$$

and

$$\begin{aligned} \max_{\beta} & e_+^T \beta - \frac{1}{2} \beta^T P (Q^T Q)^{-1} P^T \beta \\ \text{s.t.} & 0 \leq \beta \leq d_2 e_+ \end{aligned} \tag{10}$$

where

$$\begin{aligned} G &= [B e_-], H = [A e_-], P = [A e_+], Q = [B e_-] \\ e_+ &= (1, \dots, 1)^T \in R^p, e_- = (1, \dots, 1)^T \in R^q \\ A &= (x_1, \dots, x_p)^T \in R^{p \times n}, B = (x_{p+1}, \dots, x_{p+q})^T \in R^{q \times n} \end{aligned} \tag{11}$$

We can obtain the solutions (w_+, b_+) and (w_-, b_-) of the problems (9) and (10) from the above calculation. A new data $x \in \mathcal{R}^n$ can be predicted to belong to emotional category by

$$\text{Class} = \arg \min_k |(w_k \cdot x) + b_k| \tag{12}$$

where $|\cdot|$ is the perpendicular distance of point x from the planes

$$(w_k \cdot x) + b_k = 0, k = -, +$$

TWSVM improves the accuracy of the SVM, while increasing the speed of SVM, but TWSVM needs to calculate the inverse matrix during computing process, thus is still very difficult for large-scale data processing.

2.3.2 NPSVM

NPSVM is based on sustain vector machine and double support vector machine, its main goal is also to find two hyper planes fitting each category of data. However, due to its form, it is very similar to the standard SVM, we can solve the NPSVM by using the technology for solving the standard SVM.

We seek two nonparallel hyper planes $(w_+ \cdot x) + b_+ = 0$ and $(w_- \cdot x) + b_- = 0$ for the two types of data (such as (5)) by solving two convex QPPs

$$\begin{aligned}
 & \min_{w_+, b_+, \eta_i^*, \xi_j} \frac{1}{2} (\|w_+\|^2) + C_1 \sum_{i=1}^p (\eta_i + \eta_i^*) + C_2 \sum_{j=p+1}^{p+q} \xi_j, \\
 \text{s.t.} \quad & (w_+ \cdot x_i) + b_+ \leq \varepsilon + \eta_i, i=1, \dots, p, \\
 & -(w_+ \cdot x_i) - b_+ \leq \varepsilon + \eta_i^*, i=1, \dots, p, \\
 & (w_+ \cdot x_j) + b_+ \leq -1 + \xi_j, j=p+1, \dots, p+q, \\
 & \eta_i, \eta_i^* \geq 0, i=1, \dots, p, \\
 & \xi_j \geq 0, j=p+1, \dots, p+q,
 \end{aligned} \tag{13}$$

and

$$\begin{aligned}
 & \min_{w_-, b_-, \eta_i^*, \xi_j} \frac{1}{2} (\|w_-\|^2) + C_3 \sum_{i=p+1}^{p+q} (\eta_i + \eta_i^*) + C_4 \sum_{j=1}^p \xi_j, \\
 \text{s.t.} \quad & (w_- \cdot x_i) + b_- \leq \varepsilon + \eta_i, i = p+1, L, p+q, \\
 & -(w_- \cdot x_i) - b_- \leq \varepsilon + \eta_i^*, i = p+1, L, p+q, \\
 & (w_- \cdot x_j) + b_- \geq -1 + \xi_j, j = 1, L, p, \\
 & \eta_i, \eta_i^* \geq 0, i = p+1, L, p+q, \\
 & \xi_j \geq 0, j = 1, L, p,
 \end{aligned} \tag{14}$$

where $C_i \geq 0, i = 1, \dots, 4$ are the penalty parameters. First, we obtain the dual problems of (13) and (14) by using the Lagrangian equation and KKT conditions:

$$\begin{aligned}
 & \min_{\alpha_+^*, \beta_-} \frac{1}{2} (\alpha_+^* - \alpha_+)^T A A^T (\alpha_+^* - \alpha_+) \\
 & \quad - (\alpha_+^* - \alpha_+)^T A B^T \beta_- + \frac{1}{2} \beta_-^T B B^T \beta_- \\
 & \quad + \varepsilon e_+^T (\alpha_+^* + \alpha_+) - e_-^T \beta_-, \\
 \text{s.t.} \quad & e_+^T (\alpha_+ - \alpha_+^*) + e_-^T \beta_- = 0 \\
 & 0 \leq \alpha_+, \alpha_+^* \leq C_1 e_+, \quad 0 \leq \beta_- \leq C_2 e_-.
 \end{aligned} \tag{15}$$

and

$$\begin{aligned}
 & \min_{\alpha_-^*, \beta_+} \frac{1}{2} (\alpha_-^* - \alpha_-)^T B B^T (\alpha_-^* - \alpha_-) \\
 & \quad + (\alpha_-^* - \alpha_-)^T B A^T \beta_+ + \frac{1}{2} \beta_+^T A A^T \beta_+ \\
 & \quad + \varepsilon e_+^T (\alpha_-^* + \alpha_-) - e_+^T \beta_+, \\
 \text{s.t.} \quad & e_-^T (\alpha_- - \alpha_-^*) + e_+^T \beta_+ = 0 \\
 & 0 \leq \alpha_-, \alpha_-^* \leq C_3 e_-, \quad 0 \leq \beta_+ \leq C_4 e_+.
 \end{aligned} \tag{16}$$

where

$$\begin{aligned}
 \alpha_+^* &= (\alpha_+^T, \alpha_+^{*T}) = (\alpha_1, \dots, \alpha_p, \alpha_1^*, \dots, \alpha_p^*)^T, \\
 \alpha_-^* &= (\alpha_-^T, \alpha_-^{*T}) = (\alpha_p, \dots, \alpha_{p+q}, \alpha_p^*, \dots, \alpha_{p+q}^*)^T, \\
 \beta_+ &= (\beta_1, \dots, \beta_p), \beta_- = (\beta_{p+1}, \dots, \beta_{p+q}).
 \end{aligned}$$

We can obtain the solutions (w_+, b_+) and (w_-, b_-) of the problems (13) and (14) from the above calculation. A new data $x \in \mathfrak{R}^n$ can be predicted to be positive or negative by (12)

$$\text{Class} = \arg \min_{k=-,+} |(w_k \cdot x) + b_k|$$

NPSVM is much more easily solved by computing than TWSVM, Therefore, it runs faster than the latter, and easier to implement as well. Since the main functions of these two SVMs appear in pairs, we call them the dual objective function support vector machine.

3. Experiments

The study reported in this paper mainly uses the standard database that is commonly used in the sentiment analysis ,the first data set is the Intel Movie Database (IMDB),the second dataset is the Pang[13] Database. The third is the NLPDVD database. The features of these three data sets shown in table 1.

Table 1. Database features

Database	Size	Positive	Negative	Feature
Pang2002	2000	1000	1000	8614
IMDB	4000	2000	2000	5706
NLPDVD	4000	2000	2000	2000

In the experiments, in order to obtain the precision execution efficiency, we use the data preprocessing. That is, we select the emotional word with the document frequency greater than two and with the document sets greater than ten as the feature. Then, we use TF-IDF to represent the text. Each experimental data set is randomly and equally divided into a training set and test set.

In this section, we compare NB, SVM (Lib SVM and SMO) and dual objective function on the three standard data sets. In a large number of experiments, the study found using the linear kernel on the three data sets is better than radial basis function. Therefore, all the methods in the experiments use the linear kernel. For all the methods, the optimal parameters $d_i, i=1,2$ in TWSVM and, $C_i, i=1,L,4$ in NPSVM are tuned for the best classification accuracy in the range 2^{-5} and 2^5 . Get the best parameters ε in this range $[0, 0.5]$ with the step 0.05. Lib SVM, SMO and NB taken from WEKA which is an open source collection tools. Classification accuracy of each method is measured by the standard fivefold cross-validation methodology. All methods are implemented on a PC with an Intel Core processor and 4GB RAM. The experimental results are shown in Table2.

Table 2. Mean value of benchmark date set

Data sets	NB Accuracy%	Lib SM Accuracy%	SMO Accuracy%	TWSM Accuracy%	NPSM Accuracy%
Pang2002	70.5	77.9	84.8	92.74	81.4
IMDB	79.94	83.9	86.1	90	92.69
NLPDVD	66.5	75.65	69.4	87.64	81.15

In order to illustrate the stability and accuracy of the algorithm in sentiment analysis, we performed some experiments on IMDB data sets and pang2002 data sets, to account for various percentages of the training set for testing. The results are shown in figures 1 and 2.

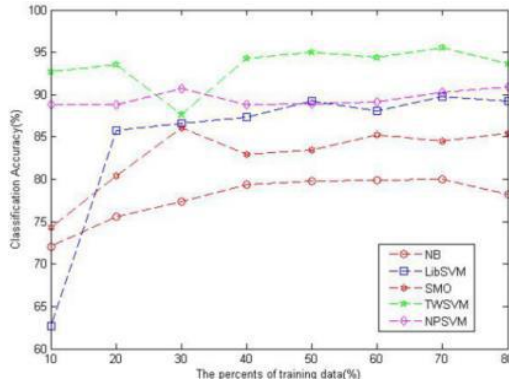


Figure 1. Average results of IMDB

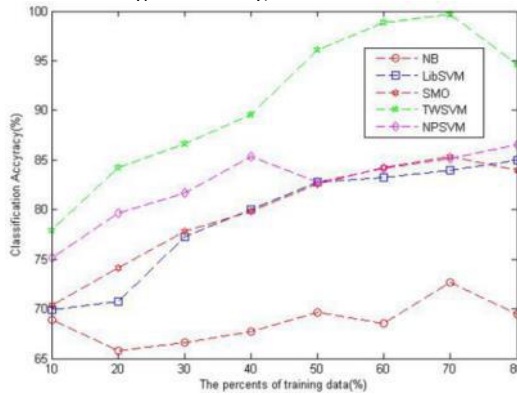


Figure.2. The average results on pang2002

The analysis of the above results according to Table 2 shows that Twin objective function SVMs are higher by approximately 1% to 10% than NB and Single objective function SVM in classification accuracy, when they run on IMDB data sets and pang2002 data sets. As clearly seen from figures 1 and 2, NPSVM and TWSVM are better than NB, Lib SVM, and SVM classifiers in stability and accuracy.

4. Conclusion

In this study, the Twin objective functions TWSVM and NPSVM were applied to sentiment analysis and good results were obtained. Experiments show that in most cases, the double objective function SVM has better classification accuracy than the other three methods. At the same time, the training speed of NPSVM is significantly faster than Lib SVM and SVM. In addition, the dual objective function SVM generates two separate hyper planes. Therefore, its application range is much wider. In sentiment analysis, the life-long objective function SVM, as a classifier, has a stronger generalization performance than a general support vector machine.

References

- [1] Loren Terveen, Will Hill et al, "PHOAKS: A system for sharing recommendations," *Communications of the ACM*, vol, 40,no.3,pp.59–62,1997.
- [2] Ellen Spertus, "Smokey: Automatic recognition of hostile messages," In *Proc. of Innovative Applications of Artificial Intelligence*, pp.1058–1065, 1997.
- [3] Polanyi L and Zaenen A, "Contextual valence shifters. In: *Computing attitude and affect in text: Theory and applications*," Springer, pp. 1–10,2006.
- [4] Thomas, M, Pang, B, and Lee, L, "Get out the vote: Determining support or opposition from Congressional debate transcripts," In *Empirical Methods in Natural Language Processing*,2006.
- [5] Finn A and Kushmerick N, "Learning to classify documents according to genre.,"*Journal of the American Society for Information Science*, vol,57,pp.1506–1518, 2003.
- [6] Bai X, Padman R and Airolidi E, "On learning parsimonious models for extracting consumer opinions," In: *Proceedings of HICSS-05, 38th Annual Hawaii International Conference on System Sciences*, IEEE Computer Society, Washington, DC, pp. 75–82, 2005.
- [7] Aue A and Gamon M, "Customizing sentiment classifiers to new domains," a case study, Technical report, Microsoft Research,2005.
- [8] Tao Li, Ti Zhang, Vikas Sindhwani, "A Non-negative Matrix Tri-factorization Approach to Sentiment Classification with Lexical Prior Knowledge" *Proceedings of the 47th Annual Meeting of the ACL and the 4th IJCNLP of the AFNLP*, pp.244–252, Suntec, Singapore, August .2009.
- [9] Ahmed Abbasi, Hsinchun Chen, Arab Salem, "Sentiment Analysis in Multiple Languages: Feature Selection for Opinion Classification in Web Forums," *ACM Tranon IS*, Vol,26, no.3, pp. 12:1-12:34,2008.
- [10] Bo Pang and Lillian Lee, "Sentiment Analysis Using Subjectivity Summarization Based on Minimum Cuts," In *Proceedings of the 42nd Annual Meeting of the Association for Computational Linguistics*, Barcelona, Spain.2004.
- [11] Berger A L, Pietra S D and Pietra V J D, "A maximum entropy approach to natural language processing," *Computational Linguistics*, vol, 22, pp.39–71, 1996.
- [12] V. N. Vapnik, "An overview of statistical learning theory," *IEEE Trans. Neural Networks*, vol ,10, no.1, pp.988–999, 1999.
- [13] Pang B, Lee L, et.al,"Thumbs up? Sentiment classification using machine learning technique," *Proc.of the EMLP* vol,7, pp.79-86, 2002.
- [14] Yingjie Tian, Xuchan Ju, Zhiqian Qi, "Nonparallel support vector for pattern classification,"*IEEE Trans. Cybernetics*, Vol,44,no.7,pp.1067-1076, 2014.
- [15] Salvetti F, Lewis S and Reichenbach C, "Impact of lexical filtering on overall opinion polarity identification," *Proceedings of the AAAI Spring Symposium on Exploring Attitude and Affect inText: Theories and Applications*, Stanford, CA.2004.
- [16] Pedro Domingos and Michael J. Pazzani, "On the optimality of the simple Bayesian classifier under zero-one loss," *Machine Learning*, vol, 29, no. (2-3): pp.103–130,1997.

Robust H-infinity State Estimation of Uncertain Neural Networks with Two Additive Time-Varying Delays

Xiaoping HU¹, Yajun WANG, Jiakai DING and Dongming XIAO

Engineering Research Center of Advanced Mine Equipment, Ministry of Education, Hunan University of Science and Technology, Xiangtan, 411201, China

Abstract. This study is mainly concerned with the problem of robust H_∞ state estimation of uncertain neural networks with two additive time-varying delays. A novel linear matrix inequalities (LMIs) is constructed based on Lyapunov-Krasovskii functionals (LKFs) which contains two additive time-varying delays components. LMIs method are used to estimate the derivative of LKFs, it is calculated that the derivative of the LKFs is smaller than zero, which proved that uncertain neural networks with two additive time-varying delays is globally asymptotically stable. Meantime, a stability criterion of error system is presented such that the H_∞ performance is guaranteed. Finally, two numerical simulation examples have been performed to demonstrate the effectiveness of developed approach.

Keywords. linear matrix inequalities (LMIs), H_∞ state estimation, two time-varying delays, uncertain neural networks

1. Introduction

Time-varying delays has received much attention which exists in many industrial and engineering systems [1, 2]. Time-varying delays frequently cause oscillations, divergence or instability of the systems, so the systems stability is the main consideration for time-varying delay in many applications.

The stability problems for time-varying delays systems have been researched in recent years. For example, Wu [3] studied the stability of an uncertain systems, which contains multiple consecutive delay components. By considering the relationship between the time-varying delays and its upper bound, but there is no system state estimation, it is difficult to discover the change of the system, and the random interference items has not been added. Whether the delay is included according to the stability criteria, the delay systems is divided into two classes by Liu [4] which are delay-independent and delay-dependent. Two additive time-varying delays system is studied by Xiong [5] and establishes two novel integral inequalities, but the calculation process of this research is too cumbersome. The stability of continuous linear system with two additive time-varying delays is studied by Xu [6], which utilized the

¹ Corresponding author. Email: hxp210@163.com

reciprocally convex combination technique. Subramanian [7] and Samidurai [8] studied robust analysis for uncertain neural networks with two additive time-varying delays by constructing Lyapunov-Krasovskii functions (LKFs) and making use of linear matrix inequalities (LMIs) technique. Liang [9] and Yuan [10] studied complex-valued neural networks (CVNNs) with two additive time-varying delays by constructing LKFs and making use of LMIs to solve the stable problems. The results show the convergence of the real and imaginary parts. The above literatures are for the stability analysis of two additive time-varying delays system. These literatures did not refer to the system by constructing state estimator.

Bao [11], Hou [12], Liu [13], Shen [14], Zhang [15] and Zhao [16] studied with H_∞ state estimation problem for time-varying artificial neural networks. The aim of these papers is to design a time-varying H_∞ estimator, such that the dynamics of the estimation error satisfy the given H_∞ performance requirement. Liu [17] studied the problem of H_∞ state estimation of a static neural network with time-varying delay by constructing a suitable Lyapunov function and ensuring the attenuation of the results as early as possible, which indicates that the system is progressively stable under Lyapunov's conditions. H_∞ state estimation had been constructed in above literatures. Shao [18] studied an H_∞ control problem with delay-dependent stability condition, a new stability criteria is obtained, and finally the system is proved to be asymptotically stable by constructing a Lyapunov functional. Zhou [19] studied the robust H_∞ control problem for a delay singular system with parameter uncertainties. The influence of the system on the disturbance is attenuated, which indicate that the system is in the state of Lyapunov asymptotically stable by constructing a robust state feedback control law. A non-fragile state estimator of the recurrent delayed neural networks is designed by Yang [20] to ensure the existence of the desired estimators. The inadequacy of these researches is that they do not consider two additive time-varying delays issues. Among them, the H_∞ control essence is an optimization design using the H_∞ norm as the objective function. The H_∞ norm is a norm defined on the Hardy space. In the H_∞ control theory, it refers to the maximum singular value of the rational function matrix parsed in the right half plane of S.

Zhou [21] studied the problem of robust finite-time state estimation for a class of discrete-time neural networks with two delay components and Mrakovian jump parameters. And a new LKFs is constructed. Duan [22] investigated the state estimation for H_∞ control static neural network with two additive time delays. Time-varying delays often occur in engineering systems, network control, and biological network control. It is a factor that must be considered in the practice problems. In the problems of robust control and nonlinear asymptotically stable control, time-varying delays are factors that must be considered [23]. The difference between the leakage delays and the time-varying delays is that it will cause the instability of the system, and the time-varying delays can cause system delays [24].

There is less research in uncertain neural networks with two additive time-varying delays, in this paper, we investigate the H_∞ state estimation for uncertain neural networks with two additive time-varying delays. Based on LKFs method, a novel LMIs method has been established to ensure the global asymptotic stability of uncertain neural networks with two additive time-varying delays. Finally, two numerical simulation examples are used to illustrate the effectiveness of the proposed design method.

2. System model and Preliminaries

Considering the following uncertain neural networks with two additive time-varying delays:

$$\begin{cases} \dot{x}(t) = -A_1(t)x(t) + A_2(t)f(x(t)) + A_3(t)x(t - \tau_1(t) - \tau_2(t)) + C_1(t)w(t) + u(t) \\ y(t) = B_1(t)x(t) + B_2(t)x(t - \tau_1(t) - \tau_2(t)) + D_1(t)w(t) \\ z(t) = Hx(t) \\ x(t) = \varphi(t), t \in [-h, 0] \end{cases} \quad (1)$$

Where $x(t) \in [x_1(t), x_2(t), \dots, x_n(t)]^T \in \mathbb{R}^n$ is the state vector, $y(t)$ is the measurement, $u(t)$ is the control input, $z(t) \in \square$ to be estimated is a linear combination of the state, $w(t)$ is disturbance input and $w(t) \in L^2[0, \infty]$, and $A_1, A_2, A_3, C_1, B_1, B_2, D_1$ and H are known real constant matrices, $f(x(t)) = [f_1(x_1(t)), f_2(x_2(t)) \dots f_n(x_n(t))]^T$ denotes the neuron activation function and a constant input vector, and. $A_1(t) = A_1 + \Delta A_1, A_2(t) = A_2 + \Delta A_2, A_3(t) = A_3 + \Delta A_3, B_1(t) = B_1 + \Delta B_1, B_2(t) = B_2 + \Delta B_2, C_1(t) = C_1 + \Delta C_1, D_1(t) = D_1 + \Delta D_1, \Delta A_1, \Delta A_2, \Delta A_3, \Delta B_1, \Delta B_2, \Delta C_1$ and ΔD_1 is real matrix. These parameters represent the uncertainty of the system. $\varphi(t)$ is represents a given initial condition.

Assumption 1. The parameter uncertainties $\Delta A_1, \Delta A_2, \Delta A_3, \Delta B_1, \Delta B_2, \Delta C_1$ and ΔD_1 are of the form:

$$[\Delta A_1 \ \Delta A_2 \ \Delta A_3 \ \Delta B_1 \ \Delta B_2 \ \Delta C_1 \ \Delta D_1] = HF(t)[M_{11} \ M_{12} \ M_{21} \ M_{22} \ S_1 \ S_2 \ S_3] \quad (2)$$

Among these parameters, $H, M_{11}, M_{12}, M_{21}, M_{22}, S_1, S_2, P_1$ is known real matrix of appropriate dimension. $F(t) \in \mathbb{R}^{i \times j}, i > 0, j > 0$ is the real unknown time-varying matrix. So,

$$F(t)^T F(t) \leq I, \forall t \geq 0 \quad (3)$$

Assumption 2. The time-varying delays $\tau_1(t), \tau_2(t)$ satisfy

$$0 \leq \tau_1(t) \leq \tau_1 \leq \infty, 0 \leq \tau_2(t) \leq \tau_2 \leq \infty \quad (4)$$

$$\dot{\tau}_1(t) \leq \mu_1, \dot{\tau}_2(t) \leq \mu_2 \quad (5)$$

$$\tau = \tau_1 + \tau_2, \mu = \mu_1 + \mu_2 \quad (6)$$

where τ_1, τ_2 are positive constants.

Assumption 3 [25]. Each neuron activation function $f_i(\cdot), i = 1, 2, \dots, n$ satisfies the following condition:

$$0 \leq \frac{f_i(\alpha) - f_i(\beta)}{\alpha - \beta} \leq l_i, \forall \alpha, \beta \in \mathbb{R}, \alpha \neq \beta \quad (7)$$

where $l_i, i = 1, 2, \dots, n$ are constants, and diagonal matrix $L = \text{diag}\{l_i\}$. Then, constructing a state estimator for estimation of $Z(t)$ as follows

$$\begin{cases} \dot{\hat{x}}(t) = -A_1(t)\hat{x}(t) + A_2(t)f(\hat{x}(t)) + A_3(t)\hat{x}(t - \tau_1(t) - \tau_2(t)) \\ \quad + K(y(t) - B_1(t)\hat{x}(t) + B_2(t)\hat{x}(t - \tau_1(t) - \tau_2(t))) + u \\ \hat{z}(t) = H\hat{x}(t) \\ x(t) = 0, t \in [-h, 0] \end{cases} \quad (8)$$

where $\hat{x}(t) \in \mathbb{R}^n$ denotes the estimated state, $\hat{z}(t) \in \mathbb{R}^m$ denotes the estimated measurement of $z(t)$, and K is the state estimator gain matrix.

Denoting the errors by $e(t) = x(t) - \hat{x}(t)$ and $\tau(t) = \tau_1(t) + \tau_2(t)$, $\bar{z}(t) = z(t) - \hat{z}(t)$, $e(t - \tau_1(t) - \tau_2(t)) = x(t - \tau_1(t) - \tau_2(t)) - \hat{x}(t - \tau_1(t) - \tau_2(t))$

Then, based on (1) and (8), the error system of the form is obtained:

$$\begin{cases} \dot{e}(t) = -(A_1(t) + KB_1(t))e(t) + A_2(t)g(e(t)) + (A_3(t) - KB_2(t))e(t - \tau_1(t) - \tau_2(t)) \\ \quad + (C_1(t) - KD_1(t))w(t) \\ \bar{z}(t) = He(t) \end{cases} \quad (9)$$

Where $g(e(t)) = f(x(t)) - f(\hat{x}(t))$,

$$\begin{cases} \dot{e}(t) = -(A_1(t) + KB_1(t))e(t) + A_2(t)g(e(t)) + (A_3(t) - KB_2(t))e(t - \tau(t)) \\ \quad + (C_1(t) - KD_1(t))w(t) \\ \bar{z}(t) = He(t) \end{cases} \quad (10)$$

In this paper we will study the stability of (10) so that guaranteed its H_∞ performance. Moreover, it is proved by numerical simulation that the state estimation of the system error equation tends to zero, which proves that the system is asymptotically stable.

Definition 1 [25]. Given a prescribed level of noise attenuation $\gamma \geq 0$, a proper state estimator (8) is founded, the equilibrium point of the result error system (10) with $w(t) = 0$ is globally asymptotically stable, and

$$\|\bar{z}(t)\|_2 < \gamma \|w(t)\|_2 \quad (11)$$

under zero-initial conditions for all nonzero $w(t) \in L_2[0, \infty)$, where

$$\|x(t)\|_2 \square \sqrt{\int_0^\infty x^T(t)x(t)dt} \quad (12)$$

In this case, error system (10) is globally asymptotically stable with H_∞ performance γ .

3. Results and Proof

In this section, the global asymptotic stability of the model (10) with the initial condition (2)~(7) is discussed and the main results are given as follows:

Theorem 1. Considering the uncertain neural networks with two additive time-varying delays (1), If there is an appropriate dimension of the matrix Y , for given

scalars $h_1, h_2 > 0$, let γ to be a prescribed constant, then the guaranteed H^∞ performance state estimation problem is solvable, if there exist real matrices $Q_1, Q_2, Q_3, Q_4, Q_5, Q_6 > 0$ $T > 0, P_1 > 0, R_i > 0, S_i > 0 (i=1,2,3), M_{je} > 0 (j=1,2, e=1,2)$; positive diagonal matrices $P_2 = \text{diag}\{\eta_i\}$, so the following LMIs are satisfied:

$$\Omega = \begin{pmatrix} \Omega_{11} & \Omega_{12} & \Omega_{13} & \Omega_{14} & \Omega_{15} & \Omega_{16} & \Omega_{17} & \Omega_{18} & \Omega_{19} & \Omega_{10} \\ * & \Omega_{22} & \Omega_{23} & \Omega_{24} & \Omega_{25} & \Omega_{26} & \Omega_{27} & \Omega_{28} & \Omega_{29} & \Omega_{210} \\ * & * & \Omega_{33} & \Omega_{34} & \Omega_{35} & \Omega_{36} & \Omega_{37} & \Omega_{38} & \Omega_{39} & \Omega_{310} \\ * & * & * & \Omega_{44} & \Omega_{45} & \Omega_{46} & \Omega_{47} & \Omega_{48} & \Omega_{49} & \Omega_{410} \\ * & * & * & * & \Omega_{55} & 0 & 0 & 0 & 0 & 0 \\ * & * & * & * & * & \Omega_{66} & 0 & 0 & 0 & 0 \\ * & * & * & * & * & * & \Omega_{77} & 0 & 0 & 0 \\ * & * & * & * & * & * & * & \Omega_{88} & 0 & 0 \\ * & * & * & * & * & * & * & * & \Omega_{99} & 0 \\ * & * & * & * & * & * & * & * & * & \Omega_{1010} \end{pmatrix} < 0 \tag{13}$$

where

$$\begin{aligned} \Omega_{11} &= -A_1 P_1 + B_1 Y - A_1^T P_1^T + B_1^T Y^T + Q_1 + Q_2 + Q_3 + Q_4 + Q_5 + Q_6 + M_{11}^T + M_{11} + M_{12}^T + M_{12} - T + H^T H, \Omega_{1010} = -\gamma^2 I \\ \Omega_{12} &= A_3 P_1 - B_2 Y - A_1^T P_1^T + B_1^T Y^T + R_2 - T - M_{11}^T + M_{21}, \Omega_{13} = A_2 P_1 - A_1^T P_1^T + B_1^T Y^T - M_{12}^T + M_{22}, \Omega_{14} = R_2, \\ \Omega_{15} &= A_2 P_2 - A_1^T P_1^T + B_1^T Y^T, \Omega_{16} = \tau_1 (A_3^T P_1^T + B_1^T Y^T) - A_1^T P_1^T + B_1^T Y^T, \Omega_{17} = \tau_2 (A_3^T P_1^T + B_1^T Y^T) - A_1^T P_1^T + B_1^T Y^T, \Omega_{18} = \tau_1 S_1 \\ \Omega_{19} &= \tau_2 S_2, \Omega_{110} = P_1 C_1, \Omega_{22} = A_1 P_1 - (1-\tau) M_{11}^T + (1-\tau) M_{22}, \Omega_{23} = 0, \Omega_{24} = R_2 - T, \Omega_{25} = (A_1 P_1)^T, \Omega_{26} = \tau_1 (A_1 P_1)^T, \\ \Omega_{27} &= \tau_2 (A_1 P_1)^T, \Omega_{28} = \tau_2 S_3, \Omega_{29} = 0, \Omega_{210} = 0, \Omega_{33} = -Q_2 - (1-\tau) M_{11}^T + (1-\tau) M_{22}, \Omega_{34} = 0, \Omega_{35} = (B_1 P_1)^T, \\ \Omega_{36} &= \tau_1 (B_1 P_1)^T, \Omega_{37} = \tau_2 (B_1 P_1)^T, \Omega_{38} = 0, \Omega_{39} = \tau_2 R_3, \Omega_{310} = 0, \Omega_{44} = -P_1, \Omega_{45} = B_1^T, \Omega_{46} = B_2^T, \Omega_{47} = \tau_2 B_1^T, \\ \Omega_{48} &= \tau_2 B_2^T, \Omega_{49} = 0, \Omega_{410} = D_1, \Omega_{55} = -\tau R_3, \Omega_{66} = -\tau_1 R_1, \Omega_{77} = -\tau_2 R_2, \Omega_{88} = -\tau_1 R_1^T, \Omega_{99} = -\tau_2 R_2^T, \end{aligned}$$

where $K = P^{-1}Y$

Proof 1. constructing the following Lyapunov-Krasovskii functional:

$$V(t) = V_1(t) + V_2(t) + V_3(t) + V_4(t) \tag{14}$$

$$V_1(e_t) = e^T(t) P_1 e(t) + 2 \sum_{i=1}^n P_i \int_0^{e_i(t)} g(s) ds \tag{15}$$

$$\begin{aligned} V_2(e_t) &= \int_{t-\tau_1(t)}^t e^T(s) Q_1 e(s) ds + \int_{t-\tau_2(t)}^t e^T(s) Q_2 e(s) ds + \int_{t-\tau(t)}^t e^T(s) Q_3 e(s) ds \\ &+ \int_{t-\tau_1}^t e^T(s) Q_4 e(s) ds + \int_{t-\tau_2}^t e^T(s) Q_5 e(s) ds + \int_{t-\tau}^t e^T(s) Q_6 e(s) ds \end{aligned} \tag{16}$$

$$V_3(e_t) = \int_{t-\tau(t)}^t \begin{bmatrix} e(s) \\ g(e(s)) \end{bmatrix}^T \begin{bmatrix} R_1 & R_2 \\ * & R_3 \end{bmatrix} \begin{bmatrix} e(s) \\ g(e(s)) \end{bmatrix} ds, \tag{17}$$

$$V_4(e_t) = \tau_1 \int_{-1}^0 \int_{+\theta}^t e^T(s) R_1 e(s) ds d\theta + \tau \int_{-1}^0 \int_{+\theta}^t e^T(s) R_2 e(s) ds d\theta + \tau_2 \int_{-1}^0 \int_{+\theta}^t e^T(s) R_3 e(s) ds d\theta \tag{18}$$

Taking the time-derivative of $V(t)$ along the trajectories of yields. Robust stabilization of the system. In this case, error system (10) is globally asymptotically stable with H^∞ performance γ .

$$\begin{aligned} \dot{V}_1(e(t)) &= 2e^T(t)P_1\dot{e}(t) + 2g^T(e(t))P_2\dot{e}(t) \\ &= 2e^T(t)P_1[-(A_1(t) + KB_1(t))e(t) + A_2(t)g(e(t)) + (A_3(t) - KB_2(t))e(t - \tau(t)) \\ &\quad + (C_1(t) - KD_1(t))w(t)] + 2g^T(e(t))P_2[-(A_1(t) + KB_1(t))e(t) + A_2(t)g(e(t)) \\ &\quad + (A_3(t) - KB_2(t))e(t - \tau(t)) + (C_1(t) - KD_1(t))w(t)] \end{aligned} \tag{19}$$

$$\begin{aligned} &= -2e^T(t)P_1(A_1(t) + KB_1(t))e(t) + 2e^T(t)P_1(A_3(t) - KB_2(t))e(t - \tau(t)) + 2e^T(t)P_1A_2(t)g(e(t)) \\ &\quad + 2e^T(t)P_1(C_1(t) - KD_1(t))w(t) - 2g^T(e(t))P_2(A_1(t) + KB_1(t))e(t) + 2g^T(e(t))P_2(A_3(t) - KB_2(t))e(t - \tau(t)) \\ &\quad + 2g^T(e(t))P_2A_2(t)g(e(t)) + 2g^T(e(t))P_2(C_1(t) - KD_1(t))w(t) \end{aligned}$$

$$\begin{aligned} \dot{V}_2(e(t)) &\leq e^T(t)(Q_1 + Q_2 + Q_3 + Q_4 + Q_5 + Q_6)e(t) - e^T(t - \tau_1)Q_1e(t - \tau_1) \\ &\quad - e^T(t - \tau_2)Q_2e(t - \tau_2) - e^T(t - \tau)Q_3e(t - \tau) - (1 - \mu_1)e^T(t - \tau_1(t))Q_4e(t - \tau_1(t)) \\ &\quad - (1 - \mu_2)e^T(t - \tau_2(t))Q_5e(t - \tau_2(t)) - (1 - \mu)e^T(t - \tau(t))Q_6e(t - \tau(t)) \end{aligned} \tag{20}$$

$$\begin{aligned} \dot{V}_3(e(t)) &= \begin{bmatrix} e(t) \\ g(e(t)) \end{bmatrix}^T \begin{bmatrix} R_1 & R_2 \\ * & R_3 \end{bmatrix} \begin{bmatrix} e(t) \\ g(e(t)) \end{bmatrix} - (1 - \dot{\tau}(t)) \begin{bmatrix} e(t - \tau(t)) \\ g(e(t - \tau(t))) \end{bmatrix}^T \begin{bmatrix} R_1 & R_2 \\ * & R_3 \end{bmatrix} \begin{bmatrix} e(t - \tau(t)) \\ g(e(t - \tau(t))) \end{bmatrix} \\ &\leq e^T(t)R_1e(t) + g^T(e(t))R_3g(e(t)) + 2e^T(t)R_2g(e(t)) - (1 - \mu)e^T(t - \tau(t))R_1e(t - \tau(t)) \\ &\quad - (1 - \mu)g^T(e(t - \tau(t)))R_3g(e(t - \tau(t))) - 2(1 - \mu)e^T(t - \tau(t))R_2g(e(t - \tau(t))) \end{aligned} \tag{21}$$

$$\begin{aligned} \dot{V}_4(e(t)) &= \tau_1^2 \dot{e}^T(t)R_1\dot{e}(t) - \tau_1 \int_{t-\tau_1}^t \dot{e}^T(s)R_1\dot{e}(s)ds + \tau_2^2 \dot{e}^T(t)R_3\dot{e}(t) - \tau_2 \int_{t-\tau_2}^t \dot{e}^T(s)R_3\dot{e}(s)ds \\ &\quad + \tau^2 \dot{e}^T(t)R_2\dot{e}(t) - \tau \int_{t-\tau}^t \dot{e}^T(s)R_2\dot{e}(s)ds \end{aligned} \tag{22}$$

Among the equation of the $\dot{V}_4(e(t))$, Then $\dot{V}_4(e(t))$ is obtained through above formulas:

$$\begin{aligned} \dot{V}_4(e(t)) &\leq \tau_1^2 \dot{e}^T(t)R_1\dot{e}(t) - \tau_1 \int_{t-\tau_1}^t \dot{e}^T(s)R_1\dot{e}(s)ds + \tau_2^2 \dot{e}^T(t)R_3\dot{e}(t) - \tau_2 \int_{t-\tau_2}^t \dot{e}^T(s)R_3\dot{e}(s)ds \\ &\quad + \tau^2 \dot{e}^T(t)R_2\dot{e}(t) - \tau \int_{t-\tau}^t \dot{e}^T(s)R_2\dot{e}(s)ds \\ &= \tau^2 \dot{e}^T(t)R_2\dot{e}(t) + \alpha^T(t) \begin{bmatrix} R_2 & T \\ * & R_2 \end{bmatrix} \alpha(t) + \tau_2^2 \dot{e}^T(t)R_3\dot{e}(t) + \alpha^T(t) \begin{bmatrix} R_3 & T \\ * & R_3 \end{bmatrix} \alpha(t) \\ &\quad + \tau_1^2 \dot{e}^T(t)R_1\dot{e}(t) + \alpha^T(t) \begin{bmatrix} R_1 & T \\ * & R_1 \end{bmatrix} \alpha(t) \\ &= \xi^T(t)\Sigma_1\xi(t) + \xi^T(t)\Sigma_2\xi(t) + \xi^T(t)\Sigma_3\xi(t) = \xi^T(t)(\Sigma_1 + \Sigma_2 + \Sigma_3)\xi(t) \\ &\leq \xi^T(t)\Sigma_1\xi(t) \end{aligned} \tag{23}$$

$$\Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} & \Sigma_{13} & \Sigma_{14} & \Sigma_{15} & \Sigma_{16} & \Sigma_{17} & \Sigma_{18} & \Sigma_{19} & \Sigma_{110} \\ * & \Sigma_{22} & \Sigma_{23} & \Sigma_{24} & \Sigma_{25} & \Sigma_{26} & \Sigma_{27} & \Sigma_{28} & \Sigma_{29} & \Sigma_{210} \\ * & * & \Sigma_{33} & \Sigma_{34} & \Sigma_{35} & \Sigma_{36} & \Sigma_{37} & \Sigma_{38} & \Sigma_{39} & \Sigma_{310} \\ * & * & * & \Sigma_{44} & \Sigma_{45} & \Sigma_{46} & \Sigma_{47} & \Sigma_{48} & \Sigma_{49} & \Sigma_{410} \\ * & * & * & * & \Sigma_{55} & 0 & 0 & 0 & 0 & 0 \\ * & * & * & * & * & \Sigma_{66} & 0 & 0 & 0 & 0 \\ * & * & * & * & * & * & \Sigma_{77} & 0 & 0 & 0 \\ * & * & * & * & * & * & * & \Sigma_{88} & 0 & 0 \\ * & * & * & * & * & * & * & * & \Sigma_{99} & 0 \\ * & * & * & * & * & * & * & * & * & \Sigma_{1010} \end{bmatrix} < 0 \tag{24}$$

Where

$$\begin{aligned} \Sigma_{11} &= -A_1P_1 + B_1Y - A_1^T P_1^T + B_1^T Y^T + Q_1 + Q_2 + Q_3 + Q_4 + Q_5 + Q_6 + M_{11}^T + M_{11} + M_{12}^T + M_{12} - T + H^T H \\ \Sigma_{12} &= A_3P_1 - KB_2Y - A_1^T P_1^T + B_1^T Y^T + R_2 - T - M_{11}^T + M_{21} \end{aligned}$$

$$\begin{aligned}
 &\Sigma_{13} = A_2 P_1 - A_1^T P_1^T + B_1^T Y^T - M_{12}^T + M_{22} \quad , \quad \Sigma_{14} = R_2 \quad , \quad \Sigma_{15} = (A_3 P_1 + B_1 Y)^T - A_1^T P_1^T + B_1^T Y^T \quad , \\
 &\Sigma_{16} = \tau_1 (A_3 P_1 + B_1 Y)^T - A_1^T P_1^T + B_1^T Y^T \quad \Sigma_{17} = \tau_2 (A_3 P_1 + B_1 Y)^T - A_1^T P_1^T + B_1^T Y^T \quad , \quad \Sigma_{18} = \tau_1 S_1 \quad \Sigma_{19} = \tau_2 S_2 \quad , \quad \Sigma_{110} = P_1 C_1 \quad , \\
 &\Sigma_{22} = A_1 P_1 - (1-\tau) M_{11}^T + (1-\tau) M_{22} \quad , \quad \Sigma_{23} = 0 \quad , \quad \Sigma_{24} = R_2 - T \quad \Sigma_{25} = (A_1 P_1)^T \quad , \quad \Sigma_{26} = \tau_1 (A_1 P_1)^T \quad \Sigma_{27} = \tau_2 (A_1 P_1)^T \quad , \\
 &\Sigma_{28} = \tau_2 S_3 \quad , \quad \Sigma_{29} = 0 \quad , \quad \Sigma_{210} = 0 \quad , \quad \Sigma_{33} = -Q_2 - (1-\tau) M_{22}^T - (1-\tau) M_{22} \quad , \quad \Sigma_{34} = 0 \quad , \quad \Sigma_{35} = (B_1 P_1)^T \\
 &\Sigma_{36} = \tau_1 (B_1 P_1)^T \quad , \quad \Sigma_{37} = \tau_2 (B_1 P_1)^T \quad , \quad \Sigma_{38} = 0 \quad , \quad \Sigma_{39} = \tau_2 R_3 \quad , \quad \Sigma_{310} = 0 \quad , \quad \Sigma_{44} = -P_1 \quad , \quad \Sigma_{45} = B_1^T \quad , \quad \Sigma_{46} = B_2^T \quad , \quad \Sigma_{47} = \tau_2 B_1^T \quad , \\
 &\Sigma_{48} = \tau_2 B_2^T \quad , \quad \Sigma_{49} = 0 \quad , \quad \Sigma_{410} = D_1 \quad , \quad \Sigma_{55} = -\tau R_3 \quad \Sigma_{66} = -\tau R_1 \quad , \quad \Sigma_{77} = -\tau_2 R_2 \quad , \quad \Sigma_{88} = -\tau_1 R_1^T \quad , \quad \Sigma_{99} = -\tau_2 R_2^T \quad , \\
 &\Sigma_{1010} = -\gamma^2 I \quad , \quad \alpha(t) = [\alpha_1^T(t), \alpha_2^T(t), \alpha_3^T(t), \alpha_4^T(t)]^T
 \end{aligned}$$

$$\alpha_1(t) = e(t - \tau(t)) - e(t - \tau), \alpha_2(t) = e(t - \tau(t)) - e(t - \tau) - \frac{2}{t - \tau(t)} \int_{t-\tau}^{t-\tau(t)} e(s) ds$$

$$\alpha_3(t) = e(t) - e(t - \tau(t)), \alpha_4(t) = e(t) - e(t - \tau(t)) - \frac{2}{\tau(t)} \int_{t-\tau(t)}^t e(s) ds$$

$$\xi(t) = [e^T(t), e^T(t - \tau(t)), e^T(t - \tau), e^T(t - \tau_1(t)), e^T(t - \tau_1), e^T(t - \tau_2(t)), e^T(t - \tau_2), g^T(e(t)), \hat{x}(t - \tau(t)), w^T(t)]^T$$

Combining the above results, $\dot{V}(e(t))$ can be obtained:

$$\begin{aligned}
 \dot{V}(e(t)) &\leq -e^T(t) [P_1(A_1(t) + KB_1(t)) + (A_1(t) + KB_1(t))^T P_1] e(t) + 2e^T(t) P_1(A_3(t) - KB_2(t)) e(t - \tau_1(t) - \tau_2(t)) \\
 &\quad + 2e^T(t) P_1 A_2(t) g(e(t)) + 2e^T(t) P_1(C(t) - KD(t)) w(t) + 2g^T(e(t)) \Lambda \dot{e}(t) + e^T(t) (Q_1 + Q_2 + Q_3) e(t) \quad (25) \\
 &\quad - e^T(t - \tau_1) Q_1 e(t - \tau_1) - e^T(t - \tau_2) Q_2 e(t - \tau_2) - e^T(t - \tau) Q_3 e(t - \tau) - (1 - \mu_1) e^T(t - \tau_1(t)) Q_1 e(t - \tau_1(t)) \\
 &\quad - (1 - \mu_2) e^T(t - \tau_2(t)) Q_2 e(t - \tau_2(t)) - (1 - \mu) e^T(t - \tau(t)) Q_3 e(t - \tau(t)) + e^T(t) R_1 e(t) + g^T(e(s)) R_3 g(e(s)) \\
 &\quad + 2e^T(t) R_2 g(e(s)) - (1 - \mu) e^T(t - \tau(t)) R_1 e(t - \tau(t)) - (1 - \mu) g^T(e(t - \tau(t))) R_3 g(e(t - \tau(t))) \\
 &\quad - 2(1 - \mu) e^T(t - \tau(t)) R_2 g(e(t - \tau(t))) + \xi^T(t) (\Sigma_1 + \Sigma_2 + \Sigma_3) \xi(t) \\
 &\leq \xi^T(t) (\Sigma_1 + \Sigma_2 + \Sigma_3) \xi(t) \\
 &\leq \xi^T(t) \Sigma \xi(t) < 0
 \end{aligned}$$

Therefore, if LMIs is to be work, then $\dot{V}(e(t)) < 0$. The neural networks (1) is asymptotically stable. This completes the Proof 1.

Since the function $f(x(t))$ satisfy (7). Then, for any $e(t) \neq 0$.

$$0 \leq \frac{g_1(e(t), \hat{x}(t))}{e(t)} = \frac{f(x(t)) - f(\hat{x}(t))}{x(t) - \hat{x}(t)} \leq l_i \quad (26)$$

Under the zero-initial condition, it is obvious that $V(e(t))|_{t=0} = 0$. For dealing easily. Let

$$J_\infty = \int_0^t [\bar{z}^T(s) \bar{z}(s) - \gamma^2 w^T(s) w(s)] ds, t > 0 \quad (27)$$

Then,

$$J_\infty \leq \int_0^t [\bar{z}^T(s) \bar{z}(s) - \gamma^2 w^T(s) w(s)] ds + V(e(t)) - V(e(t))|_{t=0}, t > 0 \quad (28)$$

Then for any $w(t) \in L^2[0, \infty]$,

$$J_\infty \leq \int_0^t p[\bar{z}^T(s) \bar{z}(s) - \gamma^2 w^T(s) w(s) + V(\dot{e}(s))] ds \quad (29)$$

Based on the equalities and inequalities, we can deduce that:

$$\bar{z}^T(t)\bar{z}(t) - w^T(t)w(t) + \dot{V}(e(t)) \leq \xi^T(t) [\Sigma_1 + \tau^2 \Sigma_2^T R_2 \Sigma_2] \xi(t) \tag{30}$$

So if $\Sigma_1 + \tau^2 \Sigma_2^T R_2 \Sigma_2 < 0$, then there must exist a sufficiently small scalar σ , so $\Sigma_1 + \tau^2 \Sigma_2^T R_2 \Sigma_2 + \sigma I \leq 0$. Then, it is easy to obtain

that for any $w(t) \neq 0$,

$$\begin{aligned} \bar{z}^T(t)\bar{z}(t) - w^T(t)w(t) + \dot{V}(e(t)) &\leq \xi^T(t) [\Sigma_1 + \tau^2 \Sigma_2^T R_2 \Sigma_2] \xi(t) \\ &\leq -\sigma \xi^T(t) \xi(t) \leq -\sigma w^T(t)w(t) < 0 \end{aligned} \tag{31}$$

Proof 2. The system is robustly, asymptotically stable if the following inequalities is satisfied:

$$\Omega + \varphi_1 F(t) \varphi_2^T + \varphi_2 F(t) \varphi_1^T + \varphi_3 F(t) \varphi_4^T + \varphi_4 F(t) \varphi_3^T + \varphi_5 F(t) \varphi_6^T + \varphi_6 F(t) \varphi_5^T < 0 \tag{32}$$

Where $\varphi_1 = [H^T P_1 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0]^T$, $\varphi_2 = [0 \ H^T R_3]^T$,
 $\varphi_3 = [M_{11} \ 0 \ M_{12} \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0]^T$, $\varphi_4 = [M_{21} \ 0 \ M_{22} \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0]^T$,
 $\varphi_5 = [S_1 \ 0 \ S_2 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0]^T$, $\varphi_6 = [S_3 \ 0]^T$

If (13) is satisfied, then the following inequalities:

$$\begin{aligned} \Omega + \varepsilon_1^{-1} \varphi_1 \varphi_1^T + \varepsilon_1 \varphi_2 \varphi_2^T + \varepsilon_2^{-1} \varphi_3 \varphi_3^T + \varepsilon_2 \varphi_4 \varphi_4^T + \varepsilon_3^{-1} \varphi_5 \varphi_5^T + \varepsilon_3 \varphi_6 \varphi_6^T &\equiv \Omega + \Psi \\ = \Omega + \begin{bmatrix} \varepsilon_1^{-1} P_1^T H H^T P_1 + \varepsilon_2^{-1} M_{11} M_{11}^T & 0 \\ + \varepsilon_1 M_{21} M_{21}^T + \varepsilon_2^{-1} S_1 S_1^T + \varepsilon_3 S_3 S_3^T & 0 \\ 0 & 0 & \varepsilon_2^{-1} M_{12} M_{12}^T + \varepsilon_3 M_{22} M_{22}^T & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ + \varepsilon_3^{-1} S_2 S_2^T & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ \varepsilon_1 R_1 H^T H R_1^T \end{bmatrix} \end{aligned} \tag{33}$$

where $\varepsilon_1 > 0, \varepsilon_2 > 0, \varepsilon_3 > 0$.

Then, the inequalities given in (33) is equivalent to the LMIs (25). Thus, if the LMIs given in (25) is satisfied, then the system (10) is robust asymptotically stable. This completes the Proof 2.

Corollary 1 Considering the neural networks with two additive time-varying delays system (1), for given scalars $\mu < 1$ and K , let γ be a prescribed constant, the guaranteed H^∞ performance state estimation problem is solvable if there exist real matrices $Y > 0$, $M_{je} > 0 (j=1,2, e=1,2)$; and diagonal matrices $P_2 = \text{diag}\{\eta_i\}$ with appropriate dimensions, then the following LMIs are satisfied:

$$\Phi = \begin{bmatrix} \Phi_{11} & A_3 Y & & \Phi_{13} & P_1 & P_1 B_1 - Y B_2 & H^T \\ & -B_2 Y & & & & & \\ * & -(1-\mu)M_{11} & 0 & 0 & 0 & 0 & 0 \\ * & * & -(1-\mu)M_{12} & \Phi_{34} & 0 & 0 & 0 \\ * & * & * & -(1-\mu)M_{21} & 0 & 0 & 0 \\ * & * & * & * & -(1-\mu)M_{22} & 0 & 0 \\ * & * & * & * & * & * & -\gamma I \end{bmatrix} \tag{34}$$

$$\Phi_{11} = -P_1 A_1 - A_1^T P_1^T - B_1 Y - Y^T B_1^T + M_{11} + M_{12} + M_{21} + M_{22}$$

$$\Phi_{13} = -(1-\mu)(M_{11} + M_{12}) - 2P_2, \Phi_{34} = -(1-\mu)(M_{21} + M_{22}) - 2P_2$$

Moreover, the gain matrix K of the state estimator of (8) can be designed as $K = P_1^{-1} Y$.

Remark 1. Theorem 1 provides a novel robust H_∞ state estimation of uncertain neural networks stability criterion for system (1) with two additive time-varying delays components, it has been verified by a form of less complex LMIs.

Remark 2. Based on the Lyapunov stability criterion, a novel LMIs is constructed to prove that the derivative of the LKFs is smaller than zero, and that the system (1) is asymptotically stable.

Remark 3. The state estimation uncertain neural networks methods are listed in table 1.

Table 1. state estimation of uncertain neural network methods.

numbers	1	2	3	4	5	6	7
methods	H_∞ State estimation [22]	Robust finite-time state estimation [21]	State estimation [26]	Extended dissipative state estimation [27]	Delay dependent state estimation [28]	Non-fragile state estimation [29]	Recursive state estimation [30]

Compared with other theories, the H_∞ theory can give delay-dependent criteria, so that the error system has globally asymptotic stability H_∞ performance.

Remark 4. The structure of paper is organized as follows: In Section 2, problem model are given. In Section 3, a new theorem and three corollaries are established. In Section 4, two simulation results are provided to demonstrate the effectiveness of the developed approach. Finally, Section 5 summarizes this work.

4. Numerical Examples

Two numerical simulation examples are to be presented to show the feasibility of the developed approach.

Example1. Considering the system (10) with following parameters:

$$A_1 = \begin{bmatrix} -1 & 0 \\ -1 & -1 \end{bmatrix}, A_2 = \begin{bmatrix} -1 & 0 \\ -1 & -1 \end{bmatrix}, A_3 = \begin{bmatrix} -2 & 0 \\ 0 & -1 \end{bmatrix}, B_1 = \begin{bmatrix} 0.1 & 0 \\ 0 & 0.1 \end{bmatrix}, B_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, C = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, D = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, x_{delay} = [0 \ 0 \ 0.2],$$

$$x = \begin{bmatrix} 0.2 \\ 0.3 \\ 1 \end{bmatrix}, \tau_1 = 0.15, \tau_2 = 0.24, t = 0, dt = 0.001.$$

In addition, the activation function is chosen as $f(x) = \tanh(x)$, the time-varying delay by $\tau(t) = 0$. It is easy to get $K = \begin{bmatrix} -4.0107 & -1.7896 \\ -1.7896 & -4.6767 \end{bmatrix}$, $\tau = 0.39$. the noise disturbance is assumed be

$$w(t) = \frac{1}{1.2t + 0.8}.$$

$$P_1 = \begin{bmatrix} 0.6749 & -0.0197 \\ -0.0197 & 0.5245 \end{bmatrix}, P_2 = \begin{bmatrix} 0.1972 & 0 \\ 0 & 0.1972 \end{bmatrix}, Y = \begin{bmatrix} -4.0107 & -1.7896 \\ -1.7896 & -4.6767 \end{bmatrix}, S_1 = \begin{bmatrix} -2.8839 & 0.0969 \\ 0.0969 & -2.4532 \end{bmatrix}, S_2 = \begin{bmatrix} -0.0163 & 0.0291 \\ 0.0291 & 0.0200 \end{bmatrix},$$

$$S_3 = \begin{bmatrix} 0.0250 & -0.0874 \\ -0.0874 & -0.0408 \end{bmatrix}, Q_1 = \begin{bmatrix} 581.417 & 58.2298 \\ 58.2298 & -436.2641 \end{bmatrix},$$

$$Q_2 = \begin{bmatrix} 0.8732 & 0.1111 \\ 0.1111 & 1.0659 \end{bmatrix}, Q_3 = Q_4 = Q_5 = Q_6 = \begin{bmatrix} -436.4433 & -57.6238 \\ -57.6238 & 64.8827 \end{bmatrix}, M_{11} = \begin{bmatrix} 282.9611 & 44.2305 \\ 44.2305 & 37.1869 \end{bmatrix}, M_{12} = \begin{bmatrix} 300.9950 & 43.6859 \\ 43.6859 & 53.3194 \end{bmatrix},$$

$$M_{21} = \begin{bmatrix} 280.9950 & 42.1068 \\ 42.1068 & 33.6494 \end{bmatrix}, M_{22} = \begin{bmatrix} 300.6336 & 43.8862 \\ 43.8862 & 53.7267 \end{bmatrix}$$

With the option H^∞ performance index $\gamma = 1.3811$. Fig 1 shows the responses of the estimation error curves which generated by random initial value, it confirms the feasibility of the developed LMIs method through the designed state estimator of uncertain neural networks with two additive time-varying delays.

By applying the MATLAB LMI toolbox, it is found that LMIs (13) is feasible. As Fig.1 illustrates, the initial matrix of the example1 is a two-dimensional matrix, there are four estimation error curves. The state estimation tends to 0 rapidly. Therefore, it is proved that uncertain neural networks with two additive time-varying delays is globally asymptotically stable through proposed LMIs. The calculated minimum H^∞ performance index γ with different $\mu < 1$ values are listed in Table 2. One can know clearly that the results obtained by Corollary 1 can provide smaller H^∞ performance index γ than recently existing method in [1].

Table 2. Minimum H^∞ performance index γ with different μ .

μ	0.4	0.5	0.7
Reference [1]	0.4632	0.5301	1.3819
Corollary 1	0.0035	0.1544	0.4402

Example2. Considering a delayed neural network of Corollary 1 with following parameters:

$$A_1 = \begin{bmatrix} -1 & 1 & 0 \\ 2 & 1 & -1 \\ 1 & -1 & 2 \end{bmatrix}, B_1 = \begin{bmatrix} -1 & 1 & 0 \\ -1 & 1 & -1 \\ -1 & -1 & 1 \end{bmatrix}, B_2 = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & -1 \\ -1 & 0 & 1 \end{bmatrix}, A_3 = \begin{bmatrix} -1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & -1 & -1 \end{bmatrix}, J = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \gamma = 0.5, \mu = 0.5$$

In addition, the activation function is chosen as $f(x) = \tanh(x)$, the time-varying delay by $\tau(t) = 0$. It is easy to get

$$K = 10^{-9} * \begin{bmatrix} 0.0120 & 0.0394 & 0.0110 \\ 0.0394 & 0.2970 & -0.1812 \\ 0.0110 & -0.1812 & 0.2188 \end{bmatrix}, P_1 = 10^{-9} * \begin{bmatrix} -0.0245 & 0.0744 & 0.0188 \\ 0.0744 & 0.2718 & -0.1799 \\ 0.0188 & -0.1799 & 0.2296 \end{bmatrix}, P_2 = 10^{-10} * \begin{bmatrix} -0.1238 & 0 & 0 \\ 0 & -0.1238 & 0 \\ 0 & 0 & -0.1238 \end{bmatrix},$$

$$Y = 10^{-9} * \begin{bmatrix} 0.0120 & 0.0394 & 0.0110 \\ 0.0394 & 0.2970 & -0.1812 \\ 0.0110 & -0.1812 & 0.2188 \end{bmatrix}, H = 10^{-11} * \begin{bmatrix} 0.6127 & 0 & 0 \\ 0 & 0.6127 & 0 \\ 0 & 0 & 0.6127 \end{bmatrix}, M_{11} = 10^{-9} * \begin{bmatrix} 0.0244 & -0.0197 & -0.0132 \\ -0.0197 & 0.0939 & -0.1353 \\ -0.0132 & -0.1353 & 0.1524 \end{bmatrix},$$

$$M_{12} = 10^{-9} * \begin{bmatrix} 0.0400 & 0.0266 & -0.0520 \\ 0.0266 & 0.1108 & -0.1033 \\ -0.0520 & -0.1033 & 0.1156 \end{bmatrix}, M_{21} = 10^{-9} * \begin{bmatrix} 0.0727 & 0.0332 & -0.0076 \\ 0.0332 & 0.1481 & -0.0368 \\ -0.0076 & -0.0368 & 0.1161 \end{bmatrix}, M_{22} = 10^{-10} * \begin{bmatrix} 0.1040 & -0.1395 & -0.1606 \\ -0.1395 & -0.3238 & 0.0840 \\ -0.1606 & 0.0840 & -0.0893 \end{bmatrix}.$$

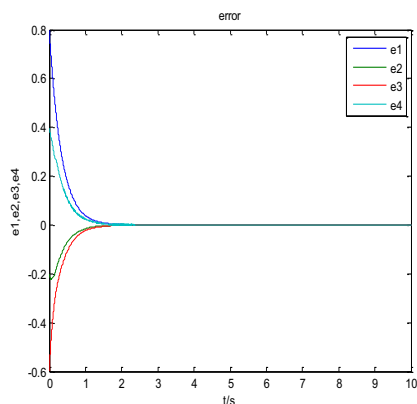


Figure 1. Estimation errors e_1, e_2, e_3 and e_4

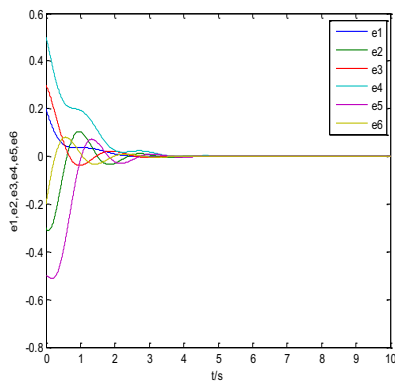


Figure 2. Estimation errors e_1, e_2, e_3, e_4, e_5 and e_6

The global asymptotically stable simulation results of the system (10) with the above parameters is illustrated in Fig.2.

the initial matrix of the example 2 is a three-dimensional matrix, there are six estimation error curves.

5. Conclusions

In this paper, the problem of H_∞ state estimation of uncertain neural networks with two additive time-varying delays has been studied. Based on LKFs method, a novel LMIs method has been established to ensure the global asymptotic stability of uncertain neural networks with two additive time-varying delays. It has fine convergence speed through constructed LMIs. Two numerical simulation examples have been performed to demonstrate the feasibility of the developed approach.

We would like to point out that this work did not include the H_∞ state estimation of uncertain neural networks with leaking time delay.

Acknowledgment

This work is supported by National Natural Science Foundation of China (No. 51875195) and National Natural Science Foundation of China (No. 61572185).

References

- [1] S. Arik, "Global asymptotic stability of a class of dynamical neural networks," *IEEE Transactions on circuits and systems I: Fundamental theory and applications*, vol. 47, no. 4, pp. 568-571, 2000.
- [2] D. Xiao, X. Li, and K. He, "Power Balance of Starting Process for Pipe Belt Conveyor Based on Master-Slave Control," *IEEE Access*, vol. 6, pp. 16924-16931, 2018.
- [3] H. Wu, X. Liao, W. Feng, S. Guo, and W. Zhang, "Robust stability analysis of uncertain systems with two additive time-varying delay components," *Applied Mathematical Modelling*, vol. 33, no. 12, pp. 4345-4353, 2009.
- [4] Y. Liu, S.-M. Lee, and H. Lee, "Robust delay-depnt stability criteria for uncertain neural networks with two additive time-varying delay components," *Neurocomputing*, vol. 151, pp. 770-775, 2015.
- [5] L. Xiong, J. Cheng, J. Cao, and Z. Liu, "Novel inequality with application to improve the stability criterion for dynamical systems with two additive time-varying delays," *Applied Mathematics and Computation*, vol. 321, pp. 672-688, 2018.

- [6] H.-T. Xu, C.-K. Zhang, L. Jiang, and J. Smith, "Stability analysis of linear systems with two additive time-varying delays via delay-product-type Lyapunov functional," *Applied Mathematical Modelling*, vol. 45, pp. 955-964, 2017.
- [7] K. Subramanian, P. Muthukumar, and S. Lakshmanan, "Robust stabilization of uncertain neural networks with additive time-varying delays*," *IFAC-PapersOnLine*, vol. 49, no. 1, pp. 154-159, 2016.
- [8] R. Samidurai and R. Sriraman, "Robust dissipativity analysis for uncertain neural networks with additive time-varying delays and general activation functions," *Mathematics and Computers in Simulation*, vol. 155, pp. 201-216, 2019.
- [9] J. Liang, K. Li, Q. Song, Z. Zhao, Y. Liu, and F. E. Alsaadi, "State estimation of complex-valued neural networks with two additive time-varying delays," *Neurocomputing*, 2018.
- [10] Y. Yuan, Q. Song, Y. Liu, and F. E. Alsaadi, "Synchronization of complex-valued neural networks with mixed two additive time-varying delays," *Neurocomputing*, 2018.
- [11] H. Bao, J. Cao, J. Kurths, A. Alsaedi, and B. Ahmad, " H^∞ state estimation of stochastic memristor-based neural networks with time-varying delays," *Neural Networks*, vol. 99, pp. 79-91, 2018.
- [12] N. Hou, H. Dong, Z. Wang, W. Ren, and F. E. Alsaadi, " H^∞ state estimation for discrete-time neural networks with distributed delays and randomly occurring uncertainties through fading channels," *Neural Networks*, vol. 89, pp. 61-73, 2017.
- [13] B. Liu, X. Ma, and X.-C. Jia, "Further results on H^∞ state estimation of static neural networks with time-varying delay," *Neurocomputing*, vol. 285, pp. 133-140, 2018.
- [14] H. Shen, M. Xing, S. Huo, Z.-G. Wu, and J. H. Park, "Finite-time H^∞ asynchronous state estimation for discrete-time fuzzy Markov jump neural networks with uncertain measurements," *Fuzzy Sets and Systems*, vol. 356, pp. 113-128, 2019.
- [15] S. Zhang, D. Ding, G. Wei, Y. Liu, and F. E. Alsaadi, " H^∞ state estimation for artificial neural networks over redundant channels," *Neurocomputing*, vol. 226, pp. 117-125, 2017.
- [16] Z. Zhao, Z. Wang, L. Zou, and H. Liu, "Finite-horizon H^∞ state estimation for artificial neural networks with component-based distributed delays and stochastic protocol," *Neurocomputing*, vol. 321, pp. 169-177, 2018.
- [17] Y. Liu, S.-M. Lee, O. Kwon, and J. H. Park, "A study on H^∞ state estimation of static neural networks with time-varying delays," *Applied Mathematics and Computation*, vol. 226, pp. 589-597, 2014.
- [18] H. Shao, X. Zhu, Z. Zhang, and G. Zhong, "Delay-dependent H^∞ control for systems with two additive time-vary delays," *Journal of Control Engineering and Technology*, vol. 3, no. 2, pp. 90-97, 2013.
- [19] S. Zhou and W. X. Zheng, "Robust H^∞ control of delayed singular systems with linear fractional parametric uncertainties," *Journal of the Franklin Institute*, vol. 346, no. 2, pp. 147-158, 2009.
- [20] F. Yang, H. Dong, Z. Wang, W. Ren, and F. E. Alsaadi, "A new approach to non-fragile state estimation for continuous neural networks with time-delays," *Neurocomputing*, vol. 197, pp. 205-211, 2016.
- [21] X. Zhou and Q. Zhu, "Robust finite-time state estimation for uncertain discrete-time Markovian jump neural networks with two delay components," *Neurocomputing*, vol. 283, pp. 64-72, 2018.
- [22] Q. Duan, H. Su, and Z.-G. Wu, " H^∞ state estimation of static neural networks with time-varying delay," *Neurocomputing*, vol. 97, pp. 16-21, 2012.
- [23] R. Saravanakumar, M. S. Ali, J. Cao, and H. He, " H^∞ state estimation of generalised neural networks with interval time-varying delays," *International Journal of Systems Science*, vol. 47, no. 16, pp. 3888-3899, 2016.
- [24] H. Zhu, R. Rakkiyappan, and X. Li, "Delayed state-feedback control for stabilization of neural networks with leakage delay," *Neural Networks*, vol. 105, pp. 249-255, 2018.
- [25] H. He, F. Gang, and J. Cao, "Guaranteed performance state estimation of static neural networks with time-varying delay," *Neurocomputing*, vol. 74, no. 4, pp. 606-616, 2011.
- [26] L. Qian, Q. Zhu, S. Zhong, X. Wang, and J. Cheng, "State estimation for uncertain Markovian jump neural networks with mixed delays," *Neurocomputing*, vol. 182, no. C, pp. 82-93, 2016.
- [27] Q. Li, Q. Zhu, S. Zhong, and F. Zhong, "Extended dissipative state estimation for uncertain discrete-time Markov jump neural networks with mixed time delays," *Isa Transactions*, vol. 66, pp. 200-208, 2017.
- [28] H. J. Yu, H. Yong, and W. Min, "Delay-Dependent State Estimation for Neural Networks with Time-Varying Delay," *Neurocomputing*, vol. 275, p. S0925231217315448, 2018.
- [29] R. Li, J. Cao, A. Alsaedi, and T. Hayat, "Non-fragile state observation for delayed memristive neural networks," *Neurocomputing*, vol. 245, no. C, pp. 102-113, 2017.
- [30] J. Hu, Z. Wang, S. Liu, and H. Gao, "A variance-constrained approach to recursive state estimation for time-varying complex networks with missing measurements," *Automatica*, vol. 64, no. C, pp. 155-162, 2016.

Dynamic Space-Covered Broadcast Algorithm Based on Neighbor-Degree in Mobile Ad Hoc Wireless Networks

Qiaoqiao LOU^{a,1} and Zhijin ZHAO^b

^a*School of Telecommunication Engineering of Hangzhou Dianzi University, China*

^b*State Key Lab of Information Control Technology in Communication System of No.36 Research Institute, China Electronic Technology Corporation, China*

Abstract. Broadcast is an important link in mobile ad hoc wireless networks (MANETs). In order to improve broadcast coverage, reduce forwarding probability and broadcast collision, a dynamic space-covered broadcast algorithm based on neighbor-degree in MANETs is proposed. The concepts of neighbor-degree, available-angle and available-distance are introduced. The neighbor-degree is used to generate the initial forwarding probability of a node. Based on the available-angle and available-distance, the node weight is proposed to calculate the final forwarding probability, thereby realizing the dynamic selection of forwarding node and reflecting the dynamic space-covered of node. The forwarding strategy is proposed to reduce broadcast collision. Simulation results show that, compared with dynamic probability broadcast and node location based space-covered broadcast, proposed broadcast algorithm reduces the broadcast collision, improves the broadcast coverage, and reduces the forwarding probability when the number of nodes is large.

Keywords. MANETs, broadcast coverage, forwarding probability, broadcast collision, forwarding strategy

1. Introduction

Broadcast [1, 2] is an important link in MANETs (mobile ad hoc wireless networks), which controls the information exchange between nodes. Compared with the broadcast with infrastructure, the broadcast in MANETs is more likely to generate broadcast storm [3] and unreliable broadcast. The specific manifestations are information redundancy, channel contention and signal collision, which seriously affect network performance. Therefore, many scholars have researched the broadcast algorithm [4-6] in MANETs.

Flooding is the simplest approach to broadcast, which ensures a complete coverage but may cause serious redundancy in dense networks, and even broadcast storm. In order to suppress this phenomenon, there are three mainly broadcast algorithms: neighbor-based, probability-based [7], and area-based [8]. The simplest broadcast among neighbor-based is self-pruning algorithm [9]. This algorithm selects forwarding nodes based on whether to cover new nodes, regardless of node density, and reduces redundancy generally. Probability-based broadcast algorithm is easy to implement, and

¹ Corresponding Author, School of Telecommunication Engineering of Hangzhou Dianzi University, Hangzhou 310018, China; Email:781398337@qq.com

optimizes by compromising overhead and coverage. References [10, 11] combined with the self-pruning algorithm to obtain a new dynamic probability broadcast algorithm. The forwarding probability of node is dynamically adjusted with the distance between the sending node and the receiving node, the node density and the transmission range. However, the forwarding strategy of this algorithm results in low broadcast coverage when the number of nodes is small and high redundancy when the number of nodes is large. The area-based broadcasting algorithm considers the additional coverage area of the forwarding node, aiming to reduce redundancy and obtain reliable coverage. References [12, 13] proposed a space-covered broadcast algorithm, which divides the coverage of sending nodes into three equal parts, corresponding to three ideal positions. Each area only forwards one node closest to the ideal position in the area. By optimizing the spatial distribution of the forwarding node, the redundancy can be reduced and the broadcast coverage can be ensured, but the broadcast delay is too high. Aiming at the problems of the above algorithms, a dynamic space-covered broadcast algorithm based on neighbor-degree in MANETs is proposed to improve the broadcast coverage when the number of nodes is small and reduce the forwarding probability when the number of nodes is large by optimizing spatial distribution.

2. Dynamic Space-Covered Broadcast Algorithm Bases on Neighbor-Degree

Assume that the rectangular coverage area of the network is $10 \times 10 \text{ km}^2$, where N nodes are evenly distributed, and the transmission radius of each node is r . In the following, the "sending node" is the node that sends the broadcast packet, which can be divided into "source node" (the first node that sends the broadcast packet) and "non-source node"; the "receiving node" is the node that receives the broadcast packet; the "parent node" is the sending node which receives the broadcast packet for the first time.

We proposed dynamic space-covered broadcast algorithm based on neighbor-degree, which includes forwarding probability calculation and forwarding strategy. Assume that the source node is uniquely determined in each broadcast; each node knows the information of all its two-hop neighbors; each node is equipped with a GPS positioning system, which can obtain its own position information.

2.1. Ideal Coverage Model

Assume that the nodes in the network are ideally distributed [13]. Considering each vertex as the sending node, the disk is the transmission range of the sending node. Each sending node is connected to three other sending nodes on the circle. When the sending node is the source node and its coordinate is (x_1, y_1) , the coordinates of the forwarding nodes on the corresponding circle (denoted as $Ls_i, i = 1, 2, 3$) [13] are respectively,

$$Ls_1 = (x_1 + r, y_1), Ls_2 = (x_1 - \frac{r}{2}, y_1 + \frac{\sqrt{3}r}{2}), Ls_3 = (x_1 - \frac{r}{2}, y_1 - \frac{\sqrt{3}r}{2}) \quad (1)$$

When the sending node is a non-source node, assuming the coordinate of the parent node is (x_0, y_0) , the coordinates of the forwarding nodes on the corresponding circle (denoted as $Lns_i, i = 1, 2$) [13] are respectively,

$$\begin{cases} Lns_{-1} = (\frac{3}{2}x_1 - \frac{1}{2}x_0 + \frac{\sqrt{3}}{2}y_1 - \frac{\sqrt{3}}{2}y_0, \frac{3}{2}y_1 - \frac{1}{2}y_0 - \frac{\sqrt{3}}{2}x_1 + \frac{\sqrt{3}}{2}x_0) \\ Lns_{-2} = (\frac{3}{2}x_1 - \frac{1}{2}x_0 - \frac{\sqrt{3}}{2}y_1 + \frac{\sqrt{3}}{2}y_0, \frac{3}{2}y_1 - \frac{1}{2}y_0 + \frac{\sqrt{3}}{2}x_1 - \frac{\sqrt{3}}{2}x_0) \end{cases} \quad (2)$$

Eqs. (1) and (2) represent the ideal node coordinates.

2.2. Forward Probability Calculation

Assume that the sending node is v_i and the receiving node is v_j . The neighbor-degree δ is defined as:

$$\delta = N(v_j) / \pi r^2 \quad (3)$$

Where $N(v_j)$ represents the number of adjacent nodes of the receiving node v_j .

Using the neighbor-degree δ instead of the node density of forwarding probability in reference [11], the initial forwarding probability $P_0(v_j)$ of the receiving node v_j is:

$$P_0(v_j) = 1 - e^{-\delta r^2 \left[2 \sin^{-1} \left(\frac{d(v_i, v_j)}{2r} \right) + \frac{d(v_i, v_j)}{2r^2} \sqrt{4r^2 - d(v_i, v_j)^2} \right]} \quad (4)$$

Where $d(v_i, v_j)$ is the distance between the sending node v_i and the receiving node v_j .

Further considering the ideal coverage model, the final forwarding probability $P(v_j)$ can be obtained by combining Eqs. (1) and (2). In the case of the ideal distribution of nodes in the network, the coordinates of forwarding nodes can be calculated according to the ideal coverage model in Section 2.1 and denoted as the ideal node $v_{opt(l)}$. However, when N is small, there may be a long distance between the ideal node $v_{opt(l)}$ and the receiving node v_j , and forwarding node cannot be directly selected from the receiving node v_j . In order to select the sending node better, the node weight w_j is introduced.

Assume that the coordinate of the receiving node v_j is (x_j, y_j) and the coordinate of the ideal node $v_{opt(l)}$ is $(x_{opt(l)}, y_{opt(l)})$. The connection between two nodes is denoted as L , so the angle γ_{jl} between $L(v_i, v_{opt(l)})$ and $L(v_i, v_j)$ is:

$$\gamma_{jl} = \left| \tan^{-1} \left(\frac{y_j - y_{opt(l)}}{x_j - x_{opt(l)}} \right) \right| \quad (5)$$

The available-angle of the receiving node v_j is defined as $\gamma_j = \min(\gamma_{jl}) \in (0, 60^\circ)$. The distance between the receiving node v_j and the ideal node $v_{opt(l)}$ is denoted as D_{jl} , and the available-distance of v_j is defined as $D_j = D_{j_c} \in (0, r)$, where $c = \arg \min(\gamma_{jl})$.

The node weight w_j is represented by the available-angle γ_j and the available-distance D_j .

$$w_j = \frac{\cos\left(\frac{3}{2}\gamma_j\right)}{\sin\left(\frac{\pi D_j}{2r}\right)+1} \in (0,1) \quad (6)$$

The smaller γ_j and D_j , the closer the receiving node v_j and the ideal node $v_{opt(l)}$, the greater the node weight w_j . Select v_j with a large w_j as the forwarding node.

Combining $P_0(v_j)$ and w_j , the final forwarding probability $P(v_j)$ of the receiving node v_j is:

$$P(v_j) = \begin{cases} 0, & U(v_j) = \emptyset \\ \alpha \cdot P_0(v_j) + (1-\alpha) \cdot w_j, & U(v_j) \neq \emptyset \end{cases} \quad (7)$$

Where, $U(v_j)$ represents the uncovered neighbors of the receiving node v_j , set $P(v_j)$ to zero when $U(v_j) = \emptyset$; $\alpha \in (0,1)$ is a parameter.

In MANETs, nodes are dynamic, and parameters such as the neighbor-degree δ , the available-angle γ_j and the available-distance D_j will change accordingly. Therefore, $P(v_j)$ also changes dynamically, which can achieve dynamic space-covered of the network.

2.3. Forwarding Strategy

When the sending node v_i is the source node, three nodes with the largest $P(v_j)$ are selected for forwarding; otherwise, two nodes with the largest $P(v_j)$ are selected for forwarding.

According to $P(v_j)$, set the forwarding delay T_j for each forwarding node to alleviate broadcast collision; the larger $P(v_j)$, the smaller T_j . Assume that the sending node v_i needs a time slot to send broadcast packet to the receiving node v_j . Sort forwarding nodes in descending order according to $P(v_j)$. When v_i is the source node, T_j is 0, 1 and 2 in order; otherwise, it is 0, 1.

2.4. Broadcast Algorithm

Based on the ideal node coordinates and two-hop position information, the sending node calculates the forwarding probability $P(v_j)$ of adjacent nodes by Eq. (7) when it first receives a broadcast packet. Then the forwarding strategy is implemented, and the forwarding node and the corresponding forwarding delay are added to the broadcast

packet. Finally, the receiving node determines whether it is a forwarding node by extracting broadcast packet. If yes, calculate the forwarding probability of adjacent nodes, determine the forwarding node, and forward after the delay expires; otherwise, do not forward.

3. Algorithm Simulation and Performance Analysis

Assume that the broadcast radius of each node is $r = 3$ km. The nodes are considered stationary during broadcast. And each node knows the information of its all two-hop neighbors.

When $\alpha = 0.7$ and $N = 10, 20, 30, 40, 50, 60, 70, 80, 90$, the simulation results of proposed broadcast, dynamic probability broadcast [11] and node location based space-covered broadcast [13] are shown in Figures 1 to 3, respectively.

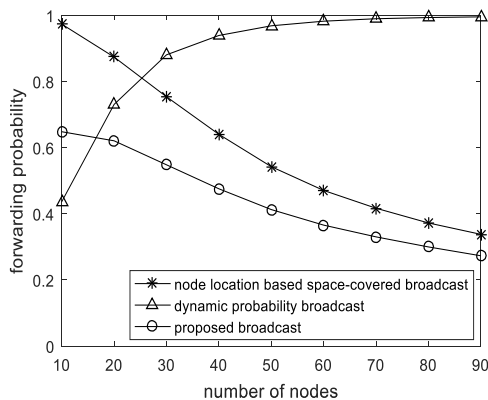


Figure 1. Forwarding probability of different number of nodes

As shown in Figure 1, when N is greater than 20, the forwarding probability of proposed broadcast is lower than dynamic probability broadcast and node location based space-covered broadcast.

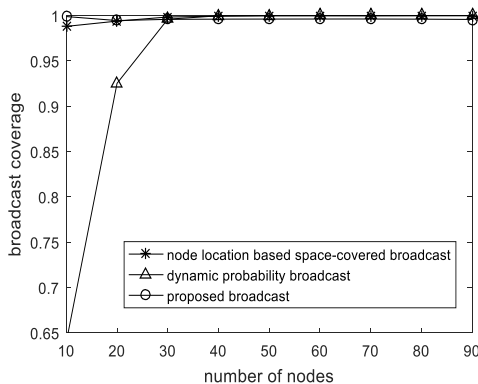


Figure 2. Broadcast coverage of different number of nodes

As shown in Figure 2, the broadcast coverage of proposed broadcast is almost 1 and has nothing to do with N which is better than dynamic probability broadcast and node location based space-covered broadcast.

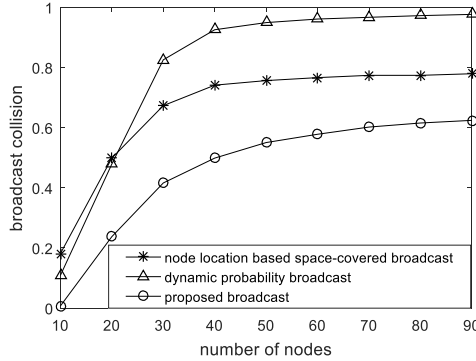


Figure 3. Broadcast collision of different number of nodes

As shown in Figure 3, the broadcast collision of proposed broadcast is the smallest.

4. Conclusion

Compared with dynamic probability broadcast and node location based space-covered broadcast, the performance of proposed broadcast is better in forwarding probability, broadcast coverage and broadcast collision.

References

- [1] Y. Ren, A study on broadcasting with low overhead in MANET, Xidian University, 2016.
- [2] D. Wang, Y. Song, ECCO: A novel end-to-end congestion control scheme in multi-hop cognitive radio ad hoc networks, *IEEE Transactions on Cognitive Communications and Networking* **5(1)**, (2019), 93-102.
- [3] B. Yuan, A. Jie, H.B. Zhang, Location aided probabilistic broadcast algorithm for mobile ad-hoc network routing, *The Journal of China Universities of Posts and Telecommunications* **24(2)**, (2017), 66-71.
- [4] B. Gomathy, D. Sathiy, Scrutiny of broadcasting efficiency with and without clustering in Manet, 2017 International Conference on Innovative Mechanisms for Industry Applications, Bangalore, (2017), 101-105.
- [5] R. S. R, C. Rajmohan, M. I. H, A modified broadcast algorithm for multi-hop relay MANETs, 2015 Fifth International Conference on Advances in Computing and Communications, Kochi, (2015), 219-222.
- [6] A. Fengyang, Research on broadcast algorithms based on wireless network, Qufu Normal University, 2015.
- [7] M. B. Yassein, A. Y. Al-Dubai, Inspired counter based broadcasting for dynamic source routing in mobile networks, 2015 International Conference on Computer and Information Technology, (2015), 1455-1459.
- [8] B. M. S. Waheed, B. M. Saeed, A. M. Hasan, A. J. Salim, Design of enhanced distance-based neighbor protocol for MANET with the aid of ELECTRE algorithm, 2019 2nd Scientific Conference of Computer Sciences, Baghdad, Iraq, (2019), 57-62.
- [9] Y. Saadi, B. Nassereddine, A. Haqiq, CBF evaluation through simulation, 2014 International Conference on Next Generation Networks and Services, Casablanca, (2014), 182-186.
- [10] Zohra, F. Tuz, Rahman, Ashikur, Mathematical analysis of self-pruning and a new dynamic probabilistic broadcast for MANETs, 2015 International Conference on Networking Systems and Security, Dhaka, (2015), 1-9.

- [11] R. Rab, A. Rahman, Fatema Tuz Zohra, Analytical modeling of self-pruning and an improved probabilistic broadcast for wireless multihop networks, *Ad Hoc Networks* **52**, (2016), 106-116.
- [12] R. Song, J. D. Brown, P. C. Mason, M. Salmanian, H. Tang, HMS: Holistic MPR selection and network connectivity for tactical edge networks, 2015 IEEE Military Communications Conference, Tampa, FL, (2015), 726-731.
- [13] L. Q. Bai, Y. Tian, Node location based space-covered broadcast algorithm in ad hoc networks, *Journal of Shenyang University (Natural Science Edition)* **26(2)**, (2014), 134-138.

Secure Server Key Management Designs for the Public Cloud

Kevin FOLTZ¹ and William R. SIMPSON

Institute for Defense Analyses, 4850 Mark Center Drive, Alexandria, VA 22311

Abstract. The Enterprise Level Security (ELS) model focuses on designing secure, distributed web-based systems starting from basic principles. One area of ELS that poses significant design challenges is protection of web server private keys in a public cloud. Web server private keys are of critical importance because they control who can act as the server to represent the enterprise. This includes responding to requests as well as making requests within the enterprise and to its partners. The cloud provider is not part of this trusted network of servers, so the cloud provider should not have access to server private keys. However, current cloud systems are designed to allow cloud providers free access to server private keys. This paper proposes design solutions to securely manage private keys in a public cloud. An examination of commonly used approaches demonstrates the ease with which cloud providers can currently control server private keys. Two designs are proposed to prevent cloud provider access to keys, and their implementation issues are discussed.

Keywords. enterprise, security, public cloud, key management, hardware security module, web server, cloud computing

1. Introduction

Web security has become more important as more aspects of business, government, and everyday life are put online. A challenge that occurs at the design level is securely managing the private key of a web server while utilizing the public cloud. The goal is a server key management design in which keys cannot be duplicated, keys can only be used by authorized individuals, and key operations are timely. At the time of this writing, no known solution has been implemented to provide all of these. This presents a problem that affects all public cloud hosted web servers.

This paper presents a review of common existing solutions' security shortcomings, an analysis of improvements to these solutions, and two proposals for a secure design.

2. ELS in a Private Cloud

The approach to security in this paper is based on Enterprise Level Security. Many security solutions patch together components that were developed independently, which leaves legacy and insecure protocols and interfaces as easy attack points [1]. Others use third parties to provide security as part of another non-security service [2],

¹ Corresponding Author, email: kfoltz@ida.org

which puts security in the hands of non-professionals and leads to easy security flaws. Still others use third parties that assume privileged positions in the architecture to perform security-specific functions [3]. The Zero Trust Architecture is a new approach, and ELS is an architecture that is consistent with ZTA ideas and has moved toward implementation. As such, it holds promise for enterprise security.

ELS enumerates a number of tenets and concepts that guide its requirements and implementation details [4]. One important idea in ELS is that all communication and sharing is done with end-to-end integrity and encryption between two endpoints. No other entities are allowed to view or modify such communication. Also, when two entities communicate, they have strong assurance that they are communicating with each other. This relies on strong end-to-end authentication of both entities [5] and hardware for storing and using private keys for high security applications.

The ELS private cloud hosting approach is shown in Fig. 1, where a normal request flows from a browser on an enterprise machine to a server hosted in the private cloud on a virtual machine. The browser uses the smart card certificate and private key to authenticate through TLS to the server. The smart card private key is accessed through a universal serial bus (USB) smart card reader, and the smart card is protected by a personal identification number (PIN). The server uses its certificate and HSM key to authenticate to the browser. The connection from the server to the HSM is secured through a separate TLS connection. This HSM connection may be local and long-lived. Session keys are used for the encryption, decryption, and validation of TLS transmitted data.

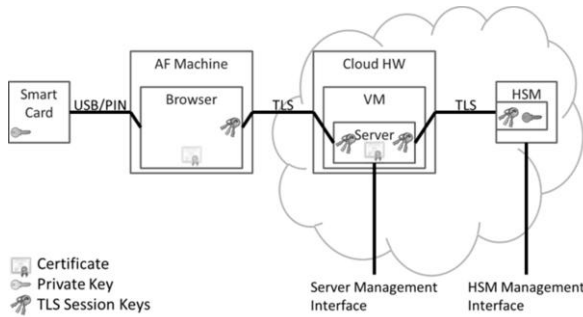


Figure 1. Normal web request flow.

3. The Public Cloud Challenge

The simplest method to move to the public cloud is to directly apply the private cloud model to the public cloud. For higher security, an administrator from the enterprise can visit the cloud provider and generate hardware key pairs on a validated HSM. A further security measure is to create only direct connections between cloud servers and the HSM and eliminate the use of virtual HSM interfaces. A final security measure is for an administrator to create all of these direct server-to-HSM connections while visiting the cloud provider.

Using all of these security measures, an administrator can create long-lived TLS connections between the servers and the HSM. Special tokens or keys that are used to create HSM connections are held by the enterprise and not shared with the cloud

provider. This prevents the cloud provider from creating additional connections. The server private keys are in hardware, and the only connections to the HSM are from the associated servers.

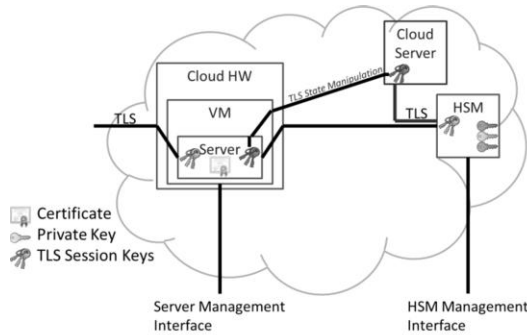


Figure 2. Web request flows in an untrusted cloud – direct TLS state manipulation to allow additional HSM connections.

However, the cryptographic keys for the secure HSM connection, along with any other necessary information including sequence numbers and other internal states, are stored in software on the server as part of the TLS connection state. This sensitive information can be extracted by an untrusted cloud provider by either copying the virtual machine on which the server is running or probing it through hypervisor interfaces. With access to the memory of the machine, methods exist to directly extract TLS keys from an executing application [6]. With these keys and the appropriate state information, the cloud provider can inject new private key usage requests to the HSM by creating the proper messages and updating the internal state of the server TLS connection, such as sequence numbers, initialization vector values, and other encryption state. This is shown in Fig. 2.

It requires specialized skills to probe a virtual machine image and manipulate TLS session keys, but the cloud provider has the required level of access to perform these actions. Even with dedicated effort to hide the TLS state in the virtual machine, the protection is only obfuscation, and such an approach fails to satisfy ELS security principles.

4. Proposed Secure Solutions

The main problems discussed in this paper have stemmed from the separation of the server and HSM and the challenge of establishing a secure connection between them. The approaches that follow attempt to combine the server and key.

4.1. Server in HSM

One way to combine server and keys is by implementing the entire server and its keys inside the HSM. The HSM server connects with the browser through a TLS connection, as shown in Fig. 3. The application is pre-loaded onto the HSM, associated with the proper private keys, and then shipped to the cloud provider. The private key access occurs completely within the HSM, which protects it from the cloud provider.

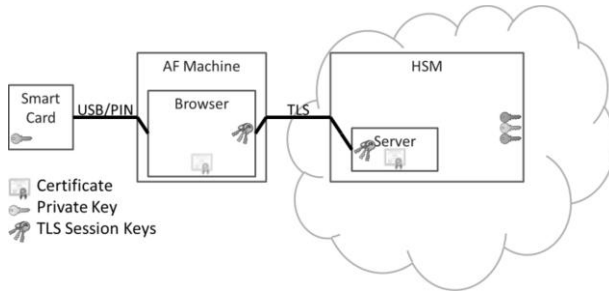


Figure 3. Proposed solution with server in HSM.

The main hurdle for this approach is that HSMs are expensive, special purpose devices designed specifically for key generation, storage, and use. They are not designed to run arbitrary software, and they typically do not have the storage or computation power required to support a full server. These are problems, but they are implementation issues within a fundamentally secure design.

4.2. Homomorphic Encryption

A second way to combine server and key is through the use of homomorphic encryption. Homomorphic encryption allows data to be processed while remaining encrypted [7]. The programs that normally operate on the unencrypted data are transformed to operate on the encrypted data.

In a homomorphic encryption solution, the requester encrypts its requests using homomorphic encryption. These requests are not decrypted at the transformed server. They are used as encrypted inputs to a program that operates on encrypted data. Part of the server processing involves authentication using an HSM. The HSM is also transformed to operate on encrypted data. The server sends encrypted key operation requests to the HSM and uses the encrypted response to authenticate to the requester. Similarly, the server can authenticate as a requester to other servers that accept encrypted requests. The design for the client to server communication is shown in Fig. 4.

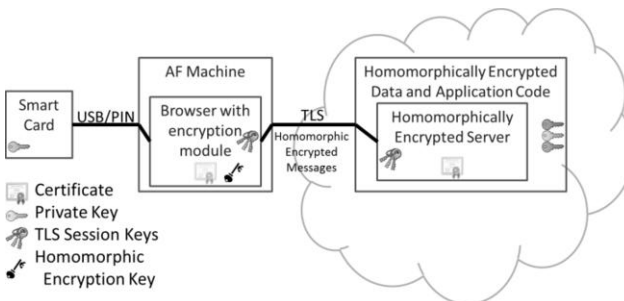


Figure 4. Proposed solution using homomorphic encryption.

The primary issue with the full homomorphic encryption (FHE) approach is performance. Current implementations of FHE are extremely slow [8]. FHE has seen significant performance improvements, but it is still far from practical.

5. Operational Considerations

The previous section discusses designs for server private key security in the public cloud. This section examines implementation issues for the secure designs.

5.1. *Cloud Vendor Support*

The cloud vendor controls the HSM and server and is centrally positioned to provide server private key security. Current cloud providers offer cloud-based key management services backed by HSM key storage, but they offer virtual HSM interfaces, and controls on key access are set at this virtual interface [9]. The cloud provider allows the enterprise administrator to set permissions on key use, but these permissions apply only to other enterprise users, not to the cloud provider itself. The cloud provider maintains direct control of the HSMs.

One promising approach for server integrity is a method to establish software root of trust unconditionally on hardware [10]. This provides assurance that no malware is present on a machine when it boots. The cloud provider performs this procedure to ensure that no malicious code is hidden on the hardware during the boot process. However, this still relies on the cloud provider to perform the procedure properly. Server private key security must come from outside the cloud provider.

5.2. *HSM Vendor Support*

An HSM can provide the property that a key cannot be duplicated. Major HSM vendors also provide some accommodations to facilitate the move to the cloud [11]. Thales, for example, provides a method to create a key on a local HSM and securely copy it to an HSM in the cloud [12]. However, this feature does not protect the cloud HSM from the attacks on the server-to-HSM connection. It also explicitly creates a copy of the private key, which is against ELS principles.

To improve security, the HSM vendor must provide a way to authenticate the server connections. However, the HSM is the key holder for server authentication, so this becomes a bootstrapping problem. Thus, the HSM vendor cannot solve this issue using advanced key management.

5.3. *Leveraging Mobile Device Management (MDM) Technologies for Cloud Assets*

Mobile device vendors embed hardware security features to enable secure communication with their devices. For example, the vendor can embed their own PKI root certificate authority (CA) public key in the hardware to enable trusted software updates [13]. The MDM services integrate with these hardware protections. The operating system, securely loaded by the hardware after signature validation, provides a set of APIs that can be used to remotely control the operation of the device.

The hardware protections of a mobile device make the device much like an HSM. The ability to remotely manage the phone, and in particular the ability to restrict user functions, makes such a device desirable for cloud hosting environments. In this scenario, the cloud provider acts as the mobile user, and the cloud client acts as the MDM administrator. The client hosts their server on a cloud-provided mobile device and manages it in the potentially hostile cloud environment through the MDM

interfaces. This allows the remote client to limit the activities of the hands-on cloud provider instead of the cloud provider limiting the client's activities.

6. Conclusions

The public cloud offers many desirable benefits in cost, efficiency, and even security, but taking advantage of these while maintaining secure server private key management is a challenge. Problems in the public cloud and hybrid cloud stem from the separation of the server from its authentication key in the HSM. An untrusted cloud provider can exploit this separation to use the private key of the server.

The proposed solutions for secure design have the common theme of encapsulating the server and HSM within a single logical entity. This eliminates the difficult problem of authenticating the server and HSM to each other. Using an HSM as the single logical entity provides a hardware encapsulation that hides the communications between the server and key management within the HSM itself. Using homomorphic encryption creates this single logical entity in software by encrypting the requests and responses and allowing the cloud provider to access and process only this encrypted data.

Current technology appears poised to be able to implement these approaches. The performance of FHE is currently poor, but it is improving. Mobile devices include technology that can allow secure remote management and software updates, which is similar to what is required of the "server in HSM" solution.

References

- [1] Ming Lu, technical discussions regarding IBM Tivoli, July-December 2014.
- [2] Farhan Saifudin, technical discussions regarding Mobile Iron mobile device management security, June 2018.
- [3] Jonn Inscoc, technical discussions regarding F5 web application firewall, F5 Government Technology Symposium, Washington, DC, December 2014.A.N. Author, Article title, Journal Title 66 (1993), 856–890.
- [4] Kevin E. Foltz and William R. Simpson, "Enterprise Level Security – Basic Security Model," 7th International Multi-Conference on Complexity, Informatics, and Cybernetics: IMCIC 2016, Orlando, Florida, March 2016.
- [5] Coimbatore Chandrasekaran and William R. Simpson, "The Case for Bi-lateral End-to-End Strong Authentication," World Wide Web Consortium (W3C) Workshop on Security Models for Device APIs, 4 pp., London, England, December 2008.
- [6] Brendan Dolan-Gavitt, T. Leek, J. Hodosh, and W. Lee, "Tappan Zee (North) Bridge: Mining Memory Accesses for Introspection," Proceedings of the ACM Conference on Computer and Communications Security (CCS), 2013.
- [7] Craig Gentry, 2009. "A Fully Homomorphic Encryption Scheme." Doctoral thesis. Stanford University.
- [8] Virgil D. Gligor, "Homomorphic Computations in Secure System Design," Final Report. Pittsburgh, PA: Carnegie Mellon University, 2014.
- [9] Amazon Web Services (AWS), "AWS CloudHSM." Available at <https://aws.amazon.com/cloudhsm/>.
- [10] Virgil Gligor and Maverick Woo, "Establishing Software Root of Trust Unconditionally," Network and Distributed Systems Security (NDSS) Symposium 2019, February 24-27, 2019, San Diego, CA, USA, <https://dx.doi.org/10.14722/ndss.2019.23170>.
- [11] Personal discussions with representatives of major HSM providers, RSA Conference, March 4-8, 2019, San Francisco, California.
- [12] Personal discussions with representatives of major HSM providers, BlackHat, August 7-8, 2019, Las Vegas, Nevada.
- [13] Apple Inc., iOS Security, iOS 12.3. May 2019. Available at https://www.apple.com/business/docs/site/iOS_Security_Guide.pdf

A Routing Protocol Based on Both of Density Variation and Distance-Aware for WSNs

Dongmei XING¹

Department of Mathematics, Nanchang University, Nanchang, China

Abstract. A hierarchical routing algorithm for wireless sensor networks (WSNs) is discussed. We select cluster heads according to related distances and residual energy. Both effects of the number of nodes dissipated and the energy consumption act on propagation distances. In addition, the related density effects on the propagation distance. We Define comprehensive influence factor and propagation influence factor, adjust the initial probability of nodes participating in cluster heads' election, make propagation distances of nodes gradually increase within a certain range. Simulation results show that both cluster heads and failure nodes are evenly distributed in the whole sensor network. The residual energy of nodes are balanced inter the living nodes, which extends the survival time of the network. The routing algorithm we have designed has the characteristics of better balanced energy consumption.

Keywords sensor, routing algorithms, density, energy reduction, failure factors.

1. Introduction

WSNs are made up of a lot of sensors with low energy. Sensors whose bodies are small can live in all sorts of environment and have several functions such as calculation, store and forward. The convenience of being carried and used let WSNs have extensive application prospect. Sensors have their own disadvantages, e.g. their limited energies, it is necessary to avoid their disadvantage. It is not convenient to add energy if any sensor has or is about to be dissipated in some application environments. When the data are forwarded in WSNs, the appropriate routing algorithm has to save energy and let the energy be consumed uniformly as much as possible. In order to extend the lifetime of WSNs, we should ensure living sensors with higher residual energy or being distributed uniformly in the network.

For the sake of convenience, sensors are named as nodes in the following, the node that finally receives data is the base station (or the sink node), and the node that sends data is called as source node. In a larger layout range, the geometric distances between the source nodes and the sink node are longer, it is not in one-hop to transport data, instead of multi-hop forwarding data. The nodes with same initial energy are uniformly distributed, but along data transmission, different nodes consume different energies and some nodes may fail in advance. The living nodes may be no longer uniformly distributed, and the residual energy of living nodes may be different. Suppose the data

¹ Corresponding Author, email: dongdongsh2000@sina.com.

is transported in distributed way in the network. With some admissible delay, the same received-data needs to be aggregated before being forwarded to the sink node. Data may be aggregated at certain relay nodes, and then be forwarded to the sink node. The hierarchical routing protocols mainly involve two phases, namely the cluster construction phase and the data transmission in the stable phase. A clustering can be divided into five steps, namely, computing the optimal number of cluster heads, selecting the optimal cluster heads, calculation of clusters' radius, constructing clusters, processing some special nodes (such as isolated nodes).

We designed a hierarchical route algorithm, the basic steps and requirements are described as follows.

1. Initial step. Set two influence factors on cluster heads' election and propagation radius. Initialize parameters.

2. Using residual energy, density change rate and related radio radius to adjust the probability of participating in the competition of cluster heads and cluster radius.

3. Selecting cluster heads in distributed way.

4. After all nodes have been joined in corresponding clusters, first special nodes (such as isolated nodes) would be processed, second the data is transported and forwarded. It is supposed that the received data could be aggregated before the data forwarded to another cluster head or to the sink node.

5. Supposed that conflicts are avoided during data transmission and information will not be lost.

2. Related Literatures

If sensors with homogeneity are arranged into a flat topology uniformly, the transmission range of WSNs is consistent with the coverage range if the propagation radius of the sensor meets certain conditions [1][2].

LEACH [3] is the first hierarchical clustering routing protocol. Many improved algorithms on WSNs have appeared [4-7] based on LEACH. LEACH protocol specifies the expected percent of the cluster heads and gives a threshold for each node to be selected as a cluster head. If the random number $p(n)$ assigned to node n is less than threshold $T(n)$, node n is selected as the cluster head, otherwise it is a non-cluster head node. Non-cluster-head nodes join in someone cluster head according to sensing strength. In each cluster, non-cluster-head nodes transmit data to their cluster heads in a single hop. Cluster heads send received data (after received data being aggregated) to the sink node in one-hop. Data forwarded between cluster heads and BS is completed in a one-hop mode, and the cluster head election is randomness and equilibrium.

Hierarchical cluster protocol HEED [8] takes into both the communication overhead and residual energy of each node. Data from source nodes to the sink node can be transported by multiple hops, and the relay nodes passing through are all cluster-head nodes. When a cluster head is selected by HEED, both primary and secondary factors are involved. The main factor is the percentage of the current residual energy E_r to the initial energy E_m of each node; the secondary factor is the communication energy consumption between the nodes in the cluster. Any cluster is constructed in HEED by three steps: initialization, loop step and final step.

Initialization. Calculate the neighboring sets, compute the minimum average communication energy consumption of each cluster (the neighboring set of a certain node). Set the initial probability CH_{prob} that any node is selected as a cluster head.

$$CH_{prob} = \max(C_p \times E_r / E_m, p_{\min}),$$

where C_p is the expected percentage of the cluster heads in the WSNs, p_{\min} is the minimum probability that any node becomes a cluster-head node, which is mainly used to ensure the convergence of HEED. The initial probability CH_{prob} changes with the network operation and residual energy of the node, which is the final probability that the node is selected as the cluster head.

Loop step. The main task is to select cluster heads from cluster heads' set S_{CH} , where cluster heads are divided into two sorts that are temporary cluster heads and final cluster heads.

In the final stage, each node determines its final state (either a member node of a cluster or a final cluster head). If the node is an isolated node, it is declared to be a cluster-head node.

Kuang zhejun [9] proposed a clustering algorithm REEM (A role, energy efficient, membership clustering) based on role membership in static network in order to eliminate "energy hole". REEM considers roles and relationships. Relationships among nodes are divided into member relationship and non-member relationship. Nodes' roles are divided into three sorts that are dominant node², relay node and source node according to their residual energy. Different roles transform identity by their own energy (residual energy of dominant node > residual energy of relay node > residual energy of source node). Unselected relay nodes are in dormant state. In REEM algorithm, dominant nodes send data to BS in one-hop, but the data is sent and forwarded from the source to a dominant node through an intermediate path composed of relay nodes. Prasenjit Chanak et al. [10] studied WSNs with uniform distribution, homogeneity and static layout. By the characteristics of the sensor network based on tasks, they proposed the problem of reducing the energy hole in the network in according to load management. The network is divided into several clusters whose scale can be determined [4]. The algorithm of load management mode [4] is completed in two steps. The first step selects the cluster head from the perspective of energy saving and complete the cluster construction. The second step would select the appropriate route path for the purpose of load balancing. Routing strategy considers the energy consumption of receiving and forwarding data. Both of clustering and selecting cluster heads are completed according to the energy consumption of the nodes, which is not completed in a probabilistic way like LEACH, HEED and other similar protocols.

The combination of optimization algorithms and cluster head selection strategies in hierarchical protocols is also a method to study wireless network routing algorithms. Tarunpreet Bhatia et al. [11] proposed routing algorithm GADA-LEACH (Genetic Algorithm based Distance-Aware LEACH) in wireless sensor networks by using genetic algorithm. First, LEACH protocol is used to select the initial cluster head. Then, define the fitness function according to three factors. Genetic algorithm is used to optimize the primary cluster head. An energy efficient clustering and routing method GEER (Genetic algorithm based Energy efficient Clustering and Routing approach) is

² For the sake of distinction, sink nodes in [9] is called dominant nodes in this paper.

designed by using genetic algorithm [12]. K. Thangaramya et al.[13] used fuzzy neural network to design an energy-sensitive clustering protocol.

In addition, some routing protocols consider other performance indexes of sensor networks and design corresponding hierarchical routing protocols. Ngoctu Nguyen et al. [14] designed the routing algorithm to maximize the utility of data aggregation and constructed a virtual aggregation tree by equilibrium adjustment in order to minimize the load consumption. By the tree, the data is received and forwarded to balance the energy consumption, so as to reduce the energy hole. Anuradha Pughat et al. [15] used the dynamic energy management model to analyze the performances. With the goal of minimizing energy consumption, Mohit Sajwan et al. [16] designed optimization algorithms of selecting multi-path routing or single-hop routing. Ansam Ennaciri et al. [17] designed the routing strategy from the perspective of load balancing with the goal of improving service quality. Surjit Singh etc. [18,19] discussed performance evaluation by optimization techniques or DMS (Dynamic Modulation Scaling).

3. Relevant Concepts

3.1. Radio Model

We use the First Order Radio Model (figure 1). The Settings of various parameters are shown in table 1~2. The energy consumption of transmitting 1bit and the energy consumption of receiving 1bit are respectively denoted as E_{Tx} (Transmitter Electronics), E_{Rx} (Receiver Electronics). The energy consumption $E_{Tx}(k, d)$, which is generated by a packet with k bits (k bits in a packet) transmitted to the target node at distance d , is consists of the energy consumption both of data transmission and data amplification (respectively denoted as $E_{Tx}(k), E_{amp}(k, d)$). The energy consumption of the packet received is denoted as $E_{Rx}(k)$.

TABLE 1 Radio model Parameters

Type	Parameter	Value	Parameter	Value
Radio model	E_{elec}, E_{Tx}, E_{Rx}	50 nJ/bit	E_{fusion} Threshold distance (d_0)	$E_{DA} = 5nJ / bit / signal$ $d_0 = 75m < \sqrt{e_{fs} / e_{amp}}$
	e_{fs}	10 pJ/bit/m ²		
	e_{amp}	0.0013 pJ/bit/m ⁴		

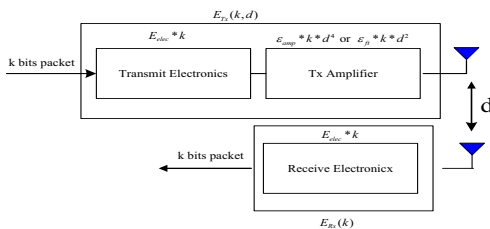


Fig 1 The First Order Radio Model

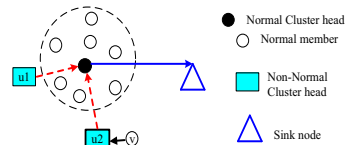


Fig 2 Disposal modal about an isolated cluster head or a cluster with only one normal member

Computing $E_{Tx}(k, d)$ [3,4] as follows.

$$E_{Rx}(k) = E_{Tx}(k, d) = E_{elec} * k, \tag{1}$$

$$\begin{aligned}
E_{Tx}(k, d) &= E_{Tx}(k) + E_{amp}(k, d) \\
&= \begin{cases} E_{elec} * k + \varepsilon_{fs} * k * d^2, & d < d_0 \\ E_{elec} * k + \varepsilon_{amp} * k * d^4, & d \geq d_0 \end{cases} \quad (1)
\end{aligned}$$

3.2. Influence Factors Design and Parameter Description

Suppose that n_m sensors with initial energy E_m are uniformly placed in a rectangular area. The remaining energy of a node at a certain time and the number of nodes alive in the network are denoted as E_r , n_{alive} , respectively. The initial probability of a node becoming a cluster head is defined as follows.

$$CH_{prob} = \max(C_p \times E_r / E_m \times f_p, p_{\min}) \quad (3)$$

where C_p is the expected percentage of the cluster heads in the network, and p_{\min} is the minimum probability that a node becomes a cluster header. f_p is the propagation influence factor with initial value 1, and its meaning is shown in the following Definition 3. CH_{prob} would change with the network operation and the residual energy of the node, which is also the final probability that the node is selected as the cluster head.

Cluster head election and radio radius usually change with the change of network energy. In the process of data transmission and forwarding, energy consumption is generally uneven. Considering the influence factors from both the whole and the part, we design the energy consumption reduction factor a_0 and the failure influence factor 1 (denoted as a_1). We also design density impact factor (denoted as b_0) and propagation impact factor f_p , failure impact factor 2 (denoted as b_1) and synthesis factor f_R .

Definition 1 Cost factor of energy reduction a_0 and failure impact factor 1 a_1 are defined as follows.

$$a_0 = 1 + \sum \frac{E_m - E_r}{n_{alive} * E_m}, \quad a_1 = 1 + \frac{n_m - n_{alive}}{n_m}.$$

Definition 2 Density impact factor b_0 and failure impact factor 2 b_1 are defined as follows.

$$b_0 = 1 - \frac{n_{TnumNbr} - n_{numNbr}}{n_m}, \quad b_1 = 1 + \frac{n_{Talive_0} - n_{alive_0}}{n_{Talive_0}},$$

where, $n_{TnumNbr}$ and n_{numNbr} are expressed respectively the number of neighbor nodes before and after data information transmission in the same round; n_{Talive_0} and n_{alive_0} are respectively the number of living nodes before and after data information transmission in the same round.

Definition 3 (energy consumption and failure) Comprehensive impact factor (denoted as f_R), and propagation impact factor (denoted as f_p) are defined as follows.

$f_R = (0.5 * a_0 + 0.5 * a_1) * b_1$, noticing $b_1 = 1$ (initial), so $f_R = 1$ (initial);

$f_p = (R_1 / d)^2$, here $f_p = 1$ (initial), $R_1 = b_0 * R$,

where R is the propagation radius of nodes (each node's R is same in the same round), d denotes the distance from non-cluster-head to cluster-head or cluster-head to BS.

Definition 4 The initial value of the communication transmission radius $R_0 = \sqrt{X^2 + Y^2} \times \sqrt{\frac{\log n_m}{n_m}}$, the adjustment formula of the transmission radius is

$R = \min(f_R * R_0, d_0)$, where d_0 is shown in table 1.

3.3. Design Idea

We believe that if a node consumes more energy in the previous round, it is less competitive in the following cluster-head competition. If the density of nodes has a larger change, the chance of selecting cluster head is smaller. The propagation distance of nodes increases with the increase of the failure node. The initial value of each node's CH_{prob} is modified according to both the residual energy and the propagation influence factor before the head election. The comprehensive influence factor is involved in the adjustment of the propagation distance of nodes. If the selected cluster head has no member node or just one member node, it is added to the nearest cluster and become a member of this cluster. Each member node send data to its cluster head by one-hop or two-hop, then cluster heads forward data to the sink node by multi-hop (or one-hop) after data aggregated.

Our routing algorithm is divided into four steps.

Step 1. Initialization. Set various parameters.

Step 2. Loop stage. Complete the function of cluster heads election and clustering.

Step 3. For nodes not in any cluster, if they are not the final cluster heads, set them as the final cluster head to complete the cluster construction; If they are final cluster heads, complete the cluster construction.

Step 4. Adjustment. A cluster head is isolated or has only one cluster member, then it is added to the nearest cluster and become a cluster member of this cluster. Adjust the member set of this cluster.

4. Algorithm Strategy and Explanations

4.1. Routing Algorithm

Our routing algorithm is an improvement of HEED routing protocol and named as Alg (in attach). It is briefly described as follows. Firstly, set parameters and define the initial radio radius R . Then select and determine the final cluster head to complete the cluster building operation (steps II~IV below). Finally, data transmission is occurred in WSNs.

4.2. Theoretical Explanation

Observation 1: Algorithm Alg is completely distributed and satisfies the conditions listed in the introduction. Any node becomes a cluster head node according to its CH_{prob} , or become a cluster member node.

The result of Observation can be obtained according to lemma 1~3 below.

Lemma 1 the Alg algorithm will terminate in $O(1)$.

Proof It is proved that Initialize, Finalize and TDMA in Alg can obviously be completed in limited steps, and the corresponding steps in Repeat are similar to those steps in Heed [8], so Alg can be completed in $O(1)$.

Lemma 2 At the end of step III, a node is either a cluster head or a normal node of a cluster (that is a member node and is a non-cluster head). In general, cluster heads are not isolated cluster heads, unless both the number of live nodes and the number of cluster heads are the same one node.

Proof Two cases.

1. If $CH1_{prob}$ and $CH2_{prob}$ of two adjacent nodes is different from each other, $CH1_{prob}$ and $CH2_{prob}$ cannot reach 1 at the same time in step III, then they cannot be selected as cluster-head nodes at the same time.

2. If $CH1_{prob}$ equals to $CH2_{prob}$, their values reach 1 at the same time, but the probability of this is small. The above result has been shown in [8](reference Lemma2 in [8]).

Lemma 3 Algorithm Alg can be used to select cluster heads in a distributed way. A node is either a cluster head node, or one or two hops can reach a cluster head node, and the cluster heads obtained are evenly distributed in the network range.

Proof Consider and Analysis step IV in Alg. According to using the conclusions of lemma 1, 2 and Lemma5 in [8], the conclusion of Lemma is easy to be drawn.

4.3. Energy Consumption Calculation

Energy consumption mainly lies in data transmission, forwarding and data aggregation. In the process of cluster head election, each node is either selected as a cluster head or join to a cluster to become a member node of a non-cluster head. In the process of each node confirming its role, the data sent or received is relatively small. Once the role is determined, the amount of aggregated data, as well as the amount of data sent or received, needs to be measured as the number of packets. There are three cases: energy consumption of one-hop (or two-hop) normal cluster members, energy consumption of cluster heads, and energy consumption of relay nodes (relay nodes refer to cluster head nodes with only one normal cluster member).

Energy consumption can be divided into three cases (in the process of data transmission and forwarding).

Case 1. Compute energy consumption of normal member nodes (non-cluster-head member nodes). Similar to formula equation (1).

Case 2. Isolated cluster heads and clusters with only one cluster member. An isolated cluster head is directly added to another nearby cluster and become its normal cluster member. Processing the clusters with only one normal cluster member, see the following description. In fig 2, Data transmission between isolated sensors and clusters with only one cluster member is expressed.

1. Isolated cluster head u join to a cluster whose cluster head is CH . u sent l bits to CH , the energy consumption is denoted as E_{u-CH} or $Tr(u)$.

$$E_{u-CH} = Tr(u) = E_{Tx}(l, d_{u-to-CH}) \tag{4}$$

$$= l * [E_{elec} + e_{fs} * d_{u-to-CH}^2],$$

2. A cluster u with only one normal cluster member v (here u is the cluster head, re-labeled u as the relay node). v sent l bit to the relay node u , then u receives and forwards data to a cluster head CH , the energy consumption E_{v-u-CH} is calculated as follows.

$$E_{v-u-CH} = Tr(v) + Re(u) + 2 * l * E_{DA} + Tr(u) \tag{5}$$

$$= l * [E_{elec} + e_{fs} * d_{v-to-u}^2] + l * E_{elec} + 2 * l * E_{DA} + \alpha * 2 * l * (E_{elec} + e_{fs} * d_{u-to-CH}^2)$$

$$= l * e_{fs} * (d_{vto u}^2 + 2\alpha * d_{u-to-CH}^2) + 2 * l * E_{DA} + 2 * (1 + \alpha) * l * E_{elec},$$

Case 3. Normal cluster-head nodes (whose size is $n_{CH} > 2$, where the size of a cluster refers to the number of nodes in the cluster).

Cluster head energy consumption = received data energy consumption + data aggregation energy consumption + data amplification and transmission energy consumption

The cluster whose size is n_{CH} ($n_{CH} > 2$) has $n_{CH} - 1$ normal member. Its cluster head CH receives $(n_{CH} - 1) * l$ bit and sends l bit at the same time. After data aggregation, data is sent to another cluster head CH_0 (CH_0 may be BS). The energy consumption E_{CH-CH_0} is calculated as follows.

$$E_{CH-CH_0} = Re(CH) + n_{CH} * l * E_{DA} + Tr(CH)$$

$$= (n_{CH} - 1) * l * E_{elec} + n_{CH} * l * E_{DA} + \alpha * n_{CH} * l * [E_{elec} + \lambda * d_{CHtoCH_0}^{\gamma}]$$

$$= \begin{cases} n_{CH} * l * [\alpha * e_{fs} * d_{CHtoCH_0}^2 + E_{DA}] + [n_{CH}(1 + \alpha) - 1] * l * E_{elec}, & \text{if } d_{CHtoCH_0} \leq d_0 \\ n_{CH} * l * [\alpha * e_{amp} * d_{CHtoCH_0}^4 + E_{DA}] + [n_{CH}(1 + \alpha) - 1] * l * E_{elec}, & \text{if } d_{CHtoCH_0} > d_0 \end{cases} \tag{6}$$

where α ($1 > \alpha \geq 0$) is the compression ratio; $Re(u)$ represents the energy consumption of node u receiving data, $Tr(u)$ represents the energy consumption of node u sending data, and E_{DA} represents the energy required for aggregation 1bit data. $d_{vto u}$ represents the distance from node to node; $d_{u-to-CH}$ represents the distance from node u to cluster head CH ; d_{CHtoCH_0} represents the distance between cluster head CH and cluster head CH_0 .

5. Simulation Experiment

5.1. Simulation Environment

TABLE 2 Simulation Parameters

Type	Parameter	Value	Type	Parameter	Value
Network	Network uniform	From (0,0) to (100,100)	Application	Data packet size	1000 bytes
	Sink	At (50, 140)		Broadcast packet size	250bytes
	Initial energy	4J/battery		Packet header size	25bytes
Application	Cluster initial radius	r_c		Gathering coefficient	0.8

Matlab is used for simulation experiments. The simulation network model and related parameters are shown in table 1~ table 2.

5.2. Results of Simulation

We use a randomly uniform network with 100 nodes. The related parameters are shown in Table 2. Both the proposed algorithm Alg and HEED are implemented by simulation. In the following figure, the red hollow point ‘o’ represents living node, solid green dot ‘•’ represents dead node, blue ‘*’ is a cluster head, ‘□’ representing sink node (or base station). Every star is a cluster (including cluster head, normal member and blue edges representing communication relationship.).

We have made an analysis and comparison from four aspects: 1) living node/round, failed node/round (Fig 3); 2) radius/round figure (Fig 4); 3) equilibrium of cluster distribution (only the case of Algorithm Alg is described; Fig 5); 4) distribution uniformity of living nodes (Fig 6-7). We also compare the rounds of algorithm Alg with HEED in five cases where the first time occurs: first_dead, 10%_dead, Half_dead, 90%_dead and All_dead (99%_dead) (table 3).

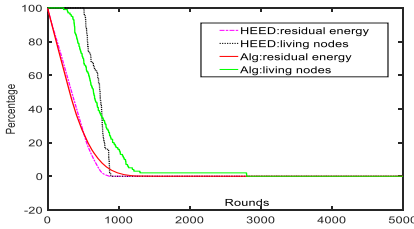
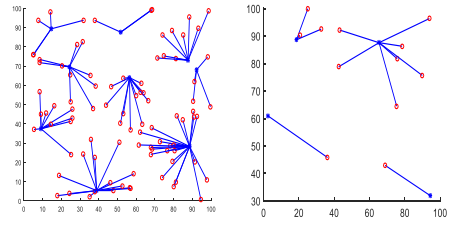


Fig3 Heed and Alg: Percentage of failure nodes, Living nodes, Rounds



a) Initial network b) most nodes failure

Fig5 Algorithm Alg: Clustering

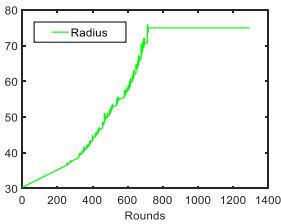
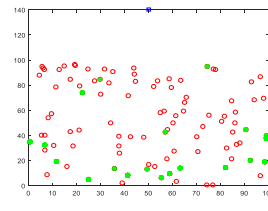
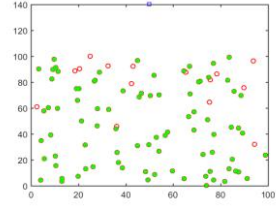


Fig4 Algorithm Alg: Radius/Rounds

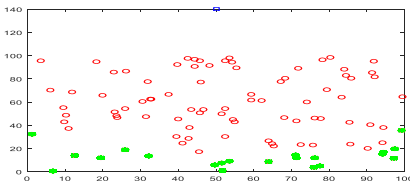


a) 20-node failure

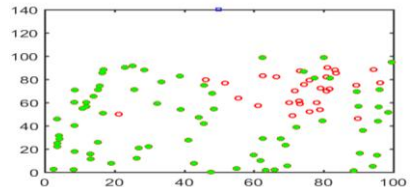


b) most nodes failure

Fig 6 Algorithm Alg: distribution about failure nodes and living nodes



a) 19-node failure



b) most nodes failure

Fig 7 Heed: distribution about failure nodes and living nodes

TABLE 3 Death Cases vs Rounds

Alg/Death	first dead	10% dead	Half dead	90% dead	All dead
Heed(rounds)	513	539	740	864	889
Alg(rounds)	255	373	656	1099	1293

As shown in Fig 5, the cluster distribution is balanced. Fig 6 and Fig 7 show that algorithm Alg can distribute failure nodes more evenly. Based on Table3 and Fig 3, we can see that the rounds are later when first_dead, 10%_dead and Half_dead appear in Heed for the first time, but the node failure rate is relatively high in the operation process with this algorithm. The first rounds of first_dead, 10%_dead and Half_dead appeared in Alg algorithm earlier, but the gap between the corresponding rounds and Heed is gradually narrowed, and the node failure rate of Alg become small during operation. When 90% of the failed nodes appears, the first round of Alg is later 235 rounds than Heed. Algorithm Alg prolongs the survival time of the network and enables the data to be transferred well in the network as a whole.

6. Conclusions

We design and implement routing algorithms Alg in wireless sensor networks with uniform distribution. The rationality of our algorithms is shown. The simulation results show that our algorithms have better performance in energy consumption uniformity.

References

- [1] Piyush Gupta and P. R. Kumar. Critical Power for Asymptotic Connectivity[C]. Proceedings of the 37th IEEE Conference on Decision & Control. Tampa, Florida USA Dec. 1998, 1106-1110.
- [2] Sanjay Shakkottai, R.Srikant, NessShroff. Unreliable Sensor Grids: Coverage, Connectivity and Diameter[C]. Proc. of the 22nd Annual Joint Conference of the IEEE Computer and Communications Societies. San Francisco, IEEE, 2003, 1073-1083.
- [3] Heinzlma, A. Chandrakasan and H. Balakrishnan. Energy-Efficient Communication Protocol for Wireless Microsensor Networks. Proceedings of the 33rd Hawaii International Conference on System Sciences, Jan. 2000. (Proc. 33rd Hawaii Int'l. Conf. Sys. Sci., Jan. 2000)
- [4] W.B. Heinzelman, A.P. Chandrakasan, H. Balakrishnan. An application specific protocol architecture for wireless microsensor networks, IEEE Trans.Wirel. Commun. 1, Oct. 2002, 4, 660-670.
- [5] G. Smaragdakis, I. Matta, A. Bestavros. SEP: A Stable Election Protocol for clustered heterogeneous wireless sensor networks. Second International Workshop on Sensor and Actor Network Protocols and Applications (SANPA), 2004, 1-11.
- [6] Dongmei Xing. Energy Consumption Balancing Algorithms on Optimizing Dynamic Topology in Wsns. Frontiers in Artificial Intelligence and Applications, Volume 320: Fuzzy and Data Mining V- Proceedings of FDSM 2019, 975-983. DOI:10.3233/FAIA 190273.
- [7] Vishal Kumar Arora (Research Scholar), Vishal Sharma, Monika Sachdeva. A survey on LEACH and other's routing protocols in wireless sensor network, Optik 127 (2016), 6590-6600.
- [8] O. Younis, S. Fahmy. HEED: A Hybrid, Energy-Efficient, Distributed Clustering Approach for Ad-hoc Sensor Networks, IEEE Transactions on Mobile Computing, 2004 , 3 (4), 366-379.
- [9] Kuang ZheJun. Research on energy saving strategies for wireless sensor networks (Thesis, in Chinese). Doctoral thesis of Jilin university, China. May 2014.
- [10] PrasenjitChanak, IndrajitBanerjee, HafizurRahaman. Load management scheme for energy holes reduction in wireless sensor networks[J]. Computers and Electrical Engineering, 48 (2015), 343-357.
- [11] Tarunpreet Bhatia,Simmi Kansal, Whivani Geol, A.K. verma . A genetic algorithm based distance-aware routing protocol for wireless sensor networks. Computers and Electrical Engineering, 56 (2016), 441-455. <http://dx.doi.org/10.1016/j.compeleceng.2016.09.016>.
- [12] Tianshu Wang, Gongxuan Zhang, Xichen Yang, Ahmadreza Vajdi. Genetic algorithm for energy-efficient clustering and routing in wireless sensor networks. The Journal of Systems and Software, 146 (2018), 196-214. <https://doi.org/10.1016/j.jss.2018.09.067>.

- [13] K. Thangaramya, K. Kulothungan, R. Logambigai, M. Selvi, Sannasi Ganapathy, A. Kannan. Energy aware cluster and neuro-fuzzy based routing algorithm for wireless sensor networks in IoT. *Computer Networks*, 151 (2019), 211–223. <https://doi.org/10.1016/j.comnet.2019.01.024>.
- [14] Ngoc-Tu Nguyen, Bing-Hong Liu, Van-Trung Pham, Yi-Sheng Luo. On maximizing the lifetime for data aggregation in wireless sensor networks using virtual data aggregation trees. *Computer Networks*, 105 (2016), 99–110. <http://dx.doi.org/10.1016/j.comnet.2016.05.022>.
- [15] Anuradha Pughat, Vidushi Sharma. Performance analysis of an improved dynamic power management model in wireless sensor node. *Digital Communications and Networks*, 3 (2017), 19-29.
- [16] Mohit Sajwan, Devashish Gosain, Ajay K. Sharma. Hybrid energy-efficient multi-path routing for wireless sensor networks. *Computers and Electrical Engineering*, 67 (2018), 96–113. <https://doi.org/10.1016/j.compeleceng.2018.03.018>.
- [17] Ansam Ennaciri, Mohammed Erritali, Jamaa Bengourram. Load Balancing Protocol (EESAA) to improve Quality of Service in Wireless sensor network. *Procedia Computer Science*, 151 (2019), 1140–1145.
- [18] ISurjit Singh, Rajeev Mohan Sharma. Optimization Techniques in Wireless Sensor Networks. CTCS '16, March 04-05, 2016, Udaipur, India. DOI: <http://dx.doi.org/10.1145/2905055.2905200>.
- [19] Maryam Bandari, Robert Simon, and Hakan Aydin. 2017. DMS based energy optimizations for clustered WSNs. *ACM Trans. Embed. Comput. Syst.* 3(16), Article 86 (April 2017), 28 pages. DOI: <http://dx.doi.org/10.1145/2998179>

Appendix

Algorithm Alg

I. Initialize

1. Compute f_R , f_P and R ; Determine the neighboring set S_{nbr} of each node

$$S_{nbr} \leftarrow \{v: v \text{ lies within my cluster range}\}$$
2. Compute and broadcast cost to $v \in S_{nbr}$
3.
$$CH_{prob} \leftarrow \max \{C_{porb} \times \frac{E_{residual}}{E_{max}} \times f_P \times P_{min}\}$$
4. $is_final_CH \leftarrow FALSE$

II. Repeat

1. **If** ($(S_{CH} \leftarrow \{v: v \text{ is a cluster head}\}) \neq \emptyset$)
2. $my_cluster_head \leftarrow least_cost(S_{CH})$
3. **If** ($CH_{prob} = 1$)
4. $Cluster_head_msg(NodeID, final_CH, cost)$
5. $is_final_CH \leftarrow TRUE$, compute R and d
6. **else**
7. **If** $my_cluster_head = NodeID$
8. $Cluster_head_msg(NodeID, tentative_CH, cost)$
9. **else If** $Random(0, 1) \leq CH_{prob}$
10. $Cluster_head_msg(NodeID, tentative_CH, cost)$
11. $CH_{previous} \leftarrow CH_{prob}$
12. $CH_{prob} \leftarrow \min\{C_{porb} \times 2, 1\}$
13. **Until** $CH_{previous} = 1$

III. Confirm

1. **If** ($is_final_CH = FALSE$)
2. **If** ($S_{CH} \leftarrow \{v: v \text{ is a cluster head}\}) \neq \emptyset$
3. $my_cluster_head \leftarrow least_cost(S_{CH})$

4. $join_cluster(cluster_head_ID, NodeID)$
5. **else** $Cluster_head_msg(NodeID, final_CH, cost)$,
6. **else** $Cluster_head_msg(NodeID, final_CH, cost)$

IV. Adjust

1. **If** ($|S_{nbr}|=1$) and ($is_final_CH = TRUE$)
2. $S_{CH} \leftarrow S_{CH} - \{NodeID\}$, $is_final_CH = FALSE$
3. $my_cluster_head \leftarrow least_cost(S_{CH})$
4. $join_cluster(cluster_head_ID, NodeID)$
5. **If** ($|S_{nbr}|=2$) and ($is_final_CH = TRUE$)
6. $S_{CH} \leftarrow S_{CH} - \{NodeID\}$, $is_final_CH = NNON_CH$
7. $my_cluster_head \leftarrow least_cost(S_{CH})$
8. $join_cluster(cluster_head_ID, NodeID)$

Real Time Digital Filter for a Front-End Electronics in Dark Matter and Neutrino Measurements

Alejandro D. Martinez R.^{a,b1}

On behalf of DarkSide Collaboration

^a *Politecnico di Torino, Italy*

^b INFN Torino, Italy

Abstract. This paper presents real-time digital filter algorithms to be applied within dark matter and neutrino measurements. The digital signal processing algorithm implements a trapezoidal pulse-shaper programmed on FPGA at 125 MHz. The real-time filter algorithm enhances the SNR of a digitized signal from a photo detection module (SiPM, cryogenic front-end electronics & 14-bits ADC). The trapezoidal filter upgrades the signal to noise ratio (SNR) from 10.4 to 15.4 with a total increment of 50%. The total on-chip power is 0.198 W.

Keywords. cryogenic, DSP, filtering, FPGA

1. Introduction

Digital signal processing (DSP) has been implemented in nuclear physics detection [1] and nuclear spectroscopy [2] in the recent decades. Digital filtering owns several advantages and a higher efficiency in comparison to traditional analog electronics methods. In fact, traditional methods find some limitations to upgrade the desired signal without risking the stability. The main aim of this study is to implement a DSP algorithm to enhance more than 30 %, in real time, the figure of merit (FOM) of the neutrino and dark matter measurements. In this case, FOM is the signal to noise ratio (SNR). In turn, this improvement will allow the worst-case condition of the front-end electronics (SNR = 6) to reach the minimum requirement, described by the experiment (SNR = 8 or 4σ resolution). As a result, a DSP algorithm was programmed in FPGA at 125 MHz, using a real-time trapezoidal filter [3]. The digital filter processes a 14-bits input signal from a photo detector module (PDM).

2. Real-time Digital Filter

2.1. Noise Contributions

The digital pulse processing implementation goal is to reduce the proportion of undesired signal over the mean signal. In other words, to minimize the different noise sources presented in cryogenic front-end electronics. The noise sources are divided into the

¹ Corresponding Author. Email: alejandro.martinez@polito.it

parallel thermal noise due the leakage current from the detector, the series thermal noise due to the input capacitance, and the flicker or low frequency noise [4]. Even though the frontend integrated electronics works at temperature of 77 K, the total detector capacitance presents a large value of 10 nF. Hence, the series and parallel thermal noise contributions become an important factor in the circuit. This condition generates a strong interest in the real time digital filter.

On the other hand, the integrated electronics (Figure 1) present a low $1/f$ noise contribution due to its large transistors area, since the transistor sizing is performed in such a way to reduce the hot carrier effect. Although the flicker noise is negligible at high frequency in this application, the CMOS technology may produce a $1/f$ noise increment at low temperatures [5]. It is important to hold the noise increment at low level, in the case of low frequency application. A digital trapezoidal pulse-shaper was chosen as the structure to build the real-time filter algorithm, since it accomplishes a high noise reduction in the case of a series and a parallel noise contribution [6].

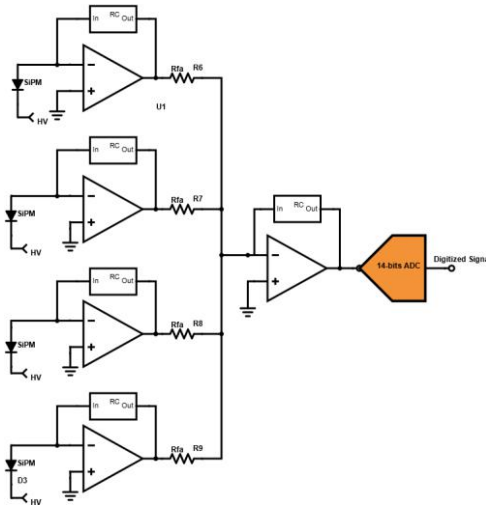


Figure 1. Front-end integrated electronics schematic

2.2. Digital Filter Application

In digital filter algorithms, the total RMS noise is significantly decreased by setting an optimal shaping/peaking time within the digital trapezoidal filter. For instance, the series noise drops as the peaking time of the trapezoidal waveform increases. However, the parallel noise presents an opposite behavior [7]. Hence, the shaping time was tuned in order to obtain an optimal performance in term of series and parallel noise contributions. Furthermore, the shaping time could be modified digitally depending on the application. For instance, a long peaking time is developed for the low rate application. In the meantime, a short shaping time is useful for the high rate application. Thus, the digital trapezoidal parameters exclusively depend on the application approach. Besides, the

trapezoidal filter provides an additional advantage in the possibility to synthesize and implement in real time using a clock frequency of 125 MHz.

The real time trapezoidal filter processes and enhances an output signal from the photo-electron detection in the DarkSide-20k experiment. The detection and readout are developed by a SiPM sensor [8] and front-end integrated electronics at 77 K [9], as shown Figure 1. The combination of the photo detection module and the digital filtering become an important support in the neutrino [10] and dark matter [11] experiments with large area detectors.

3. Design and Implementation

The real-time digital filter provides a symmetrical trapezoidal pulse at 125 MSa/s. The signal shaping goal is to obtain an RMS noise substantially lower than the digitized signal using a sampling time of 8 ns. The digitized signal is an analog input waveform, which presents a peaking time of 250 ns, a falling-edge of 900 ns, an RMS noise of 600 μ V and a photo-electron amplitude of 5.6 mV. These key parameters will be compared against the trapezoidal output signal in order to check the digital filtering efficiency.

3.1. Trapezoidal Filter Design

The trapezoidal pulse-shaper schematic is illustrated in Figure 2. As the figure illustrates, the pulse-shaper is divided into 4 stages with difference purposes within the shaping process. The first and second stage are described by the Equation 1. These are the delay units programmed with a FIFO block, which exhibits a depth of 200 samples. The FIFO depth fixes the total delay of the unit, and consequently the rising and falling edge time of the trapezoidal waveform. In this design, the Delay_a and Delay_b were set to achieve an optimal peaking time of 1.6 μ s, programming both 1st and 2nd stage with the same delay.

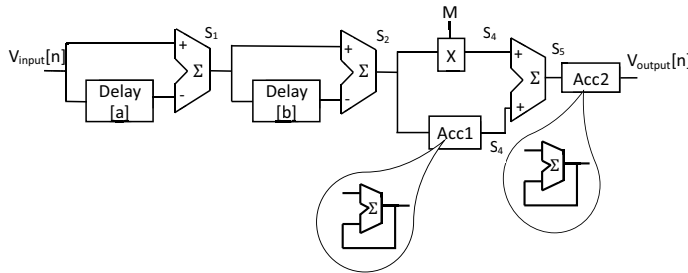


Figure 2. Digital trapezoidal filter schematic

Equation 2 (S_3, S_4) describes the operation of the third stage. This stage represents a high-pass network on the system, which includes a multiplication factor (M) in one of its branches. The M factor is implemented to cancel the exponential term of the analog input signal. For this reason, the M factor depends on the ratio between the sampling time and the decay time constant, as shown the Equation 3. The high-pass network basically deconvolutes the analog output signal, with a decay time constant (τ_{decay}), into a step

signal. The output signal of the third stage (S_4) is illustrated in the post-implementation timing simulation (Figure 3).

In this digital trapezoidal filter design, the third stage is computed by applying a sampling time (τ_{clk}) of 8 ns and a decay time constant τ_{decay} of 680 ns. Hence, the M factor computed, following the Equation 3, is equal to 85. The Delay_a , Delay_b and M were key parameters, which were set in the digital pulse processing algorithm.

$$S_1[n] = V_{in}[n] - V_{in}[n - \text{delay}1] \quad (1)$$

$$S_2[n] = S_1[n] - S_1[n - \text{delay}2]$$

$$S_3[n] = S_2[n] + S_3[n - 1] \quad (2)$$

$$S_4[n] = S_3[n] + M * S_2[n]$$

$$V_{output}[n] = S_4[n] + V_{output}[n - 1]$$

$$M = \frac{1}{e^{\frac{\tau_{clk}}{\tau_{decay}}} - 1} \quad (3)$$

3.2. Simulation and Implementation Tool

The synthesis and place & route process were carried out using the Vivado tool. The synthesis and implementation were accomplished by using the Xilinx Kintex-7 FPGA KC705 Evaluation Kit with a clock frequency of 125 MHz. The total on-chip power consumption of the digital trapezoidal filter is equal to 0.198 W. Furthermore, the digital filter implementation occupies a total area smaller than 3% of the total available area in the FPGA. The digitized signal is supplied by a 14-bits ADC, which samples the analog signal from the integrated electronics at 125 MSa/s.

The post-implementation timing simulation is shown within Figure 3. This figure presents the shaped signal waveform stage by stage described in the Figure 2. The first signal (V_{in}) represents the digitized signal. S_1 and S_2 respectively represent the output signal from the delay units, S_4 describes the high-pass network output signal. The S_4 stage shows how the digitized signal is converted into a positive and negative step signal. Finally, V_{in} outlines the trapezoidal waveform with a similar amplitude in comparison to the front-end electronics output signal.

4. Experimental Results

The experimental results were realized by processing a database with 25×10^6 samples or 200 ms of data acquisition from the photo detection module output signal. The experimental database is digitized, following this is employed as the input of the trapezoidal filter. In order to check the efficiency of the digital filtering, a peak voltage histogram was built with the output amplitude before and after real time digital filtering.



Figure 3. Signal waveform produced on the output of each stage

4.1. Peak Voltage Histogram

The peak voltage histogram represents the amplitude of the detected photoelectrons (PEs) from the cryogenic electronics, as shown Figure 4. From the histogram, the figure of merit, which analyzes the filtering efficiency, is the signal to noise ratio (SNR). The SNR is divided for two terms in this study, as shown Figure 4. This figure provides the information of SNR_{base} and SNR_{1p} . The SNR_{base} represents the ratio between a PE amplitude and the standard deviation (std) of the baseline. In the meantime, the SNR_{1p} represents the ratio between one photo-electron amplitude and its std.

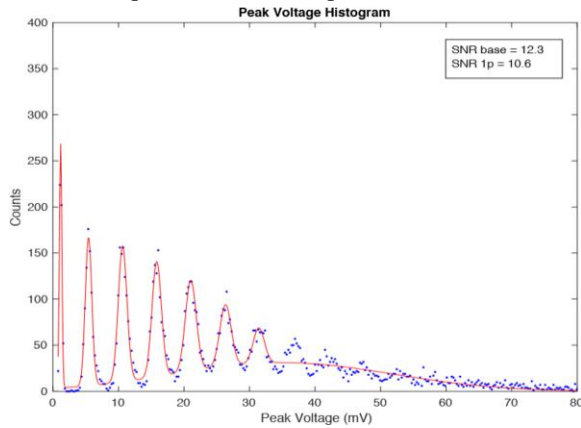


Figure 4. Peak voltage histogram of the front-end electronics

The integrated electronics results (pre-filtering) provide a SNR_{base} equal to 12.3 (21.8 dB) and a SNR_{1p} equal to 10.6 (20.5 dB). Applying the digital trapezoidal pulse shaper, the peak voltage of each photoelectron exhibits a larger separation in comparison to the pre-filtering histogram. The new histogram is described by the Figure 5. This exhibits a peak voltage distribution with a much lower std, as illustrated by a higher number of counts per peak. Hence, the figure of merit was improved considerably, as shown the post-filtering parameters SNR_{base} equals to 23 (27.3 dB) and SNR_{1p} equal to 15.4 (23.8 dB). The trapezoidal pulse shaper results describe a significant increment of SNR around 50%. This factor realizes an improvement in the photon counting resolution.

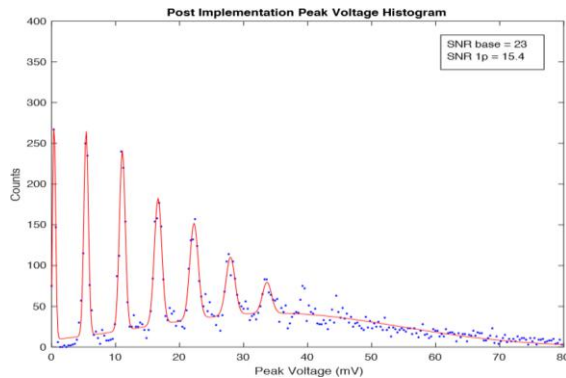


Figure 5. Peak voltage histogram after trapezoidal digital filtering

5. Conclusion

The real time digital filter algorithm for dark matter and neutrino measurements showed a high performance in terms of SNR results in comparison to the pre-filtering results. The new SNR_{base} and SNR_{1p} was enhanced around 50 %, since these reached values higher than 20 and 15 respectively. This improvement increases the photon counting resolution, as the peak voltage histograms describe. The real-time filter application is focused on low rate measurements, such as dark matter and neutrino experiments.

References

- [1] V. Jordanov et al, *Digital techniques for real-time pulse shaping in radiation measurements*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators Spectrometers Detectors and Associated Equipment **353** (1994), 10.1016/0168-9002(94)91652-7.
- [2] Z. Guzik, T. Krakowski, *Algorithms for digital -ray spectroscopy*, Nukleonika **58** (2013) 333-338.
- [3] N. Mena, P. D'Agostino et al, *Evaluation of real-time digital pulse shapers with various HPGe and silicon radiation detectors*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **652** (2011), 512-515.
- [4] A. Rivetti, *CMOS Front end electronics for radiation sensors*, CRC press, Turin, 2015.
- [5] F. J. Scholz and J. W. Roach, *Low-frequency noise as a tool for characterization of near-band impurities in silicon*, Solid-State Electron (1992), 447-452.
- [6] G. Gerardi, L. Abbene, *A digital approach for real time high-rate high-resolution radiation measurements*, Nuclear Instruments and Methods in Physics Research Section A Accelerators Spectrometers Detectors and Associated Equipment **768** (2014), 46-54.
- [7] E. Gutierrez et al, *Low Temperature Electronics: Physics, Devices, Circuits, and Applications*, Claeys Academic Press, San Diego, 2001.
- [8] F. Acerbi et al, *Cryogenic Characterization of FBK HD Near-UV Sensitive SiPMs*, IEEE Transactions on Electron Devices, Vol. 64, Iss. 2, 2017
- [9] A. Martinez, *4-Channel Front-End Integrated Circuit For Readout of Large Area of SiPM under Liquid Argon*, 15th Conference on Ph.D Research in Microelectronics and Electronics (2019), 10.1109/PRIME.2019.8787758.
- [10] J. Stewart, *The ProtoDUNE Single-Phase Detector*, IEEE Nuclear Science Symposium and Medical Imaging Conference, 2017.
- [11] C. E. Aalseth, F. Acerbi, et al, *DarkSide 20K: A Yellow book technical proporsal/pre-technical design report*, INFN, 2016.

Research on Fixture Design of NC Machine Tool Based on UG 3D Modeling

Yanxia LI¹

Nanjing Communications Institute of Technology, Nanjing, 211188, China

Abstract. Research on the design of CNC machine tool fixture based on UG 3D modeling; Firstly, the processing procedure of the back cover parts is analyzed, and the overall scheme of the fixture is determined. Then the three dimensional model of each part of the jig is established including the standard part library of jig and the design of non-standard parts of jig. Then carry on the fixture 3D assembly drawing design; Finally, it is converted into fixture technical document including assembly drawing and specific part drawing. The design process is simple and clear, which improves the quality and efficiency of fixture design.

Keywords. UG software, fixture design, assembly drawing, 3D modelling

1. Introduction

2D CAD software is adopted for fixture design. The designer should first form the three-dimensional entity of the fixture in the mind of the structure, and then express it with 2D diagram to complete the fixture design. In the design process, thinking should be converted repeatedly in the three - and two-dimensional models, and finally two-dimensional assembly drawing is drawn, but three-dimensional assembly drawing is generally not drawn. There are unfavorable factors for fixture design quality and efficiency, as well as machining and assembly.

UG software is currently the most advanced in the world, for manufacturing industry, set of 3D CAD/CAM/CAE integration of high-end software, it not only has powerful entity modeling, but also associated with three-dimensional entity model completely two-dimensional engineering drawing function of UG two-dimensional engineering graphics can be directly generated from the product design of the three-dimensional entity, and two-dimensional engineering graphics when the 3D model changes in the related view, size and other information to update.

Based on UG 3D modeling of CNC machine tool fixture design, design fixture 3D assembly drawing under UG environment, and then convert it into 2D engineering drawing, which is described in detail below.

¹ Corresponding author: Li Yanxia, professor, mainly engaged in the research of CAD/CAM application technology, NUMERICAL control processing technology and mechanical processing technology, has published more than 30 papers, tel: 13921406104, email: Yanxia200964@126.com

Research start-up Fund project for high-level talents of Nanjing Institute of Communication Technology: Design and Research of CNC Machine tool fixture based on UG 3D Modeling (No.: 2014), host: Li Yanxia.

2. Jig Information Processing

The information of fixture design includes parts geometry information, processing technology information, clamping process information and so on. Through the analysis of the structure of the parts, clamping features and technological requirements, make the fixture design scheme, that is, the positioning scheme, clamping scheme, the specific structure of the clamp layout scheme.

2.1 Analyze the Parts Process

As shown in Figure 1, (a) is the top surface of the rear cover, and (b) is the ground of the rear cover.

The underside and a side of the back cover parts have been processed. This procedure requires the processing of 8- $\Phi 10.5$ holes on the underside and 6-M12 threaded underside holes on the a side. The production scale is batch production.

According to the process requirements, the workpiece is processed on the machining center, with one side and two holes (bottom and two locating pin holes) as the positioning datum, and the $\Phi 10.5$ drill bit is selected for processing the bottom 8- $\Phi 10.5$ holes, and the $\Phi 10.25$ drill bit is selected for processing the 6-M12 threaded bottom holes, and all holes are processed by one-time clamping.

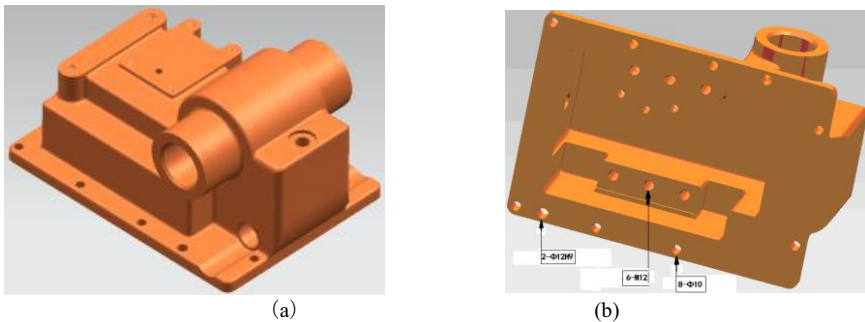


Figure 1. Back cover parts.

2.2 Determine the Overall Plan of Fixture

As the machining of the hole of the part to be machined feeds from the bottom to the top, it is necessary to place the bottom of the part upward for machining, but it is not convenient for positioning and clamping. It is comprehensively considered to place the part side vertically for positioning and clamping. After the clamping is installed, the whole clamp is turned 90° degrees, and the hole machining is carried out with the bottom upward.

According to the analysis, the overall plan of the fixture is that the workpiece is placed vertically on the side, one side and two pins are used for positioning (support plate and two positioning pins are selected), and the hinge pressing plate mechanism is used to press the workpiece from the top surface to the bottom surface for clamping; because the hole to be processed is on the positioning datum plane, the drilling template and the fixture body are combined to save space; because the fixture needs to be turned 90° , the turning mechanism needs to be designed.

3. Establish Three-dimensional Model of Each Part of Fixture

3.1 Set Up Fixture Standard Parts Library

In the design of fixture, more than 50% parts are standard parts, designers have to carry out simple repetitive work;The establishment of fixture standard parts library plays an important role in improving fixture design efficiency.

(1) Set up parameter table

The parameter table mainly expresses the parameters of each part of the standard parts. The fixture design selects the appropriate standard components according to the parameters in the table.

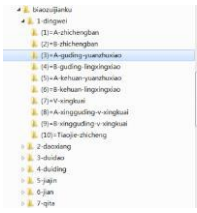
The first level heading includes positioning element, guiding element, aligning element, aligning element, clamping element, key and other elements.

There is a table for each standard part. The top of the table is the 2D graph and 3D graph of the standard part. The third line is the specification of the standard part and the bullet symbol corresponding to each part.

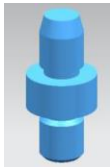
(2) Create a 3D gallery

Firstly, the catalogue of the standard element 3D gallery should be established, which should be consistent with the catalogue of the parameter table. Then, according to the parameters in the parameter table, UG software was used to create the 3d modeling of the standard parts and put them in the folder of the corresponding directory of the 3D gallery to form the 3D gallery, as shown in Figure 2 (a).

In fixture design, if you need to use the parts in the standard parts gallery, you can directly copy the fixture assembly folder; If there are no parts of corresponding size in the standard parts library, but there are similar parts, you can simply copy the files in the gallery first and then make appropriate modifications, as shown in Figure 2 (b).



(a)



(b)

Figure 2. 3D drawing library of standard parts

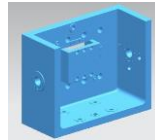


Figure 3. Clamp the specific

3.2 Design of Fixture Non-standard Parts

When jig is designed, the non-standard parts that need to be designed mainly include specific clamp, drilling template, etc., the design of non-standard parts needs to be carried out according to the specific situation, so it is more flexible.

The specific design of clamping, mainly considering the structure of the processed parts, the use of the machine tool, clamping device on the positioning element, clamping device, tool device, guide device and other dimensions and positions, clamping specific use of materials, the process of chip removal method, and so on. Therefore, it is difficult to unify the structure and size of non-standard parts, and the model library cannot be built as standard parts. Instead, the UG software can only be used to build three-dimensional models one by one, which will not be introduced here.

Figure 3 shows the details of the clamp for the processed rear cover parts.

4. Design of 3D Assembly Drawing of Fixture

4.1. Design of Fixture 3D Assembly Drawing based on UG

According to the overall fixture design scheme, copy and paste the required standard components and non-standard parts 3D graphics into the fixture assembly folder, and carry out 3D assembly design.

In 3D assembly design, the UG 3D shape of the processed part is first transferred and positioned according to the absolute origin. Then assemble positioning element, clamping element, guiding element and other devices; Then the fixture is designed and assembled according to the assembled standard components. Finally, an explosion diagram is created to facilitate the view of the relationship between the parts in the assembly diagram.

In the assembly process, if there is a position deviation between the matched parts, it can be seen directly that if there is a problem with the parts after selecting the matching conditions, it can be modified directly on the original drawing, and the size and shape of the original parts will be automatically updated on the assembly drawing.

4.2. Design of 3D Assembly Drawing of the Parts and Fixtures of the Back Cover

The design process of 3d assembly drawing of machining rear cover parts and fixtures is as follows.

1) Back cover parts: The system enters the assembly module, and the back cover parts are shown in Figure 4 (a).

2) Design positioning element: One side and two pins are adopted for positioning, limiting six degrees of freedom. Supporting plate and locating pin of positioning element are selected, and the assembly is shown in Figure 4 (b).

3) Design of auxiliary support: Since the workpiece is installed on the side and the workpiece is heavy, to prevent damage to the positioning pin, the auxiliary support is added. The auxiliary supporting elements include support plate, supporting sleeve, adjusting sleeve, etc., as shown in Figure 4 (c) after assembly.

4) Design of clamping elements: adopt the hinge pressing plate mechanism to realize clamping from the top of the workpiece to the positioning surface, select and design B-type arc pressing block, B-type hinge pressing plate, union bolt and other clamping elements, as shown in Figure 4 (d) after assembly.

5) Design guide element: the replaceable drill sleeve shall be selected for 8- Φ 10.5 holes on the bottom surface of drilling, and the fixed drill sleeve shall be selected for 6-M12 threaded bottom holes on the surface a, as shown in Figure 4 (e) after assembly.

6) Specific design clamp: the clamp body is planned to adopt casting blank, and the clamp is integrated with the drilling template. The concave part of the bottom surface is used to install the fixed drill sleeve. On the basis of the L-shaped structure, a vertical plate is added to the left and right to install the turnover bracket, as shown in Figure 4 (f) after assembly.

7) Design of turning mechanism: the hole to be processed is currently placed side by side, and the whole fixture needs to be turned 90° during processing, so the turning

mechanism should be designed. The turnover mechanism includes the supporting part and the operation positioning part, and the main parts include the bracket, shaft, shaft sleeve, rotary disk, positioning pin, guide sleeve, handle, etc., as shown in Figure4 (g) after assembly.

8) Design connection mechanism: it is mainly used to connect the bracket and workbench, and the base and positioning key need to be designed, as shown in Figure 4 (h) after assembly.

9) Create exploded view: exploded view can easily view the parts in the assembly drawing, the assembly relationship between parts, and the number of parts included. The exploded view of the fixture is shown in Figure 4 (i).

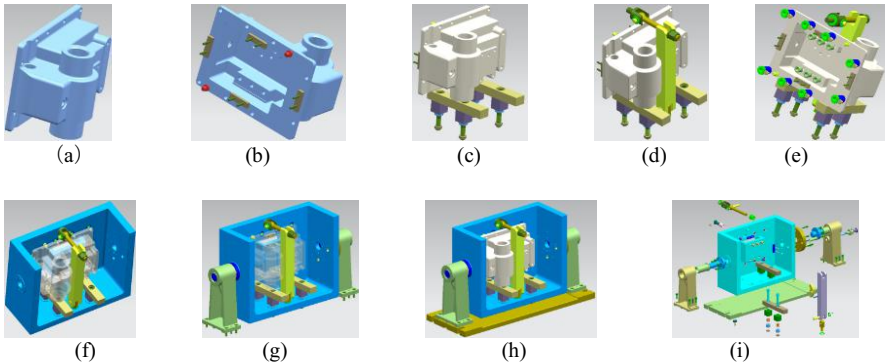


Figure 4. Design of three dimensional assembly drawing of fixture

5. Design of Fixture Technical Documents

5.1. Design Method of Engineering Drawing Based on UG

In UG software, 2D engineering drawings can be directly generated from 3D entities of product design, and when the 3D model changes, the relevant views, dimensions and other information in 2D engineering drawings can be updated simultaneously, making the generation of engineering drawings very simple. The engineering drawing design method based on UG mainly includes setting drawing, parameter presetting, view creation, view annotation and drawing output setting.

5.2 The Design of the Fixture Technical Documents for Machining the Back Cover Parts

Click "Start" to enter the "Drawing" module. First, set basic parameters, including setting drawings, drawing preferences, annotation preferences, etc., and set all contents according to national standards. Then create basic view, section view, dimension label, part number label, drawing frame, title bar, drawing detail list, marking technical requirements, etc. Fixture assembly engineering drawing, as shown in Figure 5. The specific part drawing is shown in Figure 6.

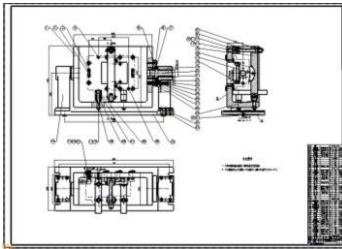


Figure 5. Fixture assembly drawing

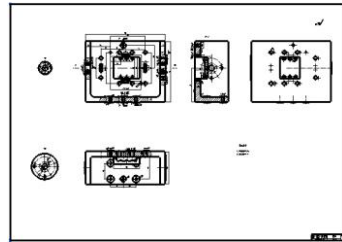


Figure 6. Clamp specific part drawing

6. Epilogue

Based on UG 3D modeling of CNC machine tool fixture design, the fixture standard library is established. During fixture design, the parts in the standard parts library can be directly called, which makes the assembly design process fast, convenient, simple and clear. The design of the fixture adopts the THREE-DIMENSIONAL model, which is not only intuitive and labor-saving, but also can check whether there is interference between the parts to improve the design quality; It is beneficial to improve the design efficiency by using the method of automatic transformation of 3D and 2D models and expert correction to generate 2D assembly drawings conforming to national standards. This method can not only free the fixture design technicians from the complicated and repetitive design labor, but also shorten the fixture development cycle, improve its design and manufacturing level and quality, so as to improve the overall productivity of the fixture industry.

References

- [1] B. Vijaya Ramnath,C. Elanchezian,S. Rajesh,S. Jaya Prakash,B. Manoj Kumaar,K. Rajeshkannan. Design and Development of Milling Fixture for Friction Stir Welding[J]. Materials Today: Proceedings,2018,5(1).
- [2] Osama Halabi,Yasmin Halwani. Design and Implementation of Haptic Virtual Fixtures for Preoperative Surgical Planning[J]. Displays,2018.
- [3] Adelaide Marzano,Parag Vichare,Patricia Muñoz-Escalona. Design and Ergonomic Analysis of the Waste Electrical and Electronic Equipment (WEEE) Remanufacturing Workcell Fixture Using Digital Mock-up Environment[J]. Procedia CIRP,2018,78.
- [4] Sun Yilin, Hu Zengrong, ZHAO Yingyi, et al. Design of Ball hole processing Fixture based on UG NX [J]. Mechanical Engineer, 2013 (10): 66-68.
- [5] Wu Xinteng. Design of special Fixture based on UG NX Platform [J]. Coal Mine Machinery, 2012 (06): 143-14
- [6] Fan Xiaoliang, LIANG Yuhong. Research on Virtual Assembly Technology of Machine Fixture Based on UG [J]. Machinery Design and Manufacturing, 2011 (8): 237-239.

Design and Implementation of Substation Operation Safety Monitoring and Management System Based on Three-Dimensional Reconstruction

Yingyi YANG^{a,b,1}, Hao WU^{a,b}, Fan YANG^c, Xiaoming MAI^{a,b} and Hui CHEN^{a,b}

^aGuangdong Diankeyuan Energy Technology Co., Ltd.

^bElectric Power Research Institute of Guangdong Power Grid Co., Ltd.

^cShaoguan Power Supply Bureau of Guangdong Power Grid Co., Ltd.

Abstract. In order to reduce operational risks and to improve the risk management and control level in substation, a substation operation safety monitoring and management system (3D2S2M) has been structured based on three-dimensional (3D) laser modeling technology. In this paper, we introduce how to build such a system and to describe its implementation details. A 3D lidar scanning technology is used to perform a holographic scan of the whole internal area in a substation to obtain color point cloud data of buildings and all equipment. Then, a novel 3D visualization safety monitoring and management system, named 3D2S2M, is developed by performing a 3D reconstruction of the point cloud data. Based on the real 3D scene model of 3D2S2M, the method of 3D distance measurement is used to replace manual on-site investigation for improving operation and maintenance efficiency. In addition, a real-time high-accuracy localization method is proposed, in order to identify and analyze the positioning and the behavior of the personnel, and the movement trajectory of the equipment. By combining positioning information and the electronic fence that used in 3D2S2M, risk levels of the personnel (or equipment) are evaluated and the corresponding alarm is issued to prevent dangerous behavior, thereby the operation risk is reduced in substation.

Keywords. substation, safety monitoring and management, three-dimensional reconstruction, personnel localization

1. Introduction

Substations are key nodes in the power grid. Their safety and stability greatly influence the whole electric power system. With the continuous development of smart substations [1], a large number of new equipment and increasingly complex building structures in substations have made the intelligent substations present the following new characteristics in substation safety monitoring and management: First, the introduction of the concept of regionalization and hierarchical security makes the construction of substations tend to be functional, which makes some safety management in the substation, such as the joint security dispatch of cyber-physical systems, remote supervision, and security dispatch, more difficult. Specifically, under the concept of regionalized

¹ Corresponding Author. Email: yangyingyi10@126.com

management, the traditional on-site investigation methods and manual dispatching methods in substation are with high security risks, multi-inspection points, and multi-dangerous points. Moreover, traditional inspection carried out artificially is complicated and time-consuming [2-3]. The predictive maintenance is difficult to implement, which makes it difficult to achieve special response strategies formulated by potential problem equipment or conditions. In addition, some special areas are difficult to supervise and personnel safety is difficult to guarantee. Second, safety monitoring and management of substations are faced with huge challenges due to the diverse operating environment and the complex composition of operating personnel. Everyone knows that traditional safety management processes place operating personnel at greater personal risks. And now, in the intelligent substation, diverse operating environment will make the operators with different job responsibilities work in different operating areas. In this case, it is difficult to know accurately whether the personnel are on the job, and there is a risk of accidents resulting in personal injury or death caused by misoperation. In addition, complex personnel entering and leaving will also make a huge threat to the safety of equipment in the operation inspection area. All these reasons can bring challenges to traditional security management.

Personnel positioning is a very essential issue in substation safety management. In some power facility monitoring systems that adopt some simple image processing algorithms, the monitoring of personnel targets stays in a situation where a small number of targets are resolved and classification is easy, and once a personnel target is detected, an alarm is issued to him in any case. It is obviously not reasonable, and these methods cannot monitor the behavior of personnel targets [4].

Therefore, there is an urgent need to develop a new safety monitoring and management system that can simultaneously control risk, prevent violations and avoid accidents. Virtual visualization technology based on three-dimensional reconstruction has become an effective solution to these problems [5].

2. Related Work

In recent years, three-dimensional real scene reconstruction of substation has been becoming one of the most popular techniques in intelligent substation safety management [6-7]. It mainly includes: geometric structure modeling method, VRM modeling method [8] and laser scanning modeling method [9]. Different methods have different principles, and their performance and application areas are also different [10]. Quintana J *et al.* [11] developed a substation simulation training system based on the geometric shapes of different electrical equipment. Wang *et al.* [12] enriched the electrical equipment material model based on the virtual reality technology to complete the 3D modeling of the substation, and added scene special effects, defect simulations, and other functions. The geometric structure modeling method does not need to separately model different electrical equipment, and can directly select a specific geometric structure, which greatly improves the modeling efficiency and the overall reconstruction effect, but the reconstruction precision is not high, and it is mostly used for simulation training [13]. Li [14] used a 3D laser scanner to scan the substation, and used an improved particle swarm algorithm to complete the automatic identification and classification of the 3D point cloud, and finally realized the 3D reconstruction of the substation, which improved the reconstruction efficiency and precision. Du *et al.* [15] also used a 3D laser scanner to scan a substation to obtain a 3D point cloud, and used

virtual simulation technology to simulate the staff's inspection and maintenance of electrical equipment in a reconstructed 3D scene. This method of acquiring and reconstructing a 3D point cloud of a substation by a 3D laser scanner has high reconstruction precision, accurate 3D positioning, and high degree of detail reduction of electrical equipment, and is suitable for practical engineering applications such as site monitoring, positioning, and measurement of a substation.

In order to improve the substation operation safety risk management level and to reduce the operation risk, accurate positioning of activity targets in the substation should be implemented in a substation operation safety monitoring and management system [16]. At present, in terms of substation personnel (or machine) safety monitoring, Zhou *et al.* [17] combined Zigbee positioning technology with fuzzy inference algorithm for real-time positioning of field workers. But this method has low positioning accuracy and is not convenient for the actual operation. Huang [18] proposed an ultrasonic orthogonal spread-spectrum coding technology based on pseudo-random modulation carrier, and built a precise positioning model for the staff in the substation. This method locates accurately. But the accuracy decays rapidly due to the decay of the ultrasonic wave propagation. Although the positioning interval is good, the accuracy is greatly reduced, so it is only suitable for positioning in specific scenarios of substations [18]. Yuan *et al.* [19] installed wireless sensor network (WSN) nodes in key areas of the substation to receive signals in real time, and completed background real-time location monitoring of field workers through signal conversion. However, this method has high equipment costs, and the monitoring methods still have a certain degree of lag. Zhang *et al.* [20] used a multi-module and multi-frequency GNSS (Global Navigation Satellite System) module equipped with GPS (Global Positioning System) or Beidou, and adopted high-precision outdoor differential positioning technology to complete real-time positioning for substation staff which has higher positioning accuracy. Du *et al.* [21] used a binocular stereo vision camera combined with an inertial measurement device to complete real-time positioning and predictive tracking of workers.

3. The Framework of 3D2S2M

The new 3D2S2M uses a 3D laser scanning technology to perform a holographic scan of the entire substation to obtain high-precision three-dimensional color point cloud data [22]. Through 3D reconstruction and importing the equipment ledger data, a 3D visualization platform for substation safety monitoring and management is constructed. Meanwhile, multiple related functional modules are developed based on the 3D visualization platform. For example, the virtual on-site survey module is developed to replace the inefficient and high-risk manual on-site investigation. The high-precise real-time localization module for personnel (or machine) are designed to intelligently identify and analyze the behavior of personnel and the trajectory of equipment in the substation, and thus the level of risk management and control of substation are improved. In risk level assessment and real-time alarm module, based on the data of equipment ledger and safety live distance, combined with the other data, such as positioning data of operator, vehicle or equipment, positioning delay, motion trajectory and inertia parameters, a corresponding motion model is established and appropriate operational risk early warning strategies are formulated. It is worth noting that the interface module is

responsible for data exchange between the new system and all external systems. The overall framework of the new system is shown in Figure 1.

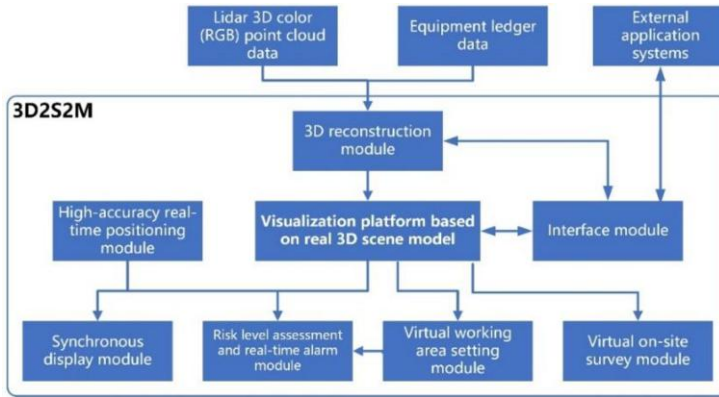


Figure 1. The framework of 3D2S2M.

4. Implementation of 3D2S2M

4.1. Three-dimensional Real Scene Reproduction of Substation

Modeling precision is the decisive factor for reconstructing three-dimensional real scene of the whole substation. To gather the 3D color (RGB) point cloud data with precise coordinates for all objects in substation, this new system adopts high-precision 3D lidar scanning device (error is $\pm 2\text{cm}$) to perform a full-substation mobile scanning. Through the holographic scanning for all objects in the substation, including various electrical equipment, site layout, actual structure, and so on, the 3D scene of the substation is accurately reconstructed based on the spatial coordinate information of latitude, longitude, height, and depth. The specification of lidar sensor used is listed in Table 1.

Table 1. The specification of lidar sensor used.

Parameters	Specifications
range type	phase
laser class	1R eye-safe laser
scan scope	150 meters
range error	$\pm 1\text{mm}$
range noise	$\leq 0.3\text{mm}$ within 25meters
field angle	vertical 300° , horizontal 360°

The core technology of the 3D2S2M is the high-precision reconstruction of the real scene inside the substation, including the coordinates, size and even the operating conditions of equipment. Therefore, the ID of the 3D virtual device in the 3D2S2M must be the same as the ID number of the real device in the substation. That is to say, 3D virtual devices need to be bound to the actual device by using the same ID to access the equipment ledger data and monitoring data. Equipment ledger data are inserted into the 3D model using offline copying method, so that the ledger information can be displayed intuitively in 3D2S2M.

4.2. Measurement Analysis in Three-dimensional Space for Virtual On-site Survey

The on-site investigation in substation consumes large costs, such as manpower and material resources. Multiple on-site investigations before the maintenance operation would be time consuming and inefficient.

In the 3D2S2M, virtual on-site investigation can be performed to replace the on-site Investigation: measurement tools in the 3D2S2M are used to obtain 3D spatial distance parameters (including manual measurements, air-ground ranging,) to achieve the goals of field inspection and measurement, and then make accurate assessments and decisions about field operations. Accurate spatial distance information can also be used to evaluate and calculate various on-site operations, which will greatly improve the work efficiency.

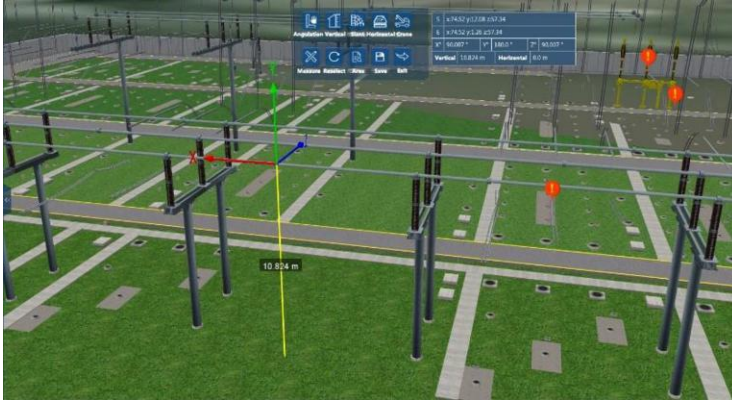


Figure 2. Illustration of distance measurement.



Figure 3. Illustration of the virtual working area setting based on three-dimensional electric fence.

For example, when ranging in our three-dimensional scene, let S be the starting point and E be the end point. In the 3D2S2M's 3D scene, the values of their X , Y , and Z coordinate axes will be displayed accordingly. Our angle of X (or Y , or Z) is defined as the angle between the line connecting these two points and the X axis (or Y , or Z). The distance between the two points is defined as the straight-line distance between the two points. Moreover, the horizontal distance is defined as the distance between the two points in the same horizontal direction. Figure 2 shows an example of distance measurement in 3D2S2M. The measurement result is 10.824m in this example.

4.3. Setting up the Virtual Working Area

The 3D2S2M sets the virtual working area by setting a 3D virtual electric fence. According to the characteristics of the operation, the electric fence is divided into [No Entry], [Ban Out], [Live Equipment] and other types. At the same time, these areas can be named, and its period of validity and the violation contents can also be set, such as illegal stay time, out time, safety distance, and so on. A virtual working area is illustrated as Figure 3.

4.4. High-accuracy Personnel (or Equipment) Positioning

4.4.1. Calculating the High-accuracy Coordinates of Personnel (or Equipment)

In this paper, the high- accuracy positioning of personnel (or equipment) in the 3D scene is achieved by using the data gathered from the GPRS (General Packet Radio Service) mobile terminal with a high-accuracy positioning device. The geographic location information obtained by the GPS receiving module is transmitted to the server (i.e. base station) through the GPRS wireless communication network, and processed on the server to achieve remote positioning [23]. Data accuracy of the high-accuracy positioning device in the GPRS mobile terminal can reach the centimeter level. The RTK differential base station (BS) receives the GPS + GLONASS (GLOBAL NAVIGATION SATELLITE SYSTEM) + BeiDou satellite signal, then decodes them by using the algorithm stored in its MCU1, and outputs RTK differential signal in RTCM format. The mobile terminal receives the GPS + GLONASS satellite signals and the RTK differential signals from the BS, and decodes them by using the algorithm stored in its MCU2. The whole positioning process can be summarized as follows: Firstly, calculating the known precise coordinates of the base station by using the following formula:

$$\tilde{\rho}_{Ri}^j = \rho_{Ri}^j + c\delta t_i + d_{ion} + d_{trop} + \varepsilon_{\rho} \quad (1)$$

where $\tilde{\rho}_{Ri}^j$ is the precise coordinates, ρ_{Ri}^j is the measured value of the coordinates, $c\delta t_i$ is the satellite clock error, d_{ion} is the ionospheric error, d_{trop} is the tropospheric error and ε_{ρ} is coordinates of the base station.

Secondly, the error from the base station to the satellite is calculated by the following formula:

$$\Delta P_{err} = \tilde{\rho}_{Ri}^j - \rho_{Ri}^j = c\delta t_i + d_{ion} + d_{trop} + \varepsilon_{\rho} \quad (2)$$

At last, the mobile terminal corrects its localization result by receiving the error from BS to improve its positioning accuracy by the formula:

$$\tilde{\rho}_{Mi}^j = \rho_{Mi}^j + \Delta P_{err} = \rho_{Mi}^j + c\delta t_i + d_{ion} + d_{trop} + \varepsilon_{\rho} \quad (3)$$

where ρ_{Mi}^j is the measured coordinates of the target, and $\tilde{\rho}_{Mi}^j$ is the corrected coordinates of the target.

4.4.2. Locating Personnel (or Equipment) in 3D Scene

In a 3D scene, the longitude and latitude of the ordinate origin is used as the reference, and the coordinates of the target (personnel or equipment) in the scene are converted from their latitude and longitude.

Let O be the origin of the coordinates, its longitude and latitude are O_{la} and O_{lo} respectively, and its coordinates in the 3D scene are $(0, 0, 0)$; the corrected longitude and latitude of the target (M) are M_{la} , M_{lo} , its coordinates in 3D scene are $(Mx, My, Mz (= 0))$. Then, the x and y coordinates of M can be calculated by the following formula:

$$\begin{aligned} Mx &= (M_{la} - O_{la}) \times K1 \\ My &= (M_{lo} - O_{lo}) \times K2 \end{aligned} \quad (4)$$

where $K1$ and $K2$ are the magnifications of the x -axis and y -axis of the 3D scene and the real scene, respectively.

For working equipment with a height such as a crane, the height must be marked by tools such as laser detection and image recognition. The localization results of personnel and equipment are shown in Figure 4.



Figure 4. Illustration of positioning of personnel and equipment.

In addition, based on the real-time positioning data of the target, the 3D2S2M can support the identity binding and real-time positioning of the on-site personnel (or equipment). Meanwhile, it also supports the review of the historical trajectory of personnel movement, and the viewing of the specific information and alarm records of a node in the historical trajectory can facilitate work performance assessment and security risk assessment.

High-accuracy positioning coordinates can be mapped to the 3D scene accurately and presented visually by adopting our satellite positioning method. Here satellite positioning we adopted is compared with other major methods, including UWB (Ultra Wide Band), RFID (Radio Frequency Identification) and WIFI. Table 2 shows the comparison among those methods for substations in terms of accuracy, cost, stability and real-time.

Table 2. The comparison among satellite, UWB, RFID and WIFI.

Methods	Accuracy	Cost	Stability	Real-time
satellite	centimetre level	relatively lower	good	good
UWB	decimeter level/ centimetre level	higher	relatively stable	good
RFID	decimeter level	very low	relatively stable	good
WIFI	decimeter level	relatively lower	poor	poor

As shown in Table 2, in the situation of cruel electromagnetism environment in the substations, satellite positioning method we adopted achieves centimetre level, the best one among all the methods. Although UWB method presents the level close to that of satellite method, it costs more. As for RFID and WIFI method, both have lower accuracy. RFID has poor real-time performance while WIFI has poor stability.

4.5. Assessing Risk Levels and Real-time Alerts

3D2S2M aims at the safety monitoring and management of the on-site operation region in substations. The new system performs big data statistical analysis and historical risk data analysis, and assesses the risk level based on the movement trajectory of personnel and equipment.

Using historical big data about massive filed tasks, 3D2S2M statistically analyzes personnel path trajectories and working areas based on positioning coordinates. The best path trajectory and safe working area for a specific task can then be predicted with Markov algorithm based on statistical analysis results. The system classifies the potential risks in actual execution, including too close distance between personnel and equipment, straying into non-working areas, diverging from the normal trajectory and sticking around abnormally for a long time, from highest to lowest priority and taking layered pre-warning accordingly. Besides, clustering analysis on coordinates and trajectory similarity comparison are also adopted by 3D2S2M for hotspot judgment, thus finding abnormal working places and raising alarm in time.

For high-risk and accident-prone locations, visually display the heat map of the working area in our 3D scene, and part of them are displayed as the work blind zone. By analyzing personnel and equipment movement trajectory in combination with virtual electronic fences, equipment operating data, our 3D2S2M can dynamically judge operating risks and trigger alarms in time. Besides, the 3D2S2M can also realize the synchronization display of the movements (attitudes) of large machine (such as cranes used in field operations, baskets, boom frames, and boom frames of curved arm climbing car) with vehicle real 3D scene models (such as cranes). Therefore, the 3D2S2M can evaluate the risk of triggering operations by dynamically judging the spatial collision trend in the live area of the equipment, so that the management personnel can use the 3D scene to realize the remote real-time grasp of on-site operations.

When the positioning coordinates of personnel (or equipment) change, for example, when the on-site personnel (or equipment) deviate from his preset routes, stray into the charged interval, or intrude into the safe distance of charged equipment, the new system will perform collision detection on the area edge of the coordinates of personnel (or equipment) according the virtual operating area (electronic fence) preset in the 3D scene. When a person (or equipment) enters a dangerous area or goes out of a preset working area by mistake, the system will perform a security risk assessment giving a corresponding real-time alarm.

5. Conclusion

In this paper, we designed and implemented a new 3D visualization safety monitoring and management system for substation, named 3D2S2M based on a 3D reconstruction technology. The new system uses a high-precision 3D lidar scanning equipment to perform an entire station holographic scanning to obtain high-precision laser point cloud data. Based on the point cloud data, a high-accuracy reconstruction of the real scene of the substation is achieved. The new system has the following main functions: virtual 3D space distance measurement is used to replace on-site investigation. Electronic fences are used to set up virtual working areas. High-precision real-time positioning of personnel (or instruments) is achieved. Positioning data and electronic fences are combined to monitor and evaluate the potential risks in the substation and give corresponding real-time warnings. In addition, the new system also realizes the intelligent visualization of operators and equipment in the real 3D scene, which is in line with the development trend of more friendly, more humane and more convenient security monitoring and management in substation, besides the visual management.

Acknowledgment

This work was supported by China Southern Power Grid Co. Ltd. Science and Technology Program under Grant No. GDKJXM20184850.

References

- [1] X.X. Gao, F. Q. Chen, Design and implementation of a remote positioning system based on GPRS + GPS, *China Science and Technology Information*, **16** (2015),85-87.
- [2] C. Zhao, Study on operation maintenance technology of intelligent substation, Tianjin University, 2011.
- [3] X. Li, Study on intelligent reconstruction of substation adapted to ubiquitous IoT, Shandong University, 2019.
- [4] L. Zhang, Y. C. Feng, W. T. Qu, Q. K. Shun, Design of aided identification and positioning software for substation operation safety management and control system, *Electrotechnical Technology*, **4** (2016), 29-30.
- [5] Y. Gao, Application of 3D digital technology in the substation design, *IOP Conference Series Materials Science and Engineering*, **782** (2020), 032-066.
- [6] J. H. Fang, W. M. Xu, X. F. Zhu, et al., Research and application of intelligent auxiliary control system for substation based on internet of things and 3D visualization technology, *Electric Automation*, **3** (2012), 71-74.
- [7] L. Bai, Research and implementation of 3D digital design of substation, Tianjin: Tianjin University, 2013.
- [8] H. Dai, Z. Yang, Z. Xue, et al., Research on key technologies of substation 3D interactive scene simulation based on virtual reality technology, *Journal of Physics Conference Series*, **1345** (2019), 062036.
- [9] X. Suo, C. Hou, L. Sun, et al., 3D reconstruction optimization algorithm based on dynamic clustering in transformer substation[J]. *Journal of Computational and Theoretical Nanoence*, 2017.
- [10] R. D. Wang, Y. Du, X. J. Shen, Status and prospect of 3D modeling methods for substation, *North China Electric Power Technology*, **2** (2015), 19-23.
- [11] J. Quintana, E. Mendoza, 3D virtual models applied in power substation projects, *IEEE 15th International Conference on Intelligent System Applications to Power Systems, ISAP'09*, Curitiba, Brazil, 2009:1-3.
- [12] X. B. Wang, X. D. Zhang, T. He, et al., Digital visual management and monitoring system for 3D virtual substation, *Journal of Wuhan University*, **6** (2012), 786-791.
- [13] X. N. ShiKang, Design and implementation of 3D panoramic view software for substation management system, Xi'an: Xidian University, 2015.

- [14] L. X. Li, Research on classification and identification of substation equipment based on 3D point cloud, Zhengzhou: Zhengzhou University, 2016.
- [15] Y. Du, Z. Chen, F. Liu, et al., Three-dimensional real-world deduction simulation of substation operation and risk management and control, *Journal of Huaqiao University*, **2** (2017), 236-240.
- [16] W. H. Hu, Design and implementation of intelligent video surveillance system for unattended substations, Changchun: Jilin University, 2015.
- [17] X. Y. Zhou, Research on rapid positioning method and positioning system for mobile personnel in digital substation, Chongqing: Chongqing University, 2013.
- [18] X. L. Huang, Research and application of safety protection and monitoring technology for substation operation and maintenance site, Baoding: North China Electric Power University, 2013.
- [19] S. K. Yuan, N. N. Gong, Z. Li, et al., Substation personnel positioning and tracking system based on wireless sensor network, *System Development and Application*, **3** (2013), 50-53.
- [20] Y. Zhang, J. H. Hu, C. G. Peng, Research on 3D real-time field operator positioning system for UHV substation based on accurate coordinates, *Science & Technology Innovation Herald*, **21** (2017), 1-2.
- [21] J. J. Yan, Design of radio frequency identification and location system for substation personnel, *Applied Technology*, **5** (2015), 94-95.
- [22] O. Vozisova, A. Egorov, D. Snegirev, E. Lyuhanov, E. Kochneva, Substation and electric systems equipment 3D laser scanning, *2016 International Symposium on Industrial Electronics (INDEL)*, Banja Luka, 2016, pp. 1-4.
- [23] B. Y. Chen, F. Y. Zhang, Application of cloud model to substation operation risk perception, *Journal of Electric Power Systems and Automation*, **7** (2019), 23-29.

Modeling of a Leg and Knee System for the Analysis of Human Gait by Means of State Feedback Control

P. A. Ospina-Henao ^{a,1}, César H. Valencia ^b, Marcelo Becker ^c, Zuly A. Mora P ^b, and S. M. Vásquez ^b

^a *Departamento de Ciencias Básicas, Universidad Santo Tomás. Carrera 18 No. 9 - 27. PBX: (+57)76985858. Bucaramanga. Colombia.*

^b *Facultad de Ingeniería Mecatrónica, Universidad Santo Tomás. Carrera 18 No. 9 - 27. PBX: (+57)76985858. Bucaramanga. Colombia.*

^c *Departamento de Engenharia Mecânica (SEM), University of São Paulo (USP), PBX: (+55)1633738646. São Carlos-SP. Brazil.*

Abstract. This paper presents the modeling, simulation and control of a human gait system, which consists of the modeling of a leg by means of Euler's classical mechanics and Lagrange's formalism, where the equations of motion of the joint are obtained both of the hip as of the knee and the solution of these. In addition, a state feedback control was implemented and the controller gains were determined by means of the Ackerman formula, based on the equations of motion rewritten in state space and simulated in simulink, where the behavior of the system can be observed with control.

Keywords. Control, Feedback of States, Formalism Euler - Lagrange, Human Gait.

1. Introduction

Human gait is a complex process that integrates the relationship of various subsystems of the human body necessary to generate biped movement. In addition, this locomotion process is marked by some phases, which are the support phase and the equilibrium phase that is selected in a series of stages that begin with the heel contact with the ground and end with the takeoff of the foot [1]. This research seeks to provide a dynamic model that describes the movement of the leg joints such as the hip and knee using the Euler-Lagrange formalism and a control, which serves as a support to reproduce human gait in an electromechanical environment. The objective was to investigate and describe a series of basic mechanical and physiological mechanisms behind human walking. The methodologies used were the biomechanical analysis of movement and physiology [2]. It is important to note that one leg dynamics is only one of many possible approaches to its study. The pathological approach integrates simple principles, clinical observations, and

¹Corresponding Author: Departamento de Ciencias Básicas, Universidad Santo Tomás. Carrera 18 No. 9 - 27. PBX: (+57) 76985858. Bucaramanga. Colombia. E-mail: paolo.ospina@ustabuca.edu.co

controlled experiments performed to explain the advantages of the method. The dynamic study and control by state feeding that we present here show how the dynamic walking approach could be applied to integrative study of gait pathologies [3]. These examples can be seen as starting points for new research, where experimental evidence will be indispensable [4]-[5]. Simple principles can't explain everything of the complexities of the march, but can help ask helpful questions what you have to answer to understand those complexities. There are two approaches to solving the direct and inverse dynamic model. A direct dynamic model is one that expresses the temporal evolution of joint coordinates as a function of the forces and torques involved, and an inverse dynamic model is one that expresses the forces and torques that intervene as a function of the evolution of joint coordinates and their derivatives [6]-[7].

It should be noted that the data obtained from this study is based on a simulation that uses computational tools. The simulation is based on a mathematical model that indicates a change between potential energy and kinetic energy, resembling the action of a double pendulum. This, with the evolution of this type of models in the future, they will be usable in the clinical context either for the training of medical personnel or to plan adjustments in the alignments that individuals may undergo, without presenting risks to patient stability [8]. The gait model produced with Solidworks was of great help in simulating the behavior of the leg system [9].

A model of the double pendulum in two dimensions is the pendulum-like motion of the swinging leg. The same conservation of mechanical energy is applied, so little work is necessary to move the swinging leg. Given a suitable starting position and velocity, the entire single limb support phase can be produced largely through the movement of 2 coupled pendulums representing the posture and swing of the leg. This modeling approach suggests that both the hip and the legs can take advantage of the pendulum dynamics during single-limb support [10] - [11].

During the present work several aspects of study were taken into account, starting in the section 2, by the dynamic modeling of the system through the Euler-Lagrange equations in a non-conservative way, the numerical model, the matrix representation of states and the graphic analysis of the numerical solutions were carried out.

Next, in section 3, the control of the dynamic system is shown, a rigorous study of the controllability and observability of the system is carried out, as well as the dominant poles and the gain matrix that allows designing the controller are calculated. The simulation of the human walking system is performed by applying the controller to an input signal. In the section 4, we present the conclusions of the work.

2. Symbolic model

In clinical routine, gait analysis identifies normal or pathological movements. From a simplified model of the structure of the human body, perform this analysis at different levels and complementary techniques that evaluate different aspects of neuromuscular function, the mechanical model is described below. From Figure 1, we can take the following measurements: x_1 and y_1 are the coordinates of the center of mass of the upper leg (femur) and in the same way, x_2 and y_2 are the coordinates of the center of mass of the lower leg (tibia and fibula).

We can start with taking some measurements from the drawing,

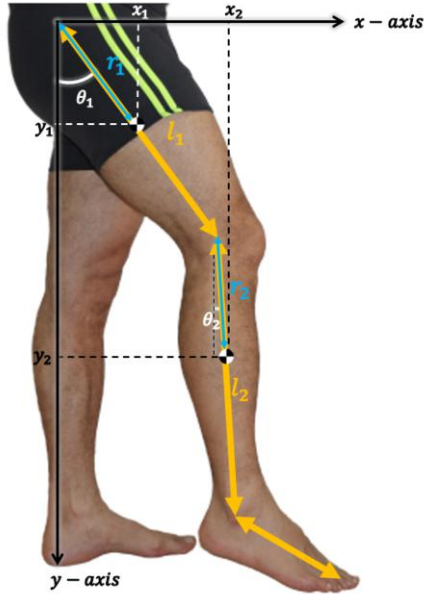


Figure 1. Initial system image

then we obtain the linear speed of the first link or the first joint,

$$v_1^2 = \dot{x}_1^2 + \dot{y}_1^2 = r_1^2 \dot{\theta}_1^2, \tag{1}$$

and also, the linear speed of the second link,

$$v_2^2 = \dot{x}_2^2 + \dot{y}_2^2 = l_1^2 \dot{\theta}_1^2 + r_2^2 \dot{\theta}_2^2 + 2l_1 r_2 \dot{\theta}_1 \dot{\theta}_2 \cos(\theta_1 - \theta_2), \tag{2}$$

2.1. Kinetic energy

Kinetic energy is defined as

$$K = \frac{1}{2} m_1 v_1^2 + \frac{1}{2} m_2 v_2^2,$$

replacing Eqs. (1) and (2), in the previous expression we obtain,

$$K = \frac{1}{2} m_1 r_1^2 \dot{\theta}_1^2 + \frac{1}{2} m_2 [l_1^2 \dot{\theta}_1^2 + r_2^2 \dot{\theta}_2^2 + 2l_1 r_2 \dot{\theta}_1 \dot{\theta}_2 \cos(\theta_1 - \theta_2)]. \tag{3}$$

2.2. Gravitational potential energy

Gravitational potential energy is defined as

$$U = m_1 g y_1 + m_2 g y_2,$$

getting,

$$U = -m_1 g r_1 \cos \theta_1 - m_2 g l_1 \cos \theta_1 - m_2 g r_2 \cos \theta_2. \tag{4}$$

2.3. Lagrangian density and equations of motion

Lagrangian density is defined [12],

$$L = K - U,$$

replacing Eqs. (3) and (4), in the previous expression we obtain

$$L = \frac{1}{2}m_1r_1^2\dot{\theta}_1^2 + \frac{1}{2}m_2[l_1^2\dot{\theta}_1^2 + r_2^2\dot{\theta}_2^2 + 2l_1r_2\dot{\theta}_1\dot{\theta}_2 \cos(\theta_1 - \theta_2)] + m_1gr_1 \cos \theta_1 + m_2gl_1 \cos \theta_1 + m_2gr_2 \cos \theta_2. \tag{5}$$

The Euler-Lagrange equations of motion have the form,

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{\theta}_i} \right) - \frac{\partial L}{\partial \theta_i} + \frac{\partial D_i}{\partial \dot{\theta}_i} = \tau_i, \quad i = 1, 2. \tag{6}$$

here, D_i is the Rayleigh dissipation function and it is only considered the viscous case [13], in other words, the friction, we can define Rayleigh function as a homogeneous quadratic form at generalized speeds, this results as a theoretical suggestion and they could be calculated in practice. With this in mind, Rayleigh’s dissipation function can take the form,

$$f_r = \frac{\partial D_i}{\partial \dot{q}_i} = \frac{1}{2}\beta_i\dot{q}_i^2. \tag{7}$$

From it, we can define the forces acting on the system, we must remember that the friction forces are proportional to a power of the instantaneous velocity, $f_r = -bv^n$. In our case, these friction forces are related to the forces delivered to each link in the system. For our case, we have two generalized coordinates that are (θ_1, θ_2) , resulting in the following equations,

resulting the motion equations,

$$\tau_1 = \ddot{\theta}_1(m_1r_1^2 + m_2l_1^2) + m_2l_1r_2\ddot{\theta}_2 \cos(\theta_1 - \theta_2) + m_2l_1r_2\dot{\theta}_2^2 \sin(\theta_1 - \theta_2) + (m_1r_1 + m_2l_1)g \sin \theta_1 + \beta_1\dot{\theta}_1. \tag{8}$$

$$\tau_2 = \ddot{\theta}_1m_2l_1r_2 \cos(\theta_1 - \theta_2) + m_2r_2^2\ddot{\theta}_2 - m_2l_1r_2\dot{\theta}_1^2 \sin(\theta_1 - \theta_2) + m_2r_2g \sin \theta_2 + \beta_2\dot{\theta}_2. \tag{9}$$

2.4. Numerical model

To define the numerical model, the constant values are replaced,

Table 1. System variables.

Variable	Value	Units
m_1	8	[kg]
m_2	3.72	[kg]
l_1	0.45	[m]
r_1	0.195	[m]
r_2	0.165	[m]
g	9.81	[m/s ²]
β_1	2.288	[kg/s]
β_2	0.175	[kg/s]

here, m_1 is thigh mass, m_2 the calf mass, l_1 thigh length, r_1 length to center of mass of thigh, r_2 length to center of mass of calf, g the gravitational acceleration, β_1 hip coefficient of friction and β_2 knee friction coefficient. The measurements were taken for a male person 1.77 m tall and 80 kg, the mass of the thigh segment and the calf next to the center of mass of each one was taken according to the synthesis of classic anthropometry works presented by Winter [14]. Replacing the values of the constants of Table 1, in Eqs. (8) and (9), we obtain,

$$\tau_1 = 31.72554 \sin \theta_1 + 0.27621\dot{\theta}_2^2 \sin(\theta_1 - \theta_2) + 0.27621\ddot{\theta}_2 \cos(\theta_1 - \theta_2) + 2.288\dot{\theta}_1 + 1.0575\ddot{\theta}_1 \tag{10}$$

$$\tau_2 = 6.021378 \sin \theta_2 - 0.27621\dot{\theta}_1^2 \sin(\theta_1 - \theta_2) + 0.27621\ddot{\theta}_1 \cos(\theta_1 - \theta_2) + 0.175\dot{\theta}_2 + 0.101277\ddot{\theta}_2 \tag{11}$$

here, Eqs. (10) and (11), represent a system of coupled partial differential equations.

2.5. Matrix representation in state variables

The matrix representation of the dynamic model is given by [15]

$$M(\theta)\ddot{\theta} + C(\theta, \dot{\theta})\dot{\theta} + G(\theta) = \tau, \tag{12}$$

then, we can rewrite the system of Eqs. (12) like this,

$$\begin{bmatrix} m_1r_1^2 + m_2l_1^2 & m_2l_1r_2 \cos(\theta_1 - \theta_2) \\ m_2l_1r_2 \cos(\theta_1 - \theta_2) & m_2r_2^2 \end{bmatrix} \begin{bmatrix} \ddot{\theta}_1 \\ \ddot{\theta}_2 \end{bmatrix} + \begin{bmatrix} \beta_1 & m_2l_1r_2\dot{\theta}_2 \sin(\theta_1 - \theta_2) \\ -m_2l_1r_2\dot{\theta}_1 \sin(\theta_1 - \theta_2) & \beta_2 \end{bmatrix} \begin{bmatrix} \dot{\theta}_1 \\ \dot{\theta}_2 \end{bmatrix} + \begin{bmatrix} (m_1r_1 + m_2l_1)g \sin \theta_1 \\ m_2r_2g \sin \theta_2 \end{bmatrix} = \begin{bmatrix} \tau_1 \\ \tau_2 \end{bmatrix} \tag{13}$$

The state variables of a system are generally critical variables, of interest to be controlled, observed or that have an important relevance in the system. It is essential to keep

in mind that these variables may not be accessible to the user; that is, they cannot be measured with sensors, so in these cases the use of state observers is necessary. In the case of leg prostheses, four variables are critical, which are the position and angular velocity of each link. For these four state variables the use of an observer is not necessary, since physically they can be measured with encoders and there is the concept of classical mechanics that the position is the integral of velocity, and velocity is the integral of acceleration.

Clearing the accelerations, we get,

$$\ddot{\theta}_1 = \frac{\tau_1 - m_2 l_1 r_2 \dot{\theta}_2^2 \sin(\theta_1 - \theta_2) - \beta_1 \dot{\theta}_1 - (m_2 l_1 + m_1 r_1) g \sin \theta_1}{m_2 l_1^2 + m_1 r_1^2 - m_2 l_1^2 \cos^2(\theta_1 - \theta_2)} - \frac{l_1 \cos(\theta_1 - \theta_2) [m_2 l_1 r_2 \dot{\theta}_1^2 \sin(\theta_1 - \theta_2) + \tau_2 - \beta_2 \dot{\theta}_2 - m_2 r_2 g \sin \theta_2]}{m_2 r_2 l_1^2 + m_1 r_1^2 r_2 - m_2 r_2 l_1^2 \cos^2(\theta_1 - \theta_2)}, \tag{14}$$

$$\ddot{\theta}_2 = \frac{(m_1 r_1^2 + m_2 l_1^2) [m_2 l_1 r_2 \dot{\theta}_1^2 \sin(\theta_1 - \theta_2) + \tau_2 - \beta_2 \dot{\theta}_2 - m_2 r_2 g \sin \theta_2]}{m_2^2 l_1^2 r_2^2 + m_1 m_2 r_1^2 r_2^2 - m_2^2 l_1^2 r_2^2 \cos^2(\theta_1 - \theta_2)} + \frac{l_1 \cos(\theta_1 - \theta_2) [m_2 l_1 r_2 \dot{\theta}_2^2 \sin(\theta_1 - \theta_2) - \tau_1 + \beta_1 \dot{\theta}_1 + (m_1 r_1 + m_2 l_1) g \sin \theta_1]}{m_2 r_2 l_1^2 + m_1 r_1^2 r_2 - m_2 r_2 l_1^2 \cos^2(\theta_1 - \theta_2)}. \tag{15}$$

Because the theoretical model obtained is nonlinear, it is required to linearize it for the design of the controller in state space. The linearization of the system was performed around its unstable equilibrium point using Taylor’s series, where only the first term of the series is taken since it is the linear term. The values of the state variables taken for the linear approximation are (0, 0, 0, 0), these values are due to the fact that the speeds must be zero for the system to remain in the desired position. Based on the above, the calculation of the linear model will be based on the state variables, which have the form,

$$\mathbf{x} = \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \\ \mathbf{x}_3 \\ \mathbf{x}_4 \end{bmatrix} = \begin{bmatrix} \theta_1 \\ \dot{\theta}_1 \\ \theta_2 \\ \dot{\theta}_2 \end{bmatrix} = \begin{bmatrix} \theta_1 \\ \omega_1 \\ \theta_2 \\ \omega_2 \end{bmatrix}, \tag{16}$$

and the equations of state are,

$$\dot{\mathbf{x}} = \begin{bmatrix} \dot{\mathbf{x}}_1 \\ \dot{\mathbf{x}}_2 \\ \dot{\mathbf{x}}_3 \\ \dot{\mathbf{x}}_4 \end{bmatrix} = \begin{bmatrix} \dot{\theta}_1 \\ \ddot{\theta}_1 \\ \dot{\theta}_2 \\ \ddot{\theta}_2 \end{bmatrix} = \begin{bmatrix} \dot{\theta}_1 \\ \dot{\omega}_1 \\ \dot{\theta}_2 \\ \dot{\omega}_2 \end{bmatrix}. \tag{17}$$

To reduce the derivatives of the differential equations by one degree Eqs. (14) and (15), which represent the dynamic solution of the system, we define the following vectors of state, $\omega_1 = \dot{\theta}_1$, $\omega_2 = \dot{\theta}_2$, $\dot{\omega}_1 = \ddot{\theta}_1$, $\dot{\omega}_2 = \ddot{\theta}_2$, which physically represent the angular velocities and accelerations of both links. Once the state vectors have been defined, the Eqs. (14) and (15), can be rewritten as follows,

$$\dot{\mathbf{x}}_1 = \mathbf{x}_2, \tag{18}$$

$$\begin{aligned} \dot{\mathbf{x}}_2 = \dot{\omega}_1 = & \frac{\tau_1 - m_2 l_1 r_2 \omega_2^2 \sin(\theta_1 - \theta_2) - \beta_1 \omega_1 - (m_2 l_1 + m_1 r_1) g \sin \theta_1}{m_2 l_1^2 + m_1 r_1^2 - m_2 l_1^2 \cos^2(\theta_1 - \theta_2)} - \\ & - \frac{l_1 \cos(\theta_1 - \theta_2) [m_2 l_1 r_2 \omega_1^2 \sin(\theta_1 - \theta_2) + \tau_2 - \beta_2 \omega_2 - m_2 r_2 g \sin \theta_2]}{m_2 r_2 l_1^2 + m_1 r_1^2 r_2 - m_2 r_2 l_1^2 \cos^2(\theta_1 - \theta_2)}, \end{aligned} \tag{19}$$

$$\dot{\mathbf{x}}_3 = \mathbf{x}_4, \tag{20}$$

$$\begin{aligned} \dot{\mathbf{x}}_4 = \dot{\omega}_2 = & \frac{(m_1 r_1^2 + m_2 l_1^2) [m_2 l_1 r_2 \omega_1^2 \sin(\theta_1 - \theta_2) + \tau_2 - \beta_2 \omega_2 - m_2 r_2 g \sin \theta_2]}{m_2^2 l_1^2 r_2^2 + m_1 m_2 r_1^2 r_2^2 - m_2^2 l_1^2 r_2^2 \cos^2(\theta_1 - \theta_2)} + \\ & + \frac{l_1 \cos(\theta_1 - \theta_2) [m_2 l_1 r_2 \omega_2^2 \sin(\theta_1 - \theta_2) - \tau_1 + \beta_1 \omega_1 + (m_1 r_1 + m_2 l_1) g \sin \theta_1]}{m_2 r_2 l_1^2 + m_1 r_1^2 r_2 - m_2 r_2 l_1^2 \cos^2(\theta_1 - \theta_2)}. \end{aligned} \tag{21}$$

The following border conditions are assumed,

$$\theta_1 = 0 \rightarrow \omega_1 = 0$$

$$\theta_2 = 0 \rightarrow \omega_2 = 0,$$

taking as a reference the upright position of a person on both feet. The equilibrium points suggested by experimental physics [16], are $(0, 0, 0, 0)$, $(0, \pi, 0, 0)$, $(\pi, \pi, 0, 0)$, $(\pi, 0, 0, 0)$. These points are where the force of gravity, g , does not influence the movement and like the acceleration, $\ddot{\theta}_1 = \ddot{\theta}_2 = 0$, the masses of the two pendulums never change places. Now as $\sin(\pi) = \sin(0) = \sin(-\pi) = 0 = \dot{\mathbf{x}}_2 = \dot{\mathbf{x}}_4 = 0$. Points $(0, \pi, 0, 0)$, $(\pi, \pi, 0, 0)$, $(\pi, 0, 0, 0)$, are unstable. Gravity is what ensures the instability of these points. The reason is that in a neighborhood of any of these points, the force of gravity will involve a change in the acceleration of the pendulum centers of mass, $\dot{\mathbf{x}}_2$, $\dot{\mathbf{x}}_4$, and this change implies a change in position, the change is in the direction of the set of points such that $\mathbf{x}_1 = \mathbf{x}_3 = 0$. This implies that for every neighborhood of the equilibrium points there is an orbit that comes out of that neighborhood, and therefore the points are unstable.

2.6. Graphical Analysis

In this section we analiced the solutions for different points.

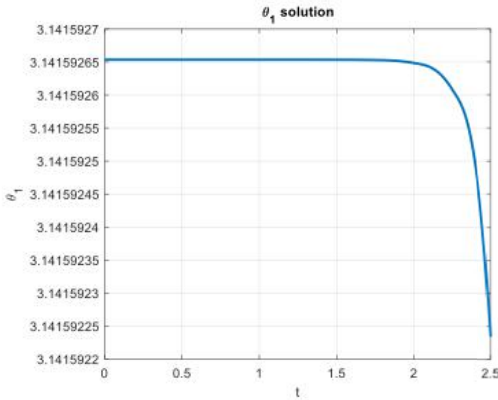


Figure 2. Solution θ_1 for $(\pi, 0, 0, 0)$

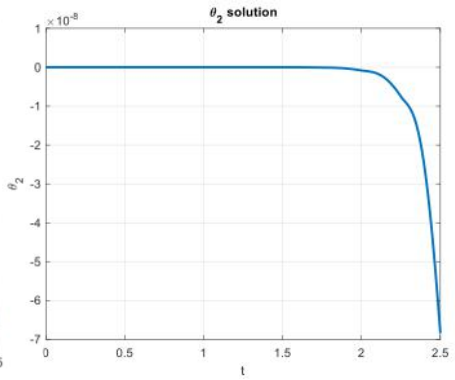


Figure 3. Solution θ_2 for $(\pi, 0, 0, 0)$

At this point the hip angle at 180° is evaluated, since the leg is not able to rotate to this position, the natural response of the system is unstable as can be seen in the Figure 2 and the Figure 3.

Analyzing the point of operation $(\pi/9, 0, 2\pi/2.5, 0)$ of the swing phase, which occurs between the take-off of the foot and the middle phase of the swing, where the hip is in neutral position and the knee rotates 20° with an angular velocity of $2.51[rad/s]$, based on the cycle period, the following numerical solution is obtained.

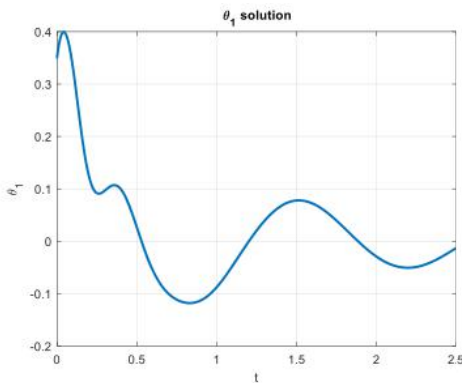


Figure 4. Solution θ_1 for $(\pi/9, 0, 2\pi/2.5, 0)$

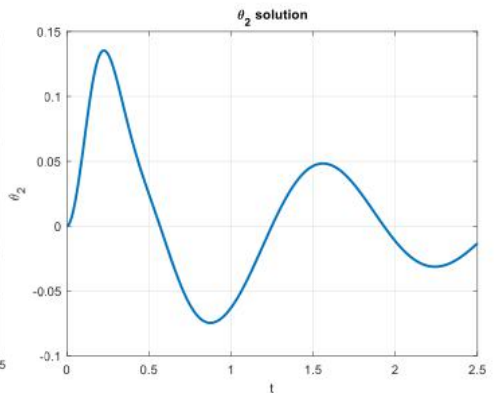


Figure 5. Solution θ_2 for $(\pi/9, 0, 2\pi/2.5, 0)$

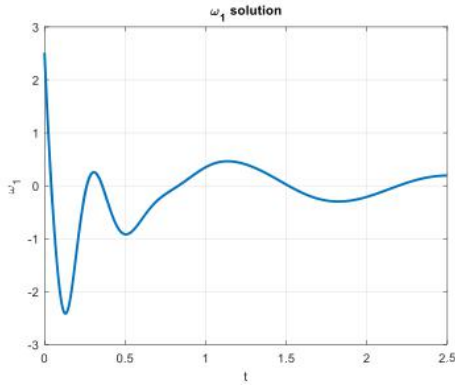


Figure 6. Solution $\dot{\theta}_1$ for $(\pi/9, 0, 2\pi/2.5, 0)$

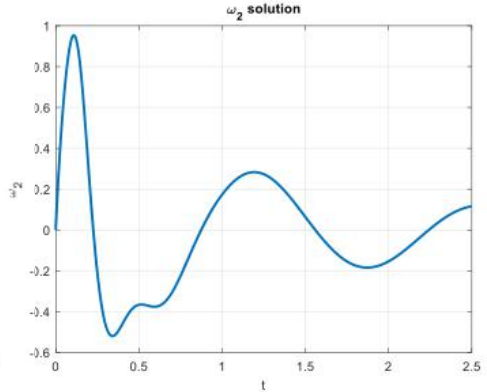


Figure 7. Solution $\dot{\theta}_2$ for $(\pi/9, 0, 2\pi/2.5, 0)$

It can be seen that in the Figure 4 and the Figure 5, the graph starts from the point designated $\pi/9$ for θ_1 , oscillating in an acceptable working range for the angles of rotation performed by the hip, as well as the angular velocity of the hip $\dot{\theta}_1$ that starts at $2.51[rad/s]$ as shown in Figure 6. Furthermore, even though the knee angle θ_2 and the angular velocity $\dot{\theta}_2$ is zero, it is observed how these vary during the 2.5 seconds of the gait cycle, which indicates that the equations of motion are coupled, that is, one depends on the other, as can be seen in Figure 7.

2.7. Linearization

An approximate linearization of the system was performed using Taylor’s series, in which the selected operating points are $(0, 0, 0, 0)$. For the system [17]-[18],

$$\begin{aligned} \dot{\mathbf{x}} &= \mathbf{Ax} + \mathbf{Bu} \\ y &= \tilde{\mathbf{C}}\mathbf{x}, \end{aligned}$$

where, A , B y \tilde{C} , are the Jacobian matrices of the system and are generally calculated and take the form,

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ -\frac{g(m_1r_1+m_2l_1)}{m_1r_1^2} - \frac{\beta_1}{m_1r_1^2} & -\frac{\beta_1}{m_1r_1^2} & \frac{m_2gl_1}{m_1r_1^2} & \frac{l_1\beta_2}{m_1r_2r_1^2} \\ 0 & 0 & 0 & 1 \\ \frac{m_2gl_1^2+m_1gr_1l_1}{m_1r_2r_1^2} & \frac{l_1\beta_1}{m_1r_2r_1^2} & -\frac{g(m_1r_1^2+m_2l_1^2)}{m_1r_2r_1^2} & -\frac{\beta_2(m_1r_1^2+m_2l_1^2)}{m_1m_2r_1^2r_2^2} \end{bmatrix} \quad (22)$$

$$B = \begin{bmatrix} 0 \\ \frac{1}{m_1r_1^2} \\ 0 \\ \frac{m_1r_1^2+m_2l_1^2}{m_1m_2r_1^2r_2^2} \end{bmatrix} \quad (23)$$

$$\tilde{C} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \tag{24}$$

Replacing the value of the constants of the Table 1, in Eqs. (22), (23), and (24), the following linearized matrices of the system are obtained,

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ -104.2917 & -7.5214 & 53.9840 & 1.5689 \\ 0 & 0 & 0 & 1 \\ 284.4320 & 20.5128 & -206.6837 & -6.0069 \end{bmatrix} \tag{25}$$

$$B = \begin{bmatrix} 0 \\ 3.2873 \\ 0 \\ 34.3250 \end{bmatrix} \tag{26}$$

$$\tilde{C} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \tag{27}$$

Given the linearized matrices Eqs. (25), (26) and (27), it can be seen that the system is a SIMO case, where there is one input (1 column in matrix B) and two outputs (2 rows in matrix \tilde{C}), from the above the transfer matrix can be obtained,

$$M.T = \frac{1}{s^4 + 13.5s^3 + 324s^2 + 627.4s + 6200.6} \begin{bmatrix} 3.3s^2 + 73.6s + 2532.4 \\ 34.3s^2 + 325.6s + 4514.8 \end{bmatrix} \tag{28}$$

The transfer matrix Eq. (28), is defined as the relationship between the inputs and outputs of the system, in the denominator is the characteristic polynomial, which for this case is fourth degree, that is, the system has four poles, this number of poles is the same that must have the desired poles to control the system.

3. Dynamic system control

A state feedback controller was designed due to its good performance in the face of modeling inaccuracies or disturbances. However, before designing the controller, the controllability and observability of the system must be verified, this is done by calculating the controllability matrix $Q = [B \ AB \ A^2B \ A^3B]$ and the observability matrix

$$\mathcal{O} := \begin{bmatrix} \tilde{C} \\ \tilde{C}A \\ \tilde{C}A^2 \\ \vdots \\ \tilde{C}A^{n-1} \end{bmatrix} = n$$

to verify the controllability and observability each of the matrices must have rank n [18], and effectively for this case, each matrix has a rank equal to 4. for our case we have,

$$Q = \begin{bmatrix} 0 & 3.2873 & 29.1289 & 1073.4 \\ 3.2873 & 29.1289 & 1073.4 & -26020 \\ 0 & 34.325 & -138.7538 & -4728.4 \\ 34.325 & -138.7538 & -4728.4 & 87384 \end{bmatrix} \tag{29}$$

and

$$\mathcal{O} := \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ -104.2917 & -7.5214 & 53.9840 & 1.3689 \\ 284.4320 & 20.5128 & -206.6837 & -6.0069 \\ 1230.7 & -15.5373 & -730.3088 & 32.7590 \\ 3847.9 & 6.9216 & 2348.9 & -138.4177 \end{bmatrix} = Rank(\mathcal{O}) = 4 \tag{30}$$

The design of the controller was carried out using the pole allocation technique, the desired poles of the system are placed under the criterion of settlement time $t_s = 4T$, (where $T = 2.5$ seconds the duration of the running cycle), therefore, the settling time is 10 seconds and the overshoot, which is defined by the maximum amount that the response exceeds the steady state value, in this case, you want to attenuate this as much as possible parameter, this is why an overshoot less than 5% is defined. With the two previous parameters the dominant poles of the controller are determined by means of the equations,

$$\xi = \frac{\ln\left(\frac{OV}{100}\right)}{\sqrt{\pi^2 + \ln\left(\frac{OV}{100}\right)^2}} = \frac{\ln\left(\frac{5}{100}\right)}{\sqrt{\pi^2 + \ln\left(\frac{5}{100}\right)^2}} = 0.69 \tag{31}$$

$$\omega_n = \frac{4}{\xi t_s} = \frac{4}{(0.69)(10)} = 0.57 \tag{32}$$

where, OV , is the overshoot value, t_s is the settlement time, then the values obtained from Eqs. (31) and (32), are replaced in the equation,

$$s^2 + 2\xi\omega_n s + \omega_n^2 = s^2 + 0.8s + 0.33, \tag{33}$$

solving the quadratic equation, we know the dominant poles of the controller

$$\mu_1 = -0.4 + 0.4195i, \tag{34}$$

$$\mu_2 = -0.4 - 0.4195i. \tag{35}$$

Since the system is of order 4, that is, it has four poles, this is deduced from the Eq. (28), four poles must be defined for the controller, from Eq. (33), two dominant poles were obtained, and the remaining two poles are insignificant, which are located 5 to 10 times farther than the dominant poles $\mu_3 = -50$ and $\mu_4 = -45$. Feedback gains are determined using Ackerman’s formula [18],

$$K = [0 \ 0 \ 0 \ 1]Q^{-1}\phi(A), \tag{36}$$

where, $\phi(A) = A^4 + \alpha_1 A^3 + \alpha_2 A^2 + \alpha_3 A + \alpha_4 \mathbf{I}$. The characteristic equation is given by,

$$\begin{aligned} |s - A + BK| &= |s\mathbf{I} - A| \\ &= (s - \mu_1)(s - \mu_2)(s - \mu_3)(s - \mu_4) \\ &= s^4 + \alpha_1 s^3 + \alpha_2 s^2 + \alpha_3 s + \alpha_4, \end{aligned} \tag{37}$$

in our case,

$$(s - 0.4 + 0.4195i)(s - 0.4 - 0.4195i)(s - 50)(s - 45) = s^4 + 95.8s^3 + 2326.3s^2 + 1831.9s + 756$$

to get feedback gains

$$K = [-105.3237 \ -9.8608 \ 57.8717 \ 3.3412], \tag{38}$$

here, the Eq. (38), there is the controller gain vector, where $K_1 = -105.3237$ is the state feedback gain x_1 , $K_2 = -9.8608$ is the state gain x_2 , $K_3 = 57.8717$ is the state gain x_3 and $K_4 = 3.3412$ is the gain of the status x_4 .

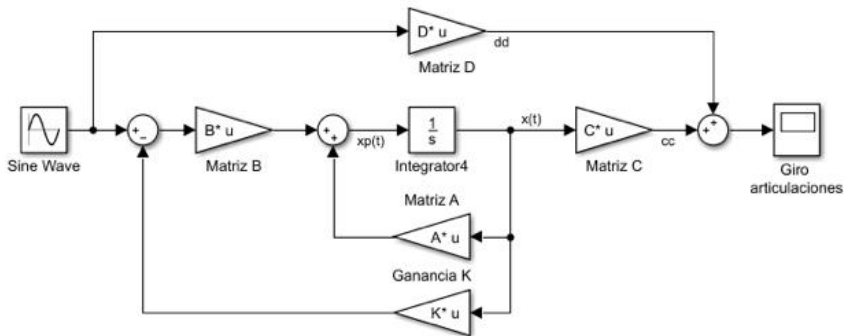


Figure 8. Block diagram with controller

The Figure 8, shows the block diagram of the system, where the matrices A , B , \tilde{C} previously described are entered respectively, additionally, the controller feedback with the gain vector Eq. (38), the result of the controller simulation is shown in Figure 9.

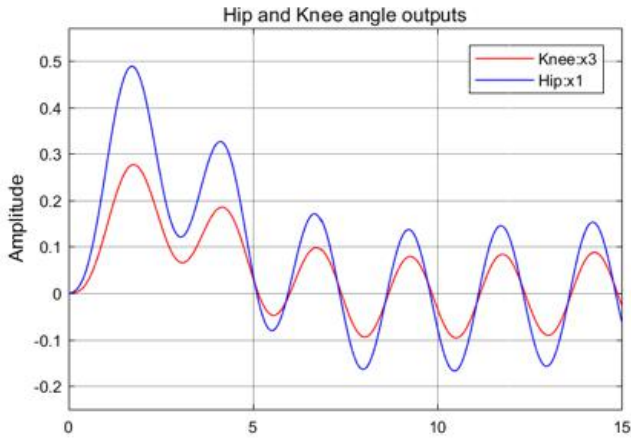


Figure 9. Hip and knee exit angles.

In the Figure 9, shows the response of the hip (θ_1) and knee (θ_2) angles with the controller to a sinusoidal input, where it can be seen that the hip joint has a greater amplitude with respect to the knee. Since the hip has a wider turning range than the knee. In addition, it is shown how each of the two outputs try to follow the sinusoidal input until both stabilize in a certain time, keeping the input signal within the operating range of each one of them.

4. Conclusions

The human gait model resembles the mechanical behavior of a double pendulum, the equations of motion are obtained by means of the Euler-Lagrange formalism based on the analysis of energies, since this allows to obtain a dynamic model of the system, these equations they are linearized using the Taylor series approximation in order to design and implement a state feedback control system, which is based on a design in state space. As can be seen in Figure 9, the controller responds satisfactorily in such a way that it satisfies the variations of the angles of both the hip and the knee, following the input signal, giving good results to the pole assignment criterion and Ackerman's formula for obtaining feedback gains. On the other hand, this study allows us to define a dynamic model of human gait, which can be taken into account for the design of electromechanical devices such as prostheses and / or recovery devices.

Acknowledgements

P. A. Ospina-Henao, we are grateful for the permanent support of the Vicerrectoría Académica of Universidad Santo Tomás, with the project of the XII Internal Call for Research Projects, with initiation act 163 of February 05, 2020.

References

- [1] Erazo Bravo, M. y Mera Otoy, E., *Diseño e implementación de un prototipo de prótesis transtibial con amortiguamiento activo mediante la investigación del comportamiento de materiales magnetoreológicos bajo diferentes campos magnéticos*, Universidad De Las Fuerzas Armadas, Latacunga, Ecuador, 2017.
- [2] Simonsen EB. Contributions to the understanding of gait control. *Dan Med J.* 2014; 61 (4): B4823.
- [3] Valencia, Cesar H; Vellasco, Marley; Tanscheit, Ricardo and Figueiredo, Karla T. Magnetorheological Damper Control in a Leg Prosthesis Mechanical. *Advances in Intelligent Systems and Computing.* 2005. Pages 805-818.
- [4] Arthur D. Kuo, J. Maxwell Donelan, *Dynamic Principles of Gait and Their Clinical Implications*, Physical Therapy, Volume 90, Issue 2, 1 February 2010, Pages 157-174.
- [5] Hunt, M. et al. (2008). Measures of Frontal Plane Lower Limb Alignment Obtained from Static Radiographs and Dynamic Gait Analysis. *Gait & posture* 27(4): 635-640.
- [6] Schiehlen, W. and García-Vallejo D. Walking dynamics from mechanism models to parameter optimization, *Procedia IUTAM* 2 (2011), pp 199-211.
- [7] González Míguez, Paula, author; Saló i Bru, Guillem, supervisor academic. Idelsohn Zielonka, Sebastian, academic supervisor; [et al.] *Biomechanical analysis of gait using an analytical model of inverse dynamics in patients with altered sagittal balance of the spine.* (121 pages). ISBN 9788449075650.
- [8] Pratt, J. E. (2000). *Exploiting Inherent Robustness and Natural Dynamics in the Control of Bipedal Walking Robots.* Ph.D. Thesis, MIT.
- [9] Pop, C. *Bondgraphs Modeling and Model Evaluation of Human Locomotion Using Experimental Data.* University Of Waterloo.
- [10] Luengas LA. *Static alignment model for transtibial prostheses.* Bogotá: Pontificia Universidad Javeriana; 2016.
- [11] Borghese NA, Bianchi L, Lacquaniti F. Kinematic determinants of human locomotion. *J Physiol.* 1996; 494 (3): 863-79.
- [12] Goldstein. *Classical Mechanics.* Editorial Addison Wesley (1980).
- [13] K. Ogata. *Dinámica de Sistemas.* Tercera edición, Prentice Hall. México D.F., 1978.
- [14] Winter, D. *Biomechanics of human movement.*, John Wiley & Sons, 1979.
- [15] R. Kelly, V. Santibáñez. *Control de Movimiento de Robots Manipuladores.* PEARSON EDUCATION, S.A., Madrid, 2003, pp. 327.
- [16] G. D. Quiroga, P. A. Ospina-Henao. *Dynamics of damped oscillations: physical pendulum.* European Journal of Physics, 2017.
- [17] S. Domínguez, P. Campoy, J. M. Sebastián, A. Jiménez. *Control en el Espacio de Estados.* Segunda edición, PEARSON EDUCATION S.A., Madrid, 2006, pp. 440.
- [18] K. Ogata. *Ingeniería de Control Moderna.* Tercera edición, PEARSON EDUCATION S.A. México D.F., 1998, pp. 1015.

Two-Dimensional Spatial Resolution in Plane Temperature Monitoring Based on Raman Distributed Temperature Sensor

Junfan CHEN^a, Ning SUN^b and Zhongxie JIN^{c,1}

^aThe Key Laboratory of Optoelectronic Technology & System, Education Ministry of China, Chongqing 400044, China; chenjf584@outlook.com

^bThe Key Laboratory of Optoelectronic Technology & System, Education Ministry of China, Chongqing 400044, China; 413807424@qq.com

^cThe Key Laboratory of Optoelectronic Technology & System, Education Ministry of China, Chongqing 400044, China; jinzhongxie@cqu.edu.cn

Abstract. Spatial resolution is an important parameter that characterizes the detection capability of a system, and there are extremely high requirements for spatial resolution in important fields such as the fossil energy industry and nuclear industry. In order to realize the high-precision distributed monitoring of the optical fiber distributed temperature sensing system (DTS), the factors affecting the spatial resolution of the DTS system were analyzed, and a two-dimensional planar temperature field distribution monitoring scheme based on Raman distributed temperature sensor (RDTS) was proposed. In this scheme, based on the layout of the two-dimensional RDTS heat source positioning system, multimode fiber was adopted. After comparing several sensing fiber routing schemes, the 45° skew 2D wiring method of sensing fiber was finally selected. According to the experimental results, the spatial resolution of the temperature field distribution in the monitoring area can break through the limitation of the system resolution. It has more application value than the traditional one-dimensional distributed temperature sensing system.

Keywords. spatial resolution, two-dimensional positioning

1. Introduction

In optics, the ability of the imaging system to separate two objects that are very close together is called the resolution of the optical imaging system. The Abbe imaging principle shows that the optical system creates a Fraunhofer diffraction pattern on the focal plane. Due to the diffraction caused by the limited aperture of the imaging system, the images of two very close points may be inseparable.

In recent years, spatial resolution [1-3] has been involved in more and more fields. The most common are optical, remote sensing satellite radar, image sensors, and medical smart wearable devices represented by artificial skin. Spatial resolution exists in two-dimensional plane space [4], which has more research significance and practical use than one-dimensional space.

¹ Corresponding Author. Email: jinzhongxie@cqu.edu.cn; Tel.: +86-186-2355-4364

In this paper, we adopt an optimized two-dimensional positioning method based on the traditional RDTS system proposed by Zhang [5] et al., using 45°-inclined-cross-type (45°-ICT) routing rules and supporting algorithms to convert one-dimensional fiber [6] measurement data to two-dimensional planar distribution results. Starting from the networking mode of the RDTS system, the advantages of the networking mode used are introduced, and the factors that affect the spatial resolution of the system are analyzed.

2. System Configuration and Spatial Resolution

2.1. Structure of RDTS system

The DTS system in this paper was mainly used for real-time monitoring of the heat source in the temperature field [7,8], which needed to meet the two requirements of precise positioning and accurate measurement simultaneously [9]. RDTS system [10-13] is mainly composed of six parts, including transmission module, filter module, photoelectric detection module, DAQ (data acquisition) module, PC (personal computer) and SFL (Sensing Fiber Link). Distributed Fiber temperature measurement system was shown in Fig. 1.

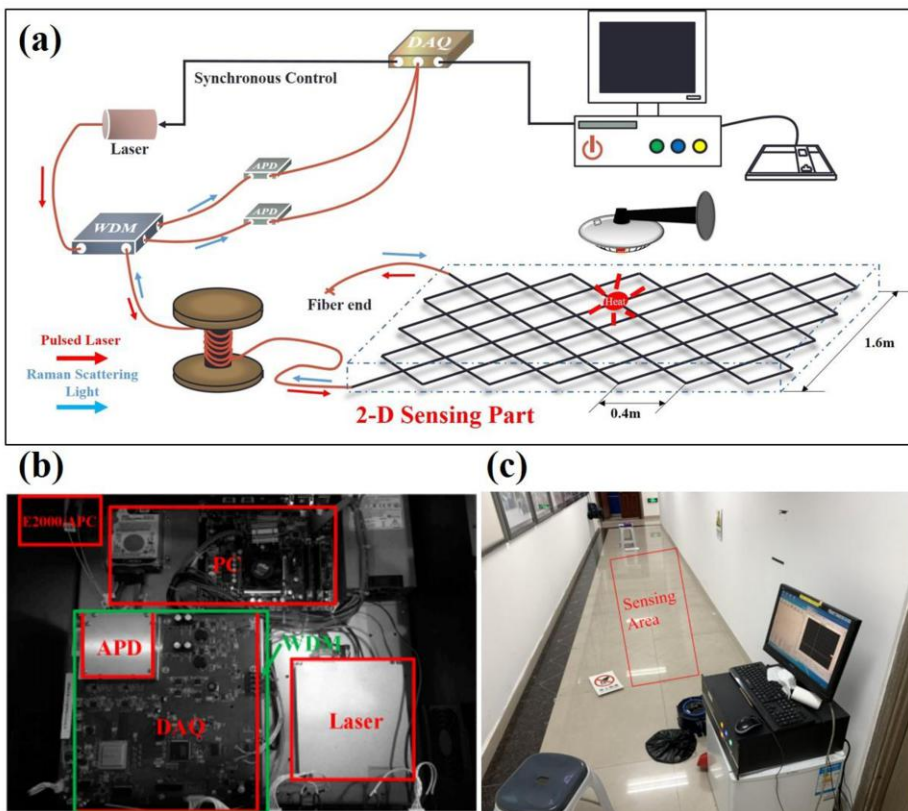


Figure 1. (a) RDTS-Based Two-Dimensional temperature field distribution monitoring; (b)(c) Experimental platform

In this experiment, we used a high-power pulsed laser with a wavelength of 1550 nm, a pulse width of 7 ns, and a frequency of 8 kHz as the light source. The pulse is transmitted into the SFL through wavelength division multiplexing (WDM), using two high-sensitivity, low-noise avalanche photodiodes (APDs) with a bandwidth of 150 MHz to simultaneously detect the weak backscattered light along the SFL. Use a 12-bit 250 MHz sampling frequency data acquisition (DAQ) card for signal processing, and transfer the processed results to a PC for graphic display. The fiber end is bent with a small radius to avoid reflection effects.

2.2. Spatial Resolution of RTDS

In a two-dimensional plane temperature distribution field, if two heat sources can be distinguished by the distributed optical fiber temperature sensing system, the minimum distance between the two heat sources is called the spatial resolution. In fact, in the distributed optical fiber temperature sensing system, there is another parameter that characterizes the resolution ability: System resolution [11,14], but the difference is, the spatial resolution is more focused on the diameter between any two temperature regions on a two-dimensional plane. The system resolution can only quantitatively indicate the one-dimensional minimum temperature resolution distance. The spatial resolution is more suitable for measuring the two-dimensional heat source discrimination accuracy of the distributed optical fiber sensing system on a two-dimensional plane.

3. Experiment

3.1. Fiber Distribution

The traditional distributed optical fiber temperature sensing system had a linear type and an S type. The linear type is suitable for scenarios where only one-dimensional spatial temperature distribution data is required, so the coordinates of the heat source in one dimension can only be determined, and the two-dimensional spatial temperature distribution cannot be obtained, and the scope of application is very narrow. The S-type sensor fiber distribution was as Fig.2(a):

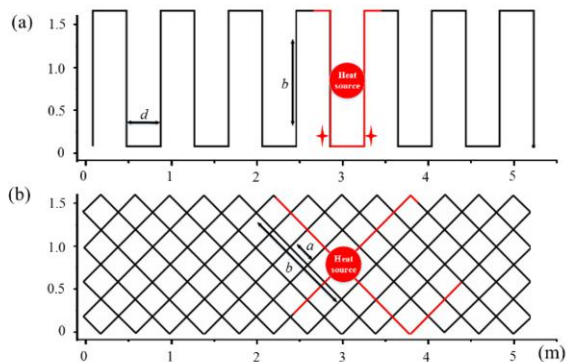


Figure 2. (a) S-type distributed sensor fiber networking; (b) 45° diagonally distributed sensor fiber networking [5]

We have introduced a new two-dimensional matrix positioning method to modify the algorithm without modifying the materials or hardware, which can reduce the cost of the RDTS system. The specific distribution is shown in Fig. 2(b). The minimum unit side length was $a= d/1.414$, and the red fiber segment corresponded to the high temperature area of the sensor fiber link. A sufficiently long sensor fiber link was distributed across the two-dimensional plane, dividing the two-dimensional plane into many Uniform and independent grid. According to the 45° diagonal span distribution we used, we kept the slope of the sensing fiber link to the two-dimensional plane coordinate system at 1 or -1. When the sensing fiber link reaches the edge of the sensing area, the slope becomes the opposite of the original. The corner of the grid, which was the intersection of the sensing fiber link, carried the temperature distribution information on the sensing fiber and was used to define the position in space. This method could significantly reduce crosstalk, because any two adjacent intersections contained at least a pair of fiber segments located far apart. And we can obtain the same positioning accuracy in both directions. Compared with the linear distribution, this fiber network can greatly increase the coverage of the optical fiber to the area to be tested, reduce the probability of missed.

The representation of the matrix is shown in Fig. 2(b). The plane was divided into many grids by optical fibers. Each vertex of the grid records the temperature information of the sensing fiber link, which is used for spatial positioning. The matrix corresponds to the intersections, the row and column of the matrix correspond to the 2D coordinates of the vertices, and the value of every element in the matrix corresponds to the temperature of the intersection point [5].

We conducted multiple tests on different locations in the experiment to verify the accuracy and effectiveness of the fiber networking. We can get that the positioning accuracy of this method is half of the diagonal of the grid ($d/2$).

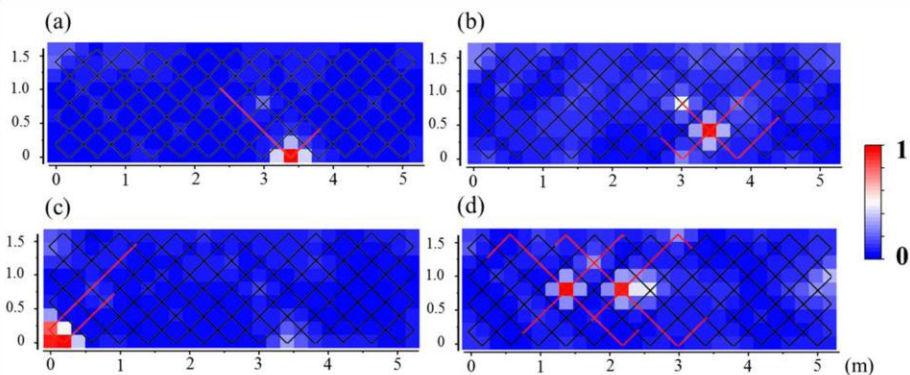


Figure 3. Temperature detection at different intersections [5]

As shown in Fig.3(a), the inflection point of the edge or corner was regarded as the intersection point. In Fig.3(b), when the heat source was located near the edge, there were high-temperature inflection points and intersection points. After setting the intersection priority to be higher than the inflection point priority, determined that the actual heat source was near the intersection. We also tested the situation of two heat sources close together, as shown in Fig.3(d). These two heat sources were difficult to distinguish because there were four intersection points of abnormal temperature. In the matrix, an additional priority was set according to the position of the crossing point of the abnormal temperature optical fiber segment (The priority would be higher if the

intersection is at the center of the high-temperature fiber segment instead of the edge position).

3.2. Determination of Spatial Resolution

The measurement of spatial resolution, that is, the measurement of the minimum distance of two heat sources which can be separated, requires the existence of a saddle point (minimum point) shown in Fig.4. Since our experimental system measures the temperature once at 0.4 m, and the grid side length is 0.28 m. When the distance between two heat sources is shorter than 0.28 m, they cannot be separated. So at this time we need a saddle point, that is, the existence of a low temperature point.

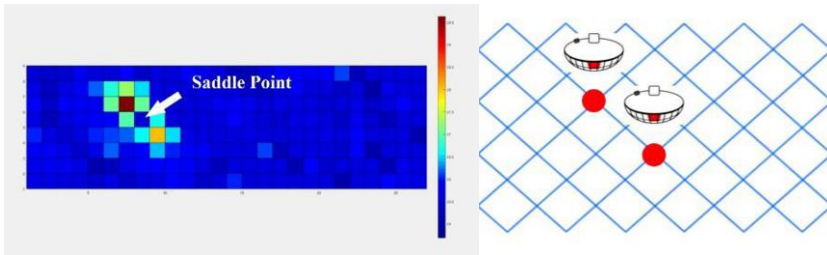


Figure 4. Temperature distribution graph displayed in MATLAB

There was an obvious low-temperature grid between the two heat sources. This is because, under the linear interpolation algorithm, when there is only one high temperature point at the intersection, the grid cannot be displayed in a high temperature state. Because the temperature within 0.4 m was displayed the same, the temperature of the same fiber where the two heat sources were located showed high temperature, while the other fiber at the cross point showed normal temperature. In this way, it can be judged that there were two heat sources.

We regarded the three heat sources as one heat source, and the gap between the three heat sources was far less than 0.4 m, so it can be ignored. At this time, after being calculated by the program, it was displayed in MATLAB as Fig.5

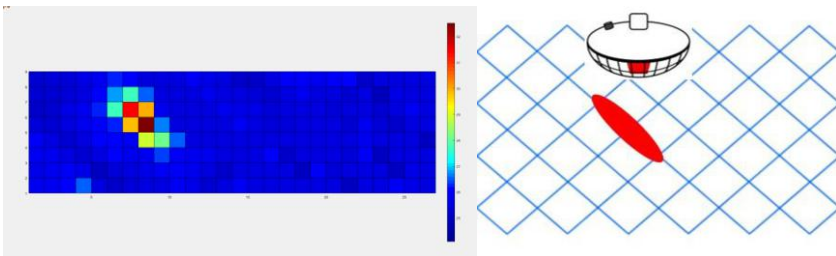


Figure 5. Temperature distribution graph of 3 connected heat sources

At this time, we can see that the temperature of the three grids has become higher together. There was no obvious low-temperature grid inside the heat source. We can judge that this heat source is just one heat source instead of two heat sources separated by a certain distance. According to the above analysis, at the intersection, the spatial resolution was 0.28 m, that is, the distance between the two heat sources at the

intersection was at least 0.28 m to be distinguished. The same conclusion can be drawn after repeated experiments.

3.3. Factors Affecting the Spatial Resolution of the System

3.3.1. Resolution of RDTS System. The resolution of the RDTS system greatly affected the spatial resolution. It refers to the minimum length that the system can distinguish by measuring the temperature field along the one-dimensional length of the fiber. It is mainly determined by the pulse width Δt , photodetector response time τ , and analog-to-digital conversion time $T_{A/D}$, which are expressed as Eq.(1) [15,16]:

$$\begin{aligned} \Delta L_1 &= \frac{v\Delta t}{2} \\ \Delta L_2 &= \frac{v\tau}{2} \\ \Delta L_3 &= \frac{vT_{A/D}}{2} \end{aligned} \tag{1}$$

v is the propagation speed of light wave in optical fiber. The actual resolution of the RDTS system depends on the largest of the three:

$$\Delta L = \max \{ \Delta L_1, \Delta L_2, \Delta L_3 \} \tag{2}$$

The system resolution of the system used in this experiment was 0.4 m, so when the area of the heat source used was smaller than 0.4 m×0.4 m, the temperature distribution diagram showed the difference between the temperature profiles of the two heat sources. Because the system resolution determined the cumulative average interval of the RDTS, it also determined the minimum area where the temperature can be detected, which in turn determined the spatial resolution in the two-dimensional plane.

3.3.2. Distribution Density. The networking density of the RDTS system can make up for the limitations caused by spatial resolution. We doubled the distribution density of optical fiber, as shown on Fig.6.



Figure 6. (a)Original distribution density of fiber; (b)Double distribution density of fiber

The heat source is changed from covering one intersection point to now covering two intersection points. There was a section of fiber with no high temperature alarm

between the intersection points as the reference fiber. The temperature diagram in MATLAB was shown in Fig.7:

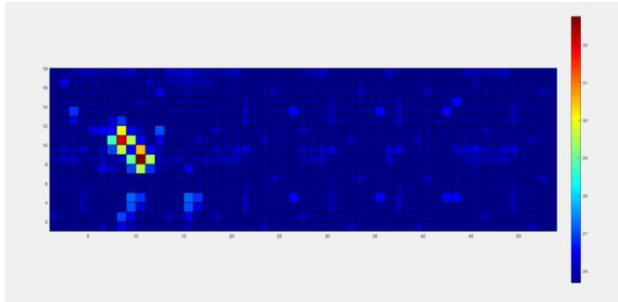


Figure 7. Temperature distribution graph of multiple heat sources in double-density distribution of optical fiber

There was a clear line of low-temperature optical fiber boundary between the two heat sources, and it was judged as two heat sources at this time. We can know that if the two heat sources were placed under the single-density fiber distribution, they cannot be separated because they are in the same temperature measurement interval, but they can be separated in the case of double density. We can be sure that density did affect spatial resolution.

For one heat source, it can be seen that the high temperature area was a connected heat source in Fig.8, which was consistent with the actual situation.

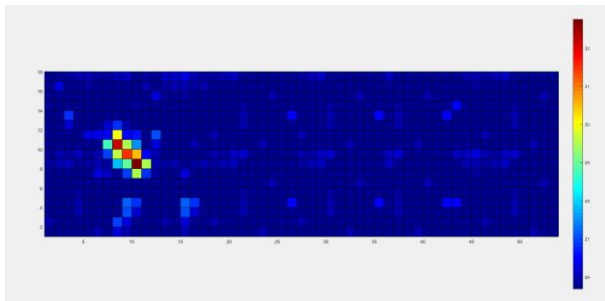


Figure 8. Temperature distribution graph of one heat source in double-density distribution of optical fiber

At this time, the spatial resolution of the system was 0.14 m, that is, at least 0.14 m between the two heat sources can they be distinguished. Therefore, it can be seen that the spatial resolution of the RDTS system is related to the laying density of the optical fiber, which can make up for the limitation caused by the spatial resolution. The denser the optical fiber is laid, the more judgment areas are available, and the higher the spatial resolution is. If the optical fiber is doubled in density, the spatial resolution will be doubled.

4. Conclusion

This paper explained the spatial resolution of the RDTS system and gave the structure of the RDTS system. The 45° diagonal span fiber laying method is used to derive a method

for converting one-dimensional sensing data into planar two-dimensional temperature field distribution data. Use MATLAB to calculate the temperature distribution, and give the block diagram and content. The spatial resolution of the RDTS system is related to the system resolution and fiber distribution density. The two-dimensional distribution of optical fibers greatly improves the spatial resolution. Through a higher distribution density of optical fibers and a DTS with 0.4 m system resolution, we obtained a spatial resolution of 0.14 m. Higher spatial resolution is vital to the practical application of the technology.

References

- [1] M.A. Soto, T. Nannipieri, A. Signorini, A. Lazzeri, F. Baronti, R. Roncella, G. Bolognini, F. Di Pasquale, Raman-based distributed temperature sensor with 1m spatial resolution over 26km SMF using low-repetition-rate cyclic pulse coding, *Opt. Lett.* **36** (2011), 2557–2559.
- [2] A.K. Sang, M.E. Froggatt, D.K. Gifford, S.T. Kreger, B.D. Dickerson, One Centimeter Spatial Resolution Temperature Measurements in a Nuclear Reactor Using Rayleigh Scatter in Optical Fiber, *IEEE Sens. J.* **8** (2008), 1375–1380.
- [3] S.D. Dyer, M.G. Tanner, B. Baek, R.H. Hadfield, S.W. Nam, Analysis of a distributed fiber-optic temperature sensor using single-photon detectors, *Opt. Express* **20** (2012), 3456–3466.
- [4] M. Sun, Y.Q. Tang, S. Yang, J. Li, M.W. Sigrist, F.Z. Dong, Fire Source Localization Based on Distributed Temperature Sensing by a Dual-Line Optical Fiber System, *Sensors* **16** (2016), 829.
- [5] C. Zhang, Z.X. Jin, RDTs-Based Two-Dimensional Temperature Monitoring with High Positioning Accuracy Using Grid Distribution, *Sensors* **19** (2019), 4993.
- [6] A.H. Hartog, A.P. Leach, M.P. Gold, Distributed temperature sensing in solid-core fibres, *Electron. Lett.* **21** (1985), 1061–1062.
- [7] J. Zhang, P. Wei, Q. Liu, Monitoring a Heatsink Temperature Field Using Raman-Based Distributed Temperature Sensor in a Vacuum and -173 °C Environment, *Sensors* **19** (2019), 4186.
- [8] A.J. Rogers, Polarisation optical time domain reflectometry, *Electron. Lett.* **16** (2007), 489–490.
- [9] K.L. Wei, Z.Y. Wen, J. Zou, et al, Based on OTDR prospective distributed optical fiber temperature alarm system, *Piezoelectrics & Acousto-optics* **32** (2010), 329–332.
- [10] I. Toccafondo, T. Nannipieri, A. Signorini, E. Guillermain, J. Kuhnhenh, M. Brugger, F.D. Pasquale, Raman Distributed Temperature Sensing at CERN, *IEEE Photonic Tech. Lett.* **27** (2015), 2182–2185.
- [11] I. Laarossi, M.Á. Quintela-Incera, J.M. López-Higuera, Comparative Experimental Study of a High-Temperature Raman-Based Distributed Optical Fiber Sensor with Different Special Fibers, *Sensors* **19** (2019), 574.
- [12] B. Yan; J. Li, M. Zhang, J. Zhang, L. Qiao, T. Wang, Raman Distributed Temperature Sensor with Optical Dynamic Difference Compensation and Visual Localization Technology for Tunnel Fire Detection, *Sensors* **19** (2019), 2320.
- [13] Z. Wang, X.H. Sun, Q. Xue, et al, an optical fiber-folded distributed temperature sensor based on Raman backscattering, *Optics & Laser Technology* **93** (2017), 224–227.
- [14] B.N. Sun, J. Chang, J. Lian, et al, Accuracy improvement of Raman distributed temperature sensors based on eliminating Rayleigh noise impact, *Opt. Commun* **306** (2013), 117–120
- [15] F. Ning, Y. Zhu, H.J. Cui, X.Q. Li, Z.X. Jin, A Linear Correcting Algorithm for Improving Space Resolution of Distributed Optical Fiber Raman Temperature Measurement System, *Acta Photonica Sinica* **4** (2012), 408–413.
- [16] X.F. Shi, Z. Li, Z.Q. Cai, Optical fiber distributed temperature measurement system and precision analysis, *Measurement and Control Technology* **21** (2002), 9–12.

Analysis for the Adoption of Security Standards to Improve the Management of Securities in Public Organizations

Segundo Moisés Toapanta Toapanta^{*a}, Madeleine Lilibeth Alvarado Ronquillo^b, Luis Enrique Mafla Gallegos^b, Alberto Ochoa Zezzatti^c

^a*Department Computer Science, Salesian Polytechnic University of Ecuador (UPS), Chambers 227 and June 5, Ecuador*

^b*Faculty of Systems Engineering, National Polytechnic School (EPN), Ladrón the Guevara E11-253, Ecuador*

^c*Industrial and Manufacturing Department, Autonomous University of Ciudad Juarez (UACJ), Av. Plutarco Elías Calles #1210, México*

Abstract. Public organizations have the ongoing task of properly managing the security of the information they handle. The objective of this research is to analyze the security standards adopted by public organizations in Ecuador to improve their management of information security. The deductive method was applied for the review and analysis of appropriate standards for public institutions. As a result, information was obtained on the different security policies, standards and guidelines that apply, national and international public organizations. A Diagram of activities for the adoption of standards for public organizations resulted; a prototype standards-based Information Security Management Model; and an Information Security Management Matrix, from which the Risk Mitigation Percentage was calculated. It was concluded that maintaining high levels of security in public organizations requires the adoption of control standards in different areas and the collaboration of the different organizational and hierarchical levels of public organizations.

Keywords. security standards, security management, public organization, information security

1. Introduction

Public organizations have the ongoing task of properly managing the security of the information they handle. For that reason, they have had to seriously consider protecting their information through normally accepted standards that seek to prevent the institution from being a victim of cyber criminals [1]. In reference to the above, a figure from the International Telecommunication Union (ITU) mentions that Ecuador is the sixth most cybersecure country in Latin America of a total of 19 countries. At the Latin American

* Corresponding Author: Segundo Moisés Toapanta Toapanta,
Email:stoapanta@ups.edu.ec

level, the most cybersecure countries are Uruguay and Brazil. While worldwide dominate Singapore and the United States [2].

Standards are written standards that are contained in a document available to the public, established by consensus and adopted by a recognized body; it establishes the rules, requirements, characteristics, guidelines or general recommendations that must be followed to achieve an optimal level of regulation in a particular area in relation to current problems or problems that may arise in the future [3].

Failure to implement security standards can bring with it a series of problems for Ecuadorian public institutions, such as theft or leakage of information, whether due to digital techniques such as embedding malicious software, viruses that affect the system, keyloggers to reveal access codes, spyware to maliciously collect information from systems and use it; SQL injection attacks, denial of service attacks [4]. Then there is the social engineering through which certain people outside or related to the organization, access the internal information of the company and use it to harm it [5]. And we must not forget natural disasters such as fires, landslides, floods or human error [6].

Ecuador's public institutions are a group of organizations designed to provide and facilitate services to the community. Currently there are 3,251 public institutions that are directly part of the central public administration or others that belong to decentralized autonomous governments. These institutions have implemented a series of Comprehensive Information Systems, which has facilitated citizen processes or procedures. They are also required to integrate their information into the National Information System and their Information Security Management is evaluated by the Ministry of Telecommunications and the Information Society (MINTEL).

Several significant flaws have been identified in current management systems that have allowed:

- The data leak of 20 million Ecuadorian data between living and dead people, which included personal, banking and even tax information.
- Alteration of citizen data in databases of the Civil Registry and the National Council for Equality of Disabilities (CONADIS) in order to obtain benefits or commit criminal acts.
- Denial of service attacks on websites of public entities.

As a result of the events that have become public knowledge in recent years, citizens distrust the capacities of public institutions to ensure the confidentiality, integrity and availability of information.

At an international level, organizations such as The International Standard Organization (ISO), have elaborated the family of the ISO/IEC 27000 standard, which contains the best practices for the development, implementation and continuous improvement of Information Security Management Systems (ISMS) in organizations [7]. Implementation of this standard allows organizations of any type to properly manage the security of their assets, their financial information, intellectual property, employee details and information entrusted by third parties related to the institution in some way in Ecuador, the Ecuadorian Standardization Service (INEN) is the entity in charge of ensuring compliance with quality requirements of public and private institutions, in this sense it has powers to regulate and comply with requirements to guarantee the safety of users and the provision of a quality service. Ecuadorian Standardization Service has adopted ISO technical standards and under the Government Information Security Scheme (EGSI) together with the Ministry of Telecommunications and the Information Society (MINTEL), periodically evaluate the Security Management of the Information from Ecuadorian public organizations.

The objective of this research is to analyze the security standards adopted by public organizations in Ecuador to improve their management of information security.

¿Does the implementation of security standards in a public organization help improve the management of information security?

The adoption of security standards based on international regulations allows organizations to establish policies, procedures and controls with the aim of reducing the risks to which the information is exposed. Due to the security controls installed, an improvement in security management can be achieved, in terms of efficiency and continuous improvement.

Assessing information security levels in each of its dimensions will ensure business continuity, as well as early identification of security risks and potential damage to information. The evaluation of security controls and metrics is necessary since they provide information for decision-making at the three decision levels of organizations: operational, tactical and strategic [8].

Related references: An approach of National and International Cybersecurity Laws and Standards to Mitigate Information Risks in Public Organizations of Ecuador [9], An Empirical Study of Information Security Management Success Factors [10], Critical Success Factors Analysis on Effective Information Security Management: A Literature Review [11], Cyber-security Policy Framework and Procedural Compliance in Public Organisations [12], Identifying factors of “organizational information security management” [13], Organizational factors to the effectiveness of implementing information security management [14], Security Related Issues In Saudi Arabia Small Organizations: A Saudi Case Study [15], Analysis of Appropriate Standards to solve Cybersecurity problems in Public Organizations [16], Adapting ISO 27001 to a Public Institution [17], Ecuadorian Standardization Service [18].

The deductive method was applied for the review and analysis of appropriate standards for public institutions that allow the improvement of information security management.

The results obtained were: a Diagram of activities for the adoption of standards for public organizations; a prototype standards-based Information Security Management Model; and an Information Security Management Matrix, from which the Risk Mitigation Percentage was calculated.

Maintaining high levels of security in public organizations requires the adoption of control standards in the different areas and the collaboration of the different organizational and hierarchical levels of public organizations.

2. Materials and Methods

In the first instance in Materials, was made a search for information from different sources that allowed defining the standards used in public government organizations and the adoption process. Secondly, in Methods, the steps to achieve the results were defined.

2.1 Materials

2.1.1 Governmental Information Security Scheme – EGSI

Ecuadorian public organizations have the obligation and responsibility to protect the information they handle, much of this information has to do with citizens and another

part strictly with the internal management of the company. In any case, with the advancement of Information and Communication Technologies (ICT), state organizations have given greater attention to the protection of their information assets and thus generate confidence in citizens that their data is safe [16].

In Ecuador, Ministerial Agreements have been issued in order to efficiently and effectively manage information security in public entities. In turn, it maintains the Commission for Computer Security and Information and Communication Technologies in charge of computer security issues for government entities [16]. This commission was created with the purpose of analyzing the situation of information security in public institutions belonging to the Ecuadorian state and has come to determine the need to apply rules and procedures for information security, and incorporate culture and institutional processes, its permanent management, for this purpose the Government Information Security Scheme (EGSI) was prepared.

The objective of the EGSI is to increase the security of information in public entities. The implementation of the EGSI is carried out through the Continuous Improvement Cycle (PDCA), which has four steps: Plan, Do, Act and Verify.

Table 1: EGSI Compliance Classification

% of milestones met	EGSI Compliance Level Weighting
90% a 100%	High
75% a 89%	Medium
50% a 74%	Regular
<50%	Low

The EGSI is divided into 11 sections and has 126 priority guidelines or milestones that must be evaluated and met by public organizations in Ecuador [16]. The final qualification of compliance with the EGSI is performed according to Table 1.

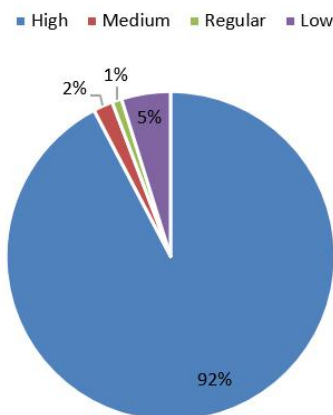


Figure 1. EGSI compliant entities

Figure 1 shows the level of compliance of public organizations in Ecuador, according to the milestones that organizations must meet.

2.1.2 ISO/IEC 27001: 2013

ISO/IEC 27001: 2013 is the international standard that governs the management of information security. It specifies the requirements that organizations must follow regarding the implementation of an Information Security Management System (ISMS) [17]. ISO/IEC 27001: 2013 defines information security as a preservation of the confidentiality, integrity and availability of information, essential dimensions that must be covered to a greater or lesser extent, controls, standards established in an organization [11].

2.1.3 ISO/IEC 27002: 2013

Describes the Code of Practice for the implementation of information security controls. ISO / IEC 27002: 2013 provides guidelines for organizational information security standards and information security management practices, including the selection, implementation and management of controls taking into account the information security risk environments of the organization [13]. Through ISO / IEC 27002: 2013, institutions will be able to select the appropriate standardized controls that will form part of the Information Security Management System for the organization. Through this standard, companies can also develop their own information security management guidelines. The standard has 114 controls, grouped into 14 domains and 35 control objectives that organizations can adapt according to their nature [6].

2.1.4 INEN ISO/IEC 27002

INEN ISO / IEC 27002 is the Ecuadorian technical standard, aimed at institutions derived from the Central Public Administration, whose objective is the adequate Management of Information Security. This standard is a translation of the ISO / IEC 27002 standard, so its structure is the same as that of the international standard. The implementation of the standard established by INEN serves as a reference for public institutions to select and adopt commonly accepted controls as part of the implementation process of the Government Information Security Scheme (EGSI) [18]. According to this standard, the assets of an organization are in constant threat of being affected, in which processes, systems, networks and people are involved. Changes in processes, business systems and other external changes can be causes of risks for information security, so it is necessary to implement a set of appropriate controls: policies, processes, procedures, organizational structures using resources software and hardware to ensure an organization meets its security objectives and strategic objectives.

2.2 Methods

The objective of this study was to analyze the security standards that Ecuadorian public organizations can adopt to improve their management of information security. For this, the deductive method was applied to review and analyze the appropriate standards for public institutions.

With the objective of evaluating the impact on the Management of information security produced by the adoption of security standard controls, scales of assessment of Confidentiality, Integrity and Availability of information were developed and used.

The creation of the prototype of the Information Security Management model was carried out based on the analysis of the ISO IEC 27001 and 27002 standard. From the 27001 standard, the dimensions of information security were considered: confidentiality, integrity and availability. In the same way, the 14 domains of the 27002 standard were considered, resulting from the analysis of 4 proposed dimensions that are part of public organizations.

For the evaluation of the impact produced by the adoption of controls in the Information Security Management, an agency of the Civil Registry of Ecuador was taken as a reference, which is part of the institutions that have completed the first phase of the EGSI.

Table 2: Scale to assess Confidentiality of information

Scale	Value	Criterion
High	3	Control is essentially important and has a crucial effect in ensuring the confidentiality of information
Medium	2	The control in place is moderately important to ensure the confidentiality of the information
Low	1	The established control is important to ensure the confidentiality of the information

Table 2 shows the rating scale of the Confidentiality of the information. The integrity and availability of the information were assessed based on the same scale. The objective of this assessment was to provide qualitative values to each quantitative value, depending on the importance of establishing control to ensure company information. The quantitative assessment of the dimensions of information security, then allowed calculating the assessment of the impact of controls, which was carried out by applying the following formula:

$$VIC = \frac{C + I + D}{n} \tag{1}$$

Where C is Confidentiality, I is Integrity, D is Availability, VIC is Control Impact Assessment and n is the Number of security dimensions evaluated, which for the case is 3.

Table 3: Control Impact Assessment Scale

Scale	Value	Criterion
High	3 - 2	Must be installed immediately
Medium	1,99 - 1,01	It must be established in the shortest possible time
Low	1	Must be established when resources are available

The results of the Control Impact Assessment were weighted using the scale in Table 3. Then the Risk Mitigation Percentage is calculated. To do this, first calculate the average VIC using the formula:

$$\bar{X} = \frac{\sum x}{n} \tag{2}$$

Finally, with the mean of X the Risk Mitigation Percentage was calculated, dividing the mean of X for the maximum score that each control can obtain, in this case 3 and multiplying by 100%.

$$PRM = \frac{\bar{X}}{3} * 100\% \tag{3}$$

3 Results

3.1 Diagram for the Adoption of Standards in Public Organizations

The improvement of Information Security Management in public organizations can be achieved through the adoption of standards, this requires that a series of basic steps be followed:

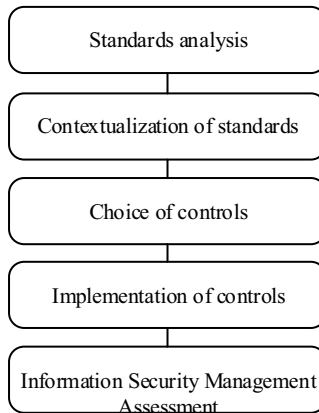


Figure 2. Adoption of standards to improve information security management

Figure 2 shows the activities that must be followed to improve Information Security Management in Public Organizations. The first step is the review and analysis of national and international standards, the second step is to contextualize the standards according to the nature of the organization to know if the controls belonging to that standard are applicable to the public organization, the third step is the choice of the information

security controls most appropriate to the nature of the organization, the fourth step is the implementation of said controls, the last step is the evaluation of the Information Security Management achieved with the implementation of the controls.

3.2 Prototype of a Standards-based Information Security Management Model

The analysis of the national and international norms and standards to which the public institutions of Ecuador must abide, allowed the creation of the following prototype of the Information Security Management Model based on standards.

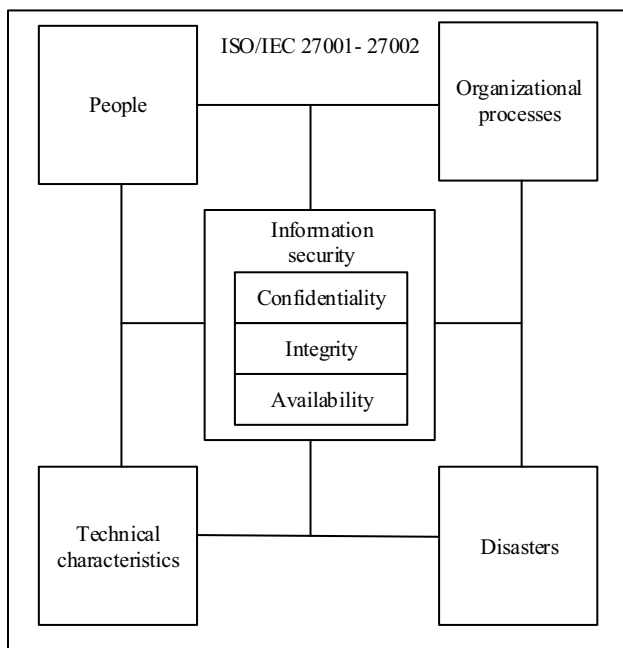


Figure 3. Standards-based information security management

This prototype integrates the domains of the ISO/IEC 27002 standard which were classified into 4 dimensions: People, Organizational Processes, Technical Characteristics, and Disasters. Information security management is carried out on these 4 dimensions to achieve satisfactory levels of information security in terms of its 3 dimensions: Confidentiality, Integrity and Availability.

The different controls of ISO/IEC 27002 that are part of the 4 dimensions represented in the model in Figure 3, can be evaluated with the Minimum Distance Method. The Minimum distance to an ideal point method is a method that allows evaluating different object parameters in a n-dimensional unit mathematical space. The minimum distance method is a methodology that can be used to assess computer security in organizations [6]. With this method, public organizations can evaluate controls from adapted standards, using the following equation:

$$d = \sqrt{(X_i - X_1)^2 + (X_i + X_2)^2 + \dots + (X_n - X_m)^2} \quad (4)$$

Eq. 4 calculates the Euclidean distance, distance from any point in n-dimensional space to the perfect point (1, 1, 1, ...) or the point that the user defines as an ideal point taking into account their requirements and available resources, from there you can decide how to spend resources to manage computer security, this is very useful when it comes to optimizing resources and it is required to give higher priority to some controls than others. Depending on whether (1, 1, 1, ...) or any other point was set as the ideal point, the parameter values can take values very close to one or the ideal point, which means that they are more important than the others; and therefore, the organization will spend more resources to control said parameters.

3.3 Information Security Management Matrix

A security management matrix was carried out in which different controls that the public organizations implement were weighted, according to the impact that said control produces in maintaining the security of the information evaluated in its 3 dimensions.

Table 4: Information Security Management Matrix

Controls	Impact on information security				Scale
	C	I	D	VIC	
Biometric	3	1	1	1,67	Medium
Security cameras	3	1	1	1,67	Medium
Firewalls	2	2	2	2	High
Antivirus	2	2	1	1,67	Medium
Wireless controller, access points	2	1	1	1,33	Medium
Hard disk backups	1	3	3	2,33	High

The generated matrix is shown in Table 4, where the backup of hard drives obtained the highest impact value in Information Security Management, followed by the implementation of firewalls. The wireless controller for access points is the control that obtained the lowest rating.

It was determined that the backup of hard drives and the implementation of firewalls, are controls that cause a high impact in the improvement of Information Security Management, so these controls must be implemented immediately in order to safeguard the information and organization assets. The installation of biometrics, security cameras, antivirus and configuration of wireless controller of access points, are controls that cause a medium impact so they must be implemented in the organization in the shortest possible time.

3.4 Calculation of Risk Mitigation Percentage

The results obtained in the Control Impact Assessment were used to calculate the Risk Mitigation Percentage.

$$\bar{X} = \frac{1,67 + 1,67 + 2 + 1,67 + 1,33 + 2,33}{6} = 1,78 \quad (5)$$

Eq. 5 calculated the average of the impact evaluations of the controls.

$$PRM = \frac{1,78}{3} * 100\% = 59,3\% \quad (6)$$

The calculation of the Risk Mitigation Percentage was carried out by applying Eq. 3 and allowed determining areas in which controls should be improved to reduce the risks related to information.

4 Discussion

This work presents an analysis for the adoption of Security Standards in order to improve Security Management in Public Organizations. Any public organization that intends to adopt national and international standards may apply the methodology proposed in this work.

The Information Security Management model developed considered the dimensions described in ISO / IEC 27002, in another study [17], and the ISO / IEC 27001.

In this study, an Information Security Management Matrix was developed based on the impact produced by the selected controls of the ISO / IEC 27002 standard according to the nature of the organization; contrary to what has been done in other studies where they analyze the impact of a risk becoming a threat, which helps determine the level of risks according to their impact and probability of occurrence; however, the proposal developed in this study is a more simplified alternative that focuses on Safety Management for risk mitigation.

The Risk Mitigation Percentage found for a public organization was 59.3%. In a similar investigation [9], this percentage was 58.6%.

5 Future Word and Conclusion

In the future, the factors that positively or negatively influence the adoption of standards in public organizations should be investigated.

Maintaining high levels of security in public organizations requires the adoption of control standards in different areas and the collaboration of the different organizational and hierarchical levels of public organizations.

The improvement of Information Security Management requires its evaluation in its three dimensions: Confidentiality, Integrity and Availability and is achieved with adequate management of people, organizational processes, technical characteristics and disasters around the controls described in the standard. ISO / IEC 27002.

The backup of hard drives and the implementation of firewalls are controls that produce a greater impact to ensure the security of the information, so they must be implemented immediately in all public institutions in order to over-safeguard the

information of the organization. The installation of biometrics, security cameras, antivirus and wireless access point controller configuration are controls that cause a medium impact, therefore they must be implemented in public organizations in the shortest time possible.

Most of Ecuador's public organizations have implemented the Government Information Security Scheme (EGSI) in conjunction with the INEN ISO / IEC 27002 standard, which demonstrates the progress of Ecuadorian public organizations in the area of protection and management of information security.

Acknowledgments

The authors thank to Universidad Politécnica Salesiana del Ecuador, to the research group of the Guayaquil Headquarters "Computing, Security and Information Technology for a Globalized World" (CSITGW) created according to resolution 142-06-2017-07-19 and Secretaría de Educación Superior Ciencia, Tecnología e Innovación (Senescyt).

References

- [1] Cannon G, Statham P, Yamada A. Biometric Security, Standardization. *Encycl Biometrics*. 2009;122–9.
- [2] Ministerio de las Telecomunicaciones y de la sociedad de la información. Ecuador ocupa sexto lugar en la región, según Índice de Ciberseguridad. [Internet]. [Consulted 10 Jul 2012] Available in <https://www.telecomunicaciones.gob.ec/ecuador-ocupa-sexto-lugar-en-la-region-segun-indice-de-ciberseguridad/>
- [3] Olifer D. Evaluation metrics for ontology-based security standards mapping. 2015 Open Conf Electr Electron Inf Sci eStream 2015 - Proc. 2015;17–20.
- [4] Heimes R. Global InfoSec and Breach Standards. *IEEE Secur Priv*. 2016;14(5):68–72.
- [5] Sedinić I, Perušić T. Security Risk Management in complex organization. 2015 38th Int Conv Inf Commun Technol Electron Microelectron MIPRO 2015 - Proc. 2015;(May):1331–7.
- [6] Felipe FD, Acevedo EM, Sanchez MM. Evaluating informatics security in an organization: The minimal distance method. Proc 2017 IEEE 24th Int Congr Electron Electr Eng Comput INTERCON 2017. 2017;17–9.
- [7] Lontsikh PA, Karaseva VA, Kunakov EP, Livshitz II, Nikiforova KA. Implementation of information security and data processing center protection standards. In: 2016 IEEE Conference on Quality Management, Transport and Information Security, Information Technologies, IT and MQ and IS 2016. 2016. p. 138–43.
- [8] Doynikova E, Fedorchenko A, Kotenko I. Ontology of metrics for cyber security assessment. *ACM Int Conf Proceeding Ser*. 2019.
- [9] Toapanta SMT, Gurumendi AJ, Gallegos LEM. An approach of national and international cybersecurity laws and standards to mitigate information risks in public organizations of ecuador. *ACM Int Conf Proceeding Ser*. 2019;(December):61–6.
- [10] Zammani M, Razali R. An empirical study of information security management success factors. *Int J Adv Sci Eng Inf Technol*. 2016;6(6):904–13.
- [11] Tu Z, Yuan Y. Critical success factors analysis on effective information security management: A literature review. 20th Am Conf Inf Syst AMCIS 2014. 2014;1–13.
- [12] Lubua EW, Pretorius PD. Cyber-security policy framework and procedural compliance in public organisations. *Proc Int Conf Ind Eng Oper Manag*. 2019;(July):1847–56.
- [13] Singh AN, Gupta MP, Ojha A. Identifying factors of "organizational information security management." *J Enterp Inf Manag*. 2014;27(5):644–67.
- [14] Chang SE, Ho CB. Organizational factors to the effectiveness of implementing information security management. *Ind Manag Data Syst*. 2006;106(3):345–61.
- [15] Almubayedh D, Khalis M Al, Alazman G, Alabdali M, Al-Refai R, Nagy N. Security Related Issues in Saudi Arabia Small Organizations: A Saudi Case Study. 21st Saudi Comput Soc Natl Comput Conf NCC 2018. 2018;1–6.
- [16] T SMT, E GSG, Enrique L, Gallegos M. Analysis of Appropriate Standards to solve Cybersecurity

problems in Public Analysis of Appropriate Standards to solve Cybersecurity problems in Public Organizations. In 2020.

- [17] Carvalho C, Marques E. Adapting ISO 27001 to a Public Institution. *Iber Conf Inf Syst Technol Cist.* 2019;2019-June(June):19–22.
- [18] Servicio Ecuatoriano de Normalización. NTE INEN-ISO/IEC 27002. 2017.

Simple Receiving Scheme Based on Balanced Detection for Half-Cycled SSB DD-OFDM Signal

Pengfei YANG^{a,1}, Xue CHEN^b

^a*The 6th Research Institute of China Electronics Corporation, Beijing, 100083, China*

^b*State Key Laboratory of Information Photonics & Optical Communications, Beijing University of Posts & Telecom, Beijing, 100876, China*

Abstract. In this letter, we proposed a simple balanced-detection reception scheme for the half-cycled single-sideband direct-detected optical orthogonal frequency division multiplexing (HSSB DD-OFDM) signal with decreased guard band. By employing this scheme, each entire OFDM symbol can be recovered perfectly, while the signal-to-signal beat interference (SSBI) can be eliminated, the guard band can be reduced greatly and the tolerance to phase noise induced inter-channel interference (PN-ICI) and potential benefit of low peak to average power (PAPR) are retained. The simulation results demonstrate that a 40 Gbps 16-QAM HSSB DD-OFDM signal was achieved successfully.

Keywords. optical OFDM, direct-detection, beat interference, guard band

1. Introduction

Optical orthogonal frequency division multiplexing (OFDM) has been proposed and intensely investigated due largely to its enhanced spectral efficiency (SE) and excellent immunity to transmission deteriorations such as chromatic dispersion (CD) and polarization-mode dispersion (PMD), and fine granularity for dynamic bandwidth allocation [1]. According to the configurations of optical receivers, the optical OFDM systems could be classified into two sorts: coherent optical OFDM (CO-OFDM) and direct-detected optical OFDM (DD-OFDM). Compared to CO-OFDM, DD-OFDM systems require a simpler and more cost-effective receiver since the desired OFDM signal is acquired by detecting the beat outcomes of the carrier and signal using a square-law photodiode (PD), and therefore DD-OFDM could be one of the promising candidates to economically supply high bandwidth in metropolitan area networks and long-reach passive optical network [2].

The single-sideband format (SSB-OFDM) is adopted to combat the immanent CD-induced power fading associated with double-sideband (DSB) OFDM and to improve the SE in DD-OFDM. Typically, SSB-OFDM signal will be influenced by inherent signal-to-signal beat interference (SSBI) after square-law PD detection in case

¹ Corresponding Author; E-mail: yhpdf@126.com

the bandwidth of guard band (GB) between the main carrier and OFDM signal is smaller than the OFDM signal [3]. As mentioned in [1, 4], the blank frequency GB between the main carrier and the data signal was proposed to protect the signal from being interfered by the SSBI, and thus the system SE will be reduced to as much as half. Many research efforts have been made to address this issue. In [5], authors proposed the turbo coding technique to combat the SSBI at the cost of SE and complexity of forward-error-correction decoding. In [3], an iterative algorithm for the SSBI cancellation at the receiver was proposed for the virtual SSB-OFDM system, which demands large carrier-to-signal power ratio (CSPR) and increases the computation complexity. Moreover, in [6], a SSBI cancellation reception is proposed and analyzed in the SSB-OFDM system with an optical filter and a balanced receiver. However, the receiver sensitivity is reduced in that half of the overall optical power is used to rebuild the SSBI without contribution to the received OFDM signal. Another scheme which was called interleaved OFDM also can effectively remove the impingement of SSBI by loading data only in odd subcarriers, while sacrificing half of the system SE. It's noted that the interleaved OFDM signal can acquire tolerance toward PN-ICI and the intrinsic defect of OFDM signal, i.e., high PAPR can be significantly ameliorated in interleaved OFDM scheme [7-10]. Further, in [11], based on the symmetry characteristic of the interleaved OFDM symbol in the time domain, the half-cycled DD-OFDM transmission and receiving scheme was proposed to fight the SSBI without SE sacrifice. Specifically, the second half of interleaved OFDM symbol is cut away during the transmission and recovered by reversing the corresponding first half after square-law detection and analog to digital conversion (ADC). However, the simple reception scheme for half-cycled OFDM signal in [11] cannot guarantee effective OFDM signal recovery especially when the system is configured with small CSPR. Actually, the CSPR is as high as 20 dB in the experimental setup of DSB half-cycled DD-OFDM transmission to maintain performance compared with the traditional interleaved OFDM signal [11].

In this letter, a simple receiving scheme for half-cycled SSB optical OFDM signal is proposed. The receiver consists of an optical interleaver (IL), a 2×2 3 dB optical coupler (OC) and a balanced PD pair. Thanks to perfect restoration of the half-cycled interleaved OFDM signal, the SSBI can be eliminated completely; on the other hand, the GB just needs to meet the requirement for splitting the optical carrier and SSB OFDM signal, so the SE can be meliorated greatly by decreasing the GB. The simulation results of the 40 Gbps 16-QAM HSSB DD-OFDM signal with considerably small GB demonstrate that the proposed scheme efficiently receives the half-cycled OFDM signal and system performance catches up with SSB DD-OFDM systems with sufficient GB or interleaved subcarrier configuration.

2. System Model

The SSB-OFDM signal may be generated by different methods, such as optical intensity modulation of radio frequency OFDM(RF-OFDM) signal with optical SSB filtering, optical IQ modulation of the baseband OFDM signal and attaching of an optical carrier, etc [1-6]. Figure 1 illustrates a conceptual diagram of the HSSB DD-OFDM system combined with the proposed receiving scheme. The output of continuous wave laser diode (LD) is split into two paths. One path passes through an

IQ modulator which is biased at null point and driven by the complex baseband OFDM signal. The other path simply provides the main optical carrier and is mixed with the modulated optical OFDM signal using a 3 dB coupler. The fiber length in both paths of the coupler is matched to eliminate the propagation delay between the two paths. Separate generation of the optical OFDM signal and the main optical carrier enables flexible adjustment of the CSPR and large dynamic range for the baseband signal. Moreover, only the positive part of the inverse fast Fourier transform (IFFT) is filled to realize SSB-OFDM modulation and a GB is placed between the main optical carrier and OFDM signal by nullifying some subcarriers.

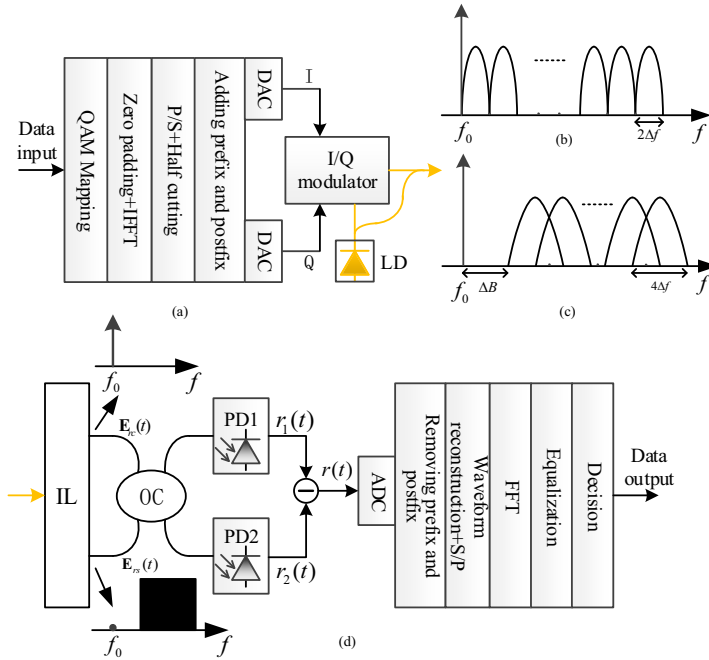


Figure 1. (a) HSSB DD-OFDM transmitter with optical I/Q modulation of the baseband OFDM signal and combination of a main optical carrier, (b) the spectral profile of conventional interleaved SSB DD-OFDM, (c) the spectral profile of HSSB DD-OFDM with a GB, (d) the proposed HSSB DD-OFDM receiver.

Let X_k be the complex-valued symbol representing the constellation point on the k th subcarrier for a given symbol. When subcarrier interleaving is introduced, symbols are only loaded in subcarriers with odd index and subcarriers with even index are loaded with “0”s. Then the IFFT realization of an interleaved OFDM signal for that symbol are given by x_m where

$$x_m = \frac{1}{N} \sum_{k=0}^{\frac{N}{2}-1} X_{2k+1} e^{j \frac{2\pi m(2k+1)}{N}}, \quad m = 0, 1, 2, \dots, N-1 \quad (1)$$

and N is the size of the IFFT. The time domain symmetry of the interleaved OFDM

signal could be represented as

$$x_{m+\frac{N}{2}} = \frac{1}{N} \sum_{k=0}^{\frac{N}{2}-1} X_{2k+1} e^{j\frac{2\pi(m+\frac{N}{2})(2k+1)}{N}} = -\frac{1}{N} \sum_{k=0}^{\frac{N}{2}-1} X_{2k+1} e^{j\frac{2\pi m(2k+1)}{N}} = -x_m. \tag{2}$$

Eq. (2) demonstrates that the first and second halves of one interleaved OFDM symbol possess the opposite amplitude in time domain. In half-cycled interleaved OFDM scheme, the waveform of one symbol is truncated by only transmitting the first half. When the second half is somehow retrieved by the receiver, the complete waveform of that symbol can be reconstructed. By this means, the time taken to transfer data is halved and data rate is doubled. Therefore, the SE will remain unchanged despite the reserved subcarriers with even index. Meanwhile, to ensure the phase continuity between the recovered first and second halves in one symbol, the guard interval of the new half-cycled symbol should be carefully designed. As shown in Figure 2, except the routine prefix cloned from the tail of the interleaved OFDM symbol, the half-cycled OFDM symbol is added a postfix which is part of the head of the second half of the symbol.

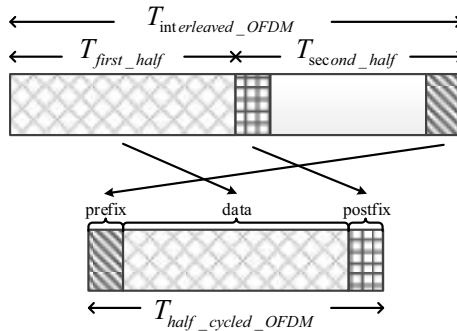


Figure 2. Half-cycled OFDM symbol with prefix and postfix.

As illustrated in Figure 1(c), the spectrum of the optical signal at the transmitter is the addition of a linear version of the electrical OFDM signal and a main optical carrier. Considering the truncating operation with OFDM symbol will not change the spectrum shape but broaden its spectral bandwidth by two subcarrier spacings, the transmitted optical HSSB DD-OFDM signal could be expressed as

$$\mathbf{E}(t) = e^{j2\pi f_0 t} + \beta e^{j2\pi(f_0 + \Delta B)t} \sum_{k=0}^{\frac{N}{2}-1} X_{2k+1} \Pi(t - \frac{T_s}{2}) e^{j2\pi f_{2k+1} t} = \mathbf{E}_c(t) + \mathbf{E}_s(t), \tag{3}$$

where f_0 is the main carrier frequency, ΔB is the GB between the main carrier frequency and OFDM signal, β is the scaling coefficient that represents the CSPR; $\Pi(t)$ is the pulse shaping function, which is 1 in $[0, T_s/2]$ and 0 otherwise; T_s is

the OFDM symbol duration. In Eq. (3), only one OFDM symbol is assumed for the sake of mathematical simplicity. If the fiber nonlinearity remains sufficiently low, the fiber channel can be modeled as a linear system. Therefore, the received optical signal at the IL still can be expressed in the form like Eq. (3). Then the IL separates the received optical HSSB DD-OFDM signal as the main optical carrier, $\mathbf{E}_{rc}(t)$, and the OFDM signal, $\mathbf{E}_{rs}(t)$. Note that the GB required by the IL could be much narrower than the bandwidth of OFDM signal. Then the main optical carrier and the OFDM signal are sent into a 2×2 3 dB OC with the following transmission matrix:

$$H = \begin{bmatrix} \frac{1}{2} & \frac{j}{2} \\ \frac{j}{2} & \frac{1}{2} \end{bmatrix}. \quad (4)$$

In Eq. (4), we assumed that this ideal 3 dB OC has no additional loss and its two optical paths introduce the same time delay. Given that the two PDs have identical sensitivity, the two output photocurrents can be written as:

$$\begin{bmatrix} r_1(t) \\ r_2(t) \end{bmatrix} = \frac{1}{4} \begin{bmatrix} |\mathbf{E}_{rc}(t) + j\mathbf{E}_{rs}(t)|^2 \\ |j\mathbf{E}_{rc}(t) + \mathbf{E}_{rs}(t)|^2 \end{bmatrix} = \frac{1}{4} \begin{bmatrix} |\mathbf{E}_{rc}(t)|^2 + |\mathbf{E}_{rs}(t)|^2 + 2\text{Re}\{j\mathbf{E}_{rc}^*(t)\mathbf{E}_{rs}(t)\} \\ |\mathbf{E}_{rc}(t)|^2 + |\mathbf{E}_{rs}(t)|^2 - 2\text{Re}\{j\mathbf{E}_{rc}^*(t)\mathbf{E}_{rs}(t)\} \end{bmatrix}. \quad (5)$$

For each photocurrent in Eq. (5), the first item is the direct current (DC) component; the second item is the so called SSBI which comes from the self-beating of the OFDM signal; and the last item is the desired carrier-signal beating products. By comparison, it's found that $r_1(t)$ and $r_2(t)$ actually correspond to the detected products of the first and second half of one interleaved OFDM symbol, respectively – that is, the discarded second half of one interleaved OFDM symbol can be retrieved from $r_2(t)$. Meanwhile, regardless of the noise, due to the SSBI that could be strengthened by small CSPR configuration, the remained alternating current (AC) components (the second term and the third term) no longer strictly have the inverted amplitude after filtering the DC components in $r_1(t)$ and $r_2(t)$. So the restoration scheme for the half-cycled interleaved OFDM signal in [11] will be confined to the large CSPR condition. Further, via the subtraction operation we obtain

$$r(t) = r_1(t) - r_2(t) = \text{Re}\{j\mathbf{E}_{rc}^*(t)\mathbf{E}_{rs}(t)\} \quad (6)$$

which is the desired signal included in the first half of one interleaved OFDM symbol. Let $S_{MS}(n)$ be the signal of $r(t)$ processed by sampling, timing and guard interval removing, then the final waveform of one symbol could be reconstructed as

$$S(n) = [S_{MS}(n), -S_{MS}(n)]. \quad (7)$$

Although the first and the second halves of one symbol can be separately retrieved from $r_1(t)$ and $r_2(t)$ (hereafter, alternative scheme for short), one more analog-to-digital converter (ADC) and additional timing and guard interval removing operation will be required. Therefore, the proposed receiving scheme allows to simple and efficient recovery of the HSSB DD-OFDM signal while maintaining the immunity to SSBI interference.

3. Results and Discussion

A proof-of-concept simulation HSSB DD-OFDM link with 40 Gbps 16-QAM signal is built using VPItransmissionMaker and Matlab to investigate the feasibility of the proposed scheme.

The simulation platform's structure and detailed DSP for transmitter and receiver are described in Figure 1. The main system parameters are as follows: the FFT size is 256 while 60 interleaved subcarriers in the positive frequency bins are for the data; subcarriers between the first subcarriers and data-bearing subcarriers are reserved for GB varying from 1 to 10 GHz; overall length of prefix and postfix is 16 points; 20 training symbols are transmitted for overall channel estimation, followed by 500 data symbols. The baseband signal is modulated on the optical wave from an output branch of LD with central frequency of 193.1 THz and linewidth of 100 MHz. The other output branch is used for optical carrier insertion. By adopting the carrier insertion, the signal can keep in linear modulation region when tweaking the CSPR by setting different split ratio of the LD's two branches. The HSSB DD-OFDM signal is transmitted through a 100 km standard single mode fiber with the power loss of 0.2 dB/km and dispersion of 16 ps/nm•km. An EDFA and tunable optical attenuator are installed before the receiver to regulate the optical power. In the receiver, the optical IL with sharp edge at 193.102 THz splits the main optical carrier and optical OFDM signal. An ideal 2×2 3 dB OC is used to combine the main optical carrier and the optical OFDM signal. Then the two mixed signals from the OC are sent into the balanced PD pair. The following DSP procedure including synchronization, guard interval removal, waveform reconstruction, FFT, channel estimation based on intra-symbol frequency domain averaging, one-tap equalization, 16-QAM de-mapping and BER/EVM computation.

Figure 3 illustrates the system performance in terms of BER as a function of the CSPR for the HSSB DD-OFDM signal received by two schemes. The CSPR of ~0 dB is discovered to be the optimum value, which suggests that the best sensitivity is achieved while the carrier power equals to the signal power. Moreover, the performance difference between the two receiving schemes can be ignored. This optimum value agrees with the past known results of 0 dB for the gapped SSB DD-OFDM systems or interleaved SSB DD-OFDM systems, in which the desired signal and the SSBI noise are distributed over different frequency zones after the PD [7]. For a given total received optical power and fixed GB, as the CSPR deviates from the optimum value, the SNR decrease because the optical noise and transmission distortions become the leading factors.

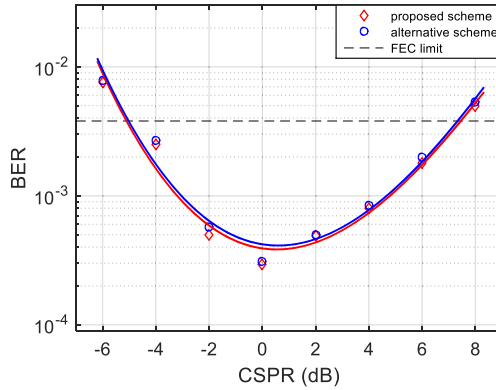


Figure 3. BER as the function of CSPR at the laser linewidth = 1MHz, GB = 5 GHz, received optical power = -15 dBm.

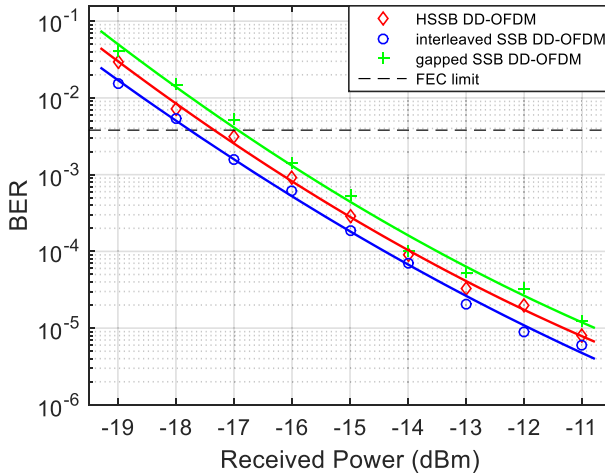


Figure 4. BER as the function of received optical power for HSSB DD-OFDM (GB = 5 GHz), gapped SSB DD-OFDM (GB = 10 GHz) and interleaved SSB DD-OFDM at the laser linewidth = 1MHz, CSPR = 0 dB.

In Figure 4 the BER performance versus the received optical power is carried out for the proposed, the previous gapped SSB DD-OFDM system and interleaved SSB DD-OFDM system at 40 Gbps with 16 QAM. The receiver sensitivities of HSSB DD-OFDM and interleaved SSB DD-OFDM are very close, which proves that HSSB DD-OFDM is practical. Compared to interleaved SSB DD-OFDM, HSSB DD-OFDM’s receiver sensitivity penalty mainly comes from the synchronization timing error which affects the phase continuity between the recovered first and second halves in one symbol. Due to low PAPR and tolerance toward PN-ICI, HSSB DD-OFDM achieves slight better receiver sensitivity than gapped SSB DD-OFDM.

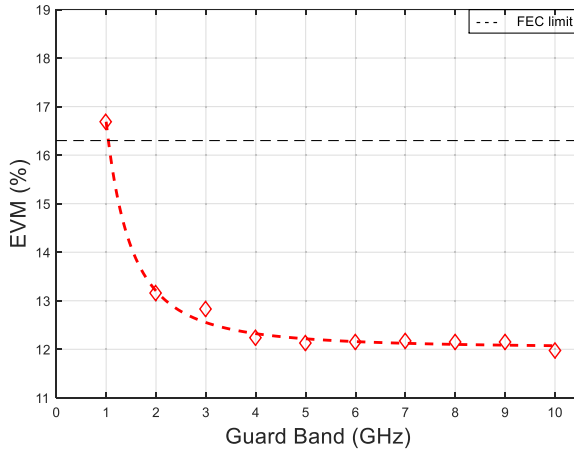


Figure 5. EVM as the function of GB with the laser linewidth = 1MHz, received optical power = -15 dBm.

To check the influence of GB on the HSSB DD-OFDM signal, the EVMs computed from the simulated symbols are in Figure 5. It's found that the EVMs are almost no change when the GB is equal to or bigger than 5 GHz. The EVM increases slowly and keeps below the FEC limit until the GB is reduced to 2 GHz, therefore the SE is meliorated greatly. As the GB is smaller than 2 GHz, the edge of the optical IL destroys the OFDM signal. Figure 6 gives the EVMs of the HSSB DD-OFDM signal with different laser linewidth from 1MHz to 10 MHz. It is found that EVM ascends almost linearly with the increment of laser linewidth and yet keeps below the FEC limit. Hence the HSSB DD-OFDM signal has good tolerance of laser linewidth.

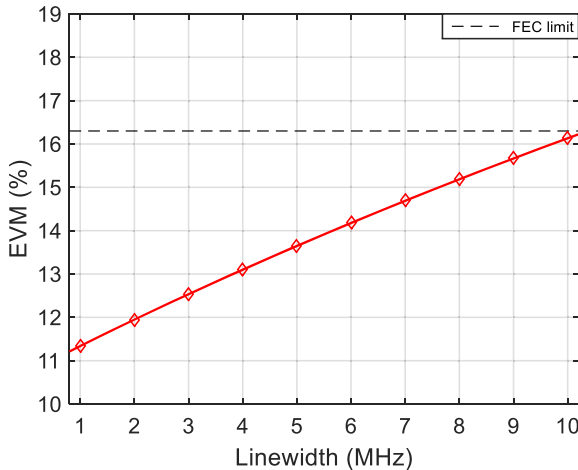


Figure 6. EVM as the function of the laser linewidth = 1MHz, received optical power = -15 dBm and GB = 5 GHz.

4. Conclusion

A simple receiving scheme for half-cycled SSB DD-OFDM signal has been proposed and demonstrated. In this scheme, since the half-cycled interleaved OFDM signal can be well retrieved, the immunity to SSBI interference is retained with the greatly meliorated SE despite the increased complexity of the receiver. We investigated the system performance with 40 Gbps 16-QAM transmission simulations. The simulation results show that the proposed scheme possesses approximate performance compared with typical SSB DD-OFDM system and interleaved SSB DD-OFDM system, inherits low PAPR and tolerance toward PN-ICI from interleaved OFDM system and lowers the requirements for GB and laser linewidth.

References

- [1] A. J. Lowery and J. Armstrong, "Orthogonal-frequency-division multiplexing for dispersion compensation of long-haul optical systems", *Opt. Express*, vol. 14, no. 6, pp. 2079–2084, 2006.
- [2] J. Armstrong, "OFDM for optical communications," *J. Lightwave. Technol.*, vol. 27, no. 3, pp. 189–204, Feb. 1, 2009.
- [3] W.-R. Peng, *et al.*, "Spectrally efficient direct detected OFDM transmission employing an iterative estimation and cancellation technique," *Opt. Express*, vol. 17, no. 11, pp. 9099–9111, 2009.
- [4] W.-R. Peng, *et al.*, "Spectrally Efficient Direct-Detected OFDM Transmission Incorporating a Tunable Frequency Gap and an Iterative Detection Techniques," *J. Lightwave. Technol.*, vol. 27, pp. 5723–5735, 2009.
- [5] Z. Cao, *et al.*, "Direct-detection optical OFDM transmission system without frequency guard band," *IEEE Photon. Technol. Lett.*, vol. 22, no. 11, 736–738, 2010.
- [6] S. A. Nezamalhosseini, *et al.*, "Theoretical and experimental investigation of direct detection optical OFDM transmission using beat interference cancellation receiver," *Opt. Express*, vol. 21, no. 13, pp. 15237–15246, 2013.
- [7] W.-R. Peng, *et al.*, "Theoretical and experimental investigations of direct-detected RF-tone-assisted optical OFDM systems," *J. Lightwave Technol.*, vol. 27, no. 10, pp. 1332–1339, 2009.
- [8] Z. Cao, *et al.*, "Interleaved and partial transmission interleaved optical coherent orthogonal frequency division multiplexing," *Opt. Lett.*, vol. 39, no. 7, pp. 2179–2182, 2014.
- [9] Mondal, M. R. H., and Panta, K., "Performance analysis of spatial OFDM for pixelated optical wireless systems." *Trans. Emerging Tel. Tech.*, vol. 28, SP. e2948, 2017.
- [10] Bharati, S., and Podder, P., "Adaptive PAPR Reduction Scheme for OFDM Using SLM with the Fusion of Proposed Clipping and Filtering Technique in Order to Diminish PAPR and Signal Distortion." *Wireless Pers. Commun.*, vol. 113, pp. 2271–2288, (2020).
- [11] F. Li, *et al.*, "SSMI cancellation in direct-detection optical OFDM with novel half-cycled OFDM," *Opt. Express*, vol. 21, no. 23, pp. 28543–28549, 2013.

Design, Simulation and Verification of a 7-DOF Joint Motion Simulation Platform

Ze CUI^{a,1}, Saishuai HUANG^a, Zenghao CHEN^a, Hongxin YANG^a and Danjie ZHU^b

^a*School of Mechatronic Engineering and Automation, Shanghai University, Shanghai, China*

^b*Orthopedics Dept, Zhejiang Provincial People's Hospital, Zhejiang, China*

Abstract. The human body has many joints, and joint injuries frequently occur in various sports. To explore the biomechanical state of ligaments or muscles in human joints before and after damage, and to help doctors judge the damage and repair of joints, this article proposes a seven-degree-of-freedom platform based on three rotations spherical parallel mechanism for simulating human joint motion. Taking the knee joint as an example, this article simplified its model, and performed kinematics simulation by ADAMS to verify the feasibility of this mechanism. And based on the TRIO motion controller, we established the physical testing system. The correctness is finally verified by experiment in kind, which proves the feasibility of the joint motion simulation platform. And in terms of accuracy, it also performs very well. For example, when it needs to rotate 30° around the Y-axis, its actual rotation angle is 29.6°, the error is less than 2%, and its translation error is also within 3%.

Keywords. seven degrees of freedom, joints, spherical parallel mechanism.

1. Introduction

To understand the function of the knee joint properly, we need to study the *In-Situ* force and the *In-Vivo* force, which need to add information in the range of ligament relaxation between full extension and deep flexion. In recent years, many studies have confirmed the role of joint in limiting tibial movement, however, due to the lack of appropriate experimental equipment to measure the total force in the ligament while the joint moving, this research has been hampered.

So far, many medical and engineering scholars have made outstanding contributions to this research. From 1996, Woo et al. [1] used a Universal Force/Robot System (UFS), which can provide a six-degree-of-freedom (DOF) manipulator when studying the knee of Anterior Cruciate Ligament (ACL). *In-Situ* force measurement used a multi-dimensional force sensor at the end of the manipulator. Atarod et al. [2-3] of the University of Calgary used a 6 DOF parallel mechanism when studying the movement of the knee joint of sheep and cooperated with a redundant DOF clamping device to fix the Femur ACL Tibia Complex (FATC) of sheep. The femur and tibia were fixed to conduct a sheep gait simulation experiment and measure the *In-Situ* force of the sheep knee ACL. Noble et al. [4] designed a universal robot based on rope drive and built a

¹ Corresponding Author: Ze Cui, School of Mechatronic Engineering and Automation, Shanghai University, 99 Shangda Road, Shanghai, China; E-mail: Cuize0421@126.com

musculoskeletal test system to design experiments to verify. Zens et al. [5] used a multi-degree-of-freedom joint motion simulator in the study of the length of the anterolateral ligament during the passive motion of the knee joint. Nesbitt et al. [6] used a KUKA robot to produce the ideal movement of the knee joint while exploring the effects of robot flexibility and bone bending on the knee kinematics simulation. In 2017, the Department of trauma, Innsbruck Medical University, developed a new type of knee biomechanical test-bed for in vitro evaluation of the knee joint [7]. The platform allows the kinematics of the knee joint to be studied in all six degrees of freedom; in 2018, the University of Aalborg proposed a new method for noninvasive and accurate measurement of knee joint relaxation in four degrees of freedom [8-12]. The joint measuring instrument combines a parallel manipulator and a 6-DOF force/torque sensor to reconstruct the relaxation measurement of tibiofemoral position and direction.

In this paper, we propose a 7-DOF joint simulation platform based on three rotations (3-RRR) spherical parallel mechanism, so that can we find a more accurate method to evaluate the mechanical function of the human joint and internal tissue. And based on the TRIO motion controller, a physical testing system has been established. The purpose of this study is to provide help for the problems related to the residual relaxation of the human joint and adverse prognosis, meniscus injury, and graft failure.

2. Design of Joint Motion Simulation Platform

2.1. Structure of Joint Motion Simulation Platform

As we all know, the six degrees of freedom of space is composed of 3 rotations and three translations. Through research on existing institutions and positive thinking, in the end, a "3+3+1" degree of freedom joint motion simulation platform was designed (as shown in Figure 1).

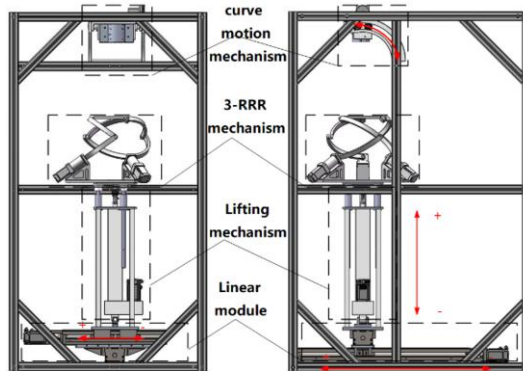


Figure 1. Structure diagram of the joint motion simulation platform

The mechanism is 1.08 meters long, 0.88 meters wide, and 1.6 meters high. It is mainly composed of an aluminum profile frame, an arc-shaped motion track, a 3-RRR mechanism, a lifting mechanism, and a two-dimensional linear module. The bottom two-dimensional linear module can realize linear motion in the X-axis and Y-axis directions. An electric cylinder drives the lifting mechanism used to complete the linear movement in the Z-axis direction. The 3-RRR mechanism can realize three-degree-of-freedom motion in space. The arc-shaped motion mechanism can manually compensate for the

pitch angle that the 3-RRR mechanism cannot reach. Through the cooperation of various mechanisms, six degrees of the freedom movement in space can be completed, thereby simulating human joints' movement. The movement space of the mechanisms is shown in Table 1.

Table 1. Movement Range of The Mechanism

	Degree of freedom	Range of motion
Rotation	Pitch	$\pm 50^\circ$
	Yaw	$-50^\circ \sim +140^\circ$
	Roll	$\pm 50^\circ$
Translation	X	$-540\text{mm} \sim 200\text{mm}$
	Y	$\pm 150\text{mm}$
	Z	$-20\text{mm} \sim +140\text{mm}$

The advantage of this mechanism is that the 3-RRR mechanism is a spherical parallel mechanism, its motion is extremely flexible, and it can rotate three degrees of freedom around the center of motion, which is very similar to the rotation of human joints, so it is very suitable for simulating joints' movement. What's more, the research on the spherical parallel mechanism at home and abroad is relatively mature. There are many kinds of research on it at present [13-14] (for example, the Agile Wrist [15]), which can help us better control the movement required to complete the experiment.

2.2. Control System of the Joint Motion Simulation Platform

The control system of the joint motion simulation platform includes a laptop, a TRIO motion controller, 6 AC servo drivers and 6 AC servo motors, three of them are 400W and the other three are 100W. Two 400W drivers and motors are used to drive the motion of linear modules along X-axis and Y-axis, the left 400W driver and motor are for the lifting mechanism, three 100W drivers and motors are used to drive the 3-RRR mechanism. The control system diagram is shown in Figure 2. The laptop, TRIO controller and AC servo drivers are connected by network cable; they communicate through Ethernet protocol, the platform can be driven by the program written on the laptop.

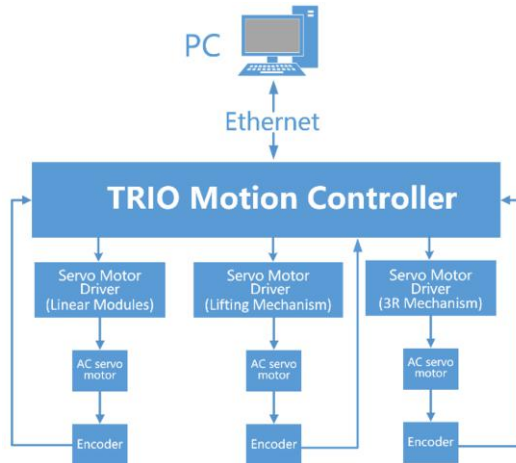


Figure 2. Control system block diagram

3. Adams Kinematics Simulation

3.1. Model Simplification of the Joint Motion Simulation Platform

To facilitate the simulation in the ADAMS software, the model is equivalently simplified. As shown in Figure 3, the 3-RRR mechanism, the center of the upper round surface of the rod, is equivalent to the center of knee motion, and the center of the lower round surface is the center of rotation of the 3-RRR mechanism; other structures have also been simplified accordingly.

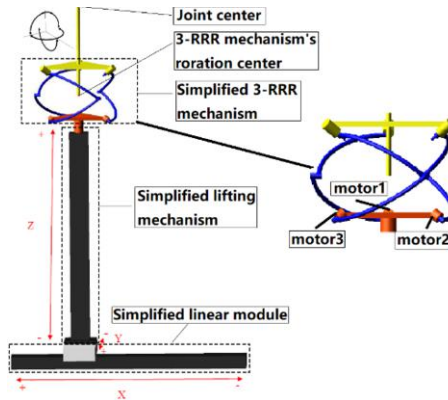


Figure 3. Simplified model of the joint motion simulation platform

3.2. Motion Simulation Analysis

According to the motion characteristics of each mechanism, the simplified model set the motion pair and drive and carried out some motion simulations. For example: Simulate the knee joint flexion movement of 30° , that is, the knee joint rotates 30° around the Y-axis to obtain the motion of each mechanism and the motion curve of the corresponding motion unit. In Figure 4, the pink dotted line represents the displacement in the Y-axis direction; the solid red line represents the displacement in the X-axis direction; the red dotted line represents the displacement in the Z-axis direction; the black dotted line represents the rotation angle of the motor 1 in the 3-RRR mechanism; The solid green line and the light blue dotted line represent the rotation angles of the motor 2 and the motor 3, respectively.

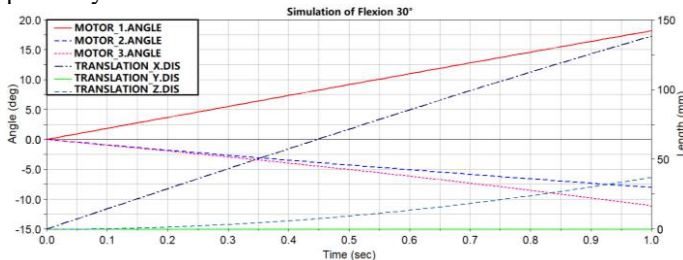


Figure 4. Simulate the knee joint flexion at 30° , the movement length/angle of each mechanism

4. Experimental Verification

To verify the feasibility of the joint motion simulation platform, a simple experiment has been done.

In this experiment, a control system has been set up based on the TRIO motion controller. The hardware mainly consists of a motion controller (TRIO, MC 4N-ECAT, England), three 400W AC servo drivers (Panasonic, MBDLN25BE, Japan), three 100W AC servo drivers (Panasonic MADLN05BE, Japan), three 400W motors (Panasonic, MSMF042L1U2M, Japan), three 100W motors (Panasonic, MSMF012L1U2M, Japan). As shown in Figure 5, the platform performed well in the experiment.

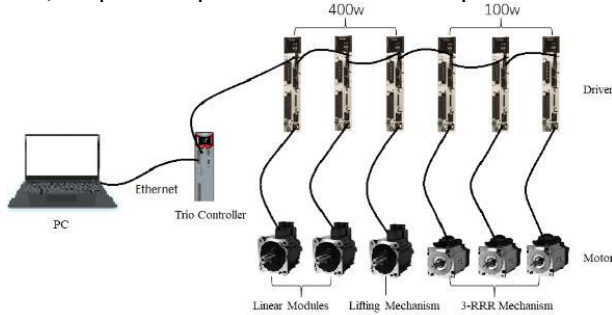


Figure 5. Control system connection diagram

The experiment process is as follows: a sheep’s FATC was fixed to the platform. The platform drove sheep’s FATC to rotate 30 degrees around Y-axis by a program. After the measurement, its actual rotation angle is 29.6°, the error is less than 2%, and its translation error is also within 3%. As shown in Figure 6, the platform performed well in the experiment.

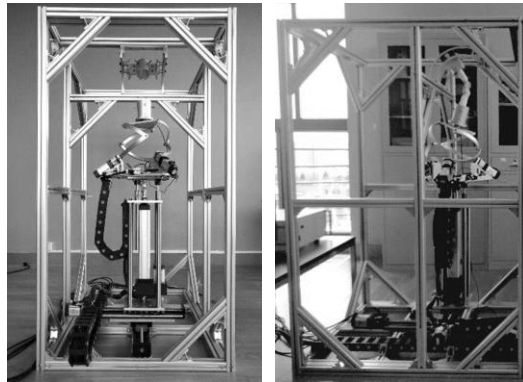


Figure 6. Experiment: sheep’s FATC rotate 30° around Y-axis

5. Conclusion

This paper proposes a joint motion simulation platform based on a 3-RRR spherical parallel mechanism. Its motion center can be controlled at the joint’s rotation center. It not only has a flexible space of seven degrees of freedom but also is easy to control. It is very suitable for studying human joint motion by simulating different movement states of human joints on the platform, which can perform *In-Situ* measurement of joints and various tissues in it. Once the mechanism is programmed, it can standardize the applied force's magnitude and direction to provide vectors (numbers and directions) to give less subjective joint measurements. And in terms of accuracy, it also performances very well, its rotation angle error is less than 2%, and its translation error is also within 3%. This

article verifies the platform's feasibility and effectiveness through kinematic analysis, ADAMS motion simulation, and experiments, which ensure that their comprehensive knowledge, diagnosis, and treatment are highly relevant to clinicians.

At present, due to limited conditions, this platform can only simulate some simple low-speed movements of the joint.

In the future, it could be extended to include further movement patterns. We hope to use this platform to research joint dynamics, impact mechanics, muscle dynamics, etc. This will promote the development of the whole limb musculoskeletal discipline and better understand the human kinematics structure. Consequently, it would be possible to investigate emerging questions, particularly in joint trauma and rehabilitation optimization.

References

- [1] Woo S L Y, Hollis J M, Adams D J, et al. Tensile properties of the human femur-anterior cruciate ligament-tibia complex: the effects of specimen age and orientation[J]. *The American journal of sports medicine*, 1991, 19(3): 217-225.
- [2] Atarod M, Rosvold J M, Frank C B, et al. Functional activity of the anterior and posterior cruciate ligaments under in vivo gait and static physiological loads[J]. *Annals of biomedical engineering*, 2013, 41(10): 2067-2076.
- [3] Atarod M, Rosvold J M, Frank C B, et al. A novel testing platform for assessing knee joint mechanics: a parallel robotic system combined with an instrumented spatial linkage[J]. *Annals of biomedical engineering*, 2014, 42(5): 1121-1132.
- [4] Noble L D, Colbrunn R W, Lee D G, et al. Design and validation of a general purpose robotic testing system for musculoskeletal applications[J]. *Journal of Biomechanical engineering*, 2010, 132(2).
- [5] Zens M, Niemeier P, Ruhhammer J, et al. Length changes of the anterolateral ligament during passive knee motion: a human cadaveric study[J]. *The American journal of sports medicine*, 2015, 43(10): 2545-2552.
- [6] Nesbitt R J, Bates N A, Karkhanis T D, et al. Impacts of Robotic Compliance and Bone Bending on Simulated in vivo Knee Kinematics[J]. *American journal of biomedical engineering*, 2016, 6(1): 12.
- [7] Green J T, Hale R F, Hausselle J, et al. A Reconfigurable Multiplanar in Vitro Simulator for Real-time Absolute Motion with External and Musculotendon Forces[J]. *Journal of Biomechanical Engineering*, 2017.
- [8] Heinrichs C H, Knierzinger D, Stofferin H, et al. Validation of a novel biomechanical test bench for the knee joint with six degrees of freedom[J]. *Biomedical Engineering/Biomedizinische Technik*, 2018, 63(6): 709-717.
- [9] Pedersen D, Vanheule V, Wirix-Speetjens R, et al. A novel non-invasive method for measuring knee joint laxity in four DOF: In vitro proof-of-concept and validation[J]. *Journal of biomechanics*, 2019, 82: 62-69.
- [10] Bates N A, Schilaty N D, Krych A J, et al. Influence of relative injury risk profiles on anterior cruciate ligament and medial collateral ligament strain during simulated landing leading to a noncontact injury event[J]. *Clinical biomechanics*, 2019, 69: 44-51.
- [11] Schall F, Seitz A M, Hacker S, et al. Proof-of-concept of a novel knee joint simulator allowing rapid motions at physiological muscle and ground reaction forces[J]. *Frontiers in Bioengineering and Biotechnology*, 2019, 7: 244.
- [12] O'Connor J D, Rutherford M, Bennett D, et al. Long-term hip loading in unilateral total hip replacement patients is no different between limbs or compared to healthy controls at similar walking speeds[J]. *Journal of biomechanics*, 2018, 80: 8-15.
- [13] Wu G, Caro S, Bai S, et al. Dynamic modeling and design optimization of a 3-DOF spherical parallel manipulator[J]. *Robotics and Autonomous Systems*, 2014, 62(10): 1377-1386.
- [14] Wu G, Caro S, Wang J. Design and transmission analysis of an asymmetrical spherical parallel manipulator[J]. *Mechanism and Machine Theory*, 2015, 94: 119-131.
- [15] F. Bidault, C.-P. Teng, J. Angeles, Structural optimization of a spherical parallel manipulator using a two-level approach, ASME Design Engineering Technical Conferences and Computers and Information in Engineering Conference, 2001, Pittsburgh, Pennsylvania, 2001.

3D Air-Ground Channel Model Reconstruction Based on the Experimental Data and Q-D Method

Alexander MALTSEV^{a,b}, Olesya BOLKHOVSKAYA^{b,1}, Vadim SERGEEV^b, Anton ELOKHIN^b and Alexander KOKAREV^b

^a LG Corp, Russia.

^b University of Nizhny Novgorod, Russia

Abstract. This paper proposes to use the Q-D (Quasi-Deterministic) method for reconstructing the angular and frequency characteristics of the 3D Air-to-Ground (A-G) channel from the available experimental data. This method allows to expand the applicability of tapped delay line (TDL) channel models for performance investigation of the aviation radio systems with directional antennas and multi-element antenna arrays. The developed 3D A-G channel models also give opportunity to take into account Doppler spread impact on the communication system performance.

Keywords. 3D Air-Ground channel, CIR, AoA, AoD, Doppler spread

1. Introduction

In recent years, it has been observed the very rapid development of unmanned aerial vehicles (UAV) technologies and their increasing use in various military and civil applications. Therefore, the number of works devoted to the study of characteristics and models of radio channels that provide UAV flight control and data transfer between them and ground stations has increased dramatically. Currently, many research organizations and standardization committees, such as [1], [2], are engaged in building radio communication networks with UAVs.

The developed analytical models [3]-[8] do not always describe the real behavior of the Air-to-Ground (A-G) channels due to insufficiently realistic assumptions. Therefore, in recent years, quite intensive experimental studies of radio communication channels for UAVs have been conducted to build their empirical models [9]-[11].

The most detail experimental studies of the A-G channel for UAV were conducted in 2013-2017 as part of a project funded by the NASA J. Glenn research center [12]-[15]. During the implementation of this project, a whole set of experimental measurements of the A-G radio channels in two frequency bands (L-band from 900-977 GHz, and C-band from 5030-5091 GHz) was carried out in various scenarios: over water, in hilly and mountainous areas, in suburban and urban environments. Based on the gathered experimental data, evaluation of path losses, power delay profiles, K-

¹ Corresponding Author: Olesya Bolkhovskaya, University of Nizhny Novgorod, Gagarina ave,23, 603109 Nizhny Novgorod, Russia; E-mail:obol@rf.unn.ru

factors, spatial correlation coefficients was made for all considered scenarios, and TDL channel models were developed.

However, for practical applications, it is desirable to have more detailed 3D A-G channel models that adequately describe the spatial, temporal, angle and frequency characteristics of the channel, and allow investigation of the modern aviation radio systems with directional and multi-element antennas.

In this paper, we, inspired by experimental data available from [12]-[15], propose to apply the Q-D method [16]-[17] to reconstruct the angular and Doppler spread characteristics of the rays. The key benefit of the Q-D approach compared to pure statistical channel models is its inherent support for spatial consistency. In this approach, the deterministic part of the channel impulse response (so called, deterministic rays) accurately takes into account the positions of the transmitter TX, receiver RX and a few strongest reflectors, always existing in considered scenario. The other part of the channel impulse response represents a number of relatively weak random reflections (so called, random rays). Simulating moving objects, the Q-D channel model can accurately reproduce fading and Doppler effects, observed in real measurements [16]. This is not possible with a purely statistical model. At the same time, compared to the fully deterministic ray-tracing method, the Q-D approach has much less computational complexity.

It should be mentioned, that the Q-D approach, used for recovering the missing parameters for building 3D A-G channel models, also allows us to expand the scope of application of the developed models to the different heights of the transmitter and receiver locations.

It is interesting to note, that the Q-D method was initially introduced for development of 3D channel models for the millimeter wave bands [16] and then justified in a number of experimental measurements [17-18]. In the present paper, we apply the Q-D approach to build 3D A-G channel models for the centimeter wave C-band, because the experimentally measured characteristics of the A-G channel for C-band [12-15] are very similar to the characteristics of the millimeter wave 57-64 GHz channel for the Open Area outdoor hotspot scenario [17]. The physical background for this similarity is approximately the same ratio of the wavelength λ to the typical scale of the environment d in both bands. Indeed, for millimeter wave channel, this ratio is about $\lambda/d=5\text{mm}/100\text{-}300\text{m}$, and for the C-band $\lambda/d=6\text{cm}/1\text{-}3\text{km}$. It results in a 1:10 scale similarity of the 3D channel impulse responses for these channels and justifies the application of the Q-D methodology for the C-band in the considered case.

2. Problem Statement

The Q-D method for recovering the unknown parameters of the channel rays exploits the combination of experimental data and some additional information about the TX and RX positions, antenna heights, properties of ground or water underlying surfaces, roughness coefficients, location of surrounding reflectors, etc. All these information should be specified in the description of the scenario. The 3D A-G channel model development is based on the experimental data provided in the set of papers [12]-[15]. The scheme of the typical experiment from these papers is shown in [Figure 1. A scheme of the experimental measurements of the Air-to-Ground \(A-D\) radio channel characteristics \[12\]-\[15\].](#)

The 10-watt sounding signal transmitter was located at the ground station (GS) with a nearly omnidirectional in azimuth antenna on height of 20 meters above ground level, while the receiver was located on an aircraft flying at an altitude of 800 to 1200 meters at various distances of 2-50 km from the GS. The Q-D approach for channel modeling is built on the representation of the channel impulse response comprised of a few quasi-deterministic strong rays (D-rays) and a number of relatively weak random rays (R-rays), originating from the surrounding objects reflections.

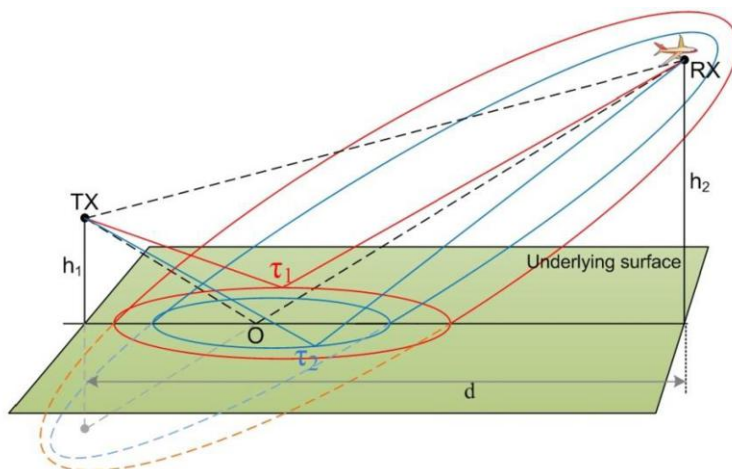


Figure 1. A scheme of the experimental measurements of the Air-to-Ground (A-D) radio channel characteristics [12]-[15].

The set of D-rays for open area scenario includes only two rays: direct LOS (line-of-sight) ray and underlying surface reflected ray (see Figure 1). The characteristics of these both D-rays for the 3D A-D channel model may be calculated from system geometry, properties of ground surface and available experimental data [12]-[15].

The R-rays, in accordance with the Q-D methodology [16]-[17], may be generated in two different ways: statistically from the pre-defined power-delay profile or as deterministic reflections from the random surrounding objects. In this paper we use statistical characteristics of the R-rays defined in [12]-[15] and, for recovering unknown angular and frequency ray parameters, convert them into reflections from the random surrounding objects as it will be shown below.

The A-G channel models, developed in [12]-[15], introduce four types of the multipath channel rays (see Figure 2):

1. Direct LOS D-ray indicated by the wave vectors $\mathbf{r}_{TX,1}$ and $\mathbf{r}_{RX,1}$.
2. NLOS (Non-line-of-sight) D-ray reflected from the underlying surface and indicated by the wave vectors $\mathbf{r}_{TX,2}$ and $\mathbf{r}_{RX,2}$.
3. R-rays that have time delays exceeding the delay of the ground reflected NLOS D-ray. These R-rays indicated by the wave vectors $\mathbf{r}_{TX,i}$ and $\mathbf{r}_{RX,i}$.
4. R-rays that have time delays less than the delay of the ground reflected NLOS D-ray. These R-rays indicated by the wave vectors $\mathbf{r}_{TX,j}$ and $\mathbf{r}_{RX,j}$.

Solving the problem of the 3D A-G channel reconstruction, it is necessary to find (generate) the values of angles of departure (AoD), angles of arrival (AoA) and frequency shifts for each of the above types of the channel rays.

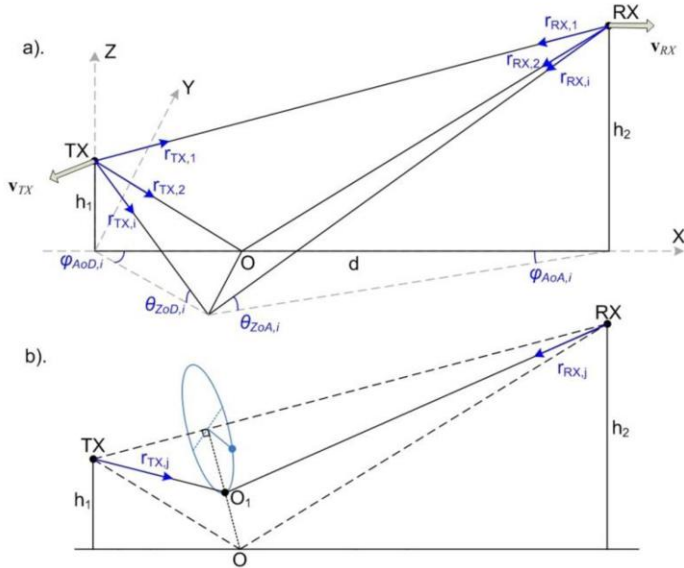


Figure 2. Four types of the A-D multipath channel model rays: LOS D-ray, NLOS D-ray reflected from the underlying surface, R-rays of type 3 and R-ray of type 4 with different delays.

3. Channel Angular Parameters Reconstruction Algorithms

The Q-D approach, used by the authors in this paper, involves calculating the AoD and AoA based on defining the position of channel reflectors in the space between the transmitter and receiver [16]-[17]. This task is solved in two steps: the first one is calculating the possible position of channel reflectors for each of the ray based on the known delays of the rays in the channel impulse response, and, the second one is calculating the AoD and AoA based on the relative locations of the TX, RX and the defined channel reflectors.

We consider the reconstruction process for the departure and arrival angles for each of the four types of rays. For the direct LOS D-ray and the NLOS D-ray reflected from the underlying surface the departure and arrival angles are calculated directly from the coordinates of the TX and RX. Note, that for these D-rays, in the azimuth plane, the angles of departure (AzoD) and the angles of arrival (AzoA) can be set to zero, without loss of generality, by selecting the appropriate coordinate system (see Figure 2a).

Two assumptions have been made for the 3rd type R-rays that delay exceeds the delay of the ground reflected NLOS D-ray. First one, it is assumed that the trajectory of each R-ray is determined by a single reflection from some GS surrounding object, and second, this object is located near the earth's surface. Under these assumptions, the coordinates of the reflector generating this type of a R-ray lies on the underlying surface on some ellipse, which is the intersection of the corresponding ellipsoid of equal delays. For example, Figure 1 shows a scheme for calculating the position of channel reflectors for the 3rd type R-rays with delays τ_1 and τ_2 .

Following these assumptions, at the first step for each 3rd type R-ray the corresponding its known delay ellipse is constructed on the ground surface. Because,

the delay of the reflected ray does not give any reason to prefer a particular point on the ellipse, as the final position of the reflector, the choice of this point is made randomly. The point on the ellipse that is visible from the reflection point O of the NLOS D-ray at the random angle (uniformly distributed from 0 to 2π) is selected as the location of the reflector generating this 3rd type R-ray.

For the 4th type R-rays those delays less the delay of the ground reflected NLOS D-ray, the reflecting object cannot lie on the underlying surface, since this D-ray has the minimum delay of all the rays reflected from this surface. Therefore, for the 4th type R-rays, the reflection point must lie above the underlying surface (earth or water). For example, it can be a reflection from a high building or from another high object that rises above the average ground level. The construction of the 4th type R-ray path, according to this assumption, is illustrated in Figure 2b. At the first step, at the height of the triangle O-RX-TX, drawn from point O, the point O₁ corresponding to the R-ray delay is selected. Then, at the second step, the final point of the 4th type R-ray reflection is determined by rotating the point O₁ in space around the LOS direction TX-RX at the random angle in the range from $-\pi / 2$ to $\pi / 2$. Since the rotation occurs around the line of sight, the final R-ray trajectory does not change the delay of the ray, but it adds additional variability to the values of the AoD and AoA of the 4rd type R-ray.

After calculating the coordinates of the reflection points for each of the R-rays from the channel impulse response (CIR), azimuth and elevation AoD and AoA in the global coordinate system are calculated using simple geometric constructions (see Figure 2).

4. Doppler Shift Calculation

The angular characteristics of the 3D A-G channel model, found by using the algorithms described in the previous section, make it possible to calculate the Doppler shifts for each ray in the CIR. Since the proposed reconstruction procedure makes it possible to set the AoD and AoA for each ray, the Doppler frequency offset can be easily calculated for each channel ray by setting the direction and speed of the transmitter and receiver relatively to the global coordinate system (see Figure 2). Therefore, in the considered approach, the Doppler spread of the transmitted signal spectrum is modeled as a set of Doppler shifts to the signal carrier frequency for each of the channel ray [17]-[18]. In the case of fixed channel reflectors and moving TX and / or RX, for each channel ray the Doppler frequency shift can be represented as the sum of frequency shifts before and after reflection (see Figure 2). So for the *i*-th ray:

$$f_{d,i} = \frac{f_0}{c} (\mathbf{v}_{TX} \cdot \mathbf{r}_{TX,i} + \mathbf{v}_{RX} \cdot \mathbf{r}_{RX,i}),$$

where \mathbf{v}_{TX} and \mathbf{v}_{RX} are the velocity vectors of the transmitter and receiver, respectively, and $\mathbf{r}_{TX,i}$ and $\mathbf{r}_{RX,i}$ are the unit wave vectors of the ray path for the TX and RX, respectively, f_0 is the carrier frequency, c is the speed of light.

5. Simulation Results

In the paper by the using the experimental data and characteristics of the TDL channel models given in [12]-[15], as well as on the basis of the developed Q-D methodology, a software model of the 3D A-G channel was implemented. At the same time, the approach to the channel modeling was slightly changed compared to the original TDL models from [12]-[15] in order to bring the methodology closer to the standards adopted for the development of wireless communication technologies by IEEE and 3GPP.

The system level simulation (SLS) platform for the 3D A-G channel model environment was created in the Matlab and C++. The software implementation includes procedures for the CIR generation, the AoD and AoA calculation for individual rays, and the ray Doppler shifts determination. The CIR generation involves calculating the power of the LOS D-ray, the NLOS mirror-reflected D-ray, and several R-rays with time delays both greater and less than the NLOS D-ray delay. For LOS and NLOS D- rays, the attenuation of their power in space is taken into account, as well as additional path losses due to the weather factors and vegetation. The R-rays generation was performed in accordance with statistical distributions of their delays and amplitudes given in [12]-[15]. After obtaining the CIRs and using the reconstruction algorithms described in the previous sections of this paper, the angular characteristics and Doppler shifts of all the rays are calculated.

To check the adequacy of the developed 3D A-G channel models, ensembles of the CIRs were generated depending on the horizontal distance between the transmitting GS and the receiver located on the aircraft. Examples of these ensembles for flight scenarios “over-sea” (the GS height 20 m and aircraft altitude 800 m) and “over near-urban environment” (the GS height 20 m and aircraft altitude 562 m) are shown in Figure 3a and Figure 3b, respectively. More detail comparison of these SLS results with experimental results, represented in [12] and [14], shows that the developed 3D A-G channel model allows us to adequately reproduce the statistics of the real CIRs.

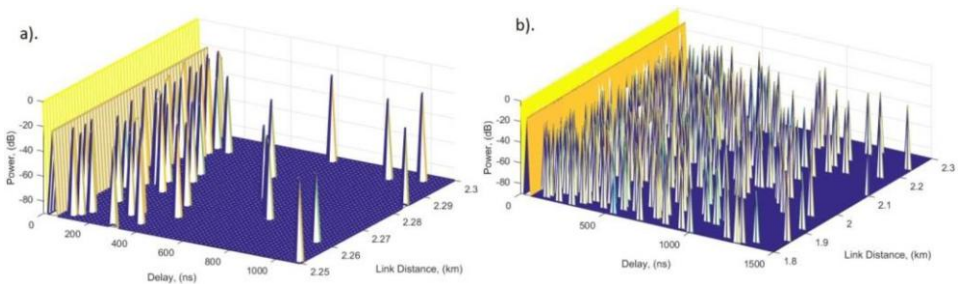


Figure 3. Ensembles of the CIRs depending on the distance between the transmitter and receiver: A-flight scenario “over-sea”; B-flight scenario “over near-urban environment”.

The histograms of the AoA and AoD distributions in the azimuth and elevation planes for the scenarios “over-sea” and “over near-urban environment”, obtained by using the reconstruction algorithms, described in the previous sections of the paper, shown in Figure 4 and Figure 5, respectively.

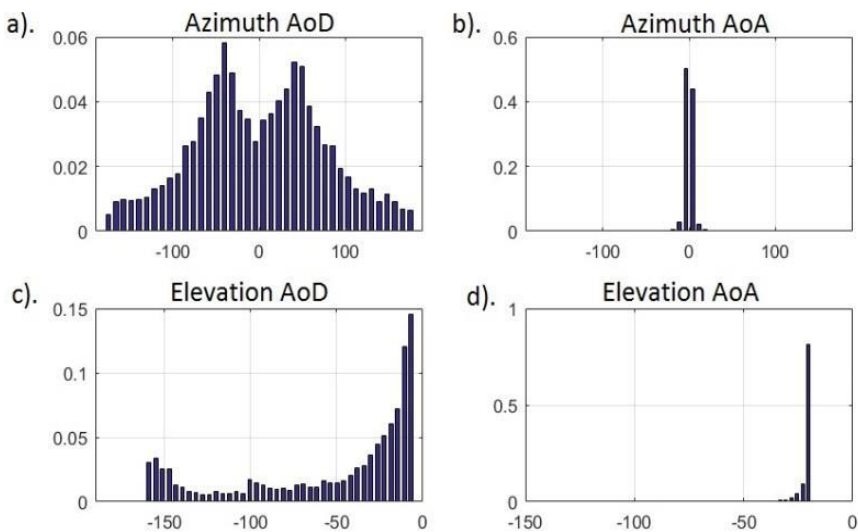


Figure 4. Histograms of the 3D A-G channel model ray angles distribution in the “over-sea” scenario: azimuth AoD and AoA on top, elevation AoD and AoA on bottom. The TX height is 20 m, the RX height is 800 m, the horizontal distance between the TX and RX is 2260 m.

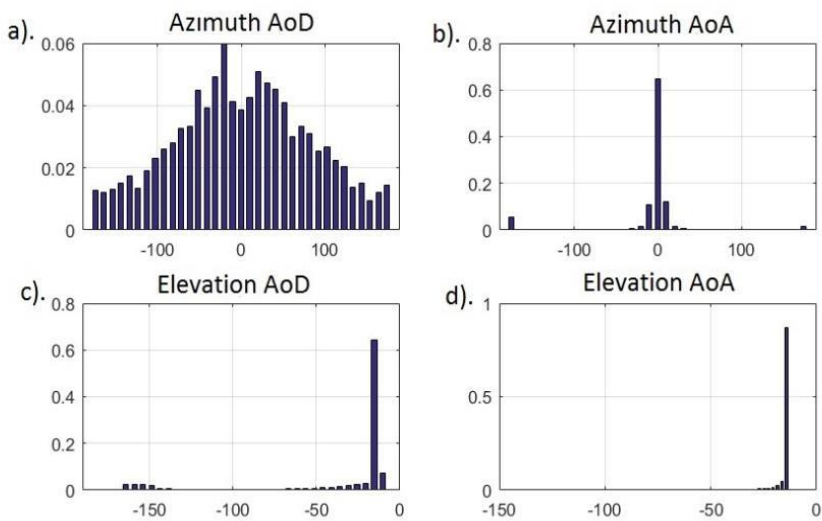


Figure 5. Histograms of the 3D A-G channel model ray angles distribution in the “over near-urban environment” scenario: azimuth AoD and AoA on top, elevation AoD and AoA on bottom. The TX height is 20 m, the RX height is 562 m, the horizontal distance between the TX and RX is 2260 m.

It can be seen that for the TX, located at the GS near the ground, the AoD distributions in azimuth and elevation planes are much wider for both scenarios than for the RX located at the high altitude aircraft. This is due to the fact that the bulk of

the random R-rays in the A-G channel are obtained as a result of reflection from the GS surrounding objects on the ground surface. At the same time, the angular size of the reflection area is significantly smaller when observed from the aircraft, than when observed from the GS.

It is interesting to note that for both scenarios, the distributions of the azimuth and elevation AoAs on the RX (aircraft) have approximately the same narrow width. The angular distributions of the departure rays for different scenarios turn out to be significantly different, which is probably determined by the various configurations of objects surrounding the transmitting GS.

6. Conclusion

In this paper, we have introduced a Q-D method for reconstructing the angular and frequency characteristics of the 3D A-D channel models based on the empirical TDL channel models [12]-[15]. The SLS platform for modeling all considered scenarios and environments was created in the Matlab. Simulations have shown that the developed 3D A-G channel model adequately reproduces the statistics of the real CIRs and provides physically reasonable AoD and AoA distributions and Doppler spreads. Thus, the authors believe that this model is useful for characteristics evaluation of modern UAV and aircraft communication systems with directional and multi-element antennas. As the further steps in this work, the authors see in introduction of a more detail cluster structure of channel rays (including flash rays F-rays), which observed in experiments, and upgrade the SLS platform for modeling both the short and long term dynamical changes of the A-G channel.

References

- [1] Radio Technical Commission for Aeronautics (RTCA), SC-228, "Minimum Operational Performance Standards for Unmanned Aircraft Systems", 2013, <http://www.rtca.org/content/sc-228>.
- [2] National Aeronautics and Space Administration, 2014, available online at <http://www.aeronautics.nasa.gov/isrp/uas/index.html>.
- [3] E. Haas, Aeronautical channel modeling, *IEEE Transactions on vehicular technology*, **2** (2002), 254–264.
- [4] J. Sykora, Tapped delay line model of linear randomly time-variant WSSUS channel, *Electronics Letters*, **36** (2000), 1656-1657.
- [5] K. Mammassis, E. Pfann, R. W. Stewart and G. Freeland, Three-dimensional channel modelling using spherical statistics for smart antennas, *Electronics Letters*, **44** (2008), 769-774.
- [6] A. Al-Hourani, S. Kandeepan and A. Jamalipour, Modeling Air-to-Ground Path Loss for Low Altitude Platforms in Urban Environments, *Proc. IEEE Global Commun. Conf.*, (2014), 2898-2904.
- [7] M. Wentz and M. Stojanovic, A MIMO Radio Channel Model for Low-Altitude Air-to-Ground Communication Systems, *Proc. IEEE Veh. Technol. Conf. (VTC-Fall'15)*, (2015), 1-6.
- [8] S. Gulfam, S. Nawaz, M. Patwary and M. Maguid, On the Spatial Characterization of 3-D Air-to-Ground Radio Communication Channels, *Proc. IEEE Int. Conf. Commun. London*, (2015), 2924-2930.
- [9] M. Simunek, F.P. Fontan and P. Pechac, The UAV Low Elevation Propagation Channel in Urban Areas: Statistical Analysis and Time-Series Generator, *IEEE Trans. Antennas Propag.*, **61** (2013), 3850-3858.
- [10] E. Yanmaz, S. Hayat, J. Scherer and C. Bettstetter, Experimental Performance Analysis of Two-Hop Aerial 802.11 Networks, *Proc. IEEE Wireless Commun. Netw. Conf.*, (2014), 3118-312.
- [11] W. Khawaja, I. Guvenc and D. Matolak, UWB Channel Sounding and Modeling for UAV Air-to-Ground Propagation Channels, *Proc. IEEE Global Commun. Conf. (GLOBECOM'16)*, (2016), 1-7.
- [12] D.W. Matolak and R. Sun, Air–Ground Channel Characterization for Unmanned Aircraft Systems–Part I: Methods Measurements and Models for Over-Water Settings, *IEEE Trans. Veh. Technol.*, **66** (2017), 26-44.

- [13] R. Sun and D.W. Matolak, Air-Ground Channel Characterization for Unmanned Aircraft Systems–Part II: Hilly and Mountainous Settings, *IEEE Trans. Veh. Technol.*, **66** (2017), 1913-1925.
- [14] D.W. Matolak and R. Sun, Air-Ground Channel Characterization for Unmanned Aircraft Systems–Part III: The Suburban and Near-Urban Environments, *IEEE Trans. Veh. Technol.*, **66** (2017), 6607-6618.
- [15] R.Sun and D.W. Matolak, Air-Ground Channel Characterization for Unmanned Aircraft Systems Part–IV: Airframe Shadowing, *IEEE Trans. Veh. Technol.*, **66** (2017), 7643-7652.
- [16] A. Maltsev et al., Quasi-deterministic approach to mmWave channel modeling in a non-stationary environment, *2014 IEEE Globecom Workshops (GC Wkshps)*, (2014), 966-971.
- [17] A. Maltsev et al. Channel Models for IEEE 802.11ay, Doc. IEEE 802.11-2015/1150r9 (2017).
- [18] J. Wang, C. Gentile, P. B. Papazian, J. Choi and J. Senic, Quasi-Deterministic Model for Doppler Spread in Millimeter-Wave Communication Systems, *IEEE Antennas and Wireless Propagation Letters*, **16** (2017).

Cost-Efficient RAA Technology for Development of the High-Gain Steerable Antennas for mmWave Communications

Olesya BOLKHOVSKAYA^{a,1} Alexander MALTSEV^{a,b}, Valentin SELEZNEV^a and Ilya BOLOTIN^c

^aUniversity of Nizhny Novgorod, Russia

^bLG Corp, Russia.

^cIntel Corp, Russia

Abstract. For the future millimeter wave wireless systems the high gain steerable antennas are the key technology to overcome large path loss and support users' mobility and reconfigurable backhauling. This paper introduces a specially designed reflect-array antenna (RAA), integrating single phased antenna array (PAA) module (as a feed source) and flat reflecting surface. The RAA provides high antenna gain by focusing the beam in vertical plane, but, at the same time, supporting beamsteering capability in the horizontal plane. The RAA technology allows creation of the very larger aperture antennas with simplified light-weight design. Two RAA prototypes with different feed source mounting demonstrated reliable multi-gigabit IEEE 802.11ad link 2.3-4.62 Gbps at distances 100-150m. These properties of the RAAs make them suitable for both millimeter-wave mobile access and reconfigurable backhauling applications.

Keywords. IEEE 802.11ay, mmWave, reflect-array antenna, RAA, 60GHz, beamsteering, 5G

1. Introduction

For satisfying the throughput requirements of the 5G system the future communication systems should exploit all possibilities, including the densification of base stations with usage of heterogeneous networks (HetNets) architecture, massive MIMO, MU-MIMO and CoMP (Coordinated Multi-Point) techniques [1]-[3] and, finally, the bandwidth increase. Millimeter wave communications allow realizing all these possibilities exploiting signals with very large bandwidth and highly directional antennas with narrow beams [3]-[5].

However, for mobile millimeter wave access and reconfigurable backhauling the narrow beam high-gain antennas must be also steerable. During the joint EU-Japan FP7 project MIWEBA [6]-[9], new antenna technologies, the modular antenna array (MAA) and the lens-array antenna (LAA), were proposed to meet these requirements. The MAA consists of multiple independent phased antenna array (PAA) modules (or subarrays), each with its own phase-shifting circuitry and RF part. The modules are

¹ Corresponding Author: Olesya Bolkhovskaya, University of Nizhny Novgorod, Gagarina ave,23, 603109 Nizhny Novgorod, Russia; E-mail:obol@rf.unn.ru

connected to a common baseband, and each module can act as a single independent phased array. This design has only light limitations on the array adaptability, and degradation in comparison with ideal fully adaptive array (FAA) is not significant [10].

The different path to develop high gain steerable antennas is to combine a single, relatively small, phased antenna array (PAA) with a passive beam focusing system. The LAA is an example of such system, wherein the special dielectric lenses [11]-[12] integrated with a small PAA are used by focusing beam in vertical plane and support the full steering possibility in horizontal plane. Corresponding combined antenna system will not be as flexible and versatile as FAA or MAA, but it can provide a cost-efficient technique to design high-gain antennas with limited possibilities for beamsteering.

In this work we consider one more approach for developing the high-gain steerable antennas for mmWave communications. This is the reflect-array antenna (RAA) technology [13]-[16] which is using the reflecting surface with special properties for the reflected beam focusing in one plane (in elevation) and allows beam sweeping in another one (in azimuth).

The RAA and LAA have high gain and simultaneously provide beamsteering in a large azimuth angular sector. Of course, both these techniques may be recommended for usage in millimeter wave 5G communication systems, but the RAA technology has an advantage over the LAA for large aperture antennas, due to very low cost and light weight of the flat passive beam focusing system which may be mounted on any wall in outdoor and indoor environments.

2. RAA Prototype

This section describes a design of the reflect-array antenna (RAA) prototype integrating single PAA module and the reflecting flat surface with special properties.

2.1. Phased Antenna Array Module

PAA uses a 16-element mmWave rectangular slit loop antenna array in the 60 GHz band, developed by Intel and outlined in [17]. The considered construction geometry and its system of axes are shown in figure 1. The PAA has 2 x 8 elements spaced by 0.5 wavelength (2.5 mm) distance in vertical and horizontal dimensions. Each element has a polarization with an E-field vector parallel to the short side of the array and a radiation pattern close to omni. Full size is $W=2$ mm width, $H=9$ mm height and $L=25$ mm length.

In the developed RAA prototypes the small PAA with 16 antenna elements was used as a primary source. This rectangular (2x8 geometry) PAA was designed by Intel for exploiting in Wi-Fi (IEEE802.11ad) systems in 60 GHz band [17]. The PAA geometry and related system of coordinates are shown in figure 1. The distance between antenna elements was 0.5λ (2.5 mm). Each element had a vertical polarization and almost omni radiation pattern.

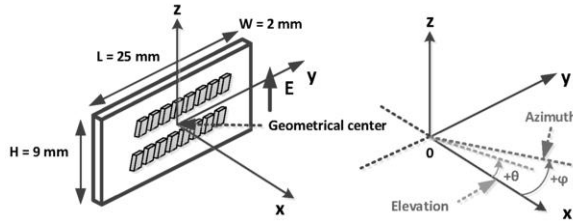


Figure 1. Phased antenna array geometry and related system of coordinates.

The main lobe of the PAA had 14.0° Half Power Beamwidth (HPBW) in azimuth, and 41.0° HPBW in elevation. The total transmit power of the PAA was $P_{PAA} = 10\text{dBm}$ and the antenna gain was near to $G_{PAA} = 15.0\text{ dBi}$.

2.2. Reflect Array Design

The RAA uses the similar principle of work as the LAA – focusing the small PAA module beam in the elevation plane, while keeping it almost unchanged (and thus, providing beamsteering capabilities) in the azimuthal plane (see Figure 2). Reflector array of the RAA consists of a number of small phase-shifting reflectors that rotate the phases of the incident wavefront. Usually elementary reflector consists of micro strip patch located at small distances from the metal ground plane (see Figure 3).

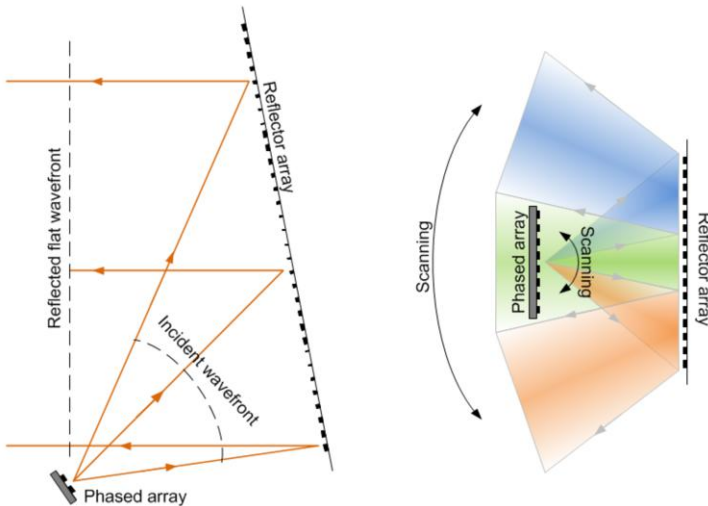


Figure 2. RAA principle illustration.

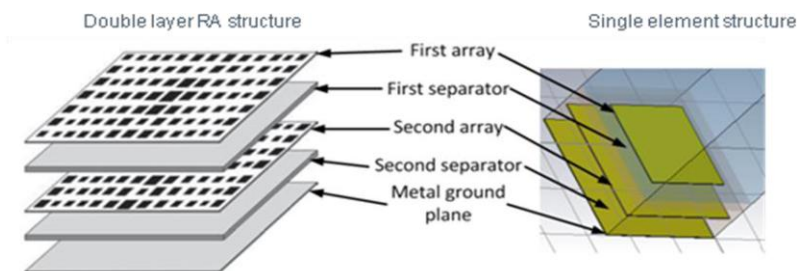


Figure 3. Reflector array structure.

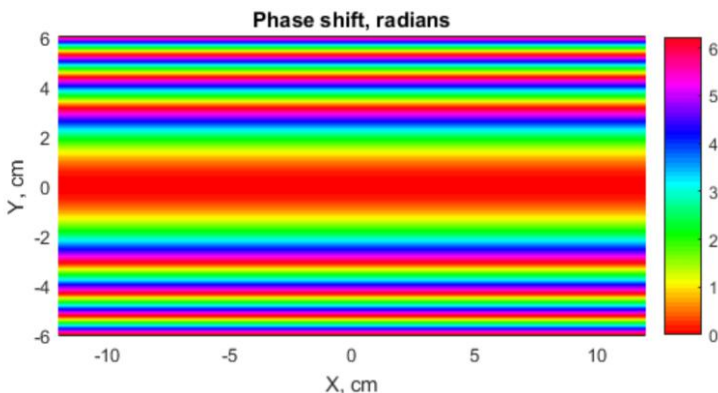


Figure 4. Phase-rotation values distribution for the developed reflector array.

The wave front phase rotation at the given value is setting by patches size and multi-layer structure parameters of the reflector. In the developed reflector array the double layer square patches were used as elementary reflectors (see Figure 3). Usage of the two-layer patch structure allows to obtain the smoother “phase-patch size” dependence, expand the range of possible phase rotations to values greater than 360° and increases the RAA frequency range up to 10% of the carrier frequency. By simple adjusting the patches sizes the reflector array can be configured to behave like any type of “curve mirror”. For the steerable in azimuth plane RAA the cylindrical-parabolic mirror representation was realized to simplify design of the reflecting surface (see Figure 4).

2.3. Reflect Array Antenna integration

Like conventional reflector antennas, axial and offset configurations are available for placement of the feeding PAA module in front of reflector array. Correspondingly, two RAA prototypes were implemented and experimentally investigated – RAA 1 with axial mount and RAA 2 with offset mount of the PAA module, see Figure 5.

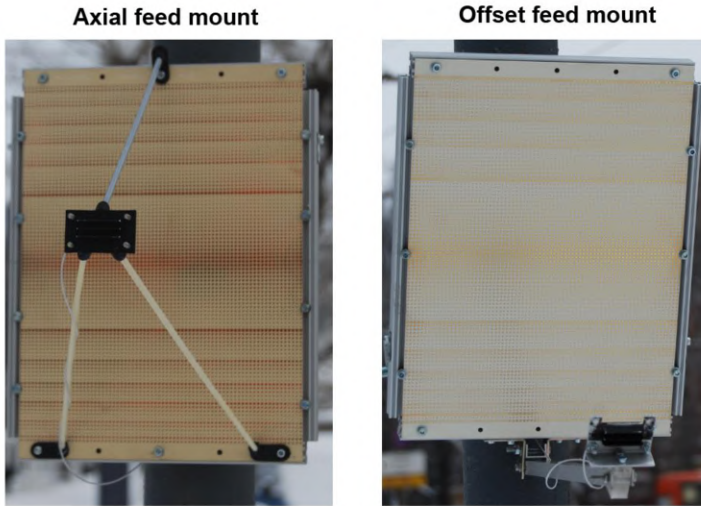


Figure 5. Two RAA prototypes external view

The main characteristics of these prototypes are provided in Table 1.

Table 1. Two RAA prototypes parameters and characteristics.

Parameter	Axial feed position	Offset feed position
Size	187 x 237 x 1 mm	84 x 154 x 1 mm
Focus	200 mm	200 mm
Antenna gain	15+9.5=24.5 dBi	15+11=26 dBi
Azimuth HPBW	18°	14°
Elevation HPBW	2°	2°
TX power	10dBm	10dBm
Operational range	100 – 150 m	100 – 150 m
Peak data rate	4.62–2.3 Gbps	4.62–2.3 Gbps

3. RAA Field Tests

3.1. Experimental Measurement Setup

Experimental measurements of the developed RAA prototypes antenna patterns and gains were carried out by using a special experimental setup mounted at the rooftops of two neighboring buildings at the University of Nizhny Novgorod campus at 30 m distance between TX and RX sides (see Figure 6).

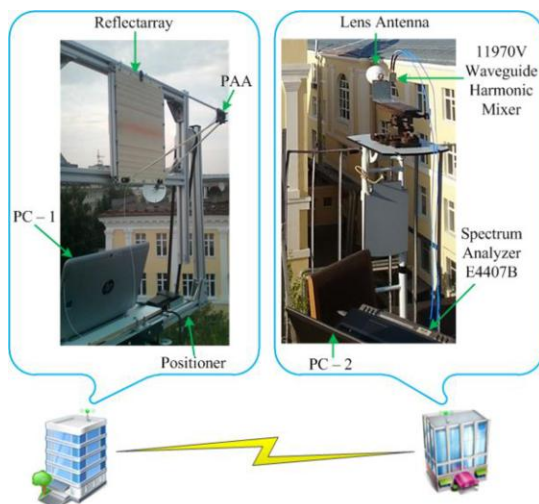


Figure 6. Experimental setup for RAA prototypes characteristics measurement.

At the transmitter side the RAA was mounted at the positioner that performed antenna rotation with angular resolution $\sim 0.4^\circ$ in horizontal plane and 0.1° in vertical plane. The RAA transmitted IEEE 802.11ad waveform signal with central frequency 60 GHz and bandwidth about 2 GHz.

At RX side the special high directional laboratory lens antenna with axial symmetry and aperture size (diameter) equal to 100 mm, HPBW equal to 3.0° , and antenna gain 34.5 dBi was used, to increase SNR and avoid spurious reflections. The RX lens antenna had static position and was perfectly aligned with the TX RAA.

The signal at the RX antenna output was fed to the frequency down converter 11970V of Agilent Technologies input and then its power and spectrum were measured by universal spectrum analyzer E4407B of Agilent Technologies. The RX setup performed signal power measurements with start and stop frequencies 59.4 and 61.56 GHz accordingly with the signal bandwidth 2.16 GHz. The sweep time was set equal to 26.0 ms, both resolution and video bandwidths were set equal to 3.0 MHz. The positioner and all measurement equipment were controlled remotely via a Bluetooth wireless channel by the software program on the PC.

3.2. Experimental Results

The RAA 1 and RAA 2 antenna patterns measurement results in elevation and in azimuth planes are shown in Figure 7. It can be seen that both RAA prototypes are well focusing the power, radiated by the PAA, in the elevation plane, forming the main lobe with HPBW about of 2 degrees. The antenna gain of RAA 1 was about 1.5-2.0 dB less than for RAA 2 that may be obviously explained by the shadowing effect due to the axial position of the PAA. The main lobe beam widths in the azimuth plane a little different for RAA 1 and RAA 2. For RAA 1 with axial feed, it is about 14° , and for RAA 2 with offset feed, it is about 16° . It can be seen that in azimuth plane the radiation pattern have almost the same HPBW as the single PAA, at the same time, in

the vertical plane the beam is focused, producing large antenna gain and narrow main lobe of the antenna pattern.

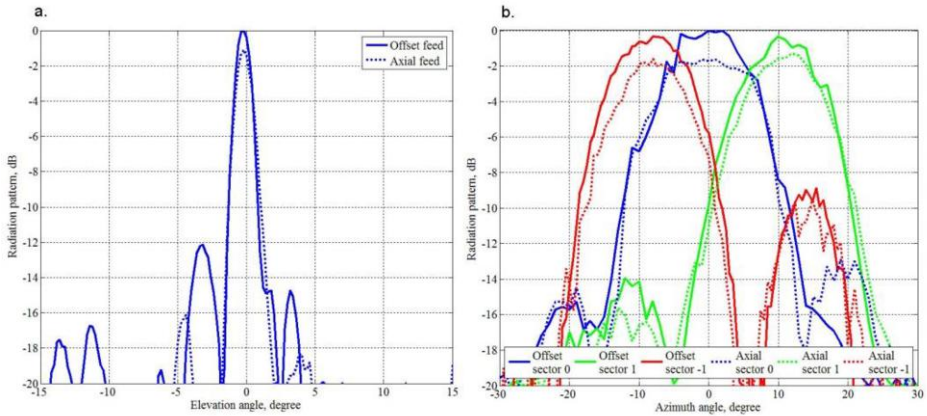


Figure 7. Antenna patterns measurement results for RAA 1 with axial feed and RAA 2 with offset feed: a – in elevation plane for azimuth sector 0; b – in azimuth plane for different sectors.

To evaluate the RAA prototypes beamsteering capabilities, the beam of the PAA feeding module was directed to three azimuthal sectors at 0° , $+10^\circ$, -10° (see Figure 7). It can be seen, that for these sectors the antenna gains are practically constant, about 24.5 dB for the RAA 1 and 26 dB for the RAA 2. It produces the effective azimuth half power steering range $[-15^\circ:15^\circ]$. But for larger beam deflections the RAAs antenna gains quickly degrades. It is clear, that the cause of this effect is not full power reflection of the steered PAA's beam, due to the limited width of the reflecting surface in horizontal plane. In next generation of the RAA prototypes, this drawback will be fixed. However, even in this first setup, the RAAs allow achieving total 25-26 dBi gain with a limited beamsteering sector, but which is enough, for example, for electronic beam alignment of the backhaul links. If used with 802.11ad standard devices, such TX-RX antenna systems allow transmission at gigabit data rates 4.62–2.3 Gbps up to 150-200m distance.

4. Conclusion

In this paper we have considered a new approach to the design of the high-gain beam steerable antennas for millimeter wave communication systems. The proposed reflect-array antenna (RAA) technology is using the reflecting surface with special properties for the beam focusing in one plane and allow beam sweeping in another one. Theoretical investigations and direct prototyping have proven the efficiency of the proposed RAA technology. Two RAA prototypes with different PAA module mounting were developed and demonstrated the characteristics which provide reliable multi-gigabit link at distances 100-150m. The RAA technology with flat and light-weight reflecting surfaces allows in cost-efficient way to create mmWave antennas with larger aperture to further increase the antenna gain, beam steering capabilities and achievable transmission range.

References

- [1] F. Boccardi, R.W. Heath, A. Lozano, T.L. Marzetta and P. Popovski, Five disruptive technology directions for 5G, *IEEE Communications Magazine*, **5** (2014), 74-80.
- [2] A. Davydov, G. Morozov, I. Bolotin and A. Papathanassiou, Evaluation of Joint Transmission CoMP in C-RAN based LTE-A HetNets with large coordination areas, *2013 IEEE Globecom Workshops (GC Wkshps)*, (2013), 801-806.
- [3] R. Heath, Jr, Role of MIMO Beyond LTE: Massive? Coordinated? mmWave?, *Workshop on Beyond 3GPP LTE-A ICC*, (2013).
- [4] E. Torkildson, U. Madhow and M. Rodwell, Indoor Millimeter Wave MIMO: Feasibility and Performance, *IEEE Transactions on Wireless Communications*, **10** (2011), 4150-4160.
- [5] T.S. Rappaport et al., Millimeter Wave Mobile Communications for 5G Cellular: It Will Work!, *IEEE Access*, **1** (2013), 335-349.
- [6] R.J. Weiler et al., Enabling 5G backhaul and access with millimeter-waves, *2014 European Conference on Networks and Communications (EuCNC)*, 2014, 1-5.
- [7] *MiWEBA Project homepage* <http://www.miweba.eu>, (FP7-ICT-2013-EU-Japan, project number: 608637, 2013).
- [8] ICT FP7 MiWEBA project #608637, Deliverable D5.3 – Highly-directional steerable mm-wave antennas prototype, <http://miweba.eu>.
- [9] A. Maltsev, A. Sadri, A. Puduev and I. Bolotin, Highly Directional Steerable Antennas: High-Gain Antennas Supporting User Mobility or Beam Switching for Reconfigurable Backhauling, *IEEE Vehicular Technology Magazine*, **11** (2016), 32-39.
- [10] A. Maltsev et al., Partially adaptive arrays application for MU-MIMO mode in a MmWave small cells, *2015 IEEE 26th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, 2015, 315-319.
- [11] A. Maltsev, A. Lomayev, A. Puduev, I. Bolotin, O. Bolkhovskaya and V. Seleznev, Millimeter-wave Toroidal Lens-Array Antennas Experimental Measurements, *2018 IEEE International Symposium on Antennas and Propagation & USNC/URSI National Radio Science Meeting*, (2018), 607-608.
- [12] O. Bolkhovskaya, V. Seleznev, A. Rulkov and A. Maltsev, Steerable Bifocal Lens-Array Antenna at 57-64 GHz, 2019 Conference on Microwave Techniques (COMITE), (2019), 1-6.
- [13] J.A. Encinar, Design of two-layer printed reflectarrays using patches of variable size, *IEEE Transactions on Antennas and Propagation*, **49** (2001).
- [14] T. Visentin, W. Keusgen, R. Weiler, Dual-polarized square-shaped offset-fed reflectarray antenna with high gain and high bandwidth in the 60 GHz domain, *Proceeding of the 2015 9th European Conference on Antennas and Propagation (EuCAP)* (2015).
- [15] Wenting Li ; Steven Gao ; Long Zhang ; Qi Luo ; Yuanming Cai An Ultra-Wide-Band Tightly Coupled Dipole Reflectarray Antenna, *IEEE Transactions on Antennas and Propagation*, **66** (2018), 533-540.
- [16] Fan Yang, Ruyuan Deng, Shenheng Xu, Maokun Li Design and Experiment of a Near-Zero-Thickness High-Gain Transmit-Reflect-Array Antenna Using Anisotropic Metasurface, *IEEE Transactions on Antennas and Propagation*, **66** (2018).
- [17] H. K. Pan, B. D. Horine, M. Ruberto, S. Ravid, Mm-Wave Phased Array Antenna and System Integration on Semi-Flex Packaging, *IEEE AP-S Conference on Antenna and Propagation for Wireless Communications*, (2011).

Detection of Specific Components in a PCA Mixture

Jin XIE ^{a,1}, Zian ZHENG^b and Jian GAO^c

^aSchool of Mathematics and Statistics, Lanzhou University, China

^bSchool of Information Science and Engineering, Lanzhou University, China

^cCollege of Chemistry and Chemical Engineering, Lanzhou University, China

Abstract. Taking a given mixture as an example, 25,000 samples were selected for the detection of 7 indicators. Firstly, the correlation between each indicator and the test result is analyzed, The T test is used to identify the main indicators that can be used to determine the existence of a specific component. Secondly, three comprehensive indexes are obtained by combining PCA. Determine whether there are specific components in the unknown mixture.

Keywords. T test, PCA, mixture, detection

1. Introduction

Judging the composition structure of mixtures is an important step in the field of physical chemistry. When facing a huge amount of information, the correlation can be analyzed through known samples and a mathematical model can be built. Principal component analysis (PCA) is a kind of multivariate analysis method, which transforms the original correlated variables into some new unrelated variables by means of variable transformation. The method is scientific and effective, and can be widely applied in many fields. For the detection of specific components in the mixture, generally there are chemical methods, professional formula analysis based on microspectrum technology, physical purification method, DAD detection and model analysis algorithm enhance the computer processing capacity and realize the qualitative identification of the mixture detection. Using mathematical modeling theory and method to integrate the detection of specific components of the mixture, applying mathematical knowledge to abstract the calculation problem in the experiment into a mathematical problem and summarized into mathematical models. Through calculation to determine whether the unknown mixture contains specific components, and get the results. Judging the detection of the composition of mixtures is an important step worth discussing in the field of physical chemistry. In the face of a huge amount of information, the correlation of known samples can be analyzed and a mathematical model can be built. The model can be used to judge whether the unknown samples contain specific components, with scientific and effective methods.

¹Corresponding Author: Xie Jin; E-mail: xijiner0211@163.com

In this paper, 25,000 samples given in the Mathematical Modeling Contest of Lanzhou University in 2020 are taken as an example ,and 7 indicators are known to be carried out on the samples of this mixture (denotable as V1,V2...,V7). The "training data" contained 20,000 samples of the mixture, with the mixture known to contain a specific ingredient, and the "test data" contained 5,000 samples of the mixture, with the mixture unknown to contain a specific ingredient. Through correlation analysis and PCA model, the main indexes for determining the existence of specific components are given.

2. Basic Theory

(1) T Test:

H0 (null hypothesis): attribute A and B are independent of each other, $\bar{x}_1 = \bar{x}_2$;

H1 (alternative hypothesis): attribute A is related to B, $\bar{x}_1 \neq \bar{x}_2$. To test whether the hypothesis is true, the sample data is processed as follows [1].

Independence test:

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{N_1} + \frac{s_2^2}{N_2}}}$$

If we assume that the variance of these two samples is the same

$$t = \frac{\bar{x}_1 - \bar{x}_2}{s_p \sqrt{\frac{1}{N_1} + \frac{1}{N_2}}} \quad s_p = \frac{(N_1-1)S_1^2 + (N_2-1)S_2^2}{N_1 + N_2 - 2}$$

Inspection level: $\alpha = 0.05$

(2) PCA

Principal component analysis (PCA) is a multivariate statistical analysis method that transforms multiple indicators into several comprehensive indicators by dimension reduction. The main objective is to explain most of the information in the original data with fewer variables.[2]

First, standardize the indicator data: $x'_{ij} = \frac{x_{ij} - \bar{x}_j}{s_j}$, $i = 1, 2, \dots, m$, \bar{x}_j is mean, s_j is standard deviation. Secondly, the correlation coefficient matrix and covariance matrix are calculated: $r_{ij} = \frac{\sum_{k=1}^n (x_{ki} - \bar{x}_i)(x_{kj} - \bar{x}_j)}{\sqrt{\sum_{k=1}^n (x_{ki} - \bar{x}_i)^2 \sum_{k=1}^n (x_{kj} - \bar{x}_j)^2}}$, $c_{ij} = \frac{\sum_{k=1}^n (x_{ki} - \bar{x}_i)(x_{kj} - \bar{x}_j)}{m-1}$

After the covariance matrix is obtained, KMO (Kaiser-Meyer-Olkin) and Bartlett ball type test are conducted to determine whether it is suitable for principal component analysis. Principal component analysis is only suitable when the KMO test value is > 0.5 and the Bartlett test value is < 0.05.

Thirdly, calculate the eigenvalue and variance contribution rate:

$$G_i = \lambda_i / \sum_{i=1}^n \lambda_i \quad , \quad TG_i = \sum_{i=1}^s \lambda_i / \sum_{i=1}^n \lambda_i$$

Finally, the principal component was extracted and the comprehensive evaluation was carried out, and the comprehensive index F was calculated and evaluated:

$$F = \sum_{i=1}^s \alpha_i F_i \quad \alpha_i = \lambda_i / \sum_{i=1}^n \lambda_i$$

3. The Process Of Problem Solving

3.1 Missing Value Detection

First, descriptive analysis of the data was carried out. Visualizing the data through Python lists the number, maximum, minimum, mean, standard deviation, and 4 quantile of each variable, as shown in Table 1. It can be seen from the table that there is no missing value in 20,000 data, so there is no need to fill in the missing value.

Table 1 Data indexing

```
In [5]: df.describe() # Each eigenvalue is 20,000
```

```
Out [5]:
```

	V1	V2	V3	V4	V5	V6	V7	label
count	20000.000000	20000.000000	20000.000000	20000.000000	20000.000000	20000.000000	20000.000000	20000.000000
mean	0.379799	0.581811	0.317147	0.881700	0.590638	0.804141	0.449740	0.694750
std	0.313877	0.331197	0.291529	0.131421	0.210565	0.126345	0.123931	0.460525
min	-0.752723	-0.062687	-1.261943	0.320840	0.134443	0.183343	0.118308	0.000000
25%	0.016551	0.220954	0.013827	0.800967	0.405840	0.721750	0.355693	0.000000
50%	0.468211	0.735289	0.347197	0.943091	0.564188	0.823860	0.434033	1.000000
75%	0.639390	0.868181	0.568549	0.987442	0.783371	0.908167	0.533794	1.000000
max	0.965002	0.985052	0.920936	0.999775	0.994328	0.999482	0.847895	1.000000

3.2 Outlier Handling

Outliers are detected by box diagram, in which outliers are usually defined as values less than $Q_L - 1.5IQR$ or greater than $Q_U + 1.5IQR$, as shown in Figure 1. It can be seen from the figure that there are some outliers in V3, V4, V6 and V7. A total of 417 outliers were calculated. As the sample size was large enough, the outliers were deleted directly. After processing, see Figure 2.

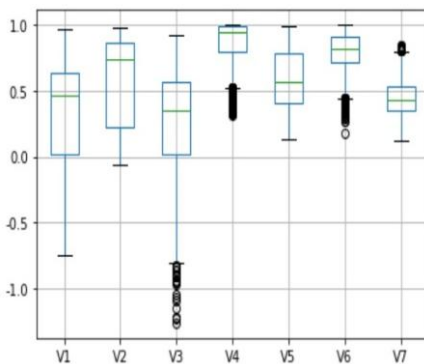


Figure 1 Data outliers

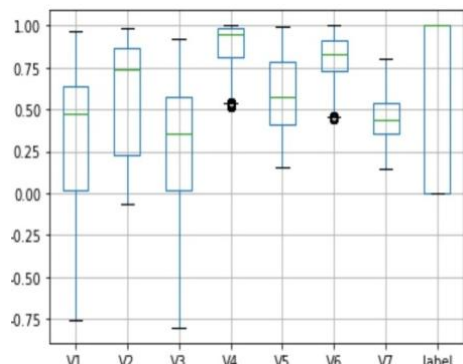


Figure 2 Delete the outliers

3.3 Category Equalization Check

In the classification task, the imbalance of sample categories will have a great impact on the training of the model, so the proportion of positive and negative samples should be checked during the training of the model.

```
In [4]: df['label'].value_counts()
Out[4]: 1    13895
        0     6105
        Name: label, dtype: int64
```

There were many samples of category 1, and the ratio of 1 to 0 was 2.27:1. There was a certain imbalance in the samples. Therefore, measures should be taken from data sampling, model selection, and algorithm evaluation criteria to reduce the impact of sample imbalance on the model. Instead of random sampling, stratified sampling is used to divide the training set and test set, so that the data of the divided training set and test set have nearly consistent distribution with the overall data in each feature.

4. Model and Solution Method

4.1 T-test Model

First, standardize the data: [3]

$$x = \frac{x - \min(x)}{\max(x) - \min(x)}$$

See Annex 1 for the distribution map, V1 in Figure 3 shows the frequency distribution of indicator V1 for y equals 0 (blue) and 1 (yellow). Considering that the randomized trial is influenced by several random factors that are independent of each other, but it is difficult to determine the leading indicators. Under the condition that the sample size is large enough, the observed values of the random experiment approximately conform to the normal distribution. When one variable is a categorical variable and the other is a continuous variable, we choose to use t test for correlation analysis. The variables significantly correlated in the T test were taken as the main influencing variable.

(1) Data Visualization

Using Python to draw the bar chart of each variable, and some of the figures are shown as follows (see Figure 3). It can be seen that it's approximately normally distributed. SPSS was used to conduct normal test respectively, and the data was normalized by its standard deviation.

Let's look at the first line, Sig=significance, if Sig>0.05, the first line shall prevail. Otherwise, the second row shall prevail. If Sig (bilateral)<0.05, it represents rejection of the null hypothesis and correlation of the variable. Otherwise it's irrelevant. T-test can compare the differences between the two groups of data and combine the analysis results of independent sample test, as shown in the Table 2 and Table 3:

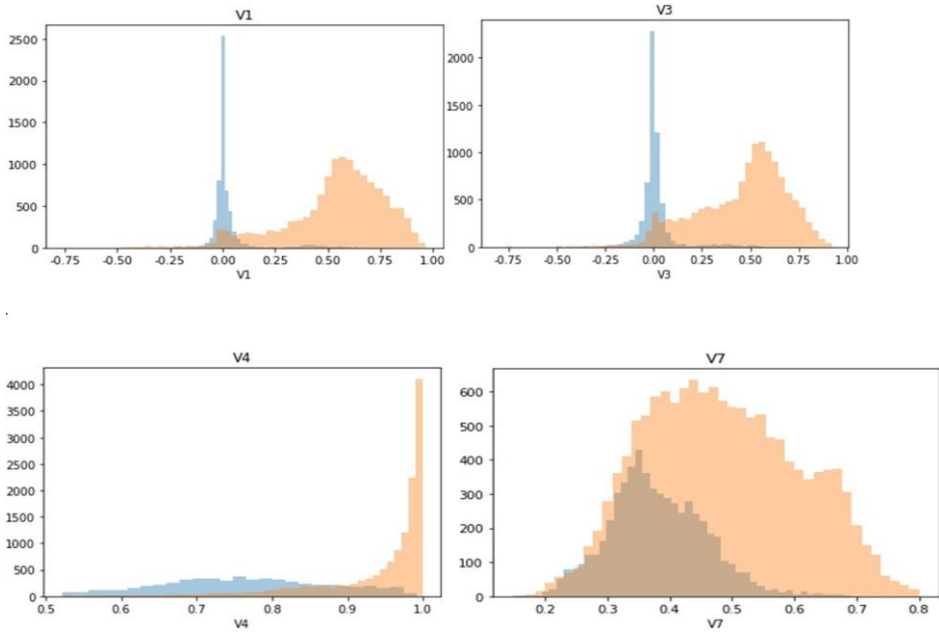


Figure 3 A distribution pattern of variables

Table 2 The set of statistics

ρ	γ	N	mean	std	standard error of mean
V1	1	13895	.74955102	.140551624	.001192358
	0	6105	.45393852	.060900144	.000779427
V2	1	13895	.78680842	.187319777	.001589111
	0	6105	.22439662	.166510506	.002131075
V3	1	13895	.78496870	.111091643	.000942437
	0	6105	.58326271	.045565254	.000583164
V4	1	13895	.91104669	.124466226	.001055899
	0	6105	.63271749	.183495732	.002348460
V5	1	13895	.63517327	.211930923	.001797898
	0	6105	.29235970	.112110461	.001434839
V6	1	13895	.79468025	.145428155	.001233727
	0	6105	.68320558	.147461188	.001887274
V7	1	13895	.49931456	.171432352	.001454332
	0	6105	.35175893	.111852070	.001431532

Table 3 Independent sample test

ρ	Levene test of variance equation ^ρ	T test for the equation of the mean ^ρ				
		F ^ρ	Sig. ^ρ	t ^ρ	df ^ρ	Sig.(bilateral) ^ρ
V1 ^ρ	The variance is equal ^ρ	3533.956 ^ρ	.000 ^ρ	157.948 ^ρ	19998 ^ρ	.000 ^ρ
	Variance inequality ^ρ			207.519 ^ρ	19994.771 ^ρ	.000 ^ρ
V2 ^ρ	The variance is equal ^ρ	65.060 ^ρ	.000 ^ρ	202.116 ^ρ	19998 ^ρ	.000 ^ρ
	Variance inequality ^ρ			211.565 ^ρ	13012.019 ^ρ	.000 ^ρ
V3 ^ρ	The variance is equal ^ρ	4940.423 ^ρ	.000 ^ρ	136.896 ^ρ	19998 ^ρ	.000 ^ρ
	Variance inequality ^ρ			182.000 ^ρ	19922.483 ^ρ	.000 ^ρ
V4 ^ρ	The variance is equal ^ρ	1617.192 ^ρ	.000 ^ρ	124.965 ^ρ	19998 ^ρ	.000 ^ρ
	Variance inequality ^ρ			108.093 ^ρ	8665.737 ^ρ	.000 ^ρ
V5 ^ρ	The variance is equal ^ρ	4017.594 ^ρ	.000 ^ρ	119.268 ^ρ	19998 ^ρ	.000 ^ρ
	Variance inequality ^ρ			149.032 ^ρ	19356.090 ^ρ	.000 ^ρ
V6 ^ρ	The variance is equal ^ρ	2.685 ^ρ	.101 ^ρ	49.708 ^ρ	19998 ^ρ	.000 ^ρ
	Variance inequality ^ρ			49.440 ^ρ	11512.010 ^ρ	.000 ^ρ
V7 ^ρ	The variance is equal ^ρ	1576.588 ^ρ	.000 ^ρ	61.726 ^ρ	19998 ^ρ	.000 ^ρ
	Variance inequality ^ρ			72.307 ^ρ	17170.562 ^ρ	.000 ^ρ

(2) Analysis Conclusion

Because the Sig values of the seven analysis items in the figure were all less than 0.05, indicating a significant difference, and the seven variables V1--V7 were all correlated. According to the comparison of the mean absolute values of the group statistics, the larger the difference was, the more correlated it was. If the absolute value is greater than 0.2, it indicates that there is a strong correlation between the two which is the main indicator. Thus, v1-V5 is the main indicator, while V6 and V7 are the secondary indicators.

(3) Test of Conclusion

Python's Pandas library is used to calculate the correlation coefficients for each indicator (V1-V7) and category (whether there is a specific indicator) and draw the correlation coefficient matrix, as shown in Figure 4.

It can be seen that the correlation coefficient of V1-V5 and label (whether containing a specific indicator) is above 0.6, which can confirm the conclusion obtained by T test that V1-V5 has a strong correlation with a specific component and is the main indicator, while the correlation of V6 and V7 is slightly weak.

4.2 PCA

Since the data has been standardized during data preprocessing, NumPY library directly calculates the standardized covariance matrix, and the visualization result of the covariance matrix is shown in Figure 5. KMO(Kaiser-Meyer-Olkin) and Bartlett ball tests were then performed on the covariance matrix. The KMO test value is 1.948 and the Bartlett test value is 1.07×10^{-12} by using the custom function. So there is a correlation between the data and principal component.

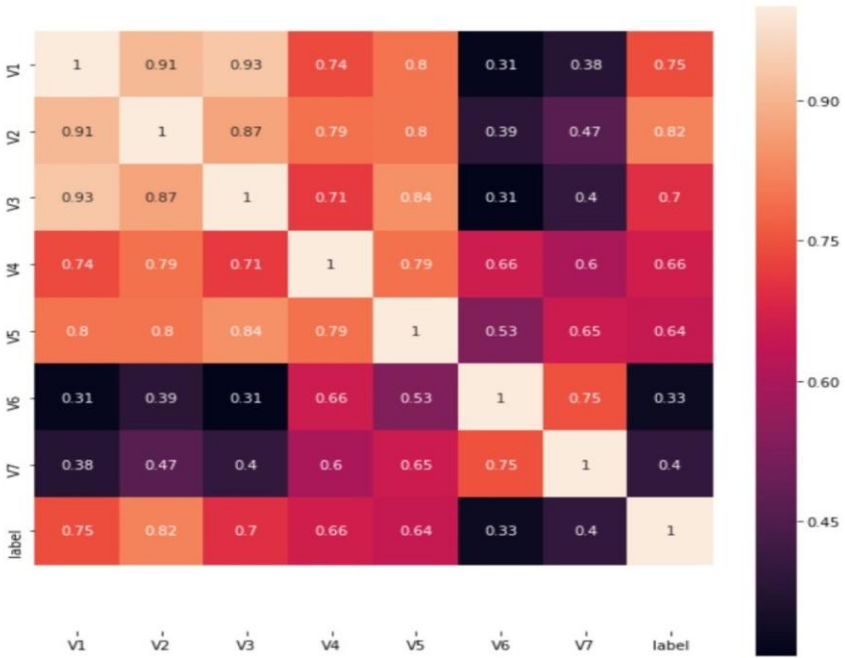


Figure 4 Correlation coefficient matrix

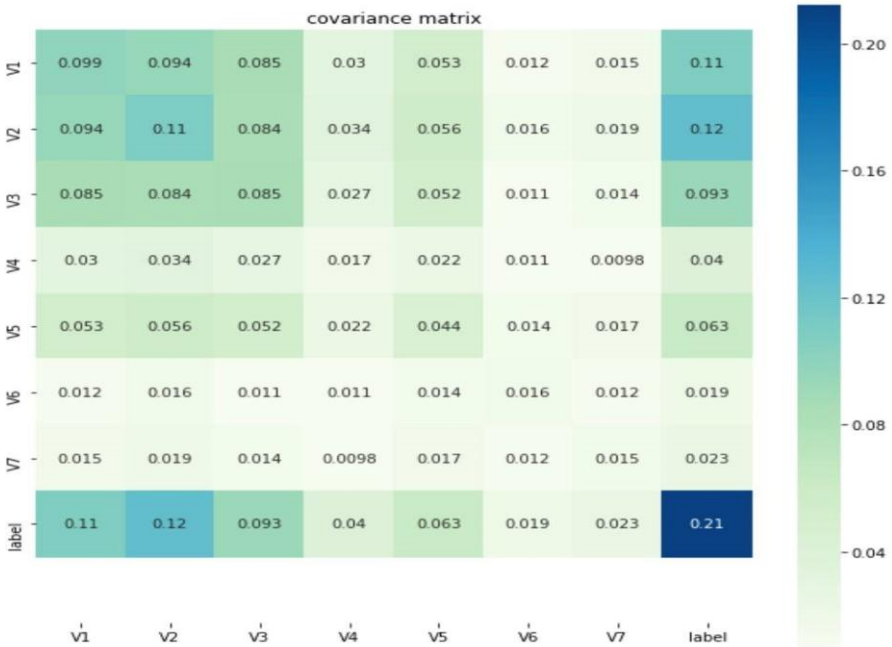


Figure 5 Covariance matrix

Step 3 and step 4 complete by calling the Sklearn library's Decomposition module. Figure 7 shows the edge information of each new variable that the principal component can create. It can be seen that the new variable f1 contains 70.6% of the original 7 variables, f2 contains 17.6% of the original 7 variables, f3 contains 4.6%, and so on.

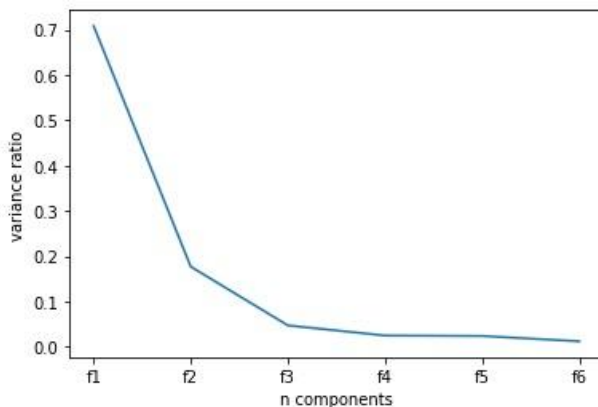


Figure 6 The edge information of the new variable in PCA

For the selection of the number of new variables, refer to the "Elbow rule", which is to find the obvious inflection point in the graph, it is easy to find the elbow position in figure 6 is at f3, that is, the use of f1, f2, and f3 variables can retain most of the information of the data set (93.2%).

Therefore, the first three features are considered as potential major components, and their relationship with the original V1-V7 is as follows:

$$F1 = -0.40043*V1 - 0.41*V2 - 0.40339*V3 - 0.4074*V4 - 0.41684*V5 - 0.27661*V6 - 0.30428*V7$$

$$F2 = 0.34374*V1 + 0.24579*V2 + 0.32363*V3 - 0.07159*V4 + 0.01153*V5 - 0.6345*V6 - 0.55574*V7$$

$$F3 = 0.04547*V1 + 0.08565*V2 - 0.13065*V3 + 0.49405*V4 - 0.3449*V5 + 0.47352*V6 - 0.6215*V7$$

As can be seen from the principal component formula, the indexes closely related to each principal component are V1, V2, V3 and V4. V5, V6 and V7 are secondary indicators. This result conforms to the correlation test obtained in Figure 4.

5. Result

In order to verify whether the three new variables F1, F2 and F3 are potential principal components, we used F1, F2 and F3 as features to train the SVM model, predict whether there are specific components, and compare the accuracy obtained with the SVM model trained by using V1-V7 as features.

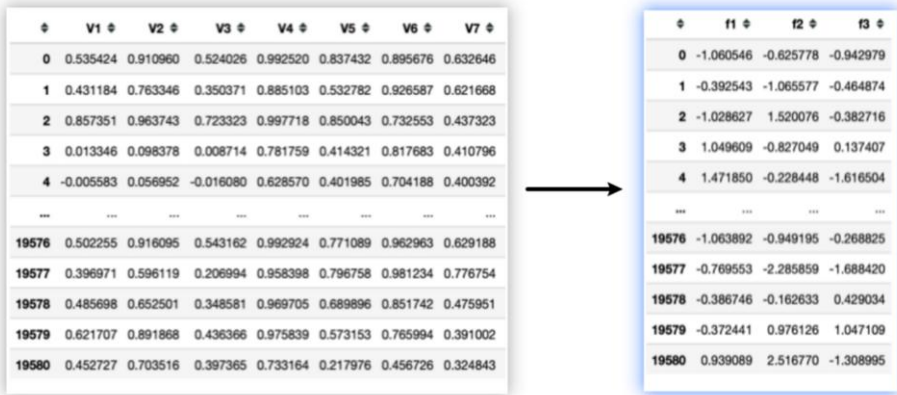


Figure 7 Feature data transforms the data of the principal component

According to the operation results, the SVM model trained with V1-V7 as the feature has an accuracy rate of 94% on the test set, and the accuracy rate of using F1-F3 as the feature is 92.6%. It can be known that the three new features can well determine whether there is a specific component, so F1, F2 and F3 are considered as the potential main components.

6. Conclusion

Since T test depends on a large sample size, a good approximation can be obtained in this paper, and T test almost uses all the data information so it's best to find the difference. In this paper, correlation analysis, PCA and other methods are comprehensively used to find the main indicators for the detection of specific components in mixtures. In the implementation process, make full use of data to illustrate problems, intuitive and visual, from the actual characteristics of data to choose the modeling strategy, scientific and objective.

References

[1] Fisher Box, Joan. Guinness, Gosset, Fisher, and Small Samples. Statistical Science. 1987, 2 (1): 45–52.
 [2] Si Shou-Kui, SUN Xi-jing. Python Mathematical Experiment and Modeling [M]. Beijing: Science Press,2020.04:335.
 [3] Si Shoukui, SUN Zhaoliang. Mathematical Modeling Algorithm and Application (2nd edition) [M]. Beijing: National Defense Industry Press, 2019, 07:223.

Investigation of the Effect of Different Covariance Estimation Methods on the Performance of Least Squares Compressive Covariance Sensing

Nuha A. S. Alwan¹

Department of Electrical Engineering, College of Engineering, University of Baghdad, Baghdad, Iraq

Abstract. Compressive covariance sensing (CCS) can recover the second-order statistics of a signal that has undergone compression, and this can be achieved without the requirement of sparsity conditions. Instead, certain structure information in the statistical domain is to be captured during compression. In particular, least squares compressive covariance sensing is considered which requires the estimation of the covariance of the available compressed signal in order to recover the covariance matrix of the original signal. Different covariance estimation methods are applied and the CCS performance compared, in the presence of white Gaussian noise, in terms of the normalized mean square error between the true and recovered covariance.

Keywords. compressive covariance sensing, least squares, linear sparse ruler, covariance matching, covariance estimation methods, preprocessing.

1. Introduction

Compressive sensing (CS) [1] is a technique that enables simultaneous signal acquisition and compression to limit sensing, storage and communication costs using sub-Nyquist sampling of the signal. Recovery of the original signal is possible provided the latter has a sparse representation in a known transformed domain. This paradigm of CS has impacted a wide range of applications in communications, networking and signal and image processing. On the other hand, compressive covariance sensing (CCS) [2] can recover the second-order statistics of a signal after compression rather than the signal itself, but without the requirement of the sparsity condition. In this case, structure forms other than sparsity are to be captured whilst the signal undergoes compression. Such structure information is present in the statistical domain enabling the reconstruction of second-order statistics of a wide-sense stationary (WSS) signal, and recently [3], even non-stationary signals via online CCS. Since many signal processing techniques and methodologies are based on second-order statistics, CCS finds applications in power spectrum estimation, frequency estimation and direction-of-arrival (DOA) estimation, among others. In CCS, it is generally assumed that the covariance matrix of the original signal has a Hermitian Toeplitz structure such that the covariance matrix can be reconstructed from the covariance sequence, so that, actually, it is the recovery of the covariance sequence that is aimed at.

¹ Corresponding Author, email: n.alwan@ieee.org

This paper focuses on least squares compressive covariance sensing (LS-CCS) in which compression is achieved by a sparse ruler that identifies the sampling pattern. The latter is such that a ruler of length $N-1$ can measure all integer lags from zero to $N-1$ according to the sparse marks of the ruler [2, 3]. This leads to a development of an overdetermined problem due to the assumed Hermitian Toeplitz structure of the covariance matrix [4, 2]. The overdetermined problem can then be solved by least squares leading to the LS-CCS method which involves the computation of an estimate of the covariance matrix of the compressed signal as a step towards recovering the covariance matrix of the original signal.

Different covariance estimates of the compressed signal exhibit different properties and thereby affect the CCS process. The purpose of the present work is to assess the performance of the LS-CCS method using different covariance estimation methods in the presence of white Gaussian noise (WGN). It is found that the unbiased covariance estimation method yields the optimum results.

The rest of the paper is organized as follows: Section 2 presents covariance and sparse ruler sampling. In Section 3, the LS-CCS method is explained, and different methods of covariance estimation are also discussed. The simulation results are presented in Section 4. Finally, Section 5 concludes the paper.

2. Covariance and Sparse Rule Sampling

The autocorrelation of a WSS random signal $x(n)$ is given by:

$$\mathbf{r}_x(\tau) = E[x(n)x^*(n-\tau)], \quad (1)$$

where the operator E is the expectation or ensemble average operator, the asterisk represents the complex conjugate, n is the time index and τ is the time shift or lag. Another second-order statistic is the autocovariance given by:

$$\mathbf{c}_x(\tau) = E[(x(n) - \mu)(x^*(n-\tau) - \mu)], \quad (2)$$

where μ is the statistical mean of the random signal $x(n)$. Clearly, if the mean is zero, autocorrelation and autocovariance, as given by Equations 1 and 2 respectively, are interchangeable. For simplicity, autocovariance will be referred to as covariance throughout the paper. Moreover, if the random process is also ergodic [5], the theoretical ensemble average can be replaced by a time average which is easily calculated as follows:

$$\mathbf{c}_x(\tau) = \frac{1}{T} \sum_{n=1}^T x(n)x^*(n-\tau) \quad (3)$$

where T is the number of terms to be averaged.

If we assume that the available number of signal samples is N , the covariance can be computed for lag values ranging from 0 to $N-1$. The covariance as a sequence is then defined by:

$$\mathbf{c}_x = [c_x(0), c_x(1), \dots, c_x(N-1)]. \quad (4)$$

In applications, covariance is generally described as a matrix and is characterized by specific structures such as positive semidefinite Hermitian Toeplitz [5]. As such, the covariance matrix can be constructed solely from its first column which is none other than the sequence given by Equation 4. In the present work, only real signals will be considered, and therefore, the covariance matrix structure is reduced to symmetric Toeplitz (ST).

As explained in the introduction, CCS recovers the covariance of a signal from a compressed version of this signal to attain the entailed compression advantages. This is possible if compression is achieved using a sparse ruler. A linear sparse ruler (LSR) is shown in Figure 1. It can be thought of as a ruler with missing marks. The existing remaining marks allow all integer distances between zero and the ruler length to be measured.

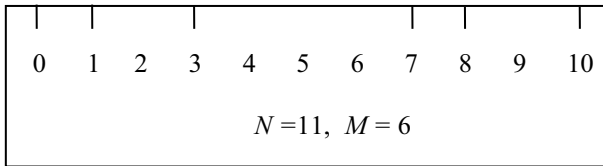


Figure 1. An example of a linear sparse ruler of length $N-1=10$.
 $K=\{0, 1, 3, 7, 8, 10\}$

As shown in Figure 1, a length- $(N-1)$ LSR consists of M marks, usually with $M \ll N$. M takes on integer values between 0 and $N-1$. The ruler can be thought of as a set $K \subset \{0, 1, \dots, N-1\}$ on the condition that for all lag values τ between 0 and $N-1$, there is at least one pair of elements in K , namely (k, k') , such that this pair satisfies $k - k' = \tau$, with $k > k'$ [2]. When this LSR of length $N-1$, represented by K , is synchronized with a signal vector of length N , it will sparsely sample the signal vector at its M marks resulting in a length- M compressed signal vector. Since the LSR can measure all distances from 0 to $N-1$, all possible lags can be derived and the covariance sequence of the original signal can be estimated from the compressed signal.

The compression ratio of the length- $(N-1)$ LSR is given by M/N , and is almost inversely proportional to the length $N-1$. Therefore, long LSRs perform better. There are also circular sparse rulers (CSR) that yield better compressibility than LSRs because of their ability to measure two different distances with each pair of marks [2]. However, the discussion in this work will be concerned with LSRs only.

Although there arise in applications covariance matrix structures other than Toeplitz such as circulant [6] or banded [7], most works on CCS deal with estimating Toeplitz matrices with non-periodic sparse samplers using the LSR, requiring at least a pair of samples in the compressed signal vector for each possible value of lag [4].

3. Least Squares Compressive Covariance Sensing (LS-CCS)

To formulate the CCS problem first, consider the recovery of the covariance of a zero-mean random signal vector $\mathbf{x} \in \mathbf{R}^N$ from the compressed signal vector $\mathbf{y} \in \mathbf{R}^M$ given by:

$$\mathbf{y} = \Phi \mathbf{x} \tag{5}$$

where the matrix $\Phi \in \mathbf{R}^{M \times N}$ is called the compression matrix or sampler, and $M \ll N$. Several realizations of the vector \mathbf{y} may be available.

The matrix Φ performs mathematically the function of the sparse ruler. It is a sparse matrix with at most one non-zero entry in each row or column. Each of the M rows of Φ contains a ‘one’ at one of the corresponding M positions of the LSR and zeros elsewhere. In contrast to sparse sampling matrices, there are applications that employ dense sampling matrices whose design relies on probabilistic arguments and that have been proven successful [2]. These, however, will not be considered further in this work.

The theoretical covariance matrix described by Equations 1 and 2 will be re-written here in terms of the vector \mathbf{x} , and given a temporary different symbol here (Σ) for notation convenience:

$$\Sigma = E[\mathbf{x} \mathbf{x}^T]. \tag{6}$$

The matrix Σ is assumed to be a linear combination of the ST matrices Σ_s that are elements of the set $\mathfrak{S} = \{\Sigma_0, \Sigma_1, \dots, \Sigma_{S-1}\} \subset \mathbf{R}^{N \times N}$. This implies that there exist real scalars α_s such that:

$$\Sigma = \sum_{s=0}^{S-1} \alpha_s \Sigma_s. \tag{7}$$

When the subspace \mathfrak{S} is assumed to be a linearly independent set of matrices, the decomposition in Equation 7 is unique such that a knowledge of the α_s leads to knowing Σ . The above ST structure for Σ that is characteristic of WSS processes enables a certain degree of compression. If no prior information about Σ is available and the latter is simply considered to be a symmetric positive semidefinite matrix, then no compression is possible that preserves the second-order statistics of the original signal. The condition $S < 2N - 1$ must be satisfied for \mathfrak{S} to be linearly independent [4].

The covariance of the vector \mathbf{y} is given by

$$\bar{\Sigma} = E[\mathbf{y} \mathbf{y}^T]. \tag{8}$$

Substituting Equations 5 and 6 in 8 yields:

$$\bar{\Sigma} = \Phi \Sigma \Phi^T = \sum_{s=0}^{S-1} \alpha_s \bar{\Sigma}_s, \tag{9}$$

where $\bar{\Sigma}_s = \Phi \Sigma_s \Phi^T,$ \tag{10}

so that $\bar{\Sigma}$ is linear combination of the symmetric matrices that are elements of the set $\bar{\mathcal{S}} = \{\bar{\Sigma}_0, \bar{\Sigma}_1, \dots, \bar{\Sigma}_{S-1}\} \subset \mathbf{R}^{M \times M}$. These symmetric matrices in the subspace $\bar{\mathcal{S}}$ are not necessarily Toeplitz. If compression preserves the second-order statistical information, that is if, for example, it is achieved according to the LSR described in Section 2, then $\bar{\mathcal{S}}$ is linearly independent and knowing Σ_s leads to knowing $\bar{\Sigma}_s$ from Equation 10. This entails knowing the α_s from Equation 9 and finally knowing Σ using Equation 7.

The problem of estimating the α_s from the linear parameterization of Equations 9 and 10 is known in the literature as structured covariance estimation or covariance matching [8, 4], and it plays an important role in CCS as will be shown by way of the following example. The example also clarifies the need for a least squares solution to an overdetermined system of equations.

3.1. CCS Example

We assume the original signal vector \mathbf{x} to have limited dimension ($N=4$) to clarify the CCS concepts discussed so far. Let the vector \mathbf{x} be given by:

$$\mathbf{x} = [x_o \ x_1 \ x_2 \ x_3]^T$$

A LSR of $M=3$ will be used to compress the above vector. Due to the small dimension ($N=4$), M cannot be made smaller than 3 if the second-order statistical information is to be preserved. This LSR is mathematically equivalent to multiplying \mathbf{x} by the following sparse sampling matrix:

$$\Phi = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Using Equation 5, the compressed vector will be:

$$\mathbf{y} = [x_o \ x_2 \ x_3].$$

Therefore, Σ has dimensions of 4×4 , whereas $\bar{\Sigma}$ has dimensions of 3×3 . The covariance matrix of the compressed signal, $\bar{\Sigma}$, can be readily computed from the vector \mathbf{y} using one of the covariance estimation methods to be discussed shortly, one of which is given by the practical computation of Equation 3. The steps for recovering the covariance matrix of the original signal, Σ , are the following:

- Choose the matrices Σ_s such that they are symmetric, Toeplitz and linearly independent. We can choose:

$$\Sigma_o = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \quad \Sigma_1 = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \quad \Sigma_2 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

$$\Sigma_3 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

We are taking $S=4$ (or N) since, as explained, it is required that $S < 2N - 1$.

- For each Σ_s , find $\bar{\Sigma}_s$ from Equation 10. The results are:

$$\bar{\Sigma}_o = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \quad \bar{\Sigma}_1 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad \bar{\Sigma}_2 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \quad \bar{\Sigma}_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

It is clear that the $\bar{\Sigma}_s$ are symmetric and linearly independent, but not necessarily Toeplitz.

- Compute the covariance matrix of the compressed signal, $\bar{\Sigma}$, using one of the covariance estimation methods to be discussed next. Naturally, the availability of more than a single realization of \mathbf{y} (that is, of \mathbf{x}) gives a more accurate result.
- Find the α_s from Equation 9, knowing $\bar{\Sigma}$ and $\bar{\Sigma}_s$.
- Finally, find Σ from Equation 7.
- The result can be verified by finding the covariance matrix Σ by the same estimation method used to find $\bar{\Sigma}$. If we assume that the covariance sequence vector of the original signal, as in Equation 4, is denoted by \mathbf{c}_x , and the recovered covariance sequence vector computed for this same signal from the above steps is denoted by $\hat{\mathbf{c}}_x$, then the normalized mean square error (NMSE) between the covariance sequence \mathbf{c}_x and its recovered version $\hat{\mathbf{c}}_x$ can be calculated as the ratio of the squared magnitude of the difference vector to that of the original sequence vector, and is given below in dB's:

$$NMSE(dB) = 10 \log_{10} \left[\frac{\|\mathbf{c}_x - \hat{\mathbf{c}}_x\|^2}{\|\mathbf{c}_x\|^2} \right] \tag{11}$$

We notice from the above example that $S=4$, so we have four values of α_s , that is, four unknowns. However, Equation 9 from which these unknowns can be computed provides nine equations as all involved matrices have nine elements each. This is an overdetermined system of equations that can be solved by least squares. In MATLAB, any of the two commands, `lsqr(.)` or `pinv(.)`, can be used.

3.2. Estimation Methods for the Covariance Sequence of the Compressed Signal

As can be seen from Equations 1 and 2, covariance involves averaging. LS-CCS can utilize an additional average as preprocessing to obtain more accurate covariance estimation [2, 9] that can be explained as follows. The NMSE computed from Equation 11 becomes larger as compression increases since the total number of available signal samples is reduced. This is true for any estimation, not just LS estimation. Therefore, to increase accuracy of the estimate, several realizations of the vector \mathbf{y} are needed over which an average is taken. This does not contradict the concept of compression since the average sampling rate is reduced anyway, and it is this reduction in sampling rate that entails the compression advantage of reduced hardware cost [2].

In accordance with Equation 4, the covariance matrix of the compressed signal, $\bar{\Sigma}$, has a covariance sequence, derived from its first column, that can be denoted by \mathbf{c}_y . Regardless of the additional averaging or preprocessing just discussed, common estimation methods to obtain \mathbf{c}_y are as follows:

- Unscaled estimation: This is exemplified by the MATLAB command `xcorr(.)`. This can be expressed as

$$\mathbf{c}_y(\tau) = \sum_{n=1}^T y(n)y(n-\tau) \tag{12}$$

where T is the number of terms to be summed and is equal to $M - \tau$.

- Biased estimation. This is given by:

$$\mathbf{c}_y(\tau) = \frac{1}{M} \sum_{n=1}^T y(n)y(n-\tau) \tag{13}$$

- Unbiased estimation. This is given by:

$$\mathbf{c}_y(\tau) = \frac{1}{T} \sum_{n=1}^T y(n)y(n-\tau) \tag{14}$$

The covariance estimates of Equations 12-14 can all be made more accurate by preprocessing averaging.

4. Simulation Results and Discussion

Simulations are carried out in MATLAB. A length-33 signal vector \mathbf{x} is chosen with the unscaled covariance shown in Figure 2. This has been computed and plotted using the MATLAB function `xcorr(.)` whose argument is set to \mathbf{x} , yielding the covariance for lag values ranging from -32 to 32 that correspond to $\mathbf{c}_x(-32)$ to $\mathbf{c}_x(32)$. The signal vector \mathbf{x} is divided into three parts or three realizations, each of length 11, denoted by \mathbf{x}_1 , \mathbf{x}_2 and \mathbf{x}_3 . Compression and CCS will be performed using one realization (\mathbf{x}_1) and thereafter the three realizations (\mathbf{x}_1 , \mathbf{x}_2 and \mathbf{x}_3) to highlight the increased accuracy advantage of preprocessing discussed in Section 3.2.

Working with length-11 realizations, the CCS results are expected to be $\hat{\mathbf{c}}_x(0)$ to $\hat{\mathbf{c}}_x(10)$. Thus, the original values of \mathbf{c}_x considered for comparison or verification will be limited to $\mathbf{c}_x(0)$ to $\mathbf{c}_x(10)$ only out of the 65 values of Figure 2.

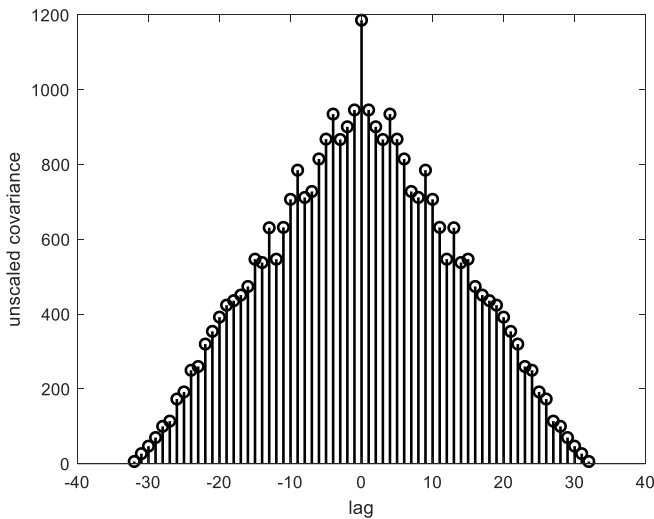


Figure 2. Unscaled covariance of the vector \mathbf{x} of length 33 obtained by the MATLAB command `xcorr(x)`.

For each length-11 realization, the LSR of length $N=11$ shown in Figure 1 is used for compression to obtain a length- M \mathbf{y} vector where $M=6$. For the first realization, CCS is carried out exactly as outlined in the example of Section 3.1, but with respect to the different vectors and matrices dimensions, and the NMSE is calculated using Equation 11.

Preprocessing and averaging over three realizations is done as follows: The LSR of Figure 1 is applied to each of \mathbf{x}_1 , \mathbf{x}_2 and \mathbf{x}_3 , and the $\hat{\mathbf{c}}_x$ values are computed and averaged over the three realizations for $\tau = 0$ to $\tau = 10$ using LS-CCS. Thereafter the NMSE is calculated.

To incorporate WGN, the noise is added to the observed compressed signal vector \mathbf{y} . The compressed signal power is computed for each compressed realization denoted by \mathbf{y}_1 , \mathbf{y}_2 or \mathbf{y}_3 . Each compressed realization is of length $M=6$ corresponding to the length-11 original realizations. The total concatenated compressed vector \mathbf{y} is of length 18 corresponding to the length-33 original signal \mathbf{x} . The standard deviation (SD) of the noise is found by specifying the values of signal-to-noise-ratio (SNR) in dB's and substituting them in:

$$SD = \sqrt{\frac{\text{compressed signal power}}{10^{0.1 \times SNR(\text{dB})}}} \quad (15)$$

In the simulations, the above value of SD is computed and added to the compressed realizations. CCS is performed and NMSE is computed in dB from Equation 11. All simulations are averaged over 1000 independent runs.

To compare between the different estimation methods of the covariance \mathbf{c}_y , all the afore-mentioned explained procedure is repeated for each of the estimation methods presented in Section 3.2, and $\hat{\mathbf{c}}_x$ is found in each case by LS-CCS. Note that the computed original \mathbf{c}_x needed in Equation 11 must be found by the same estimation method used for \mathbf{c}_y .

Figure 3 demonstrates the results of the CCS process as a plot of the NMSE versus the SNR when each of the three estimation methods are used for the computation of the covariance of the compressed signal vector. Preprocessing is employed so that the three realizations are used. We notice that the unbiased scaled estimation results in the best CCS performance for moderate and high SNR values whereas the unscaled method is best for low SNR. This is explained as follows. In general, the smaller the value of lag τ , the higher the quality of the estimate because the number of averaging terms is larger for small lags as can be observed from Equations 12-14. Moreover, we note that the biased and unbiased estimation methods in Equations 13 and 14 reduce the covariance magnitudes compared to the unscaled estimation of Equation 12, thereby reducing the importance of the high-lag covariance values even further and making the low-lag values the prominent or dominant ones. It is also noteworthy that the white uncorrelated noise variance or power is actually added only to $\mathbf{c}_y(0)$ and is absent from the rest of the covariance values. Thus, for low SNR the noise level is high and therefore the dominant low-lag covariance values are the noisy ones when the second and third estimation methods are used, leading to poor overall CCS performance. In the first method, the noisy low-lag values are not that dominant because the covariance values for other lags are also quite considerable.

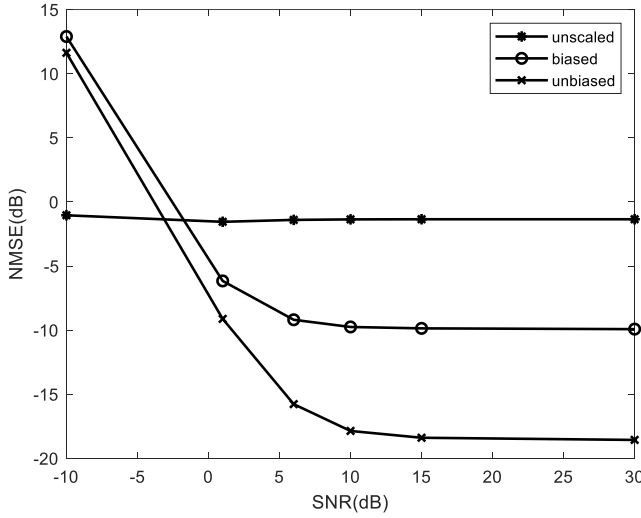


Figure 3. LS-CCS performance as a plot of NMSE (dB) vs. SNR (dB) using the unscaled, biased and unbiased covariance estimation methods. Three realizations are used.

To demonstrate the accuracy advantage of preprocessing, Figure 4 is a plot of NMSE vs. SNR using only the third estimation method (scaled unbiased) with LS-CCS carried out for one and then three realizations. Clearly, the latter case gives better results, that is, smaller NMSE.

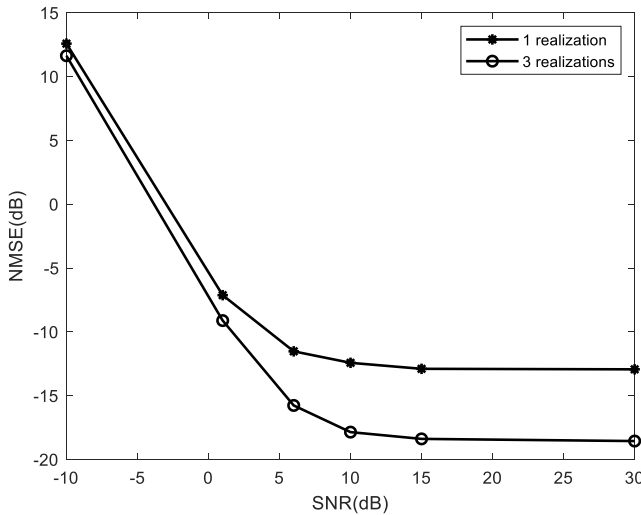


Figure 4. LS-CCS performance as a plot of NMSE (dB) vs. SNR (dB) using the scaled unbiased covariance estimation method, and for the cases of one and three realizations.

5. Conclusion

Least squares CCS is simulated using different covariance estimation methods in computing the compressed signal covariance in order to recover the covariance of the original signal through covariance matching. It is found that the scaled biased and especially the unbiased covariance estimation methods, in achieving LS-CCS, outperform the unscaled method for moderate and high SNR's. This improvement is manifested in terms of the normalized mean square error between the true and recovered covariance values. Additional averaging by preprocessing is found to result in even better CCS performance. As future work, applications that necessitate computation of correlation values from compressed measurements will be researched, such as the important field of single-tone frequency estimation in white and colored noise.

References

- [1] E. J. Candes and M. B. Wakin, An introduction to compressive sampling, *IEEE Signal Processing Magazine* **25** (2008), 21-30.
- [2] D. Romero, D. D. Ariananda, Z. Tian and G. Leus, Compressive covariance sensing: structure-based compressive sensing beyond sparsity, *IEEE Signal Processing Magazine* **33** (2016), 78-93.
- [3] C. Park and B. Lee, Online compressive covariance sensing, *Signal Processing* **162** (2019), 1-9.
- [4] D. Romero and R. Lopez-Valcarce, Compression limits for random vectors with linearly parameterized second-order statistics, *IEEE Transactions on Information Theory* **61** (2015), 1410-1425.
- [5] V. Ingle, S. Kogon and D. Manolakis, *Statistical and Adaptive Signal Processing*, Artech, 2005.
- [6] C. P. Yen, Y. Tsai and X. Wang, Wideband spectrum sensing based on sub-Nyquist sampling, *IEEE Transactions on Signal Processing* **61** (2013), 3028-3040.
- [7] D. D. Ariananda and G. Leus, Compressive wideband power spectrum estimation, *IEEE Transactions on Signal Processing* **60** (2012), 4775-4789.
- [8] B. Ottersten, P. Stoica and R. Roy, Covariance matching estimation techniques for array signal processing applications, *Digital Signal Processing* **8** (1998), 185-210.
- [9] D. Romero and G. Leus, Wideband spectrum sensing from compressed measurements using spectral prior information, *IEEE Transactions on Signal Processing* **61** (2013), 6232-6246.

Ad-Hoc Network Access Control System and Method for Edge NFC Terminal

Yiqin BAO^a and Zhengtang SUN^b

^aCollege of Information Engineering of Nanjing XiaoZhuang University, China

^bSchool of computer science and technology of Heilongjiang University, China

Abstract. With the rapid development of wireless communication technology, Near field communication (NFC) and Ad-Hoc network (ZigBee) have attracted wide attention for their security, speed and low power consumption. NFC has been used in access control research. In the past, the system only considered the switch control of the door lock, but could not achieve safety management in large building scenarios, resulting in security loopholes and poor maintenance of access control. In order to solve the above problems, Automatic networking access control system based on NFC and ZigBee technology is implemented. The access control and Internet of Things system (IoT) are fused through NFC. NFC identity information is transmitted through ZigBee to the management for identity authentication. The wireless transmission adopts the improved reduced AES advanced encryption algorithm R-AES, which ensures reliable data transmission. On this basis, the security of access control is realized. Practical results show that the Ad-Hoc network access control system and method for edge NFC terminal can improve the management efficiency and security of building access control.

Keywords. lot, edge computing, NFC, ZigBee, AES-128

1. Introduction

With the rapid development of national economy, people's living environment has changed dramatically. High rise buildings bring more and more serious safety and security problems. The increasing number of high-tech crimes directly threatens the personal safety. At the same time, people's demand for intelligent management is also increasing, so the term "intelligent access control system" merged. At present, the design of the convenience and security of access control system is a very popular topic [1]. Based on the rapid development of Internet of things (IoT) solutions, the system presents new challenges to information collection, processing [2]

In recent years, with the continuous progress of biometric technology and induction technology, access control system has grown unprecedentedly and entered a relatively mature stage. For domestic, contactless IC card, password and biological characteristics constitute the three solutions of the current intelligent access control system, but there are some disadvantages. First of all, for contactless IC card and password, the corresponding reader can only identify through the information or password in IC card, but can't confirm whether it's personal operation [3]; second, for biometrics, such as camera, fingerprint, etc., its hardware cost is relatively high, the identification time is long, and it needs to be identified repeatedly. For NFC access control, smart phones have the advantages of low cost, convenience and strong reliability. Related literature

[4] has used NFC for access control system and ECS connection management; related literature [5] realizes authentication management through 3G connection; literature [6] realizes access control identity identification management by combining NFC and fingerprint.

At present, NFC intelligent application has become a development hotspot. It has been applied in the process of vehicle charging[7], in food monitoring[8], in geological and mineral data collection[9], and in electronic payment[10].

This paper presents an intelligent access control system for mobile NFC terminals, which is connected to the Internet of things through NFC. It has the advantages of low cost, high reliability and scalability. The self-organizing intelligent access control system for mobile NFC terminal is close to the NFC access control terminal through NFC smart phone. After the access control terminal reads the identity information, it is uploaded to the server side through ZigBee network and compared with the database to give the corresponding access rights. At the same time, Managers can view the user's access records in the system and change the opening and closing status of corresponding doors according to the actual situation. The wireless information transmission adopts the customized simplified encryption algorithm R-AES to ensure the security of information transmission. This paper analyzes the technical framework of the access control system, analyzes the system design, unifies the wireless communication network communication and encryption, and realizes an intelligent access control system for mobile NFC terminal. With the functions of authentication, unified management and wireless data encryption, it greatly facilitates access control management, enhances the security of access control data transmission, and greatly facilitates management and life [11].

2. System Architecture

The structure of smart access control system for mobile NFC terminal is shown in Figure 1. The system is mainly composed of four parts, including Android application terminal, access control terminal, network forwarding terminal and server terminal. Android application terminal mainly refers to Android smart phone with NFC function, which transmits identity authentication information through contact with access control terminal; access control terminal includes NFC module, gate magnet module and ZigBee module; access control terminal integrates data collection and control, which can not only collect NFC data, but also control corresponding gate magnet module; network forwarding terminal is composed of ZigBee router, Responsible for the data interaction between the access control identification control terminal and the server; the server is mainly composed of ZigBee coordinator, WiFi router, PC and background database, responsible for the upload of the authentication information of the whole system and the release of access control commands. The system supports two-way data transmission from left to right. From left to right, first collect NFC authentication information through the access control terminal, then upload it to the server through ZigBee wireless sensor network and ZigBee network forwarding terminal to realize the docking of system access platform. From the right to the left, the server sends the access control data to the corresponding access terminal via ZigBee wireless sensor network through the network forwarding terminal to realize the access control. The whole system can not only realize NFC identity authentication and access control functions, but also has the advantages of convenient networking, strong

anti-interference performance and low cost [12], which can greatly improve the security management efficiency of the whole access control system.

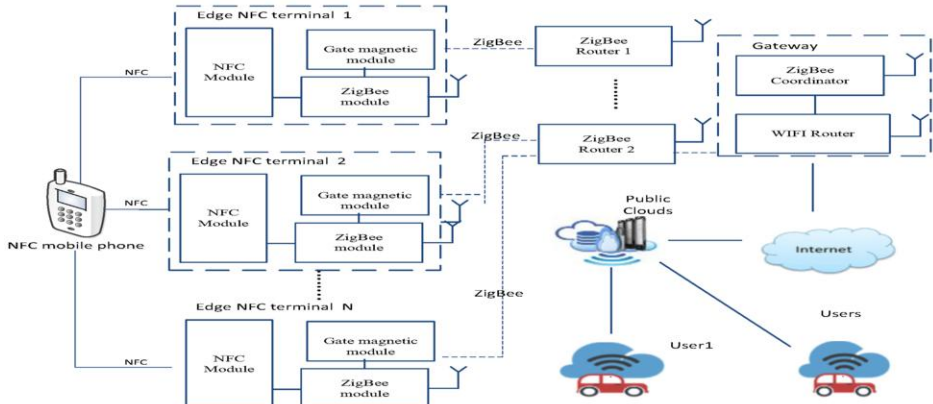


Figure 1. System architecture chart

3. System Design

3.1 NFC Module

NFC (Near Field Communication) is a short-range wireless communication technology. Mobile phones communicate with access control through NFC, and its foundation is RFID technology [13]. PN532 of NXP company is adopted as NFC access control reading module. As a short-range and high-frequency radio technology, NFC operates at a distance of 20cm in the frequency of 13.56MHz. Its transmission speed includes 106kbit / s, 212kbit / s and 424kbit / s. NFC has three modes of operation:

(1) Card emulation mode. This mode simulates the NFC device as an IC card. This has a great advantage: the card is powered through the RF domain of the contactless card reader, and can work even if the NFC device is powered down.

(2) Point to point mode. For data exchange. In this mode, the transmission distance is short, but the transmission speed is fast and the power consumption is low.

(3) Reader mode. Use the NFC device as a contactless card reader only, and read the corresponding information from the electronic label .

3.2 Zigbee Module

The ZigBee module used in the system is CC2530 chip, which is a real SOC solution based on 2.4GHz IEEE802.15.4, ZigBee and RF4CE application produced by Texas Instruments (TI). It can build powerful network nodes with little total material cost, program the chip, and perform logic operation, sequence control, timing, counting and calculation Technical operation, etc., through digital or analog input and output for environmental monitoring and equipment control signal output. As a new technology of wireless sensor network (WSN), it has a wide application prospect [14]. CC2530f256 is selected to realize the maximum function. The minimum system circuit is shown in Figure 2 [15].

NFC data needs to be uploaded to the server through ZigBee network for identity authentication. As a low-power, wireless ad hoc network protocol, ZigBee adopts

802.15.4 standard as the basis of its peer-to-peer communication. The standard was developed and managed by the ZigBee alliance. ZigBee is most commonly used in asynchronous communication, with CSMA / Ca channel intervention capability, and has all the functions in 802.15.4 standard. ZigBee can work in 2.4GHz (Universal), 868mhz (European) and 915MHz (U.S.), with the highest transmission rates of 250kbit / s, 20kbit / s and 40kbit / s respectively. Its transmission distance is in the range of 10-75m, but it can continue to increase. Because of its low power consumption, low cost, short time delay, large network capacity, reliability, security and other characteristics, it is widely used in intelligent home, intelligent medical, intelligent agriculture and other Internet of things industries.

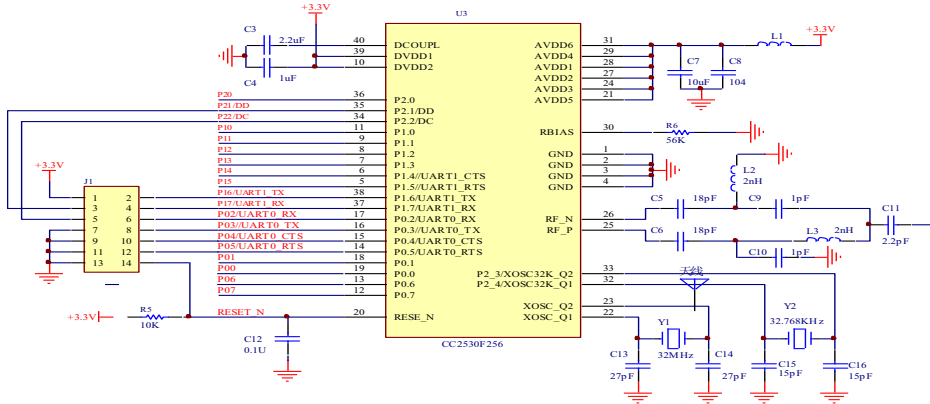


Figure 2. CC2530 system circuit diagram

3.3 Access Control Terminal Design

Access control terminal consists of three parts, NFC module, ZigBee module and door magnetic module. The NFC module Hsu_TX and Hsu_RX are respectively connected with pin P0.2 and pin P0.3 of ZigBee module to realize data interaction between NFC module and ZigBee module; the relay in pin is connected with pin P0.0 of ZigBee module to realize access control, and the corresponding pin is set in the program. CC2530 program flow chart is shown in Figure 3.

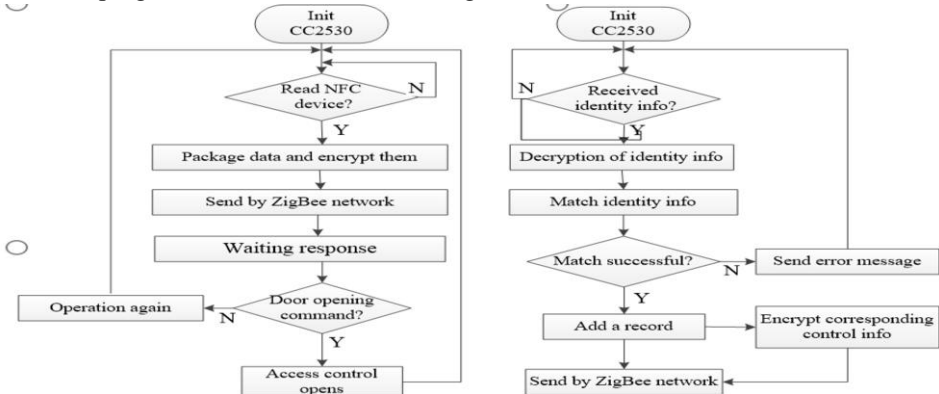


Figure 3. CC2530 program flow chart

Actual operation scenario: the system is installed on the door. When the NFC smart phone is close to the NFC reading device, the read information is transmitted to ZigBee module for data encryption and packaging. The data frame is sent to the server through ZigBee network to decrypt the package and match the data with the database. If the information matching is successful, the server will return the corresponding response command to control the opening and closing of the door; otherwise, it will remind the user to repeat the above steps; if the information matching still fails, it will send a warning message to the administrator for management.

4. System Communication

4.1 NFC Communication Protocol

The reliable transmission between NFC module and ZigBee ensures the accuracy of authentication. Therefore, the communication protocol is developed in the system, and the data frame structure is shown in Figure 4. For the communication mode between NFC communication module and ZigBee module in the system, the baud rate is set to 9600, the data bit is none, and the stop bit is 1.

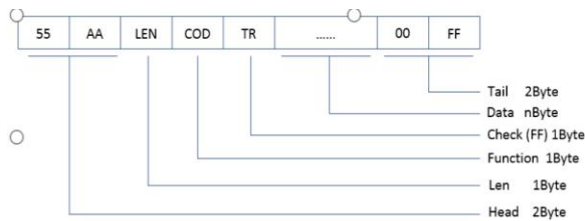


Figure 4. NFC frame format

For NFC communication frame format, frame header takes up 2 bytes, fixed as 55 and AA; len is frame length takes up 1 byte, which is the sum of function code, check sum and data; COD is function code takes up 1 byte, which is used to set the mode of NFC reading equipment, specific function code settings are shown in Table 1, fixed as 12 in the system, that is, point-to-point mode; check sum takes up 1 byte, fixed as FF; data bits take up n bytes It is used to store data; the frame end occupies 2 bytes, fixed as 00, FF.

Table 1. Function code table

Pattern	Function Code
Reader mode	0x10
Card simulation mode	0x11
Point to point mode	0x12
Query firmware version	0x25

4.2 Modbus Communication Protocol

After the ZigBee module packs the data transmitted from NFC, it needs to send the identity authentication data to the server through ZigBee. The international Modbus communication protocol is adopted in ZigBee wireless network. Modbus protocol is a general language applied to electronic controller. Through this protocol, communication between access control terminal and server can be realized, identity information of NFC can be authenticated and access control command can be sent.

Table 2. Frame format

Address	Function Code	Data Area	CRC Check
1Byte	1Byte	nByte	2Byte

The format of Modbus protocol frame is shown in Table 2, and the order of register transmission is from high to low. When the devices communicate with each other in the network, they need to know their device addresses.

4.3 Communication Encryption Algorithm

4.3.1 R_AES

R-AES is a simplified version of Advanced Encryption Standard AES. AES encryption algorithm [16] is a group encryption method, AES-128 algorithm is based on 128 bit length as the main standard to achieve encryption design. R-AES inherits the advantages of AES advanced encryption, considering the limited speed and resources of embedded system, on the basis of AES, firstly, the encryption steps that occupy large resources of embedded system are cut; at the same time, the number of encryption and decryption rounds is reduced. In the case of ZigBee security and encryption, the resource overhead of embedded system is reduced and the encryption speed is improved.

4.3.2 R_AES Workflow

AES algorithm [17] mainly consists of five steps: 1) Extended key; 2) Byte replacement ; 3) Row shift; 4) Mix Columns; 5) Round key addition, which has undergone 10 rounds of operation.

The R-AES algorithm is a simplified AES algorithm, and its workflow is shown in Figure 5. Compared with the traditional AES advanced encryption algorithm, it cuts out the column confusion that takes up a lot of resources in the algorithm steps. At the same time, the number of encryption rounds is reduced from the traditional 10 rounds to 3 rounds, which reduces the system resource occupation and greatly improves the running speed of the embedded system on the premise of ensuring the system security. The A-AES algorithm consists of four steps: extended key, byte replacement, row shift and round key addition.

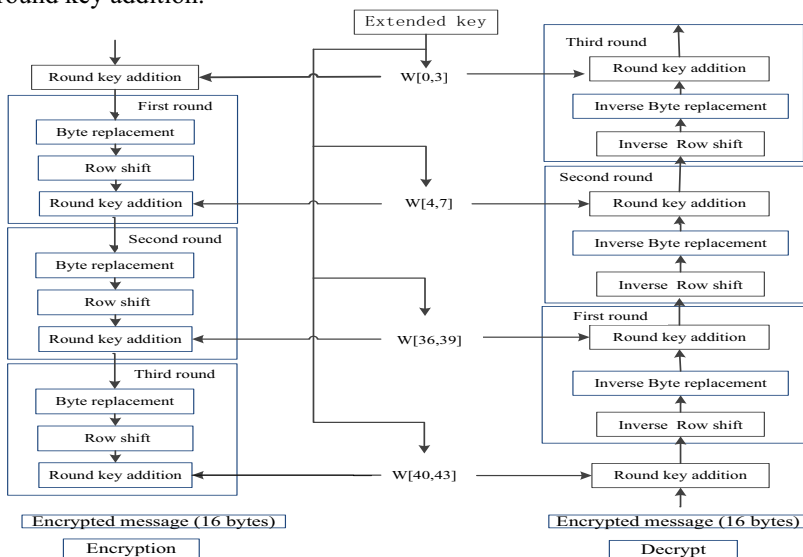


Figure 5. R-AES algorithm workflow

4.3.3 Application of R_AES

After passing the NFC demonstration, ZigBee wirelessly transmits the Modbus message, as shown in the door access command: 20050000ff008c3a, and encrypts and decrypts the data in the system as shown in Figure 6

```

The key is:
2b 7e 15 16 28 ae d2 a6 ab f7 15 88 09 cf 4f 3c
Message to be encrypted:
20 05 00 00 ff 00 8c 3a
Encrypted Message:
b7 e2 49 7e c1 cf d2 b5 c1 e6 10 00 c4 40 49 3c
Decrypted Message:
20 05 00 00 ff 00 8c 3a
    
```

Figure 6. Sample of message encryption and decryption data

R-AES key is: 0x2b, 0x7e, 0x15, 0x16, 0x28, 0xae, 0xd2, 0xa6, 0xab, 0xf7, 0x15, 0x88, 0x09, 0xcf, 0x4f, 0x3c. 8 bytes Messages is:0x20,0x05,0x00,0x00,0xff,0x00,0x8c,0x3a.Because AES adopts block encryption, 16 bytes are one group, and the Messages is insufficient to fill zero.The Messages after R-AES encryption is: 0xb7,0xe2,0x49,0x7e,0xc1,0xcf,0xd2,0xb5,0xc1,0xe6,0x10, 0x0,0xc4,0x40,0x49,0x3c.

It is almost impossible for attackers to decrypt ciphertext, so as to achieve the effect of encryption security.

5. System Testing

During the test, when the Android mobile terminal registered and authorized in the access control system is close to the NFC data reading area, the identity information is encrypted and transmitted to the server. After the server decrypts the information, it can query and confirm to open and close the corresponding door lock. Compared with the traditional IC card, NFC technology saves a lot of unnecessary troubles. At the same time, if the mobile intelligent terminal is lost or maliciously stolen by others, it can't be unlocked because it doesn't know the AES password of the terminal. In addition, it can't be unlocked if communication attack is used, which guarantees the security of the system. The physical test diagram of mobile terminal in access control system is shown in Figure 7.

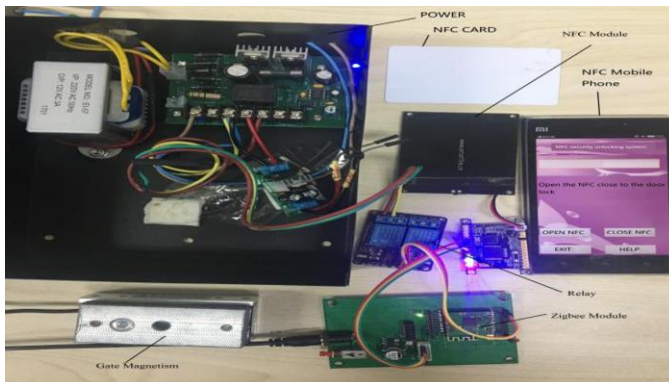


Figure 7. Physical map of mobile terminal in access control system

6. Conclusions

In this paper, the Ad-Hoc network access control system has been developed and applied to a large number of access control, The result is effective management of access control security. When the door needs to be opened, just take out the smart phone, close to the reading area, the door can be opened and closed; ZigBee wireless network transmits information to the server, which enables administrators to manage access control in batches more effectively; wireless transmission R-AES advanced encryption can ensure the safe transmission of information. Compared with other access control systems on the market, the system has the advantages of convenient operation, strong stability, safety and reliability. The smart access control system of Ad-hoc network for mobile NFC terminal has a high promotion value.

Acknowledgement

This work is supported by Natural Science Foundation Project of China (61976118), Key topics of the '13th five-year plan' for Education Science in Jiangsu Province (B-b/2020/01/18).

References

- [1] Ye Chen. Design of Intelligent Access Control System Based on RFID [J]. Information System Engineering, 2018 (06): 34-36.
- [2] Răzvan Andrei Gheorghiu, Iordache V, Minea M . Messaging capabilities of V2I networks[J]. Procedia Manufacturing, 2018, 22:476 - 484.
- [3] Pi Haitao, Jiao Huirong, Hao Kui et al. Design of bus IC card real-name system based on ZigBee [J]. Computer measurement and control, 2014, 22 (11): 3754-3756.
- [4] Sun Heng. Design and implementation of a new access control system based on NFC technology and cloud service [J]. Laboratory Research and Exploration, 2016,35(01): 114-120.
- [5] Li Manling. Design of digital intelligent access control system based on NFC [J]. Journal of Ezhou University, 2014, 21 (06): 108-110.
- [6] Jin Zhigang, Jie Bingshan. A highly secure access control system authentication protocol combining fingerprint identification and NFC technology [J]. Journal of Nankai University (Natural Science Edition), 2017, 50 (05): 1-7.
- [7] Ritrovati G , Maso-Gentile G D , Scavongelli C , et al. Active role of a NFC enabled smartphone in EV-EVSE charging process[C]// 2014 IEEE International Electric Vehicle Conference (IEVC). IEEE, 2019.
- [8] T B , Tran V T , Chung W Y . Pressure Measurement-Based Method for Battery-Free Food Monitoring Powered by NFC Energy Harvesting[J]. Scientific Reports, 2019, 9(1):17556.
- [9] Huang Ting, Liu Gang, Zhang zhiting, et al. Geological and mineral data acquisition system based on NFC and geoml [J]. Computer application and software, 2019 (9).
- [10] Chabbi S , Boudour R , Semchedine F , et al. Dynamic array PIN:A novel approach to secure NFC electronic payment between ATM and smartphone[J]. Information Security Journal A Global Perspective, 2020, 29(4):1-14.
- [11] Lu Kailang, Wang Huo, Zhao Ming. Design and Implementation of Intelligent Access Control System Based on NFC [J]. Electrical Applications, 2013, 32 (S1): 639-641.
- [12] Ran Junjun, Liu Zhiqin, Zhong Min. WiFi interference avoidance method based on channel state prediction in ZigBee networks [J]. Computer measurement and control, 2017, 25 (06): 124-127.
- [13] Li Manling. Design of digital intelligent access control system based on NFC [J]. Journal of Ezhou University, 2014, 21 (06): 108-110.
- [14] Wu Y, Liu K S, Stankovic J A, et al. Efficient Multichannel Communications in Wireless Sensor Networks[J]. Acm Transactions on Sensor Networks, 2016, 12(1):1-23.
- [15] Bao Yiqin, Xu Wenbin. Research and Design of Intelligent Network Switch Based on Human Infrared and Light Intensity Sensors [J]. Internet of Things Technology, 2018,8(05): 16-18+23.

- [16] Bi Ganbin. Research and application of encryption algorithm based on ZigBee [D]. Guizhou University, 2017:30-56.
- [17] Zhang Yao, Ye Ling. WSN Encryption Algorithms Based on AES [J]. Computer Engineering and Design, 2015, 36 (03): 619-623.

Joint User Clustering and Power Allocation for Vehicular Communications

Qian CHEN, Gongbin QIAN ¹, Chunlong HE, Rujun ZHAO and Yuping ZHENG
College of Information Engineering, Shenzhen University, Shenzhen, 518060, China

Abstract. Device-to-device (D2D) communication has emerged as a promising concept for supporting the vehicular networks, which can efficient and reliable enhance cellular network. In this paper, we discuss two different design criteria for vehicular networks. They are maximizing overall vehicle-to-infrastructure (V2I) link throughput while guaranteeing the minimum reliability for each vehicle-to-vehicle (V2V) link, and maximizing the minimum throughput of all V2V links under the constraints of minimum V2I link throughput requirements. Because both of these problems is an mixed integer non-linear programming problem, we solve these problems in two steps, i.e., by first clustering D2D users into clusters and then optimizing their respective power allocations. Specifically, we first propose a spectral clustering (SC) method for D2D users clustering. Then, two power allocation algorithms are developed to maximize the sum V2I link throughput and maximize the minimum V2V link throughput, respectively. The effectiveness of proposed resource allocation algorithms is validated by computer simulation.

Keywords. Vehicular communications, D2D communication, resource allocation

1. Introduction

In recent years, the vehicular networks have attracted more and more attention with the rapid development of the Intelligent Transport System (ITS) [1] [2]. As a potential application scenario of the fifth generation of mobile communication systems (5G), vehicular communication networks require large bandwidth, low latency and high reliability [3].

Vehicle communication networks is considered to support three applications, namely road safety, traffic efficiency and infotainment. Infotainment and traffic efficiency applications usually are based on vehicle-to-infrastructure (V2I) communications with large bandwidth requirement. Meanwhile, the road safety application is based on vehicle-to-vehicle (V2V) communications with high reliability and low delay [4] [5]. To achieve this goal, the device-to-device (D2D) communication assisted cellular networks is considering to be a feasible solution for vehicular communication networks.

Resource allocation for vehicular communication networks have been extensively researched. In [6], an efficient resource allocation scheme is studied for D2D-enabled vehicular networks. Indeed, the presented algorithms results in improved system performance, however, each resource block is used by only one CUE and DUE. Similarly, a

¹Corresponding Author: College of Information Engineering, Shenzhen University, Shenzhen, 518060, China; E-mail: qiangb@szu.edu.cn.

robust resource allocation algorithm has been proposed in [7] [8] to improve the system performance of vehicular network. By exploit graph partitioning algorithm, the authors in [9] has proposed an efficient resource allocation scheme ,when considering the multiple DUE share the same resource block (RB) with one CUE. Further, in [10], a reinforcement learning algorithm for the D2D-based vehicular networks has been developed.

In this paper, we will investigate two different design criteria for vehicular services with different needs, including maximizing the overall V2I links throughput and maximizing the minimum throughput of all V2V links. Firstly, we adjust those V2V links that cause more harmful interference to other disjoint spectrum sharing clusters by exploit spectral clustering (SC) algorithm. Then, in one scenario of infotainment applications (e.g., movie/video sharing), we will consider maximizing overall V2I link throughput while guaranteeing the minimum reliability for each V2V link. For this criterion, we can derive a power allocation algorithm to maximize the capacity of each cluster by utilizing the monotonicity of the objective function as same as [9]. In other scenario of security applications, we will consider maximizing the minimum throughput of all V2V links under the constraints of minimum V2I link throughput requirements. For this criterion, we will introducing an auxiliary variable and a sequential convex approximation based iterative method to perform power allocation.

The rest of this paper is organized as follows. In Section 2, the vehicular communication networks model is given. In Section 3, we get V2V clustering result by using a simple algorithm. Section 4 and 5 investigate the power allocation problems with different optimization problem. Section 6, the numerical results and discussion are provided. We conclusion the paper in Section 7.

2. System Model

This paper considers a D2D-based vehicular communication networks, as shown in Fig.1. The considered networks consist of M V2I high-capacity communication for some infotainment services and K V2V high-reliability communication by using D2D communications fashion for some security services, and $K \gg M$. For notational simplicity, we denote the set of V2I links and V2V links by \mathcal{M} and \mathcal{K} respectively, i.e., $\mathcal{M} \triangleq \{1, 2, \dots, M\}$ and $\mathcal{K} \triangleq \{1, 2, \dots, K\}$. We consider the case that each V2I channel is reused by multiple V2V users, and V2I links using orthogonal frequency division multiple access (OFDMA) technology to access the cellular networks. Without loss of generality, we assume that both the BS and vehicular are equipped with single antenna and the perfect channel state information(CSI) is available at both the BS and vehicular. The signal to interference plus noise ratio (SINR) of the m th V2I link can be expressed

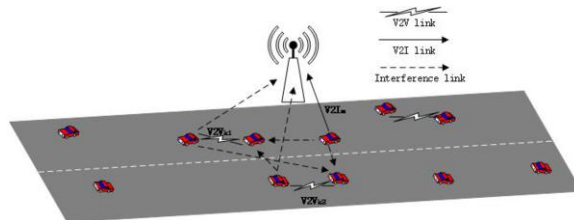


Figure 1. The considered D2D-based vehicular communication model

as

$$\gamma_m^c = \frac{P_m^c g_{m,B}}{\sigma^2 + \sum_{k=1}^K \rho_{m,k} P_k^d g_{k,B}}, \tag{1}$$

where P_m^c and P_k^d denotes the transmit power of the m th V2I transmitter and the k th V2V transmitter, respectively. $g_{m,B}$ is the channel gain between the k th V2V transmitter and the BS . $\rho_{m,k}$ is a binary parameter, $\rho_{m,k} = 1$, means that the k th V2V pair reuses the k th V2I’s channel; otherwise, $\rho_{m,k} = 0$. σ^2 represents the power of complex additive white Gaussian noise(AWGN) of both V2I links and V2V links. Similarly, the SINR of the k th V2V link can be written as

$$\gamma_k^d = \frac{P_k^d g_k}{\sigma^2 + \sum_{m=1}^M \rho_{m,k} P_m^c g_{m,k} + \sum_{k'=1, k' \neq k}^K \rho_{k',k} P_{k'}^d g_{k',k}}, \tag{2}$$

where g_k is the channel gain between the k th V2V transmitter and the k th V2V receiver. $g_{m,k}$ is the channel gain between the k th V2I transmitter and the k th V2V receiver. $g_{k',k}$ is the channel gain between the k' th V2V transmitter and the k th V2V receiver.

3. V2V Clustering

In order to solve the strong interference that is caused by multiple V2V links sharing the same V2I links channel, we need to design appropriate channel reuse strategies to assign all V2V links to different clusters. Based on minimizing the mutual interference criterion, we can design the following V2V clustering scheme.

First, we transform the V2V clustering problem into a graph partition problem, where each vehicle and channel are interpreted as the vertices and weighted edges of the graph, specifically, the problem is a NP-hard problem. Obviously, in a graph, all the sum of the edge weights is a fixed value, the cut target can be regard as the weight in all clusters that minimizes the sum, i.e., maximizing the sum of the weights between the clusters. Next, we will use spectral clustering (SC) algorithm to solve this problem.

Spectral clustering is an algorithm evolved from graph theory with small computational complexity and simple implementation. A simple spectral clustering algorithm has been used in [11]. For the proposed problem, we try to make all the reciprocal of the channel gain of all channels as the edges of graph. The goal of our algorithm is to assign K users to n clusters, C_1, \dots, C_n , in this paper, $n = M$. The V2V clustering algorithm is shown in the Table 1.

Specifically, $W = [1/w_{k',k}]$ denotes the similarity matrix of the proposed graph, $w_{k',k}$ is the weight of the graph’s edge, which is define the reciprocal of channel power gain $g_{k,k'}$. D is denoted as a diagonal degree matrix with $d_{k,k} = \sum_{k'=1}^K w_{k,k'}$.

Table 1. V2V Clustering Algorithm

Algorithm 1: SC for V2V Clustering

- 1: Input: The information of channel gains between DUE $w_{k',k}$.
- 2: Construct the undirected graph G and get the similarity matrix W and the diagonal degree matrix D .
- 3: Calculate the Laplacian matrix $L = D - W$.
- 4: Calculate the eigenvectors corresponding to all the eigenvalues of the matrix L , and take the eigenvectors corresponding to the first K eigenvalues.
- 5: Clustering eigenvectors with k-means algorithm.
- 6: Output: DUE partition result C_n .

4. Maximum Sum of V2I Links Capacity Optimization

In this section, we consider maximizing the sum capacity of V2I links under the constraints of satisfying the requirements of the minimum V2I transmit rate, maximum transmit power of each vehicle, and guaranteeing the minimum reliability for each V2V links. In addition, we transfer the reliability of each V2V links to the probability of outage events, i.e. V2V received SINR is below a predetermined threshold γ_0^d . The problem can be written as

$$\max_{P_m^c, P_k^d} \sum_{m=1}^M \log_2 \left(1 + \frac{P_m^c g_{m,B}}{\sigma^2 + \sum_{k \in C_m} P_k^d g_{k,B}} \right) \tag{3}$$

$$\text{s.t.} \quad 0 \leq P_m^c \leq P_{\max}^c, \tag{3a}$$

$$0 \leq P_k^d \leq P_{\max}^d, \tag{3b}$$

$$P_r \left\{ \frac{P_k^d g_k}{\sigma^2 + P_m^c g_{m,k} + \sum_{k'=1, k' \neq k}^K P_{k'}^d g_{k',k}} \leq \gamma_0^d \right\} \leq p_0. \tag{3c}$$

In the problem (3), constraint (3a) and (3b) represent V2I link's and V2V link's maximum power limit, where P_{\max}^c and P_{\max}^d are the maximum transmit power of each V2I transmitter and V2V transmitter, respectively. Constraint (3c) represent the reliability requirements for each V2V link, R_{\min}^c and γ_0^d denote the minimum capacity needed by V2I link and the minimum SINR requirement of V2V link. $P_r \{ \cdot \}$ evaluates the probability of the input and p_0 is the tolerable outage probability for V2V links.

From [9], the constraint (3d) can be transformed into definite boundaries. From the previous discussion, we have get a V2V clustering result, next, we want to find the maximum capacity of all possible sharing patterns of V2I and V2V clustering. So the problem can be transformed as

$$\max_{\mathbf{P}_m^c, \mathbf{P}_k^d} \sum_{m=1}^M \log_2 \left(1 + \frac{P_m^c g_{m,B}}{\sigma^2 + \sum_{k \in C_m} P_k^d g_{k,B}} \right) \quad (4)$$

s.t. The constraints in 3(a), 3(b), (4a)

$$\frac{P_k^d g_k}{\sigma^2 + P_m^c g_{m,k} + \sum_{k'=1, k' \neq k}^K P_{k'}^d g_{k',k}} \geq \frac{\gamma_0^d}{-\ln(1-p_0)}. \quad (4b)$$

Notice that the objective function (4) is monotonically nonincreasing in terms of P_k^d , from the (4b), the optimal P_k^d in the n th cluster must be achieved at

$$\tilde{P}_k^d = \frac{\tilde{\gamma}_0}{g_k} (\sigma^2 + P_m^c g_{m,k} + \sum_{k'=1, k' \neq k}^K P_{k'}^d g_{k',k}), \quad (5)$$

where $\tilde{\gamma}_0 = -\frac{\gamma_0^d}{\ln(1-p_0)}$. By substituting \tilde{P}_k^d into (5), we can obtain the optimal \tilde{P}_m^c as

$$\tilde{P}_m^c = \left\{ P_{max}^c, \left\{ \frac{P_{max}^d g_k - (\sigma^2 + \sum_{k'=1, k' \neq k}^K P_{max}^d g_{k',k}) \tilde{\gamma}_0}{g_{m,k} \tilde{\gamma}_0} \right\} \right\} \quad (6)$$

Next, we can find the best spectrum multiplexing partner between the V2I link and V2V link cluster sets using the Hungarian method [6] by constructing a bipartite graph.

The developed power allocation algorithm is showed in Table 2. For convenience, sum of V2I links capacity can be written as $R_{m,n}^c$.

Table 2. Maximum sum V2I link capacity algorithm

Algorithm 2: Maximum sum V2I link capacity

- 1: **for** $m=1:M$ **do**
 - 2: **for** $n=1:M$ **do**
 - 3: Obtain the optimal power allocation (P_m^c, P_k^d) from (6) and (7) for the single CUE-DUE cluster.
 - 4: Calculate $R_{m,n}^c$ according to equations (4).
 - 5: **end for**
 - 6: **end for**
 - 7: Use the Hungarian method to find the optimal reuse pattern ρ_m^* based on $R_{m,n}^c$.
 - 8: Return the optimal spectrum reuse pattern $\rho_{m,k}^*$ and the corresponding power allocation (P_m^c, P_k^d).
-

5. Maximum the Minimum V2V Link Capacity Optimization

The maximizing the sum V2V link capacity problem in section IV aims to ensure V2I link’s large capacity services. However, in many security applications, a balance connectivity for V2V links rather than a maximized sum V2V capacity, is more desirable. In this section, we mainly study maximizing the minimum achievable rate among all the V2V links while guaranteeing the minimum V2I link throughput requirements. The problem can be formulated as

$$\max_{\mathbf{P}_m^c, \mathbf{P}_k^d} \min_{k \in \mathcal{K}} \log_2 \left(1 + \frac{P_k^d g_k}{\sigma^2 + P_m^c g_{m,k} + \sum_{k'=1, k' \neq k}^K P_{k'}^d g_{k',k}} \right) \tag{7}$$

s.t. The constraints in (3a), 3(b). (7a)

$$\log_2 \left(1 + \frac{P_m^c g_{m,B}}{\sigma^2 + \sum_{k \in C_m} P_k^d g_{k,B}} \right) \geq R_{\min}^c. \tag{7b}$$

Constraint (7b) represents the minimum throughput requirement for V2I link. In general, problem (7) is a generalized fractional linear programming problem, which can be solved by parametric linear programming [12]. By introduce an auxiliary variable τ and add a constraint that all V2V links not less than τ , we can transform (7) into the standard form of convex optimization problem with inequality constraints as follow

$$\max_{\mathbf{P}_m^c, \mathbf{P}_k^d} \tau \tag{8}$$

s.t. The constraints in (3a), 3(b), 7(b), (8a)

$$\log_2 \left(1 + \frac{P_k^d g_k}{\sigma^2 + P_m^c g_{m,k} + \sum_{k'=1, k' \neq k}^K P_{k'}^d g_{k',k}} \right) \geq \tau. \tag{8b}$$

According to [7], the original problem (7) can be solved through solving the equivalent problem (8) by using a sequential convex approximation based iterative algorithm. Then, we can using the Hungarian method to find optimal spectrum reuse scheme.

6. Simulation Results

In this section, we compare different resource allocation criteria through simulation results. We follow the simulation setup for the freeway case in [9], and assume that $K = 10$ and $M = 30$. Firstly, we compare the performance of our proposed SC algorithm with

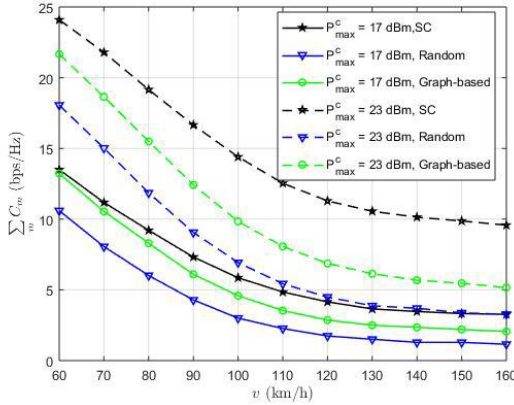


Figure 2. Sum capacity of V2I links.

the Graph-Based method in [13] and a random partitioning algorithm. And then we compare the three different power allocation methods.

Fig.2 compares the sum of V2I link throughput versus the vehicle speed when using different clustering algorithms. From Fig.2, we can obtain that the sum of V2I link throughput is decreased with the growth of vehicles speed. This is because higher speed will make the power limit of V2I link more stringent and reduce the V2I link transmission power. Fig.2 also shows the impact of different V2V clustering methods. It can be seen from Fig.2, compared with Graph-Based and random partitioning methods, the SC method has a better performance. The reason is that the SC method can reduce the intra-cluster interference.

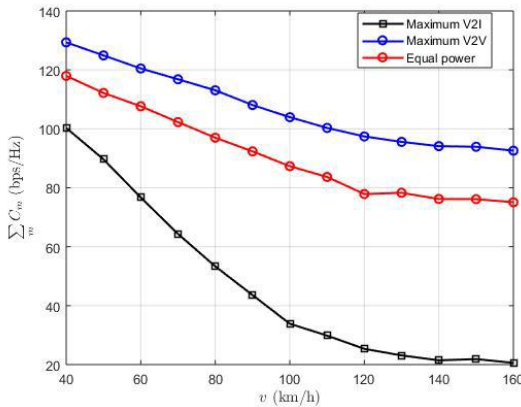


Figure 3. Minimum capacity of V2V links.

Fig.3 compares the minimum throughput of V2V link among maximizing the V2V link minimum throughput scheme, maximum sum V2I link throughput scheme and equal power allocation scheme. Compared with maximum sum V2I link throughput and equal power allocation, it is observed in Fig.3 that the proposed scheme outperforms the other

two allocation schemes in terms of minimum throughput. This illustrates the necessity of choosing the right power allocation method for different vehicular applications. Specifically, we can use maximum sum V2I link throughput scheme in infotainment and traffic efficiency applications, and using maximizing the V2V link minimum throughput scheme in safety applications.

7. Conclusion

In this paper, we have compared two different design criteria for vehicular networks. We first exploit the SC method to take D2D user into different cluster, and then two optimal power allocation schemes for maximizing the overall V2I link throughput while guaranteeing the minimum reliability for each V2V link, and maximizing the minimum throughput of all V2V links under the constraints of minimum V2I link throughput requirements, respectively. Simulation results showed effectiveness of proposed schemes.

However, because of the computational complexity of solving power allocation problem, the proposed resource allocation is difficult to implement. Future works will focus on developing low complexity resource allocation algorithms, such as machine learning method.

Acknowledgement

This work was supported in part by Shenzhen Overseas High-level Talents Innovation and Entrepreneurship under Grant KQJSCX20180328093835762, in part by the Shenzhen Basic Research Program under Grants JCYJ20190808122409660 and JCYJ20170412104656685.

References

- [1] G. S. Khekare and A. V. Sakhare, "A smart city framework for intelligent traffic system using vanet," in *International Multi-conference on Automation*, 2013, pp. 302–305.
- [2] G. S. Khekare, "Design of emergency system for intelligent traffic system using vanet," in *International Conference on Information Communication & Embedded Systems*, 2014.
- [3] G. Araniti, C. Campolo, M. Condoluci, A. Iera, and A. Molinaro, "LTE for vehicular networking: a survey," *IEEE Communications Magazine*, vol. 51, no. 5, pp. 148–157, May 2013.
- [4] L. Liang, J. Kim, S. C. Jha, K. Sivanesan, and G. Y. Li, "Spectrum and power allocation for vehicular communications with delayed CSI feedback," *IEEE Wireless Communications Letters*, vol. 6, no. 4, pp. 458–461, Aug. 2017.
- [5] L. Liang, H. Peng, G. Y. Li, and X. Shen, "Vehicular communications: A physical layer perspective," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 12, pp. 10 647–10 659, Dec. 2017.
- [6] L. Liang, G. Y. Li, and W. Xu, "Resource allocation for D2D-Enabled vehicular communications," *IEEE Transactions on Communications*, vol. 65, no. 7, pp. 3186–3197, Apr. 2017.
- [7] K. Doppler, M. Rinne, C. Wijting, C. B. Ribeiro, and K. Hugl, "Device-to-device communication as an underlay to LTE-advanced networks," *IEEE Communications Magazine*, vol. 47, no. 12, pp. 42–49, Dec. 2009.
- [8] S. Bharati and P. Podder, "Adaptive PAPR reduction scheme for OFDM using SLM with the fusion of proposed clipping and filtering technique in order to diminish PAPR and signal distortion," *Wireless Personal Communications*, no. 1, pp. 1–18, 2020.

- [9] L. Liang, S. Xie, G. Y. Li, Z. Ding, and X. Yu, "Graph-based resource sharing in vehicular communication," *IEEE Transactions on Wireless Communications*, vol. 17, no. 7, pp. 4579–4592, Apr. 2018.
- [10] G. Khekare, P. Verma, U. Dhanre, S. Raut, and S. Sheikh, "The optimal path finding algorithm based on reinforcement learning," *The optimal path finding algorithm based on reinforcement learning*, vol. 12, 2020.
- [11] B. Golkar and E. Sousa, "Adaptive partitioned interference management in cellular networks," in *in Proc. 2012 IEEE GLOBECOM*. Anaheim, CA, USA: IEEE, Dec. 2012, pp. 4261–4266.
- [12] Z. Luo and S. Zhang, "Dynamic spectrum management: Complexity and duality," *IEEE Journal of Selected Topics in Signal Processing*, vol. 2, no. 1, pp. 57–73, Feb. 2008.
- [13] R. Y. Chang, Z. Tao, J. Zhang, and C. J. Kuo, "Multicell ofdma downlink resource allocation using a graphic framework," *IEEE Transactions on Vehicular Technology*, vol. 58, no. 7, pp. 3494–3507, 2009.

Study of Work-Function Variation on Performance of Dual-Metal Gate Fin Field-Effect Transistor

Tianyu YU, Liang DAI, Zhifeng ZHAO, Weifeng LYU¹, and Mi LIN

School of Electronics and Information, Hangzhou Dianzi University, Hangzhou 310018

Abstract. The impact of work-function variation (WFV) on performance of an inversion-mode (IM) dual-metal gate (DMG) fin field-effect transistor (FinFET) was investigated for the first time. The statistical fluctuations induced by WFV on the threshold-voltage (V_{TH}), transconductance (g_m), and subthreshold slope (SS) were demonstrated and estimated utilizing a 3D technology computer-aided design (TCAD) simulator. We found that the performance variations of the DMG FinFET were affected by two different metals near the drain and near the source, respectively. Additionally, this effect of the two metals on the channel was not monotonic with the length of the channel of their own control. Our work fills a gap in the study of WFV for a DMG IM FinFET and provides a reference for optimizing the distribution of the two metals.

Keywords. Work-function variation; fin field-effect transistor; dual-metal gate; process variation

1. Introduction

With the continuous scaling of a semiconductor device size, the short channel effects (SCEs) and gate transport inefficiency become more and more serious; this means that the gate-to-channel control becomes worse and worse [1–3]. To counter these undesired barriers, high-k materials (such as HfO_2) replaced the traditional SiO_2 as the oxide layer material at sub-45 nm technology node to avoid the occurrence of gate leakage current being larger than the subthreshold leakage [1]. In addition, a metal gate has substituted for polysilicon because of the incompatibility between polysilicon and high-k materials. However, the use of metal gate inevitably introduces work-function variation (WFV). This is because the work-function (WF) of each metal grain depends on its own crystalline orientations and is difficult to control during metal gate deposition [4–7].

Meanwhile, a FinFET with silicon fin as one of the most promising candidates has attracted much attention in recent years; this is because the fin-based devices have high carrier mobility and a great control over the carriers in the channel, and are excellent in suppressing SCEs [8–10]. To further improve the FinFET's performance, the structure of the dual-metal gate (DMG) has been proposed for advanced devices [11–12]. Instead of the single-metal gate (SMG), dual-metal gates (DMG; i.e., Metal gate 1 'M1', near

¹ Corresponding Author. Email: lvwf@hdu.edu.cn

the source, and Metal gate 2 ‘M2’, near the drain) with different WFs are used on the gate stack. According to related studies, DMG can not only suppress SCEs and simultaneous transconductance (g_m) improvement, but can also improve current drive capability compared with the SMG structure due to different WFs in the metal gate [3, 10–16]. Meanwhile, different manufacturing methods for DMG structure have also been studied, and the new DMG structure requires only additional processing steps to laterally form two well-controlled contacting metal materials [12, 17].

In nanoscale integrated devices, WFV as one of major random variation sources has been of enormous interest for high-K/metal gate technology [18–21]. The effect of WFV on the SMG has already been well reported. However, knowledge of its influence on the DMG is still missing. Therefore, we investigate, for the first time, the impact of WFV on the performance of DMG FinFET. The different ratios of the length of the M1, also named as control gate to that of the total gate on device performance fluctuation caused by WFV are also explored. Therefore, we study to provide a reference for optimizing the distribution of the two metals in the design of FinFETs.

2. Device Structure and Simulations

Fig. 1 shows a bird’s eye view of a DMG IM FinFET, which is used in the present study. The detailed structural information for the DMG FinFET is as follows: a fin height of 16 nm, a channel length of 20 nm, a fin width of 5 nm, and a gate oxide thickness of 2 nm. In addition, a uniform donor doping concentration of $1 \times 10^{19} \text{ cm}^{-3}$ in the drain and source regions is used, and $1 \times 10^{16} \text{ cm}^{-3}$ acceptor particles are doped in the channel. TiN and TaN are chosen as the metals of M1 and M2, respectively. According to requirements of DMG, the WF of M1 is bigger than that of M2. As shown in Fig. 1, it illustrates three different WFs of TaN (4.0 eV, 4.15 eV, and 4.8 eV) depending on three possible grain orientations ($\langle 100 \rangle$, $\langle 200 \rangle$, and $\langle 220 \rangle$) with different probabilities of occurrences (50%, 30%, and 20%) and two possible WFs of TiN (4.6 eV and 4.4 eV), owing to two different grain orientations ($\langle 100 \rangle$ and $\langle 111 \rangle$) with two probabilities of occurrences (60% and 40%) [1]. It is noted that TiN is a well-known gate metal used in an SMG n-channel device, and TaN is not suitable as a metal gate for the SMG n-channel device. The length of the channel under control gate M1 is L_1 ; while the length of the channel under screen gate M2 is L_2 ; and the total gate length is $L = 20$ nm. Fig. 2 gives a cross section along the channel of the studied IM DMG FinFET.

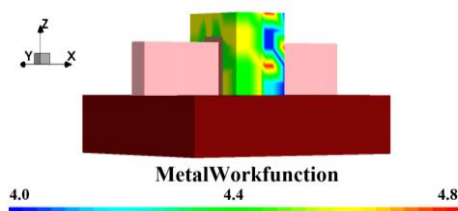


Fig. 1 Schematic diagram of the simulation of WFV of DMG FinFET.

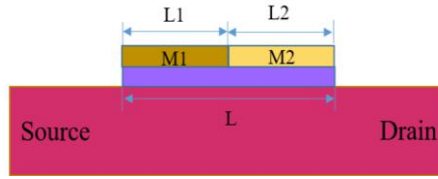
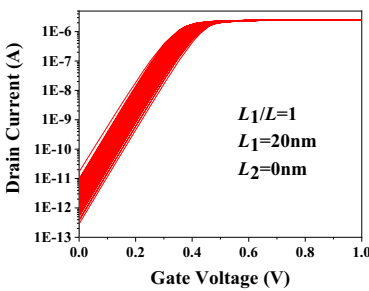


Fig. 2 Cross section along the y -axis at $y = 0$. The gate near the source is also named as the “control gate”, and the gate near the drain is also named as the “screen gate.”

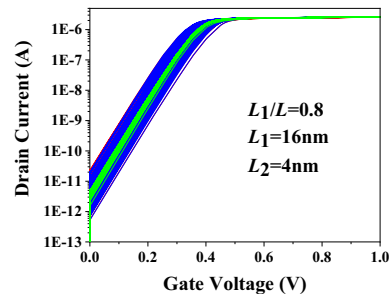
To accurately investigate the impact of the ratio of the length of different control gates on the performance fluctuation caused by WFV in the IM DMG FinFET, we carefully let L_1/L be 0, 0.2, 0.4, 0.6, 0.8 and 1. It is worth mentioning that the device can be considered to be an SMG with TaN and TiN only as gate stack at $L_1/L = 0$ and 1, respectively.

3. Results and Discussions

The impact of WFV on the electrostatic integrity of the DMG IM FinFET is shown by the dispersion of the transfer curve in Figs. 3 (a)–(f). The figures give input characteristic curves of drain current against gate voltage (V_{GS}) at drain voltage (V_{DS}) = 50 mV for each case. In these figures, the red curves represent the WFV-induced variation of performance in the case of S1; the green ones represent the case of S2; and the blue ones represent the S3 situation. As shown by the area occupied by the discrete curves that is getting larger and larger, WFV-induced electrostatic integrity fluctuations in the IM DMG FinFET becomes more serious, which is at L_1/L from 0.8 to 0.2. This means that the value of L_1/L has a bad effect on the WFV-induced variability in the DMG IM FinFET. More especially, when L_1/L is from 1 to 0.2 in steps of -0.2, the green and red curves occupy an increasing area significantly; this means that drain current’s dispersion is getting worse. Furthermore, the standard deviation of V_{TH} , the subthreshold slope (SS), the saturation current (I_{sat}), and g_m were calculated for each case. To better analyze the experimental results, the relative standard deviations of the relevant parameters of the DMG IM FinFET were also calculated.



(a)



(b)

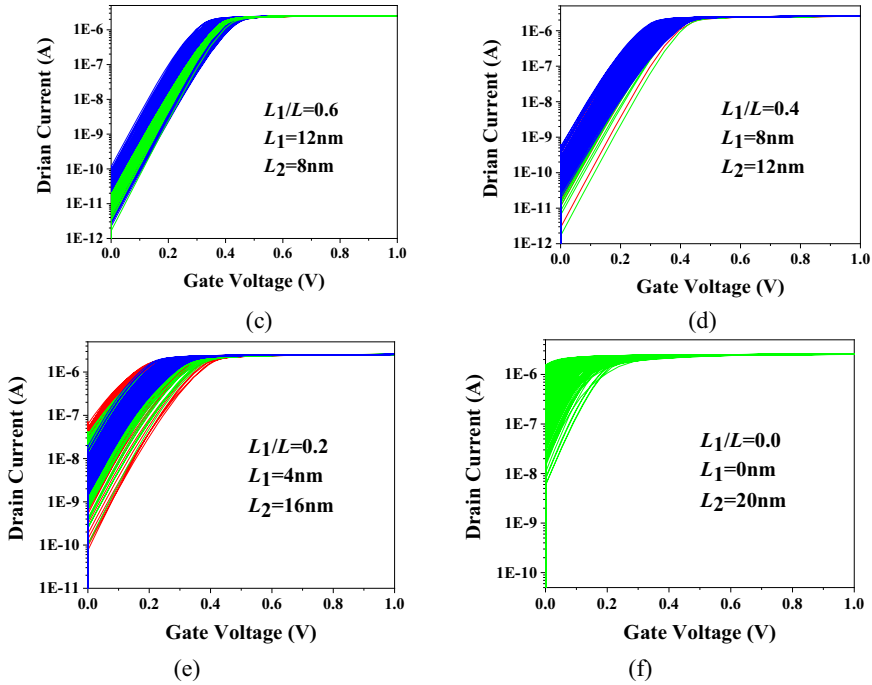


Fig. 3 WFV-induced variations of drain current versus V_{GS} for an IM DMG FinFET.

The V_{TH} has been extracted by the constant-current method at a fixed current of $0.1 \mu A/\mu m$. These values are shown in Figs. 4, 5, and 6. In these figures, it should be noted that off-state current is greater than $0.1 \mu A/\mu m$ at $L_1=0$ [22]. Therefore, V_{TH} is defined as 0 V, and the missing value of SS. Therefore, our study verifies that TaN cannot be used alone as the metal gate for an n-channel FinFET. In the Fig. 4 (a), as the value of L_1/L keeps getting smaller, the WFV-induced variability becomes worse as the average value of V_{TH} becomes smaller. In cases S1, S2 and S3, when L_1/L is from 0 to 1, the average values of V_{TH} are almost equal (see the solid line in Fig. 4 (a)). At the same time, when L_1/L is from 1 to 0.4, in the case of S1, the V_{TH} fluctuation caused by WFV is similar to that in case S3, but the V_{TH} fluctuation caused by case S1 is similar to that by case S2 (see the red circle in Fig. 4 (b)) when L_1/L is less than 0.4. At the same time, it is interesting that when L_1/L is between 0.4 and 1, the effect of WFV on the V_{TH} in case S1 is smaller than that in case S3. This indicates that the WFVs of M1 and M2 are not independent of each other; on the contrary, for the same L_1/L , the WFV of M2 has a certain inhibitory effect on the WFV of M1.

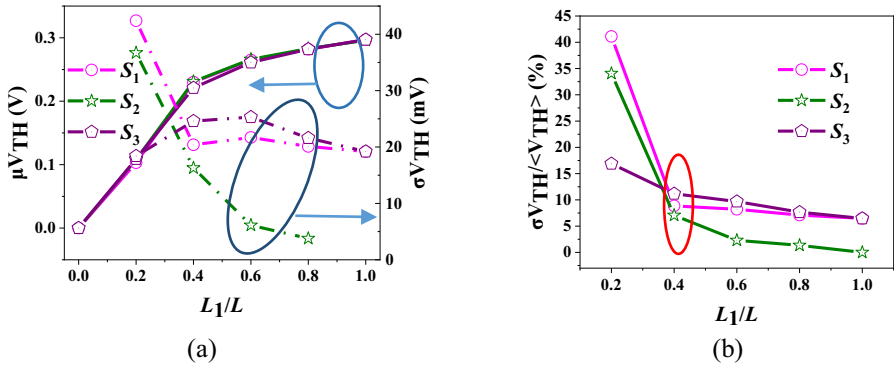


Fig. 4 The average and standard deviation of V_{TH} variations caused by WFV at different values of L_1/L .

The characteristics of g_m in Figs. 5 (a) and (b) for DMF FinFET with different L_1/L values are shown. Similar to the case of V_{TH} , when L_1/L is from 1 to 0.4 in steps of -0.2, the g_m fluctuation caused by WFV in the case of S1 is similar to that in the case of S3, but when L_1/L is less than 0.4, the g_m fluctuation caused by case S1 is similar to that by case S2 (see the red circle in Fig. 5 (b)). Moreover, the g_m characteristic of the device no longer simply increased or decreased as L_1/L decreases; when L_1/L is larger than 0.4, the g_m value of DMG is generally greater than the g_m value of SMG with TiN as the MG, which means that the ability of g_m for DMG IM FinFET is improved at this time, and g_m reaches the maximum at $L_1/L = 0.8$. Moreover, when L_1/L is greater than 0.4, the fluctuation of g_m decreases slightly as L_1/L decreases.

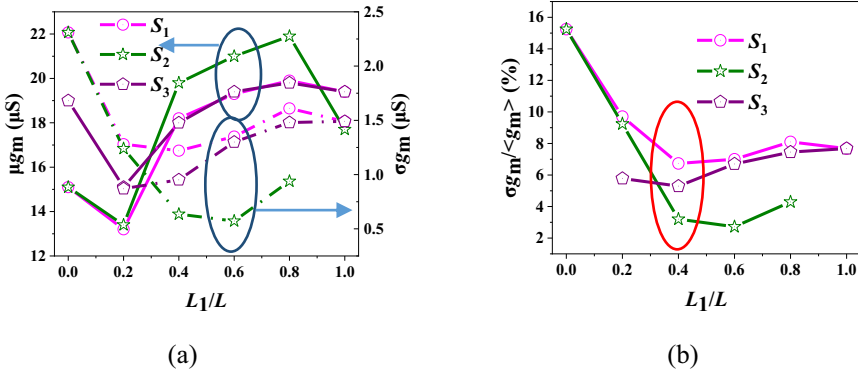


Fig. 5 Average and standard deviation of g_m variations caused by WFV at different values of L_1/L .

The SS characteristics in Figs. 6 (a) and (b) of the DMG FinFET are shown in steps of 0.2 at L_1/L from 0 to 1. It is observed that the DMG FinFET has a bigger SS compared to a corresponding SMG FinFET with TiN. It can be attributed to the SS for a transistor is inversely proportional to its effective length [23], and the effective length of the DMG device is only slightly bigger than that of L_1 in the subthreshold region. By contrast, the effective length of the SMG device is slightly bigger than the total channel length [3]. Inevitably, since we set L_1/L from 0 to 1 in the present study, L_1 is less than or equal to L . The effective length decreases with a decrease in L_1/L or L_1 . Therefore, the SS of the DMG FinFET device increases gradually (see the solid line in Fig. 5 (a)).

When L_1/L is from 1 to 0.6 in steps -0.2, in case S1, the WFV-induced SS fluctuation is similar to that in case S3, but when L_1/L is smaller than 0.6, the SS variability caused by the case S1 is similar to that of case S2 (see the red circle in Fig. 6 (b)). When L_1/L is from 0.4 to 1, the rise and SS variability is not large, but when L_1/L is less than 0.4, the value and fluctuation of SS increase sharply.

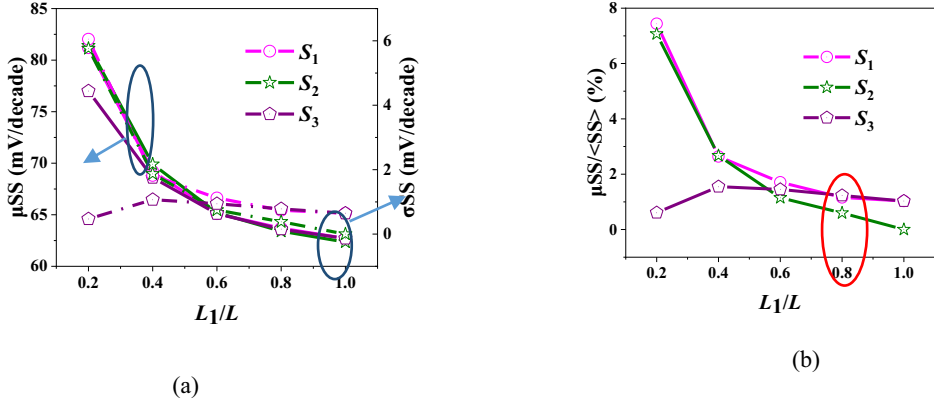


Fig. 6 The average and standard deviation of SS variations caused by WFV at different values of L_1/L .

Generally, although the WFV of M1 has a strong impact on the performance fluctuation for the DMG FinFET, this effect is weakened as L_1 decreases. Meanwhile, the effect of M2, which is relatively weak in causing device parameter variations for the DMG FinFET, gradually becomes stronger when the length L_2 increases. Even when L_1/L is less than 0.4, the WFV-induced device performance fluctuations by MG of M2 gradually dominate the device performance fluctuation caused by the entire MG WFV. Additionally, when L_1/L is less than 0.4, the influence of TaN's WFV on the device begins to dominate. Therefore, for the DMG FinFET, it is necessary to ensure that L_1/L is greater than a certain amount. Furthermore, at this time, because the metal work function of M1 is dominant, the channel region under M2 has greater optimization freedom. For example, the substrate doping can be reduced in this region, thereby increasing the benefit that source and drain capacitance can be reduced, while potentially increasing the speed of the device [12]. In addition, some studies have considered $L_1/L = 0.5$ to be a suitable amount when they do not take the effect of WFV into account. However, our observations indicate that due to the WFV, the amount of L_1/L should be slightly bigger than 0.5 when WFV is considered. Also, it can reduce the performance variation caused by misalignment of the two gates.

4. Conclusion

For the first time, the impact of WFV-induced performance variability in DMG device has been explored in this study. It can be observed that the WFV from the M1 metal gate near the source is the main factor causing the device performance fluctuation. Moreover, simulation results show that the fluctuation caused by WFV gradually becomes serious after a smaller WF metal gate is introduced near the drain. In addition, the performance fluctuations of the DMG FinFET caused by WFV are not monotonic by the change of L_1/L . Rather, as L_1/L decreases, the fluctuation of device performance

rises dramatically. From the perspective of the WFV, the introduction of a second metal in the gate has an obvious impact on the device performance of the DMG FinFET as L_1/L is smaller than 0.5. Therefore, increasing the value of L_1/L appropriately is reasonable.

Acknowledgements

This work is supported in part by Zhejiang Provincial Natural Science Foundation of China (grant LY18F040005), and in part by National Natural Science Foundation of China (grant 61571171).

References

- [1] H F Dadgour, K Endo, V K De, and K Banerjee. Grain-Orientation Induced Work Function Variation in Nanoscale Metal-Gate Transistors—Part I: Modeling, Analysis, and Experimental Validation. *IEEE Transactions on Electron Devices*, **57** (2010), 2504-2514.
- [2] L Dai, W F Lü, and M Lin. Random Dopant Fluctuation-Induced Variability in n-Type Junctionless Dual-Metal Gate FinFETs[J]. *Electronics*, **8**(2019), 282.
- [3] H Lou, L Zhang, Y Zhu, X Lin, S Yang, J He, and M Chan. A Junctionless Nanowire Transistor With a Dual-Material Gate. *IEEE Transactions on Electron Devices*, **59** (2012), 1829-1836.
- [4] G Indalecio, A J Garcia-Loureiro, N Seoane Iglesias, and K Kalna. Study of Metal-Gate Work-Function Variation Using Voronoi Cells: Comparison of Rayleigh and Gamma Distributions. *IEEE Transactions on Electron Devices*, **63**(2016), 2625-2628.
- [5] S Migita, Y Watanabe, H Ota, H Ito, Y Kamimuta, T Nabatame, and A Toriumi. Design and demonstration of very high-k ($k \sim 50$) HfO₂ for ultra-scaled Si CMOS. *Symposium on VLSI Technology*, 2008, 152-153.
- [6] K Endo, S I O'Uchi, Y Ishikawa, Y Liu, T Matsukawa, K Sakamoto, J Tsukada, H Yamachi, and M Masahara. Variability Analysis of TiN Metal-Gate FinFETs. *IEEE Electron Device Letters*, **31** (2010), 546-548.
- [7] K Ko, D Son, M Kang, and H Shin. Analysis of metal gate work-function variation for vertical nanoplate FET in 6-T SRAMs. *Silicon Nanoelectronics Workshop*. 2017, 61-62.
- [8] Y C Wu, Y R Jhan. 3D TCAD Simulation for CMOS Nanoelectronic Devices. *Springer*: New York, 2018.
- [9] R Saha, B Bhowmick, and S Baishya. Statistical Dependence of Gate Metal Work Function on Various Electrical Parameters for an n-Channel Si Step-FinFET. *IEEE Transactions on Electron Devices*, **64** (2017), 969-976.
- [10] K P Pradhan, S K Saha, P K Sahu, and Priyanka. Impact of Fin Height and Fin Angle Variation on the Performance Matrix of Hybrid FinFETs. *IEEE Transactions on Electron Devices*, **64** (2017), 52-57.
- [11] W Long, and K K Chin. Dual material gate field effect transistor (DMGFET). *International Electron Devices Meeting*. 1997, 549-552.
- [12] W Long, H Ou, J M Kuo, and K K Chin. Dual material gate (DMG) field effect transistor. *IEEE Transactions on Electron Devices*, **46** (1999), 865-870.
- [13] H Chakrabarti, R Maity, and N P Maity. Analysis of surface potential for dual-material-double-gate MOSFET based on modeling and simulation. *Microsystem Technologies*, **25** (2019), 4675-4684.
- [14] N P Maity, R Maity, S Maity and S Baishya. A New Surface Potential and Drain Current Model of Dual Material Gate Short Channel Metal Oxide Semiconductor Field Effect Transistor in Sub-Threshold Regime: Application to High-k Material HfO₂. *Journal of Nanoelectronics and Optoelectronics*, **14**, (2019), 868-876.
- [15] K S Gupta, and K Satyaveer. Analytical Modeling of a Triple Material Double Gate TFET with Hetero-Dielectric Gate Stack. *Silicon*, **11** (2019), 1355-1369.
- [16] S C Wagaj, and SC Patil. Dual material gate silicon on insulator junctionless MOSFET for low power mixed signal circuits. *International Journal of Electronics*, **106** (2019), 992-1007.
- [17] I Polishchuk, P Ranade, T J King, and C M Hu. Dual work function metal gate CMOS transistors by Ni-Ti interdiffusion. *IEEE Electron Device Letters*, **23** (2002), 200-202.

- [18] B Jena, S Dash, and G P Mishra. Electrostatic performance improvement of dual material cylindrical gate MOSFET using work-function modulation technique. *Superlattices and Microstructures*, **97**(2016), 212-220.
- [19] K Ko, M Kang, J Jeon, and H Shin. Compact Model Strategy of Metal-Gate Work-Function Variation for Ultrascaled FinFET and Vertical GAA FETs. *IEEE Transactions on Electron Devices*, **66** (2019), 1613-1616.
- [20] H Nam, Y Lee, JD Park and Shin C. Study of Work-Function Variation in High-kappa/Metal-Gate Gate-All-Around Nanowire MOSFET. *IEEE Transactions on Electron Devices*, **63** (2016), 3338-3341.
- [21] H P Lee, and P Su. Suppressed Fin-LER Induced Variability in Negative Capacitance FinFETs. *IEEE Electron Device Letters*, **38** (2017), 1492-1495.
- [22] Synopsys Inc. Sentaurus Device User Guide, Mountain View, CA, USA, 2017.
- [23] Y Lee, and Shin C. Impact of Equivalent Oxide Thickness on Threshold Voltage Variation Induced by Work-Function Variation in Multigate Devices. *IEEE Transactions on Electron Devices*, **64** (2017), 2452-2456.

A Novel Approach to Radiometric Identification

Raoul NIGMATULLIN^a, Semyon DOROKHIN^b and Alexander IVCHENKO^{b, 1}

^a*Kazan National Research Technical University named after A.N. Tupolev, (KNRTU-KAI)*

^b*Moscow Institute of Physics and Technology (MIPT)*

Abstract. This paper demonstrates that highly accurate radiometric identification is possible using CAPoNeF feature engineering method. We tested basic ML classification algorithms on experimental data gathered by SDR. The statistical and correlational properties of suggested features were analyzed first with the help of Point Biserial and Pearson Correlation Coefficients and then using P-values. The most relevant features were highlighted. Random Forest provided 99% accuracy. We give LIME description of model behavior. It turns out that even if the dimension of the feature space is reduced to 1, it is still possible to classify devices with 99% accuracy.

Keywords. Radiometric identification, wireless networks, security, RF fingerprinting, SDR, CAPoNeF, machine learning

1. Introduction

The security of a wireless network can be significantly improved if there is a device identification algorithm. Review [1] provides sufficient examples, such as authentication & flooding attack and Access Point forgery. The features for identification can be acquired from application, transport, MAC and physical layers [1]. Application and transport layer do not provide a large number of features. Therefore, they are quite rarely chosen in device identification. MAC layer provides vendor-specific features that are based on implementation of underspecified standard details. If one obtains the technical details of that realization, a new forgery can be committed.

Physical layer features can be described as channel-specific or device-specific. Channel-specific features ought to be neglected in mobile devices identification, since such devices change their location thus altering features' value. Device-specific features are based on minor imperfections in electronic components. These imperfections constitute a unique set of features thus allowing reliable identification. RF identification by device-specific features is called radiometric identification. A brief overview of radiometric identification methods is presented in Table 1.

Radiometric identification can be performed in waveform or modulation domain. A classical modulation domain algorithm was developed by Brik et al. [2]. The researchers analyzed modulation errors to obtain features. The method was tested on

¹Corresponding Author: Deputy Head, Laboratory of Multimedia Systems and Technology, Moscow Institute of Physics and Technology, Institutskiy per., 9, 141701, Dolgoprudny, Moscow Region, Russian Federation; E-mail: ivchenko.a.v@phystech.edu.

138 network interface cards and proved to be reliable. However, such methods require knowledge of modulation scheme and imply proper synchronization and channel estimation, which may be difficult to achieve.

Among the methods operating in waveform-domain, there is a class of algorithms based on transient analysis. Barbeau et al. [3] generated features using wavelet transform. This method requires an additional acquisition algorithm to track the start of transient process. Adam Polak et al. [4] resorted to a more fundamental approach. The researchers modeled DAC and amplifier nonlinearities and used the models' coefficients for further classification. Though this method illustrates the nature of radiometric fingerprint, the amplifiers need to be analyzed on evaluation boards. Several research groups tried to use features based on high-order statistics. The general approach is to split the signal into several regions of interest (ROI) and to calculate variance, skewness and kurtosis for amplitude, frequency and phase of each ROI [5], [6]. Trevor Bihl et al. [6] proved that statistical features extracted from the signal phase play the major role. However, such approach has low average true verification rate. Recently there were several attempts to solve the task of radiometric identification with neural networks. O'Shea et al. [7] used high-order statistics to generate features for CNN. This resulted in quite a good classification at high signal-to-noise ratio (SNR).

Table 1. Comparison of existing radiometric identification methods

Authors	Features' nature	Advantages	Disadvantages
Brik et al. [2]	Modulation accuracy	Tested on large (138) number of devices	Requires demodulation
Barbeau et al. [3]	Transient wavelet coefficients	Tested on 10 devices	Requires acquisition of transient
Polak et al. [4]	Nonlinearity model	Error probability expressed analytically	Additional equipment needed
Bihl et al. [6]	High-order statistics	Importance of features analyzed	Low average true verification rate (71%)
O'Shea et al. [7]	Convolutional neural network	High accuracy for different constellations	High computational complexity

This paper focuses on feature engineering for radiometric identification. A novel method is proposed based on the recent paper [8]. The new parameters allow to achieve high accuracy (97%) without the need for computationally complex feature generation and complicated classification algorithms. The designed algorithm is tested on experimental data. Since it is the first attempt to apply methods described in [8] to radiometric identification, we consider a simple task of binary classification, with one receiver and only two transmitters. We subtract etalon signal from the received one and analyze only the phase of the resulting signal, as it was shown to be more relevant [6].

The rest of the paper is organized as follows: section 1 provides the detailed description of the features used for classification, section 2 describes the experimental setup and the process of dataset generation, followed by section 3, where the results are demonstrated and important comments are given.

2. The Proposed Features

In this section we outline the basic principles of the CAPoNeF (Comparative Analysis of the Positive and Negative Fluctuations) method. Usually, in the conventional statistics only the mean value and standard deviation are frequently used for description

of trendless sequences (TLS). In the recent paper [8] it was shown that a new set of quantitative parameters can be suggested for quantitative description of various TLS. These parameters have different sensitivity to the hidden random factors. More important, they are free from treatment errors and model assumptions. In this paper, we choose the following parameters to characterize a specific competition between positive and negative fluctuations of the given rectangle matrix $y_j(m)$ ($j = 1, 2, \dots, N$ – the number of data points/matrix rows; $m = 1, 2, \dots, M$ the number of columns/successive measurements). For the selected column of the given rectangle matrix (we omit the column index m for simplicity) the chosen quantitative parameters are the following:

P_1 = arithmetic mean(y) coincides with mean value of the given TLS.

P_2 = $\max(Y_{up}) - \min(Y_{dn})$ – the range between positive and negative fluctuations.

P_3 = $\max(Y_{up}) - |\min(Y_{dn})|$ – the relative fluctuation intensity.

P_4 = $\max(J_{up}) - \min(J_{dn})$ – the maximal cumulative intensity between positive/negative fluctuations. (J_{up}/J_{dn} corresponds to the integrals taken from initial TLS(s) $y_j(m)$).

P_5 = $[\max(y) - \text{mean}(y)] / [\text{mean}(y) - \min(y)]$ – measure of asymmetry. If $P_5 = 1$, then the TLS can be considered as “ideally” symmetrical sequence relatively its mean value.

P_6 = $\text{last}(x_{up}) - \text{last}(x_{dn})$ – this parameter serves as an asymmetrical measure in the horizontal direction.

P_7 = $\max(J_S)$ – maximal value of the bell-like curve formed from sequence of the range amplitudes (SRA). This SRA, when the amplitudes ($y_1 > y_2 > \dots > y_N$) are organized in the descending order, is integrated and forms the bell-like curve. The maximal value separates the positive/negative branches of the ordered fluctuations.

P_8 = $\text{Range}[J(y_n)]$, $y_n = (y - \langle y \rangle) / \text{Range}(y)$, $\text{Range}(y_n) = 1$, where $\text{Range}(f) = \max(f) - \min(f)$. This parameter is also important, because it allows to compare the integral ranges taken from equivalent/normalized columns having the unit range value.

We would like to emphasize that any TLS $Dy_j = y_j - \langle y \rangle$ oscillates near zero value. We noticed that the distribution of the roots numbers is generally described by the segment of the straight line $R_k \cong a \cdot k + b$, where a and b are regression parameters. This observation allows to find the mean oscillation frequency and phase from the equation $\cos(\langle \omega \rangle R_k - \Phi) = 0$ or $\langle \omega \rangle R_k - \Phi = \pi/2 + \pi \cdot k$. It means that one can add two additional parameters invariant to scaling, mean frequency $\langle \omega \rangle$ and phase Φ :

$P_9 = \langle \omega \rangle$.

$P_{10} = \Phi$.

3. Experimental Setup

The data was collected using Adalm Pluto software-defined radio (SDR). One SDR was used as a receiver, while the two other devices were used as transmitters. The signal of $L = 1024$ samples was broadcasted from one transmitter at a time, repeated in a cyclic fashion. The signal was also known at the receiver side, which allowed to perform time-domain tracking and compensate for sampling clock offset. The synchronized signal was normalized and the etalon signal was subtracted from every group of L samples. Thus, we obtained the error signal. Its phase was used for further feature generation.

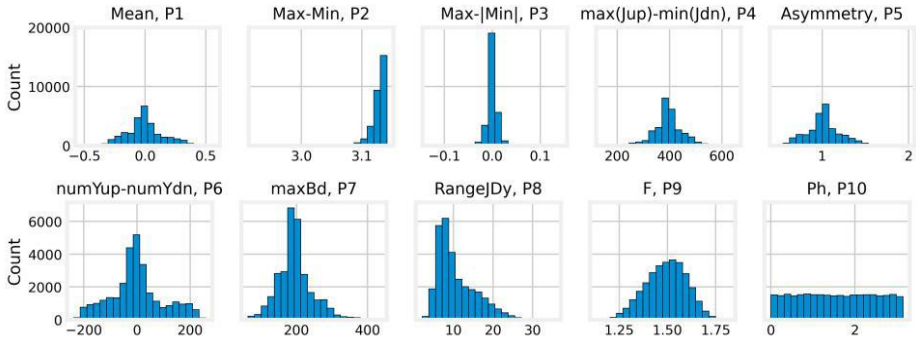


Figure 1. Histograms of the features

We assume that the choice of the signal itself does not affect the statistical properties of the error signal significantly. Therefore we felt free to use “trans-noise” – a noise-like signal, generated from the first and the second L digits of transcendental number π . Every digit was mapped to certain amplitude, so that 0 and 9 corresponds to minimal and maximal DAC level.

Experimental data was collected in the following setup. First, the main lobes of TX and RX antennas radiation patterns where facing each other, the distance between the devices changed from 1 to 5 meters. Then the distance between the devices was fixed (2 meters), but the angle between the main lobes changed from $-\pi/2$ to $\pi/2$ with the step of $\pi/6$. The resulting dataset contains 30000 labeled entries.

4. Results and Discussion

Prior to classification, we analyzed statistical properties of the proposed features. Histograms were calculated for every feature to study their distributions (see Figure 1). After that we studied correlation properties. We calculated the Pearson Correlation Coefficient (PCC) for every pair of features. As it can be seen in Figure 2, there is high correlation between the parameters P_1, P_5 and P_6 (PCC reaches 0.96). This peculiarity should be taken into consideration. We will show that eliminating these parameters improves the accuracy. We also calculated Point Biserial Correlation Coefficient (PBCC) and P-value for every feature. The P-value of the features’ significance level is

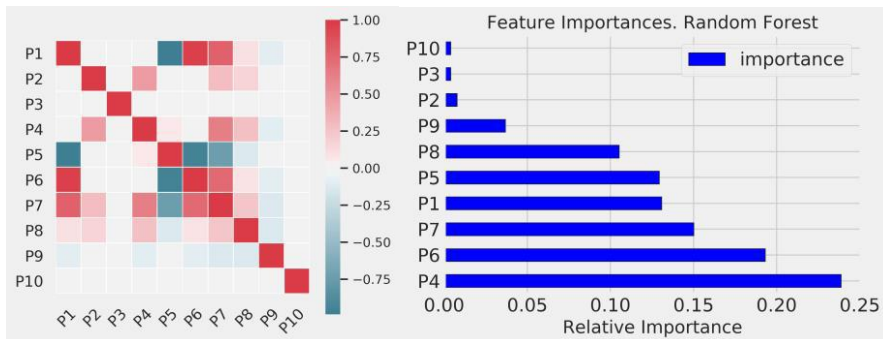


Figure 2. Heatmap (left) and feature importances (right)

0.05. If the P-value level is less than 0.05, the parameter is considered statistically significant. Table 2 contains PBCCs and P-values for the dataset. The most interesting parameters are P₉, P₈ and P₂. The parameters P₄, P₃, P₁₀ have P-values higher than 0.05.

Table 2. Point Biserial Correlation Coefficients and P-values

Parameter	PBCC	P-value
$\langle \omega \rangle, P_9$	0.3025	0.0
$\text{Range}[J(y_{in})], P_8$	0.4293	0.0
$\max(Yup) - \min(Ydn), P_2$	0.1227	0.0
$\max(Jup) - \min(Jdn), P_4$	0.0053	0.3602
$\max(Yup) - \min(Ydn), P_3$	-0.0008	0.8839
Φ, P_{10}	-0.0003	0.9517
$\max(Js), P_7$	-0.0173	0.0027
$\text{last}(x_{up}) - \text{last}(x_{dn}), P_6$	-0.0163	0.0047
$\text{mean}(y), P_1$	-0.0179	0.002
$[\max(y) - \text{mean}(y)] / [\text{mean}(y) - \min(y)], P_5$	-0.0662	0.0

We compared basic ML classification algorithms. Among all classifiers, such as Linear Regression, Decision Tree and KNN, Random Forest showed the best results with the accuracy of 0.99.

4.1. Feature Importances and LIME

Feature importances show the relevance of every parameter. They are defined as the total decrease in node impurity, weighted by the probability of reaching that node, averaged all over the trees of the ensemble. Probability of reaching the given node is approximated by the proportion of samples reaching that node [9]. The most predictive features are P₃, P₆, and P₇. Locally Interpretable Model-agnostic Explanations (LIME) [10] shows how the model makes decisions by approximating the area around the forecast using a linear model. We use it to explain the reason for wrong predictions and proper behavior (Figure 3).

The plot from LIME shows the contribution to the final forecast of each of the example parameters. In case of correct classification (Figure 3, left) P₈, P₂, P₁ decrease the accuracy, while P₉ has a positive impact. In case of wrong prediction P₁ decreases the accuracy, but P₉, P₈, P₄ have a positive impact (Figure 3, right).

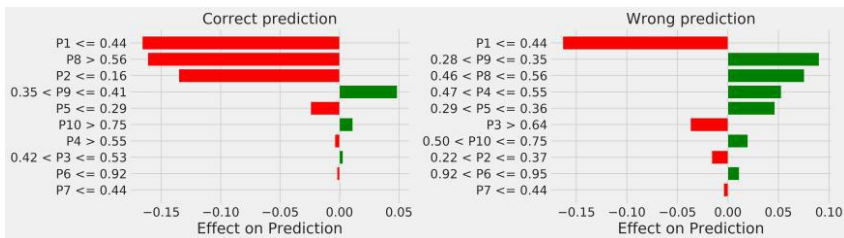


Figure 3. LIME description of wrong (right) and correct (left) prediction

4.2. Feature Selection

In the experiments, the P8 parameter was the most significant. With P8 only, a simple threshold of 0.5 is sufficient to achieve 99% accuracy. We assume that other features will contribute more in multiclass case under more complex conditions. It is difficult to compare with existing methods since the results depend on the dataset and the task. However, the dataset is available online [11] for further comparative analysis.

4.3. Conclusion

This paper demonstrates the application of the CAPoNeF method for radiometric identification (binary classification). We collected data using SDR with various distances and angles between the devices. We subtracted etalon signal from the received one and extracted 10 parameters from the phase of the resulting signal. The parameters were further analyzed using the ML. P2 (the range between positive and negative fluctuations), P8 (the range of the integral of the normalized signal) and P9 (mean frequency) have the highest predictive ability. Classification task was solved: Random Forest classifier showed 99% accuracy.

In further research, we are planning to solve multiclass classification task. Current research did not take into account possible fading. Additional study should be carried out to determine whether the proposed method is robust to fading environment.

References

- [1] Q. Xu, R. Zheng, W. Saad and Z. Han, "Device Fingerprinting in Wireless Networks: Challenges and Opportunities," in *IEEE Communications Surveys & Tutorials*, vol. 18, no. 1, pp. 94-104, Firstquarter 2016
- [2] Vladimir Brik, Suman Banerjee, Marco Gruteser, and Sangho Oh. 2008. Wireless device identification with radiometric signatures. In *Proceedings of the 14th ACM international conference on Mobile computing and networking (MobiCom '08)*. Association for Computing Machinery, New York, NY, USA, 116–127.
- [3] Barbeau, M., Hall, J., &Kranakis, E. (2006, October). Detection of rogue devices in bluetooth networks using radio frequency fingerprinting. In *proceedings of the 3rd IASTED International Conference on Communications and Computer Networks, CCN* (pp. 4-6).
- [4] A. C. Polak, S. Dolatshahi and D. L. Goeckel, "Identifying Wireless Users via Transmitter Imperfections," in *IEEE Journal on Selected Areas in Communications*, vol. 29, no. 7, pp. 1469-1479, August 2011
- [5] C. K. Dubendorfer, B. W. Ramsey and M. A. Temple, "An RF-DNA verification process for ZigBee networks," *MILCOM 2012 - 2012 IEEE Military Communications Conference*, Orlando, FL, 2012, pp. 1-6
- [6] T. J. Bihl, K. W. Bauer and M. A. Temple, "Feature Selection for RF Fingerprinting With Multiple Discriminant Analysis and Using ZigBee Device Emissions," in *IEEE Transactions on Information Forensics and Security*, vol. 11, no. 8, pp. 1862-1874, Aug. 2016
- [7] T. J. O'Shea, T. Roy and T. C. Clancy, "Over-the-Air Deep Learning Based Radio Signal Classification," in *IEEE Journal of Selected Topics in Signal Processing*, vol. 12, no. 1, pp. 168-179, Feb. 2018
- [8] Nigmatullin, R. R., &Vorobev, A. S. (2019). The "Universal" Set of Quantitative Parameters for Reading of the Trendless Sequences. *Fluctuation and Noise Letters*, 18(04)
- [9] Breiman, L., Friedman, J., Stone, C. J., & Olshen, R. A. (1984). *Classification and regression trees*. CRC press.
- [10] Ribeiro, M. T., Singh, S., &Guestrin, C. (2016, August). "Why should I trust you?" Explaining the predictions of any classifier. In *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining* (pp. 1135-1144).
- [11] Dorokhin, S., Radiometric Identification, <https://github.com/svdpriima/Radiometric-Identification>

Order-Preserving and Efficient One-to-Many Search on Encrypted Data

Dongping HU, Aihua YIN¹, Huaying YAN, Tao LONG

School of Software and Internet of Things Engineering, Jiangxi University of Finance and Economics, Nanchang 330023, Jiangxi, China

Abstract. Order-preserving encryption (OPE) is an useful tool in cloud computing as it allows untrustworthy server to execute range query or exact keyword search directly on the ciphertexts. It only requires sub-linear time in the data size while the queries are occurred. This advantage is very suitable in the cloud where the data volume is huge. However, the order-preserving encryption is deterministic and it leaks the plaintexts' order and its distribution. In this paper, we propose a one-to-many OPE by taking into account the security and the efficiency. For a given plaintext, the encryption algorithm firstly determines the corresponding ciphertext gap by performing binary search on ciphertext space and plaintext space at the same time. An exact sample algorithm for negative hypergeometric distribution is used to fix the size of the gap. Lastly a value in the gap is randomly chosen as the mapping of the given plaintext. It is proven that our scheme is more secure than deterministic OPE with realizing efficient search. In particular, a practical and exact sampling algorithm for the negative hypergeometric distribution (NHGD) is first proposed.

Keywords. Order-preserving Searchable Encryption, One-to-many Mapping Encryption, Range Query, Negative Hypergeometric Distribution

1. Introduction

In cloud computing, the sensitive data are always encrypted before outsourcing for the security concern. It is an effective approach to protect the confidentiality of the data. However, the encryption operation reduces the utilization of data. For example, comparison operations, such as MIN, MAX and range queries, are common operations in encrypted SQL database which requires the server to perform operations over ciphertexts. For the server, one approach is to decrypt the whole database first and then runs the operations for correct results. Obviously, it requires trusting server and the process is time-consuming. Recently, the fully homomorphic encryption shows it can perform compare computations directly on encrypted data. However, its performance cost is extremely high, on the order of 10^9 times [1].

For tackling this problem, searchable encryption was proposed and it allows the cloud sever (in our work, the cloud server is believed to be an untrustworthy server) to execute queries on ciphertexts. And then Order-preserving Encryption was proposed in [2]. This primitive ensures the order between ciphertexts is as the same as the order between plaintexts. It permits efficient range queries or comparison operations. In the paper, "efficient" means at least in sub-linear time in the database size, for executing

¹ *Corresponding author. E-mail: Aihuayin@jxufe.edu.cn.

linear work (such as homomorphic encryption) is prohibitively slow in real world applications for huge database. Due to this advantage, OPE is prevalent in cloud computing.

Agrawal *et al.* [2] initialized OPE for supporting range queries over ciphertexts. The encryption of the scheme takes all the plaintexts as input before encrypting. In real applications, it is difficult for users to know all plaintexts in advance. Furthermore, the work of [2] did not provide the security definition.

Boldyreva *et al.* formally defined the security of OPE in [3]. The ideal security of OPE, named IND-OCPA, was first given. Authors then proved that IND-OCPA was unachievable in practical because of huge ciphertexts space required. Instead a 'weaker' security notion POPF-CCA was proposed, which meant that adversary cannot distinguish between OPE and ideal order-preserving function (OPF). Furthermore, a deterministic OPE scheme satisfied POPF-CCA is proposed. Here we called it BCLO scheme.

To satisfy IND-OCPA, Popa *et al.*[4] proposed a new OPE model by building a balanced tree to store ciphertext values. The tree would be rebalanced after a new ciphertext inserted. A secure client decrypted the ciphertexts first and then returns the ordering after compares the plaintexts. Although Popa's scheme satisfied the ideal security; it was not a real OPE in nature as the order of plaintexts was not preserved. As discussed in [5,6], the computation cost of Popa's scheme was linear time and not suitable the applications with large data (e.g. cloud computing). Some other OPE schemes have been proposed in [7-10] and they either provided weaker security guarantee or are not efficient.

Recently, Liu *et al.* [11] focused on how to conceal the distribution values. They inserted many letters into ciphertext space to expand it and then divided it nonlinearly. Then a plaintext value was mapped to a value in the expanded space. It helped enhance security and order comparison could be executed directly over the ciphertexts.

In this work, we improve the Boldyreva's work and design a one-to-many OPE (Otm OPE) that achieves higher security and efficiency.

We organized our paper as follows. Our Otm OPE is introduced in Section 2. An exact sampling algorithm for generating the negative hypergeometric random variables is proposed in Section 3. Security proof is presented in Section 4. We evaluate the performance in Section 5. A conclusion is given at last.

2. Our Otm OPE

The BCLO scheme is deterministic and it leaks not only the order but also the plaintext space's distribution. To overcome its weaknesses, we adapt the BCLO scheme by raising its randomness. We execute binary search directly on plaintext space for achieving efficient search. For a plaintext, we first determine its corresponding interval of ciphertext space. This is achieved by employing an exact sample algorithm for NHGD we propose. Finally, a value is randomly chosen from the interval. In this way, same plaintext can be mapped to different ciphertexts with different random seeds. The Notations List of the paper is given in Table 1. We first describe the Algorithm Binary Search in Algorithm 1. And then we present our Otm OPE in Algorithm 2.

Table 1. Notations List

Notation	Descriptions
SK	Symmetric Key
x	a Plaintext
y	a Ciphertext
PS	Plaintext Space $\{1, \dots, P\}$
CS	Ciphertext Space $\{1, \dots, C\}$
$NHGD$	Negative Hypergeometric distribution
NBD	Negative Binomial Distribution
Γ	Gamma Distribution
$F_T(t)$	The Cumulative Distribution Function of variable t , where t follows distribution T

Algorithm 1 Binary SearchInput: SK, PS, CS, x .Output: PS, CS .

- 1 Let $A = |PS|$; $B = |CS|$;
- 2 Let a be the minimum value of PS and b is the minimum value of CS ;
- 3 Let $s = a + \lceil A/2 \rceil - 1$;
- 4 Let $rd = PRndGen(SK, l, (PS, CS, s))$;
- 5 Let $t = b + NHGDEaxtSample(A, B, s, rd)$;
- 6 If $x < s$ then
 - 7 $PS = \{a, \dots, s-1\}$;
 - 8 $CS = \{b, \dots, t-1\}$;
- 9 If $x \geq s$ then
 - 10 $PS = \{s+1, \dots, a+A\}$;
 - 11 $CS = \{t+1, \dots, b+B\}$;
- 12 Endif
- 13 Return $\{PS, CS\}$.

In Algorithm 1, $PRndGen()$ is a pseudorandom function (PRF) whose input length and output length are both variable. It take as input SK , output length l and $\{0,1\}^*$ and then return $\{0,1\}^l$, where the input $\{0,1\}^*$ depends on PS, CS and the plaintext.

The $NHGDEaxtSample()$ is a sampling algorithm for generating negative hypergeometric variables. It takes as input $|PS|, |CS|$ and a plaintext x and return a ciphertext y with probability $NHGD(y-b; |PS|, |CS|, x-a)$, where a is the minimum value of PS and b is the minimum value of CS .

Algorithm 2 Otm OPEInput: SK, PS, CS, x, fid .Output: y .

```

1 While  $PS$  contains more than one plaintext do
2    $\{PS, CS\} = \text{Binary Search}(SK, PS, CS, x)$  ;
3 End While
4 Let  $coin = \text{PRndGen}(SK, l, (PS, CS, x, cc))$  ;
5 Choose a number  $y$  in  $CS$  with  $coin$  ;
6 Return  $y$  .

```

In Algorithm 2, for a given plaintext, it first determines the mapping interval of ciphertexts by calling Binary Search () (Step 2). Once the interval is fixed, a random seed $coin$ is then generated by PRndGen () (Step 4). With the $coin$, a number in the interval is randomly chosen as the ciphertext (Step 5). We stress that the result of Step 5 is not deterministic. In this way, the goal of one-to-many can be achieved, that is the same plaintext will be mapped to different ciphertexts.

3.Sampling Following the Negative Hypergeometric Distribution

To the best our knowledge, the exact sampling algorithm for generating variables following NHGD has not appeared. And it is believed as an open problem. In this section, we give a solution for this problem.

It is impossible to generate NHGD variables by direct sampling method as it has four parameters. We then turn to the acceptance-rejection sample method, requiring a continuous approximation for NHGD. Here a Gamma approximation is given. In this section, we first introduce the acceptance-rejection sample method. Then the upper bound for Gamma approximation is given. Based on above work, an exact sampling method for NHGD is proposed.

3.1. Acceptance-Rejection Method

This method can generate variables following distribution function f by utilizing another distribution g , where g is a continuous and there exists sampling method following distribution g .

The method runs as follows [12]:

(1) Sample ν following g ;

(2) Generate $\omega \sim U[0,1]$;

(3) If $\omega \leq \frac{f(\nu)}{l \cdot g(\nu)}$, where $l > 1$, then output ν , else nothing is output and new ν be generated.

Note that distribution g and distribution f should be independent of each other.

With this method, we know that the probability of outputting ν is $\frac{1}{l}$. That means

$P\left(\omega \leq \frac{f(v)}{l \cdot g(v)}\right) = \frac{1}{l}$. Obviously, the smaller the value of l , the higher the sampling

efficiency. In general, let l be maximum of $\frac{f(v)}{g(v)}$.

We wish to select variables according to negative hypergeometric distribution $NHGD_{k,A,B}$, which we bound with $NBD_{k,q}$, where $q = A/B$. The $NBD_{k,q}$ is being bounded by the Gamma distribution $Gamma_{k,\vartheta}$, where $\vartheta = -\ln(1-q)$. Here NBD is discrete, whereas Gamma is continuous.

$$NHGD_{k,A,B}(v) = \frac{\binom{v-1}{k-1} \binom{B-v}{A-k}}{\binom{B}{A}} \tag{1}$$

$$NBD_{k,q}(v) = \binom{v-1}{q-1} q^k (1-q)^{v-k} \tag{2}$$

$$Gamma_{k,\vartheta}(v) = \frac{\vartheta^k v^{k-1} e^{-\vartheta v}}{\Gamma(k)} \tag{3}$$

3.2. Upper-Bounding Distribution

Theorem 1. If A, B, k are integers, let $q = A/B$, where $A \leq B, k \in \{1, 2, \dots, A\}$, then:

$$\frac{NHGD_{k,A,B}(v)}{NBD_{k,q}(v)} \leq \left(1 - \frac{k}{B-k}\right)^{B-A} \tag{4}$$

Theorem 2. If $k > 0, q > 0$, Let $\vartheta = -\ln(1-q)$, then:

$$\frac{NBD_{k,q}(v)}{Gamma_{k,\vartheta}(v)} \leq \left(\frac{q}{\vartheta(1-q)}\right)^k \tag{5}$$

With Theorem 1 and Theorem 2, then:

$$\frac{NHGD_{k,A,B}(v)}{Gamma_{k,\vartheta}(v)} \leq \left(1 - \frac{k}{B-k}\right)^{B-A} \cdot \left(\frac{q}{\vartheta(1-q)}\right)^k \tag{6}$$

3.3. Gamma Distribution

If $\omega \sim U[0,1]$, the variable v following the Gamma distribution and $\omega \leq NHGD(v) / l \cdot Gamma(v)$, the integer part of v we need, and then we have NHGD.

3.4. Description of Exact Sample Method of Negative Hypergeometric Distribution

Algorithm 3 NHGDExactSample

Input: A, B, k, rd .

Output: v .

- 1 Let $q = A / B$; $\mathcal{G} = -\ln(1 - q)$;
 - 2 Let $l = \left(1 - \frac{k}{B - k}\right)^{B - A} \cdot \left(\frac{q}{\mathcal{G}(1 - q)}\right)^k$;
 - 3 Let $x = Gamma(k, \mathcal{G}, rd)^{[13]}$;
 - 4 Let $v = \lfloor x + 0.5 \rfloor$;
 - 5 If $v > B + k - A$ or $v < k$, then go to step 3;
-

6 Let $NHGD(v) = \frac{\binom{v-1}{k-1} \binom{B-v}{A-k}}{\binom{B}{A}}$;

7 Let $Gamma(v) = \frac{\mathcal{G}^k v^{k-1} e^{-\mathcal{G}v}}{\Gamma(k)}$;

8 Generate $\omega \sim U[0,1]$;

9 If $\omega \leq NHGD(v) / l \cdot Gamma(v)$ then return v ;

10 Go to step 3.

According to the description of Algorithm 3, variable v will be returned when $\omega \leq NHGD(v) / l \cdot Gamma(v)$ with probability $1/l$. Next we will proof the correctness of Algorithm 3. In another words, we will illustrate v follows exactly NHGD by proving $F(v)$ is equal to F_{NHGD} .

$$\begin{aligned}
 P\left(\omega \leq \frac{NHGD(v)}{l \cdot Gamma(v)} \mid V \leq v\right) &= \frac{P\left(\omega \leq \frac{NHGD(v)}{l \cdot Gamma(v)}, V \leq v\right)}{P(V \leq v)} = \frac{P\left(\omega \leq \frac{NHGD(v)}{l \cdot Gamma(v)}, V \leq v\right)}{F_{Gamma}(\mathcal{X})} \\
 &= \int_{-\infty}^v \frac{P\left(V \leq \frac{NHGD(v)}{l \cdot Gamma(v)} \mid V = \lambda \leq v\right)}{F_{Gamma}(v)} \cdot Gamma(\lambda) d\lambda \\
 &= \frac{1}{l \cdot F_{Gamma}(v)} \int_{-\infty}^v \frac{NHGD(\lambda)}{Gamma(\lambda)} \cdot Gamma(\lambda) d\lambda \\
 &= \frac{1}{l \cdot F_{Gamma}(v)} \int_{-\infty}^v NHGD(\lambda) d\lambda \\
 &= \frac{F_{NHGD}(v)}{l \cdot F_{Gamma}(v)}
 \end{aligned}
 \tag{7}$$

Algorithm 3 generates variable v while $\omega \leq NHGD(v) / l \cdot Gamma(v)$ is satisfied. Thus we have

$$\begin{aligned}
 P\left(V \leq v \mid \omega \leq \frac{NHGD(v)}{l \cdot Gamma(v)}\right) &= \frac{P\left(V \leq v, \omega \leq \frac{NHGD(v)}{l \cdot Gamma(v)}\right)}{P\left(\omega \leq \frac{NHGD(v)}{l \cdot Gamma(v)}\right)} \\
 &= \frac{F_{Gamma}(v) \cdot \frac{F_{NHGD}(v)}{l \cdot F_{Gamma}(v)}}{1/l} = F_{NHGD}(v)
 \end{aligned} \tag{8}$$

4. Security Analysis

We first evaluate security by analyzing the security of $PRndGen$ in Algorithm 1.

$RndGen$ is a PRF with variable length inputs and outputs. For adversary A , its advantage against $PRndGen$ is defined as

$$\begin{aligned}
 &Adv_{PRndGen}^{PRF}(A) \\
 &= \Pr\left(A^{PRndGen(SK, l, \{0,1\}^*)} = 1\right) - \Pr\left(A^{O(l, \{0,1\}^*)} = 1\right)
 \end{aligned} \tag{9}$$

Where SK is key, l is output length, and O is an oracle which outputs random numbers $\{0,1\}^l$.

For computing $Adv_{PRndGen}^{PRF}(A)$, we set

$$\begin{aligned}
 &Adv_{PRndGen}^{PRF}(A) \\
 &= \Pr\left(A^{PRndGen(SK, l, \{0,1\}^*)} = 1\right) - \Pr\left(A^{O_F(SK, l, \{0,1\}^*)} = 1\right) \\
 &\quad + \Pr\left(A^{O_F(SK, l, \{0,1\}^*)} = 1\right) - \Pr\left(A^{O(l, \{0,1\}^*)} = 1\right)
 \end{aligned} \tag{10}$$

where oracle $O_F(SK, l, \{0,1\}^*)$ take as inputs key SK , output length l and $\{0,1\}^*$ and then returns $G(l^l, s)$ with random $s \in \{0,1\}^k$.

Construct an adversary B as following, where B is PRF.

Adversary $B^{O(\cdot)}$

- (1) Choose $m \in [1, k]$ randomly;
- (2) Run A and answer its queries as:
 - a. Deliver $G(l, F(SK, r))$ to A with first $q-1$ queries, where r is seed;
 - b. Deliver $G(l, O(r))$ to A with the q th queries;
 - c. Deliver $G(l, s)$ to A with $q+1, \dots, k$ queries;

(3) Return the outputs of A .

Then,

$$\begin{aligned}
 \Pr\left(A^{PRndGen(SK, I, \{0,1\}^*)} = 1\right) - \Pr\left(A^{O_E(SK, I, \{0,1\}^*)} = 1\right) &= \sum_{t=1}^k \left(\Pr\left(B^{F(SK, \cdot)} = 1 | E\right) - \Pr\left(B^{R(\cdot)} = 1 | E\right) \right) \\
 &= k \cdot \frac{1}{k} \cdot \sum_{t=1}^k \left(\Pr\left(B^{F(SK, \cdot)} = 1 | E\right) - \Pr\left(B^{R(\cdot)} = 1 | E\right) \right) \\
 &= k \cdot \sum_{t=1}^k \left(\Pr\left(B^{F(SK, \cdot)} = 1 \text{ and } E\right) - \Pr\left(B^{R(\cdot)} = 1 \text{ and } E\right) \right) \\
 &\leq k \cdot \left(\Pr\left(B^{F(SK, \cdot)} = 1\right) - \Pr\left(B^{R(\cdot)} = 1\right) \right) \\
 &= k \cdot \text{Adv}_F^{prf}(B)
 \end{aligned} \tag{11}$$

where we denote by $E : m = t$.

Construct an adversary C against blockcipher O_E , where C is PRG. We then have

$$\begin{aligned}
 \Pr\left(A^{O_E(SK, I, \{0,1\}^*)} = 1\right) - \Pr\left(A^{O(I, \{0,1\}^*)} = 1\right) &= \sum_{t=1}^k \left(\Pr\left(C^{O_E(SK, I, \cdot)} = 1 | E\right) - \Pr\left(C^{R(\cdot)} = 1 | E\right) \right) \\
 &= k \cdot \frac{1}{k} \cdot \sum_{t=1}^k \left(\Pr\left(C^{O_E(SK, I, \cdot)} = 1 | E\right) - \Pr\left(C^{R(\cdot)} = 1 | E\right) \right) \\
 &= k \cdot \sum_{t=1}^k \left(\Pr\left(C^{O_E(SK, I, \cdot)} = 1 \text{ and } E\right) - \Pr\left(C^{R(\cdot)} = 1 \text{ and } E\right) \right) \\
 &\leq k \cdot \left(\Pr\left(C^{O_E(SK, I, \cdot)} = 1\right) - \Pr\left(C^{R(\cdot)} = 1\right) \right) \\
 &= k \cdot \text{Adv}_{O_E}^{prg}(C)
 \end{aligned} \tag{12}$$

With Equation (1) and (2), then

$\text{Adv}_{PRndGen}(A) \leq k \cdot (\text{Adv}_F^{prf}(B) + \text{Adv}_{O_E}^{prg}(C)) + \lambda$, where the value of λ depends on the sampling accuracy of NHGD.

As analysis in Section 2, random seed *coin* is used to choose the ciphertext in final selection. Therefore different ciphertexts will be chosen to be the different mapping of the same plaintext. Our Otm OPE is not deterministic and has higher security than BCLO scheme.

5. Efficiency Analysis

From Algorithm 1 and 2, it is easy to observe the computation cost of the scheme depends on the running time of Binary Search.

There is P elements in the plaintext space. To reach one element, on average, the expected recursions is $M = \frac{1}{P} \cdot \left[\sum_{k=1}^{\log P} 2^{k-1} k + (\log P + 1) \right]$. Next we estimate the value of M . Let $A = \log P, A \geq 1$, then

$$M = M_A = \frac{1}{2^A} \cdot \left[\sum_{k=1}^A 2^{k-1} k + (A+1) \right] \quad (13)$$

Mathematical induction is employed on A to prove $A-1 \leq M_A \leq A$.

(1) When $A=1$, then $M_1=1$ and $0 \leq M_1 \leq 1$;

$$M_w = \frac{1}{2^w} \cdot \left[\sum_{k=1}^w 2^{k-1} k + (w+1) \right] \text{ and } w-1 \leq M_w \leq w \quad (14)$$

(2) Suppose $A = w$ and assume

(3) when $A = w+1$, it gets

$$M_{w+1} = \frac{1}{2^{w+1}} \cdot \left[\sum_{k=1}^{w+1} 2^{k-1} k + (w+2) \right] = \frac{1}{2} M_w + \frac{1}{2} w + \frac{1}{2} + \frac{1}{2^{w+1}} \quad (15)$$

Obviously, $w \leq M_{w+1} \leq w+1$.

As analysis above, the procedure Binary Search () would be called about $\log P$ rounds to map a plaintext to a randomized interval. While in BCLO scheme, $\log C$ recursive calls are required to determine the mapping interval. In real applications, the value of P is much smaller than that of C , so our Otm OPE has higher efficiency.

6. Conclusions

In this paper, we focus on how to enable OPE to support effectively searching with higher security. To this end, we derive a practical sampling algorithm following NHGD distribution. Based on this work, an efficient order-preserving and one-to-many searchable encryption is proposed. In our scheme, a value in the ciphertext bucket is chosen as the mapping of a given plaintext. Comparing with BCLO scheme, we prove that our Otm scheme is more secure. We also show that our work is more efficient by performance analysis.

Acknowledgement

This work was supported by the National Natural Science Foundations of China (No. 61702238, No.61862027), Jiangxi Provincial Natural Science Foundation (20202BABL202041, 20192BAB207008), The Technology Project of Jiangxi Education Department (No. GJJ170316, GJJ180264) and The 15th Student Research Project of JXUFE (20200611170950596).

References

- [1] C. Gentry. Fully Homomorphic Encryption Using Ideal Lattices[C]. Proceedings of the 41st Annual ACM Symposium on Theory of Computing, ACM Press, 2009, 169-178.
- [2] Agrawal R, Kiernan J, Srikant R, et al. Order preserving encryption for numeric data. Proceedings of the 2004 ACM SIGMOD international conference on Management of data. 2004:563-574.
- [3] Boldyreva A, Chenette N, Lee Y, et al. Order-preserving symmetric encryption. Advances in Cryptology-EUROCRYPT 2009. Springer Berlin Heidelberg, 2009: 224-241.
- [4] Popa R A, Li F H, Zeldovich N. An ideal-security protocol for order-preserving encoding. Security and Privacy (SP), 2013 IEEE Symposium on. IEEE, 2013: 463-477.
- [5] Xiao L, Yen I L. Security analysis for order preserving encryption schemes[C]. Information Sciences and Systems (CISS), 2012 46th Annual Conference on. IEEE, 2012: 1-6.
- [6] Z. Liu et al. New Order Preserving Encryption Model for Outsourced Databases in Cloud Environments. Journal of Network and Computer Applications. 59 (2016) :198-207.
- [7] Lewi K, Wu D J. Order-revealing encryption: New constructions, applications, and lower bounds[C]. Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security. ACM, 2016: 1167-1178.
- [8] Boldyreva A, Chenette N, O'Neill A. Order-preserving encryption revisited: Improved security analysis and alternative solutions[C]. Annual Cryptology Conference. Springer, Berlin, Heidelberg, 2011: 578-595.
- [9] Popa R A, Redfield C, Zeldovich N, et al. CryptDB: protecting confidentiality with encrypted query processing[C]. Proceedings of the Twenty-Third ACM Symposium on Operating Systems Principles. ACM, 2011: 85-100.
- [10] Wang C, Cao N, Li J, et al. Secure ranked keyword search over encrypted cloud data[C]. 2010 IEEE 30th international conference on distributed computing systems. IEEE, 2010: 253-262.
- [11] Liu D, Wang S. Nonlinear order preserving index for encrypted database query in service cloud environments[J]. Concurrency and Computation: Practice and Experience, 2013, 25(13): 1967-1984.
- [12] Schmeiser B W, Shalaby M A. Acceptance/rejection methods for beta variate generation[J]. Journal of the American Statistical Association, 1980, 75(371): 673-678.

The Application of Improved Sparrow Search Algorithm in Sensor Networks Coverage Optimization of Bridge Monitoring

Yueqi PENG, Yunqing LIU¹ and Qi LI

Changchun University of Science and Technology, Changchun, China

Abstract. Aiming at the uneven random coverage distribution of wireless sensor networks sensor nodes, the sparrow search algorithm (SSA) is used to optimize the node coverage of wireless sensor networks. To improve the global search capability of SSA, the algorithm is improved and applied to the wireless sensor networks of bridge monitoring. For enhancing the coverage of the wireless sensor networks, this article uses two improved methods. One is to use the good point set theory to make the initial population evenly distributed; another is to introduce a weight factor to speed up its convergence. The experiments have proved the reliability and rationality of the algorithm. The improved algorithm is superior to other meta-heuristic algorithms and provides a new idea for optimizing bridge monitoring wireless sensor networks coverage.

Keywords. sparrow search algorithm, wireless sensor networks, bridge monitoring, coverage

1. Introduction

Bridges play an essential role in the transportation industry. The safety of bridge construction is directly related to the security of people's lives and property. Monitoring bridges have significant practical value. Wireless sensor networks are widely used in various fields such as national defense and military, space exploration, medical and health, traffic management, smart home, etc. Wireless sensor networks have been widely used in bridge monitoring due to their convenient installation, flexible deployment, and low maintenance costs. Coverage is an essential criterion for measuring wireless sensor networks.

The coverage optimization of wireless sensor networks is a multi-objective optimization problem. A bridge monitoring system requires an appropriate deployment of monitoring sensors to achieve good monitoring results. At present, many scholars at home and abroad have conducted in-depth research on the coverage optimization of wireless sensor networks. The key to the problem is to optimize the deployment of sensor nodes in the area by adopting appropriate optimization strategies for different monitoring situations, so as to ensure sensor coverage. With the rise of meta-heuristic algorithms,

¹ Corresponding Author: Yunqing Liu, Changchun University of Science and Technology, Changchun, China; E-mail: mzliuyunqing@163.com.

the use of meta-heuristic algorithms to optimize the coverage of wireless sensor networks has attracted the attention of many researchers. Literature [1] proposed a particle swarm optimization algorithm with dynamic weight optimization to uniformly cover sensor network nodes. Literature [2] proved the feasibility of the improved artificial fish swarm algorithm to optimize the coverage of wireless sensor network nodes. Literature [3] proposed an improved cuckoo search algorithm for wireless sensor networks coverage multi-objective optimization. Literature [4] used the improved glowworm swarm optimization algorithm to optimize the coverage of wireless sensor networks. Literature [5] applied a genetic algorithm to offshore wireless sensor networks. Literature [6] combined the improved heuristic ant colony algorithm with chaos operator to optimize wireless sensor networks' layout. Literature [7] designed a dichotomous fruit fly optimization algorithm applied in sensor deployment to improve senso networks coverage.

This paper uses the improved algorithm to deploy the nodes of the wireless sensor networks. By updating the position of the node, the optimal solution for each iteration is quickly calculated and the convergence speed is improved. This paper takes into account the complex and changeable environment of the bridge. It uses the characteristics of good mobility of sensor nodes, takes coverage as the final optimization goal, and combines neighbor node information to achieve uniform deployment of sensor nodes to ensure maximum coverage and high deployment reliability. The simulation experiment results show that the improved SSA has a significant improvement in the coverage optimization problem of bridge monitoring wireless sensor networks, which verifies the effectiveness of the improved algorithm.

2. Bridge Monitoring Sensor Coverage Model

In the deployment of bridge monitoring sensor networks, the number of sensors is often increased to increase the coverage of the sensor network. However, too many sensors deployed will generate a large number of redundant nodes, leading to data conflicts, waste of resources, and affecting network stability. Therefore, when deploying sensor networks, both network coverage and the number of nodes must be considered. This paper uses the sparrow search algorithm (SSA) in the two-dimensional plane of the bridge monitoring area.

Initialize the model, assuming that each sensor node has the same sensing radius and communication radius. Let $s = \{s_1, s_2, s_3 \dots s_n\}$ be a set of sensor nodes. The coordinate of any sensor node s_i in the set is (x_i, y_i) , and the coordinate of any sensor node s_j in the set is (x_j, y_j) . Then the distance between node $s_i(x_i, y_i)$ and node $s_j(x_j, y_j)$ can be obtained by the Euclidean distance calculation formula:

$$d(s_i, s_j) = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} \quad (1)$$

Discretize the monitoring area A into $m \times n$ grid points. With $s_i(x_i, y_i)$ as the center of the circle, and r_i as the sensor's sensing range radius. The coordinate of point

K is (x_k, y_k) , and the probability that the sensor node s_i covers point K is $P(K, s_i)$. From the Boolean perception model:

$$P(K, s_i) = \begin{cases} 1 & , d(s_i, s_j) \leq r_i \\ 0 & , d(s_i, s_j) > r_i \end{cases} \tag{2}$$

For any coordinate point (x, y) , as long as there is a sensor node in the node set C such that $P_{cov}(x, y, c_i) = 1$. The joint measurement coverage $P_{cov}(C)$ of the sensor node set is:

$$P_{cov}(C) = 1 - \prod_{c_j \in C} (1 - P_{cov}(x, y, c_j)) \tag{3}$$

The wireless sensor network coverage rate can be expressed as the ratio of the total monitoring coverage rate to the monitoring area:

$$P_{area} = \frac{\sum_{x=1}^m \sum_{y=1}^n P_{cov}(C)}{m \times n} \tag{4}$$

3. Sparrow Search Algorithm

Xue and Shen [8] propose the SSA. It is inspired by the foraging and anti-predation behavior of sparrows. It has the characteristics of fast convergence, strong stability and good robustness. This algorithm can solve practical problems such as nonlinear optimization and reducer design, and has broad development prospects.

The algorithm divides sparrows into two different types through different foraging strategies, namely producers and scroungers. Producers are responsible for finding food sources, and scroungers obtain food through producers. Sparrows with abundant food resources are more vulnerable to attack. Sparrow's choice of foraging strategy plays an important role in the individual's energy reserve, while most sparrows have low energy reserves. Sparrows on the outer edge of the group are more vulnerable to attack and gather in the middle. Sparrows in the middle of the group will approach their neighbors to reduce the probability of encountering danger. When the sparrow finds danger, it will scream and the whole sparrow flock will fly away.

Assuming a sparrow is x , in the search process, producers with high adaptability are given priority to get food. The producers are responsible for finding food and guiding the migration of the entire group. The position of producers is updated as follow:

$$X'_{i,j} = \begin{cases} X'_{ij} \cdot \exp\left(\frac{-i}{\alpha \cdot iter_{max}}\right) & , \text{if } R_2 < ST \\ X'_{i,j} + Q \cdot L & , \text{if } R_2 \geq ST \end{cases} \tag{5}$$

Where t is the current iteration. The value of j is $1, 2 \dots d$. $X'_{i,j}$ is the value of the j th dimension during the iteration of the i th sparrow. $iter_{max}$ is the maximum number of iterations. α is a random number in $(0,1]$. $R_2 \in [0,1]$ is the alarm value, and $ST \in [0.5,1.0]$ is the safety threshold. Q is a random number that obeys normal distribution. L is a $1 \times d$ matrix, and each element in the matrix is 1.

When $R_2 < ST$, it means that there is no danger around, and the producers perform a global search. When $R_2 \geq ST$, it means that sparrows have found a danger, and the entire sparrow group is evacuated to a safe area.

When the scroungers found that the producers had found the food, they immediately got the food. If the scroungers do not get food, they follow the producers to find food. The position update formula for scroungers is expressed as follows:

$$X'_{i,j} = \begin{cases} Q \cdot \exp\left(\frac{X'_{worst} - X'_{i,j}}{i^2}\right) & , \text{if } i > n/2 \\ X'_p + |X'_{i,j} - X'_p| \cdot A^+ \cdot L & , \text{otherwise} \end{cases} \tag{6}$$

Where X'_p is the best position occupied by the producer, and X'_{worst} is the worst position in the current global situation. A is a $1 \times d$ matrix, each element in the matrix is randomly assigned as 1 or -1, satisfying $A^+ = A^T(AA^T)^{-1}$. When $i > n/2$, it indicates that the i th scrounger has poor adaptability and may starve to death.

Assume that sparrows that perceive danger account for 10% to 20% of the total number of sparrows. The initial positions of these sparrows are randomly generated in the population. The mathematical model can be expressed as follows:

$$X'_{i,j} = \begin{cases} X'_{best} + \beta \cdot |X'_{i,j} - X'_{best}| & , \text{if } f_i > f_g \\ X'_{i,j} + K \cdot \left(\frac{|X'_{i,j} - X'_{worst}|}{(f_i - f_w) + \varepsilon}\right) & , \text{if } f_i = f_g \end{cases} \tag{7}$$

Where X'_{best} is the best position in the current global situation. β is the parameter that controls the step length. $K \in [-1,1]$ is a random number. f_i is the fitness value of the current sparrow. f_g is the current global best the fitness value. f_w is the current global worst fitness value. ε is the minimum constant to avoid zero-division-error.

When $f_i > f_g$, it means that the sparrows are at the edge of the group and are safe. When $f_i = f_g$, it means that the sparrows in the group are aware of the danger. K is the direction of the sparrow's movement, which is the parameters control step length.

4. Improved Sparrow Search Algorithm

In order to design a reasonable initial population of SSA, it is necessary to consider how to use as few individuals as possible to comprehensively characterize the distribution of all potential solutions in the entire space. If the initial population is randomly generated, it is difficult to find a solution to the problem. The good point set theory is used to improve the algorithm for population initialization so as to solve the problem of uneven population distribution. The good point set theory was first proposed by Hua Luogeng [9]. The good point set theory selects points with less deviation than randomly selected points, and its deviation is equivalent to the square root level of random selection [10].

The basic definition and structure of the good point set theory are as follows: Let G_s be the unit cube of s dimensional Euclidean space, that is $x \in G_s$, $x = (x_1, x_2, x_3, \dots, x_s)$. Where $0 \leq x_i \leq 1, i = 1, 2, \dots, s$, the point $r = (r_1, r_2, \dots, r_s)$ in G_s , let $r \in G_s$, then:

$$P_n(k) = \{(\{r_1k\}, \dots, \{r_s k\}), k = 1, 2, \dots, n\} \tag{8}$$

$P_n(k)$ is a good point set, and r is a good point. The deviation $\varphi(n)$ of the point set satisfies:

$$\varphi(n) = C(r, \varepsilon)n^{-1+\varepsilon} \tag{9}$$

where $C(r, \varepsilon)$ is a constant related only to r, ε (ε is any positive number), take $rk = 2\cos(2\pi k / p), 1 \leq k \leq D \cdot p$ is the smallest prime of satisfied $(p - D) / 2 \geq D$.

When generating the initial population, in the case of the same number of points, the points selected by the good point set are more uniform than the points randomly selected. Therefore, mapping the best points of G_s to the target solution space makes the initial population more ergodic, effectively avoiding the algorithm from falling into the local optimal value and stagnating, and improving the global convergence.

As the number of algorithm iterations continues to increase, the algorithm's optimization capability gradually tends to converge, and once it converges to the local optimum, it is difficult to jump out. Adding a weight factor to the algorithm can effectively improve the ability of the algorithm to jump out of the local optimal solution. This paper uses the weight factor to improve the original algorithm. The weight factor calculation formula is:

$$\omega = 2 - t \cdot \frac{2}{iter_{max}} \tag{10}$$

Where t is the number of iterations. $iter_{max}$ is the maximum number of iterations.

5. Simulation and Experiment Analysis

Aiming at the problem of low coverage of the random deployment of sensors for bridge monitoring, the improved SSA is used to optimize the deployment of sensor nodes. Based on the wireless sensor network's sensor node coverage optimization strategy for simulation, Matlab 2018b programming is selected to implement the sensor node coverage optimization program, and the environmental parameters are listed in Table 1.

Table 1. Set the same environmental parameters for the three algorithms

parameters	value
monitoring area	100m×100m
number of nodes	25
communication radius	10m
population size	25
maximum number of iterations	1000

To further compare the performance of the algorithm, this paper compares the particle swarm algorithm, the SSA, and the improved SSA. The coverage rate of sensor nodes is optimized by the particle swarm algorithm, and the coverage rate is relatively low, only 62.3%. The sensor nodes optimized by the original SSA are evenly distributed, with a coverage rate of 74.6%. The improved SSA algorithm has a more uniform sensor coverage, reduces the overlap of sensors, and the coverage rate is 77.1%, thereby optimizing the deployment of bridge monitoring. Sensor nodes can effectively improve sensor performance. The sensor node coverage optimization is shown in Figure 1.

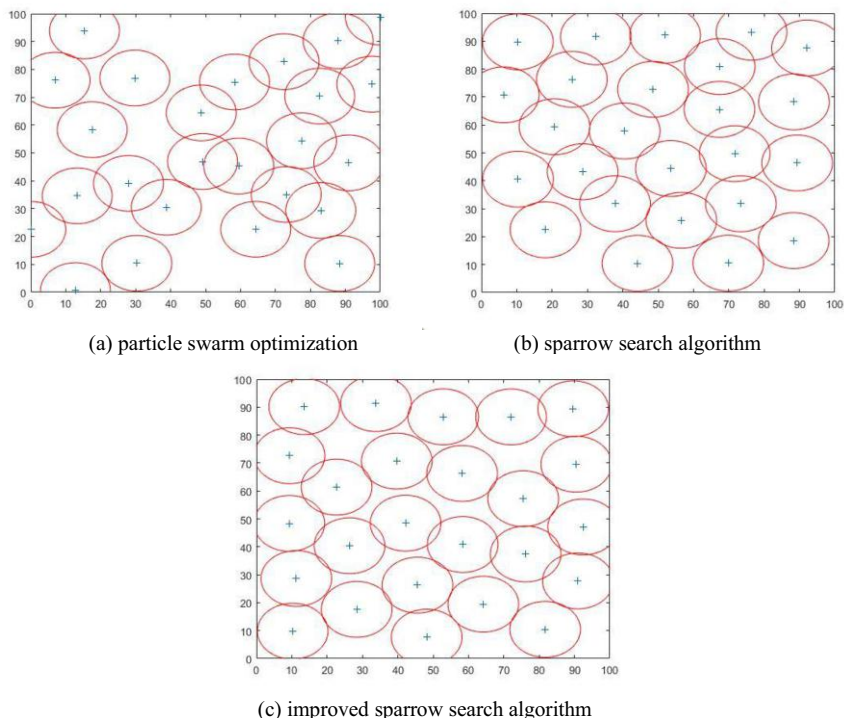


Figure 1. Sensor node coverage optimization

Figure 2 is an experimental comparison diagram of the particle swarm optimization, the original SSA and the improved SSA in this paper. The experiment deployed the same number of sensor nodes in the same monitoring area. Taking into account the interference of randomness on the experimental results, the three algorithms were tested ten times, and the average coverage rate was taken. It can be seen from Figure 2 that under the same experimental conditions, the improved SSA in this paper can effectively improve the node coverage of wireless sensor networks.

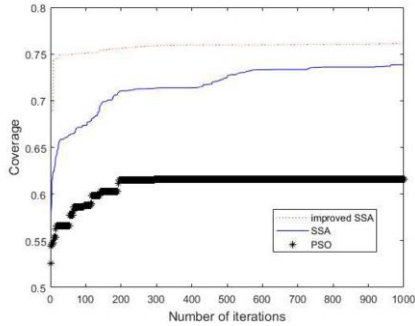


Figure 2. Comparison of coverage.

6. Conclusion

As a new algorithm, SSA has fewer parameters, high search efficiency, stable optimization results, and strong robustness, which provides an effective way for bridge monitoring sensor coverage. The original SSA has defects such as slow convergence speed and insufficient population vitality in the evolution process. This paper proposes an improved SSA and applies it to the problem of bridge monitoring sensor coverage optimization. The experiment verified the feasibility and effectiveness of SSA. In addition, the improved strategy proposed in this paper can further improve the search performance of the algorithm. The improved SSA search performance is higher and the optimization result is more stable.

References

- [1] Xu Yiming, Peng Yong, Zheng Chuhong, Liao Yi, Node energy balanced coverage strategy in WSNs based on improved PSO algorithm[J], *Transducer and Microsystem Technologies* **39** (2020), 29-32.
- [2] Huang Yuyue, Li Keqing, Coverage optimization of wireless sensor networks based on artificial fish swarm algorithm[J], *Application Research of Computers*, **30** (2013), 554-556.
- [3] Pan Hao, Fui Hua, Wireless sensor coverage multi-objective optimization improved based on cuckoo search algorithm[J], *Journal of Jilin Normal University (Natural Science Edition)* **38** (2017), 125-129.
- [4] Liu Zhouzhou, Wang Fubao, Zhang Kewang, Performance analysis of improved glowworm swarm optimization algorithm and the application in coverage optimization of WSNs[J], *Chinese Journal of Sensors and Actuators* **26** (2013), 675-682.
- [5] Zhang Guihong. Application of genetic algorithm in optimal configuration of maritime wireless sensors[J], *Ship Science and Technology* **41** (2019), 148-150.
- [6] Duan Yujun, Wang Yaoli, Chang Qing, Liu Xin, Wireless sensor deployment optimization based on improved IHACA-CpSPIEL algorithm [J], *Journal of Computer Applications* **40** (2020), 793-798.
- [7] Jia Zhe, Novel dichotomy fruit fly optimization algorithm in sensor deployment[J], *Microcontrollers and Embedded Systems* **19** (2019), 36-39.

- [8] Xue Jiankai, Shen Bo, A novel swarm intelligence optimization approach: sparrow search algorithm[J], *Systems Science and Control Engineering* **8** (2020), 22-34.
- [9] Hua Luogeng, Wang Yuan, *Application of number theory in approximate analysis* [M], Science Press, Beijing, 1978.
- [10] Fang Xian, Tie Zhixin, Li Jingming, et al, A parallel search good-point set glowworm swarm optimization of re-created population after clustering of multi-modal functions[J], *Journal of Zhejiang Sci-Tech University (Natural Sciences)* **37** (2017), 843-850.

Navigation Signal Design and Ranging Performance Evaluation of Cn Band Based on Satellite-to-Ground Link

Xiaofei CHEN^{a,b}, Xue WANG^{a,b,1}, Xiaochun LU^{a,b}, Jing KE^{a,b} and Xia GUO^b

^a*University of Chinese Academy of Sciences*

^b*National Time Service Center, Chinese Academy of Sciences*

Abstract. As the only priority frequency band for navigation services except L protected by ITU, the Cn band could provide navigation services to solve problems of spectrum congestion and vulnerability to interference faced in L using global navigation satellite systems. However, Cn band navigation still faces some problems such as limited-bandwidth and link uncertainty. To solve these problems, an orthogonal MSK signal is designed in this paper under Cn limited bandwidth constraint. The analysis results show that although its ranging performance of narrow correlation spacing has been deteriorated, the performance of wide correlation spacing has been improved, and it can reduce 98.7% power interference to adjacent radio astronomy band. On the other hand, the Cn band navigation signal test based on the satellite-to-ground link is carried out in this paper. The test results show that the trend of designed signals' ranging performance is consistent with the simulation results and its rain attenuation is 0.5-1dB.

Keywords. Cn band, navigation, ranging performance, orthogonal MSK

1. Introduction

Global navigation satellite systems (GNSS), including GPS, GLONASS, Galileo, BDS and other region navigation systems, is the infrastructure in modern society, which provides navigation, positioning and timing services. Although it plays an essential role in many fields, it still has some drawbacks that are difficult to be solved. One of which is vulnerable to interference, and another is spectrum congestion. Due to the low receiving signal power, GNSS is easily interfered with strong signals. Moreover, it is not easy to provide more diversified services because of the L band frequency resources' limitation. As the only priority frequency band for navigation services other than L protected by ITU, the Cn band (5010MHz-5030MHz) could provide navigation services to alleviate the above problems effectively.

In 2000, Cn band navigation was initially proposed and intended to be used by the Galileo system [1, 2], and its benefits and drawbacks were studied [3]. However, due to the limitation of the satellite's power and equipment complexity and cost of Cn band compared with L, the Cn band navigation was given up, and the development of the Cn

¹ Corresponding Author: Researcher on Navigation. National Time Service Center, Shuyuan East Road 3, 710600 Lintong distinct Xian Shaanxi, China; E-mail: wangxue@ntsc.ac.cn.

band navigation were also restricted [4, 5]. With the development of technology, these problems are no longer restrictions, and as a critical resource, Cn band navigation has been paid attention again. Hein, G.W. etc. analyzed the feasibility of developing navigation services on the Cn band [6]. Guoxiang Ai etc. studied the feasibility of Cn band to provide high-precision services [7]. Thomas Jost etc. compared the navigation performance of L band and Cn band with ground test [8]. Yinjiang Yan etc. provided the rain attenuation analysis of C-band navigation signal [9]. Rui Xue etc. proposed continuous phase modulation to reduce signal sidelobes and analyzed its performance [10, 11]. Ying Wang etc. studied the technology and prospect of Cn band navigation [12]. Yanbo Sun etc. discussed the scheme for L/C dual-frequency combined navigation signal [13]. The compatibility of the Cn band with its adjacent service has also been studied [14]. However, this paper presents the current lack of the ranging performance evaluation results of the Cn band navigation based on the satellite-to-ground link.

This paper designed orthogonal MSK signal under Cn limited bandwidth constraint and provided the test results of designed signal based on satellite-to-ground link, revealing the ranging and link performance of Cn band used for navigation.

2. Signal Design

2.1. Carrier Frequency Selection

Cn band and its adjacent telecommunication services are shown in fig.1. It can be seen that the Cn bandwidth is only 20MHz, its upper frequency sideband is the microwave landing service (MLS), and its lower frequency sideband is the radio astronomy service (RAS) and Galileo’s uplink service. Generally, RAS is easily interfered and has stringent restrictions on in-band interference. Therefore, considering maximum utilization efficiency and reducing the interference to RAS, the carrier frequency should be selected at the center or the upper sideband close to it.

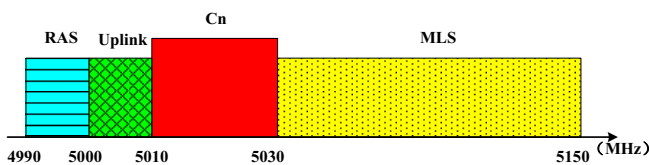


Figure 1. The frequency allocation of the Cn band and its adjacent service.

On the other hand, considering forward compatibility with GNSS and carrier and ranging code with the same frequency source to improve positioning accuracy, the carrier frequency must be an integer multiple of the ranging code frequency, and they all must be an integer multiple of 1.023MHz. Several alternative carrier frequencies are listed in table 1.

Table 1. Several alternative carrier frequencies.

Code Frequency (MHz)	Multiple relationship with Code	Multiple relationship with 1.023M	Carrier Frequency (MHz)
10.23	490	4900	5012.70
10.23	491	4910	5022.93
10.23	492	4920	5033.16

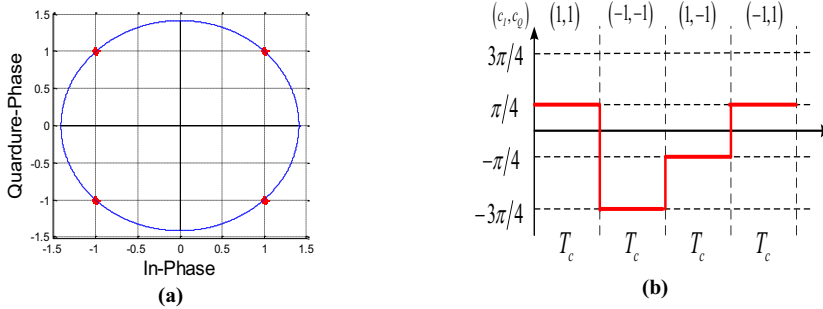


Figure 3. (a) BPSK (5) signal constellation; (b) BPSK (5) signal phase change.

Because of Cn band limited-bandwidth and compatibility, and to reduce the interference to adjacent services, especially RAS, it is necessary to reduce the power spectrum’s sidelobes. From Fig.3 (b), it can be seen that phase jumping causes larger sidelobes. To reduce the sidelobes of the power spectrum, a kind of orthogonal MSK modulation is proposed based on the traditional MSK modulation. It delays c_Q half chip compared with the c_I and is denoted as

$$c_d = c_Q(t - T_c/2) \tag{2}$$

where T_c is the duration time of one chip. Then the orthogonal MSK modulation could be expressed as

$$s_{O-MSK} = c_I \sin\left(\frac{\pi t}{T_c}\right) \cos(2\pi f_c t) - c_d \sin\left(\frac{\pi(t - T_c/2)}{T_c}\right) \sin(2\pi f_c t) \tag{3}$$

Its power spectrum and baseband waveform are shown in Fig 4. Fig.4 (a) describes that compared with BPSK, orthogonal MSK’s power spectrum sidelobe has a significant drop. The power in the band of RAS is denoted as

$$P_{RAS} = \int_{4990MHz}^{5000MHz} G(f) df \tag{4}$$

where $G(f)$ is the power spectrum. Through calculation, the power in band of RAS for BPSK(5) is -45.69dBW, and for orthogonal MSK(5) is -64.5dBW, it is found that compared with BPSK, the power of orthogonal MSK in the RAS band is reduced by 98.7%. The baseband waveform of orthogonal MSK is shown in Fig.4 (b). It can be seen that the orthogonal MSK replaces the rectangular chip with a sinusoidal chip. The orthogonal MSK’s phase could be expressed as

$$\theta(t) = \begin{cases} \left(-\frac{\pi t}{T_c} c_l + \frac{\pi}{2}\right) c_d, & 0 \leq t < \frac{T_c}{2} \\ \left(\frac{\pi t}{T_c} c_l - \frac{\pi}{2}\right) c_d, & \frac{T_c}{2} \leq t < T_c \end{cases} \quad (5)$$

Its signal constellation and phase change as Fig.5 shown. It can be seen that its phase changes linearly with time and there is no phase jump.

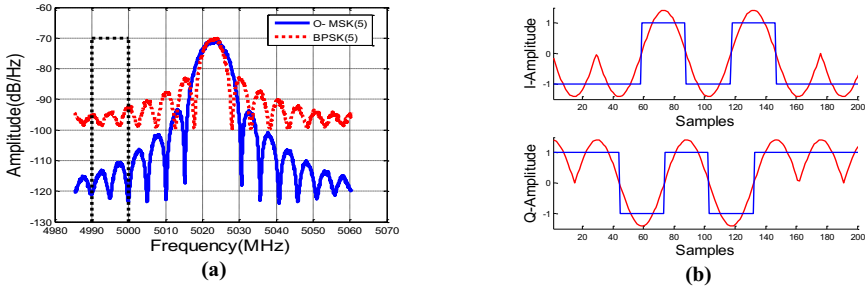


Figure 4. (a) Orthogonal MSK (5) signal power spectrum; (b) Orthogonal MSK (5) signal baseband waveform.

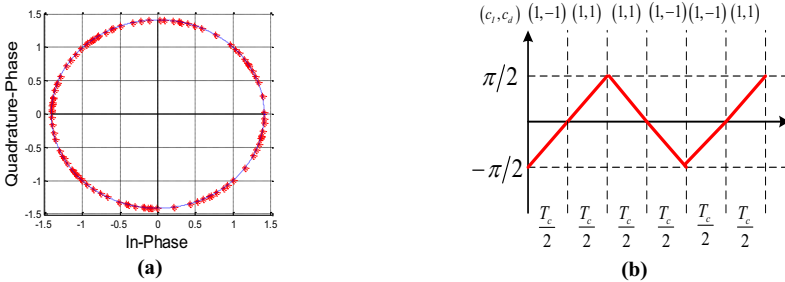


Figure 5. (a) Orthogonal MSK (5) signal constellation; (b) Orthogonal MSK (5) signal phase change.

The autocorrelation function curve determines the signal’s receiving performance. Fig.6 reveals the autocorrelation function of the two signals. It can be seen that compared with BPSK, the correlation function of orthogonal MSK becomes smoother especially near the correlation peak.

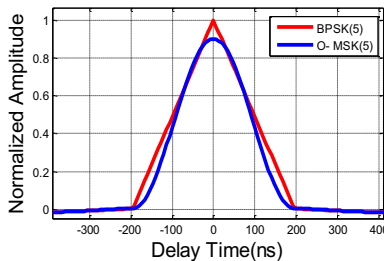


Figure 6. The autocorrelation function of BPSK (5) and orthogonal MSK (5).

2.3. Ranging Accuracy Analysis

With the ranging accuracy equation, signal’s ranging theoretical performance can be easily to get [16].

$$\sigma^2 = \frac{B(1-0.25BT) \int_{-\beta/2}^{\beta/2} G_{s0}(f) \sin^2(\pi f \Delta) df}{(2\pi)^2 (C/N_0) \left(\int_{-\beta/2}^{\beta/2} f G_{s0}(f) \sin(\pi f \Delta) df \right)^2} \times \left(1 + \frac{\int_{-\beta/2}^{\beta/2} G_{s0}(f) \cos^2(\pi f \Delta) df}{T(C/N_0) \left(\int_{-\beta/2}^{\beta/2} G_{s0}(f) \cos(\pi f \Delta) df \right)^2} \right) \tag{6}$$

where β is the radio frequency bandwidth and B is the bandwidth of the code tracking loop; $G_{s0}(f)$ is the normalized power spectrum; Δ is the early-to-late spacing; C/N_0 is the carrier-to-noise ratio; T is the correlation integration time. With the parameter of $\beta = 20\text{MHz}$, $B = 5\text{Hz}$, $\Delta = 200\text{ns}$, $T = 1\text{ms}$ and C/N_0 from 35dBHz to 50dBHz , the ranging accuracy curve can be drawn as shown in Fig.7 (a). Changing its early-to-late spacing from 200ns to 20ns , the ranging accuracy curve can be drawn as shown in Fig.7 (b).

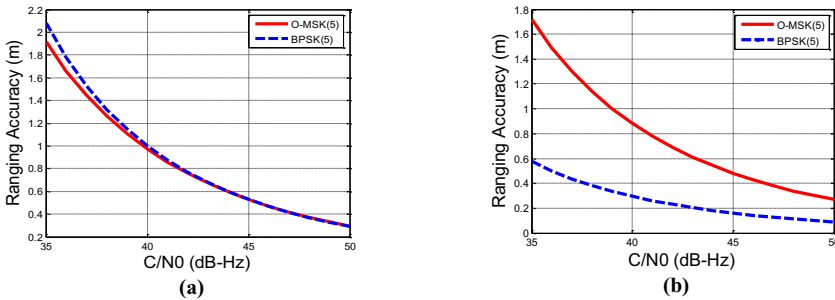


Figure 7. (a) The ranging accuracy of BPSK(5) and orthogonal MSK(5) with C/N_0 from 35 to 50 dB-Hz and early-to-late spacing 200ns; (b) The ranging accuracy of BPSK(5) and orthogonal MSK(5) with C/N_0 from 35 to 50 dB-Hz and early-to-late spacing 20ns.

Because of bandwidth limitation, orthogonal MSK (5) has more power in the bandwidth, and it could bring higher ranging accuracy in width correlation spacing -the early-to-late spacing more than 160ns- as Fig.7 (a) shown. While in narrow correlation spacing-the early-to-late spacing less than 160ns, the ranging accuracy of orthogonal MSK(5) becomes worse than BPSK(5), and that is because its autocorrelation curve becomes smoother and its anti-noise ability will be reduced as Fig.7(b) and Fig.6 shown. The ranging performance loss of narrow correlation spacing is the price to reduce the interference with out-of-band service.

3. Test Equipment and Methods

The testing system is shown in Fig.8, which includes signal generating equipment, transmitting antenna, on-orbit satellite, receiving antenna and equipment, and analysis equipment.

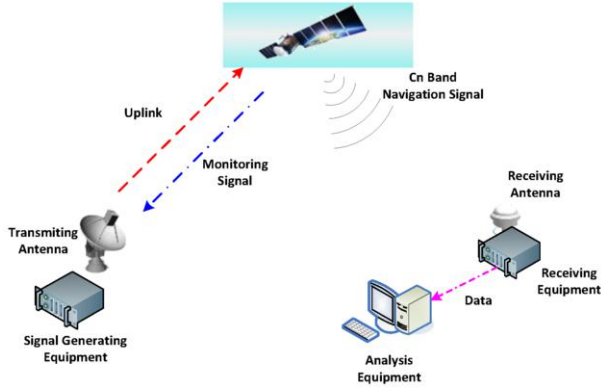


Figure 8. The testing equipment and interface between them.

The signal generating equipment generates the signal designed, including BPSK (5) and orthogonal MSK (5). The antenna transmits the generated signal by uplink and receives the monitoring signal to evaluate signal status. The satellite retransmits the Cn band navigation signal, and the receiving equipment receives it with an omnidirectional receiving antenna. Then the analysis equipment collects the pseudorange to analysis the signal's ranging performance. The signal generating equipment and receiving equipment are shown as Fig.9.

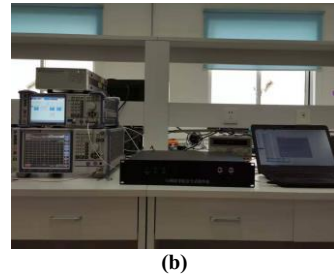
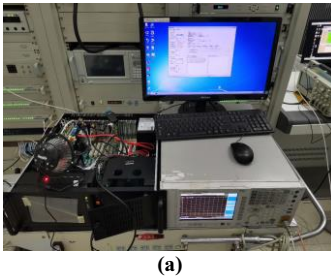


Figure 9. Testing equipment: (a) signal generating equipment; (b) receiving equipment.

During the test, the BPSK (5) signal is transmitted by the signal generating equipment. At the same time, the monitored signal is correct and stable, receiving equipment records the carrier-to-noise ratio and pseudorange. Then altering the signal to orthogonal MSK (5) and repeat the above work. The monitoring spectrum of BPSK (5) and orthogonal MSK (5) are shown in Fig.10. It can be seen that compared with BPSK (5), orthogonal MSK (5)'s sidelobe have significantly reduced.

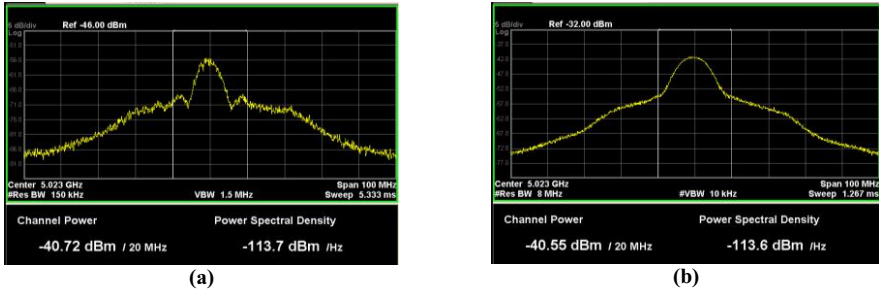


Figure 10. (a) BPSK(5) monitoring signal; (b) Orthogonal MSK(5) monitoring signal.

4. Results Analysis

Two scenarios were designed to test the ranging performance and rain attenuation during the test respectively.

4.1. Ranging Performance

Because the noise is a random variable, which is the main factor affecting the ranging performance and the satellite's motion following the Kepler equations, other variables are determined and changed slowly, the method based on high-order polynomial fitting can be used to measure the ranging performance. The receiving pseudorange y_d is

$$y_d = \rho + \Delta\rho + n \tag{7}$$

where ρ is the real range, and $\Delta\rho$ is the range bias, n is the thermal noise. Because $\rho + \Delta\rho$ could be fitted with an n -order polynomial,

$$y = a_0 + a_1x + a_2x^2 + \dots + a_nx^n \tag{8}$$

then choose the polynomial to minimize the Eq. (9).

$$\min\{E[y_d - y]\} = \min\{E[\rho + \Delta\rho - y + n]\} \tag{9}$$

where E expresses the mean operation. As the polynomial fitting eliminates the bias, what is left is the variation that affects the ranging performance, and then the Eq. (10) could be used to reflect the ranging performance.

$$\sigma = \sqrt{\frac{\sum_{i=1}^n (y_d[i] - y)^2}{n - 1}} \tag{10}$$

where $y_d[i]$ is the i^{th} receiving pseudorange.

The parameter of the receiver was chosen as $\beta = 20\text{MHz}$, $B = 5\text{Hz}$, $\Delta = 200\text{ms}$, $T = 1\text{ms}$, and data time is 10 minutes, the test result of ranging performance as shown in Fig.11 and Fig.12.

The carrier-to-noise ratio (upper) and the ranging performance (lower) of orthogonal MSK(5) and BPSK(5) are shown in Fig.11 (a) and Fig.11 (b) respectively. It can be seen that the carrier-to-noise ratio changes around 45.5dB-Hz, the ranging performance of orthogonal MSK(5) is 0.7607m, BPSK(5) is 0.7919m. Fig.12 shows the test results with carrier-to-noise ratio changes around 42.5dB-Hz, and their ranging performance reduces to 1.0518m and 1.0595m respectively.

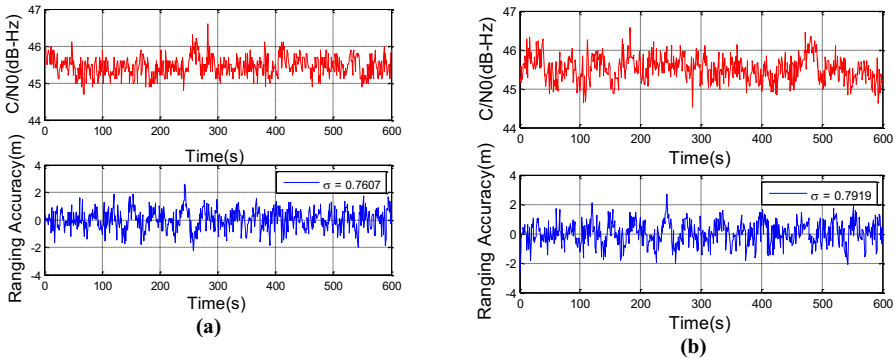


Figure 11. (a) The ranging performance of orthogonal MSK(5); (b) The ranging performance of BPSK(5).

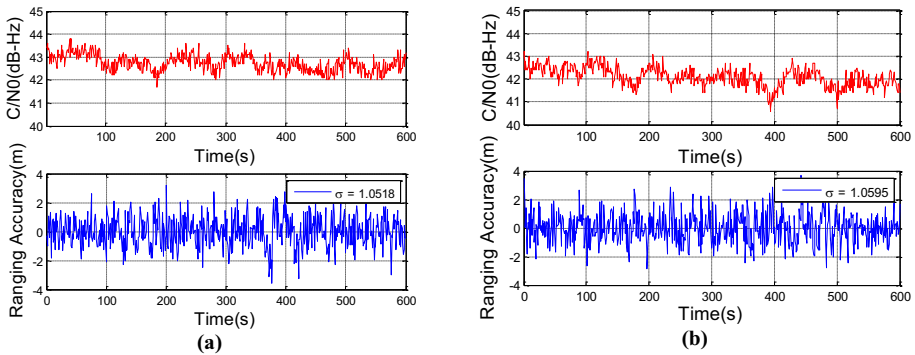


Figure 12. (a) The ranging performance of orthogonal MSK (5); (b) The ranging performance of BPSK (5).

Table 2 lists the ranging performance of simulation and test results with different carrier-to-noise to compare with test results and simulation results. It can be seen that the overall trend of test results is consistent with the simulation.

Table 2. The ranging performance of the two signals with different carrier-to-ratio.

Carrier-to-noise Ratio	Orthogonal MSK (5)		BPSK(5)	
	Simulation	Test Results	Simulation	Test Results
45.5 dB-Hz	0.4973m	0.7607m	0.4978m	0.7919m
42.5 dB-Hz	0.7136m	1.0518m	0.7222m	1.0595m

4.2. Rain Attenuation

In order to test the rain attenuation of the Cn band, choose the same testing time and location on different days when it rains and not. The test environment is shown in Fig.13.



Figure 13. (a) The omnidirectional receiving antenna when it rains; (b) The omnidirectional receiving antenna when it not.

Fig.14 shows the rain attenuation of the Cn band. There is the carrier-to-noise ratio of the receiver at the same testing time and location with orthogonal MSK (5) modulation on different days, the weather of the upper one is rain and the lower not. It can be seen that the rain attenuation is 0.5-1dB, which should be considered at the Cn band navigation system design.

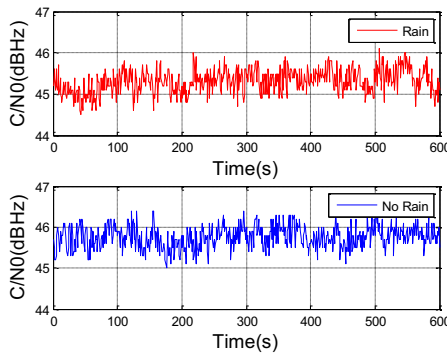


Figure 14. The test result of rain attenuation on the Cn band.

5. Conclusion

In this paper, a Cn band navigation signal test is carried out based on the satellite-to-ground link. It tested the ranging performance and link performance of the designed Cn band signals. The test results show that the ranging performance trend of the signals is consistent with the simulation, and its rain attenuation is 0.5-1dB. As an important frequency resource, the Cn band can be used for navigation service, but its rain attenuation should be considered at the Cn band navigation system design.

References

- [1] S.Loddo. Galileo Overall Architecture Definition Executive Summary, Issue 06, Dec 08, 2000.
- [2] Lindenthal, W.etal. GalileoSat. C-Band Navigation Payload Space Segment Assessments, WP 3300, Astrium, 2001.
- [3] Irsigler, M., Hein, G.W.&Schmitz-Peiffer,A. "Use of C-Band Frequencies for Satellite Navigation: Benefits and Drawbacks". GPS Solutions (2004): 119-139.
- [4] J.A. Avila-Rodriguez, S.Wallner, J.H.Won,B.Eissfeller,A.Schmitz-Peiffer, J.J.Fl-och,E.Colzi,J.-L.Gerner. Study on a Galileo Signal and Service Plan for C-band, Proceedings of ION GNSS 2008, Savannah, Georgia, USA, 2008.
- [5] Schmitz-Peiffer, A.,D.Felbach, F.Soualle, R.King, S.Paus, A.Fernandez, R.Jorgens-en, B.Eissfeller, J.A.Avila-Rodriguez, S.Wallner, T.Pany, J.H.Won, M.Anghileri, B.Lankl, E.Colzi. Assessment on the Use of C-Band for GNSS within the European GNSS Evolution Programme, Proceeding of ION GNSS 2008, Savannah, Georgia, USA, 2008.
- [6] Hein, G.W., J.A.Avila-Rodriguez, S.Wallner, J.H.Won. "Architecture for a Future C-band/L-band GNSS Mission-Part2: Signal Considerations and Related User Terminal Aspects". Inside GNSS (2009):47-56.
- [7] Guoxiang Ai, Lihua Ma, Huli Shi, Guanyi Ma, Ji Guo, Zhigang Li, Xiaohui Li, Haitao Wu, and Yujing Bian. "Achieving Centimeter Ranging Accuracy with Triple-frequency Signals in C-band Satellite Navigation Systems". Navigation 58, no. 1 (2011):59-68.
- [8] Thomas Jost, Wei Wang, Uwe-Carsten Fiebig, and Fernando Perez-Fontan. "Comparison of L-and C-band Satellite-to-indoor Broadband Wave Propagation for Navigation Applications". IEEE transactions on antennas and propagation 59, no. 10 (2011): 3899-3909.
- [9] Yinjiang Yan, Shuyu Zhang. "Rain Attenuation Analysis of C-Band Satellite Navigation Signal". Modern Navigation (2012): 157-159.
- [10] Rui Xue, Qingming Cao, Qiang Wei. "A Flexible Modulation Scheme Design for C-Band GNSS Signals[J], Mathematical Problems in Engineering". 2015.pt.11.(2015): 165097.1-165097.8.
- [11] Rui Xue, Huan Yu, and Qinglin Cheng. "Adaptive Coded Modulation Based on Continuous Phase Modulation for Inter-satellite Links of Global Navigation Satellite Systems". IEEE Access 6 (2018): 20652-20662.
- [12] Ying Wang, Yansong Meng, Xiaoxia Tao, Lei Wang, Zhe Su. Study on Cn Band Satellite Navigation, China Satellite Navigation Conference 2016, Chang Sha, China, 2016.
- [13] Yanbo Sun, Rui Xue, Dun Wang, et al. "General Modulation Scheme for L/C Dual-frequency Combined Navigation Signal". Journal of Harbin Engineering University (2018).
- [14] Zhiyun Li, Yan Bai, Xiaochun Lu. "Research on Radio Frequency Compatibility Evaluation Method of Cn Band Navigation Signal". Journal of Time and Frequency (2019): Vol. 42, 0087-0096.
- [15] Markus Irsigler, Günter W. Hein, Bernd Eissfeller, Andreas Schmitz-Peiffer etc. Aspects of C-band Satellite Navigation: Signal Propagation and Satellite Signal Tracking, Proceedings of the European Navigation, Conference ENC-GNSS 2002, Copenhagen, Denmark, 27-30 May 2002.
- [16] Betz, J.W. Design and Performance of Code Tracking for the GPS M Code Signal. Proceedings of the 13th International Technical Meeting of the Satellite Division of the Institute of Navigation ION GNSS 2000, Salt Lake City, UT, 19-22 September 2000, 2140-2150.

Prototype Design of a Ship Intelligent Integrated Platform

Fuxin WANG^{a,1}, Hao LUO^a, Yuan YU^b and Liyong MA^b

^aAutomation Engineering Department, Shanghai Marine Diesel Engine Research Institute, Shanghai, China

^bSchool of Information Science and Engineering, Harbin Institute of Technology, Weihai, China

Abstract. Recently, more and more attention has been paid to ship intelligence. However, in the sensor data acquisition network represented by the integrated ship bridge system, sensor data is collected and transmitted point to point, and the data coupling is strong, which is not conducive to the hierarchical utilization of data. To provide more effective data communication methods and flexible support to applications, intelligent integrated platform is needed by modern ship. Prototype design and practice of a ship intelligent platform is proposed, and the key technologies of the platform is discussed. The overall architecture of the platform is described. A hybrid network architecture with fieldbus, real-time ethernet and ethernet information network is introduced. And data storage architecture using NoSQL and hadoop distributed file system is described. The system can meet the real-time performance requirement of the control and information communication. An energy efficiency application based on the designed platform is developed, machine learning based method is employed to predict the heavy oil fuel consumption for ship navigation.

Keywords. intelligent ship, integrated platform, sensor, energy efficiency

1. Introduction

Recently, more and more attention has been paid to ship intelligence. The most typical traditional data collection system is the integrated ship bridge system [1]. The system collects sensor information centrally. And the functions of dispersed equipment such as navigation, driving control, collision avoidance, and surveillance are integrated into different systems with information fusion, which includes radar system, information system, and central driving control system. In this data collection method, the point-to-point communication method is used between the platform and shipboard, equipment, and subsystems. In this connection architecture, the coupling of the data is strong, and it brings inconvenience to the data update. It can be seen that modern ships need a smart application-oriented integrated intelligent platform design that can provide more effective data communication methods and more flexible support.

In recent years, some ships have carried out intelligent practice to change traditional data collection methods [2-5]. In 2015, South Korea's Hyundai Heavy Industry Group

¹ Corresponding author, Fuxin WANG, Automation Engineering Department, Shanghai Marine Diesel Engine Research Institute, Shanghai, China; E-mail: 18801909355@163.com.

and Accenture jointly developed a new intelligent ship interconnection platform. The platform collects real-time dynamic information about the ship's position, equipment operation information, shipboard cargo information, and surrounding sea area information through sensors, and visually displays the analysis results through big data analysis. Danish Maersk selected the ShipManager system developed by DNVGL Group to collect and analyze information on its ships. GE Power Conversion has developed Visor Insight, a management system that can be applied to remotely monitor drive equipment, power distribution facilities, electric motors, and navigation systems.

Intelligent ships are also being rapidly developed in China. In 2015, the i-DOLPHIN ship was released by China State Shipbuilding Corporation, which means that China's first intelligent ship has started the design and construction. After that, China's intelligent ships flourished. For example, in 2017, the world's first intelligent ship "Dazhi" was officially delivered. This ship has the first ship smart operation and maintenance system with autonomous learning. In 2018, the 400,000-ton intelligent super-large ore ship "Mingyuan" was delivered to use. It is the world's largest intelligent ship with decision-making functions such as assisting collision avoidance. In 2019, the "Zhiteng" ship successfully carried out a demonstration of autonomous navigation and autonomous collision avoidance. The ship is equipped with an intelligent situational awareness system, autonomous navigation decision-making system, and autonomous driving control system. These are all necessary parts for ocean-going autonomous ships.

As more and more intelligent ships are built and used, research on intelligent ship integrated platforms is an urgent task. There are many sensors in intelligent ships, different types, and different communication methods. Big data and data monitoring applications for intelligent ship are proposed in [1,2], but the simple network cannot support complex intelligent applications. Navigation based on information fusion is discussed in [3-5], these applications do not consider the control data network. In this paper, network architecture considering both fieldbus and information network with high performance data storage method is introduced. With the development of data driven method [6-8], energy efficiency applications are focused by many researches [9,10], and intelligent methods have been employed to energy efficiency applications. Neural network is the common method for energy consumption prediction [11,12]. In this paper an energy consumption prediction application based on machine learning is reported.

The main contribution of this paper is as follows. First, a ship integrated platform is designed and tested in Shanghai Marine Diesel Engine Research Institute. The key technologies of the intelligent integrated platform for ship is discussed. Second, a hybrid network architecture of fieldbus and real-time ethernet for sensor data collection is introduced, and hierarchical database architecture design is discussed. Third, an application for energy efficiency is implemented based on the integrated platform.

2. Architecture of Intelligent Integrated Platform

The China Classification Society officially released the Intelligent Ship Specifications on December 1, 2015, and the specifications have come into effect on March 1, 2016. This specification is the world's first intelligent ship specification and proposes to develop an intelligent integrated platform. In this specification, the intelligent integrated platform is one of the seven important components of intelligent ships. The intelligent integrated platform can realize data collection and unified storage, and it can also provide support for data analysis and intelligent applications.

The overall architecture of an intelligent integrated platform is divided into seven parts, they are data acquisition, data storage, data integration, information application, information display, standard specification system, and information security system, as shown in Figure 1.

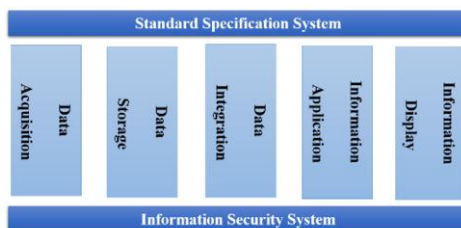


Figure 1. Architecture of intelligent integrated platform for ship.

In data acquisition, sensing devices, such as sensors, controllers, signal acquisition devices, and data acquisition devices, are used to collect the required data. In data storage, the dataset needs to be established to support the decision-making and management process, which is subject-oriented, integrated, relatively stable, and reflects historical changes. In data integration, the data is cleaned up and processed systematically, classified and summarized, analyzed, and integrated. In information applications, business applications are provided for customers. In information display, corresponding data mining results are provided human-computer interaction. The standard specification system and information security system are two other key requirements for data application. For the design and construction of a prototype platform, data acquisition, data storage, and data integration are the most important parts. These will be discussed in detail as follows.

3. Network and Data Storage Design for Data Acquisition

Due to the large number of ship information systems and the huge amount of information transmission between each other, the ship communication network has the characteristics of complexity and multi-channel, so the ship data communication network is the key problem for the design of an intelligent integrated platform. Ship communication must meet the needs of parallel multi-tasks. The main types of data transmission include audio, video, and other sensor data, the timeliness, and accuracy of data communication are also required.

In the design of the communication network, the method of separating the control layer and information layer is adopted. Because the control layer requires higher response time and synchronization time, CAN Fieldbus is applied. And the information layer does not require such high response time and synchronization time, real-time ethernet with better compatibility is adopted. EPA (Ethernet for plant automation) is a new open real-time ethernet standard for industrial field equipment, which is suitable for real-time communication of industrial measurement and control systems. EPA adds CSME between the data link layer and the network layer specified in ISO/IEC8802.3 protocol to control the transmission of real-time and non-real-time data packets, so as to ensure the timely transmission and processing of EPA periodic information and non-periodic information such as alarm. In addition, the EPA network divides the control network into

several micro-network segments separated by a bridge in order to avoid the occurrence of conflict, and the communication in each micro-network segment does not interfere with each other. EPA based distributed network control system has been successfully applied in the factory. The real-time response time of EPA is less than 1 millisecond. EPA control module CON21 and ECS-100 are used in field control terminals and system center stations respectively. The control layer adopts distributed control, which is connected to dual redundant optical fiber Ethernet through a redundant gateway to realize the interconnection with the upper information management network. The network architecture is illustrated in Figure 2. As far as we know, this is the first integrated platform network design and practice for fieldbus, real-time ethernet and information ethernet, and its network real-time performance can reach 1 millisecond.

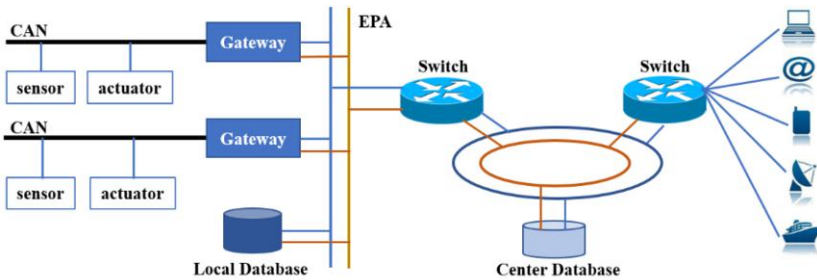


Figure 2. Network architecture for integrated platform.

The database is the key design of data storage. The real-time database used in the integrated platform consists of two parts: the field database and the central database. The on-site database is responsible for data acquisition, compressed storage, and processing to provide on-site monitoring data services. The central database is responsible for synchronous storage and processing of all field data to provides a data query interface. AutoNet, which is widely used in ship engineering, is used as the field real-time database system of the platform.

Different from the traditional solution to achieve high performance of the system through the high reliability of hardware, some NoSQL distributed databases can use low-cost servers to build large-scale clusters and provide unified external memory access services. In the current data explosion and cost considerations, more and more distributed database products are used to replace RDBMS. Some studies have compared the performance of some NoSQL and SQL databases in reading, write, delete and instantiation operations. Although for different databases, their performance will change with different operations, but in the case of large operands, NoSQL shows its advantages. As a typical representative of the distributed NoSQL database, HBase has been widely used in sensor data storage. HBase is a distributed NoSQL open source database, it has the characteristics of high availability, high performance, and scalability of the distributed storage system. HDFS (Hadoop distributed file system) is a kind of distributed file system, which is used to solve the storage problem of massive data. It can manage the files on multiple machines at the same time and realize the file sharing on multiple machines. It has strong fault tolerance and stability. And it can be combined with HBase to achieve efficient data file support. In the prototype design of the intelligent integrated platform in this paper, HBase is used as the central database and HDFS as the file system. The database in the platform is shown in Figure 3. This database architecture with the previous network architecture, can fully guarantee the real-time performance of

data. Compared with the schemes in the literature [1-3], this scheme not only considers the hybrid network architecture, but also considers the database architecture. The scheme ensures the real-time performance of the integrated platform more completely.

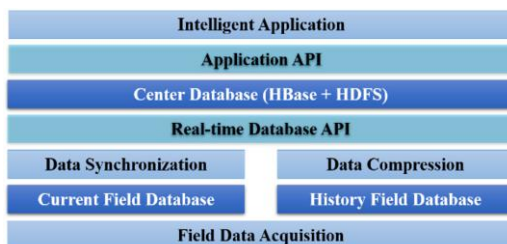


Figure 3. Database in the integrated platform.

4. Application of Ship Energy Efficiency with Data Integrated Platform

The purpose of ship intelligent integrated platform is to provide a unified data interface and support for data analysis in different subsystems. The value of these integrated data can only be fully exploited by using statistics, machine learning, and other artificial intelligence technologies. Various data analysis applications are the further expansion and extension of the data in the intelligent integrated platform. With the emergence of emerging technologies such as big data and artificial intelligence, as well as the continuous maturity of data mining technology, these data analysis applications will bring more opportunities and challenges. These applications can extract hidden and potentially useful information and knowledge from a large number of practical application data. This is the characteristic and construction purpose of an intelligent platform. The ship's intelligent integrated platform provides strong support for these applications.

Using the intelligent integrated platform, some preliminary applications have been developed by our team. In this paper, we introduce the application of energy efficiency. Energy-saving has always been an important goal in ship design and operation management [6]. A basic ship navigation energy consumption model is established, which is based on the marine environment, ship navigation, and energy consumption data. The marine environment includes wind speed, wind direction, air temperature, air pressure, and sea wave height. These data come from the marine environment sensing module of the ship, which is obtained through its measurement or shore-based communication. The ship navigation module can provide the ship's actual speed and heading data information. The diesel engine real-time monitoring module can provide speed and diesel consumption data. Although these data come from different subsystems, a unified interface can be used to access and query historical data through the integrated platform.

When the diesel engine working condition and the ship's heading are stable, the data within 20 minutes is collected and preprocessed. Finally, the shipping speed, ship heading course, wind speed, wind direction, air temperature, air pressure, wave height, and diesel consumption are used for data analysis. All this data comes from a newly constructed cargo ship undergoing sea trials. A total of 328 pieces of valid data were collected, of which 295 were used as training dataset for model building, and the remaining 33 were used as test dataset for model testing.

The support vector machine (SVM) is a classifier developed with the statistical learning theory. The feature of SVM is that a largest separation hyperplane is found for classification. The hyperplane is established by maximizing the interval between the two classes in the feature space. This problem can be solved with optimization algorithm. The support vector machine method based on statistical learning theory can also be used to deal with regression problems, this method is called support vector regression (SVR). In our application based on the integrated platform, SVR is used to estimate heavy oil fuel consumption. At first, a model is established with the training dataset employing SVR. Then the trained model is used to estimate the heavy oil fuel consumption according to other known parameters. The experimental results show that the maximum relative error is 9.11%, the average relative error is 3.35%, the maximum absolute error is 3.5 kg/km, and the average absolute error is 1.4 kg/km. The results are shown in Figure 4. This shows that the model can be used to predict fuel consumption. And the performance of this method is similar to that reported in other literatures [9-12]. Moreover, the model can be used to select the best ship speed based on energy efficiency under known navigation conditions. More accuracy prediction model can be established when more data are collected or some other intelligent algorithms are employed.

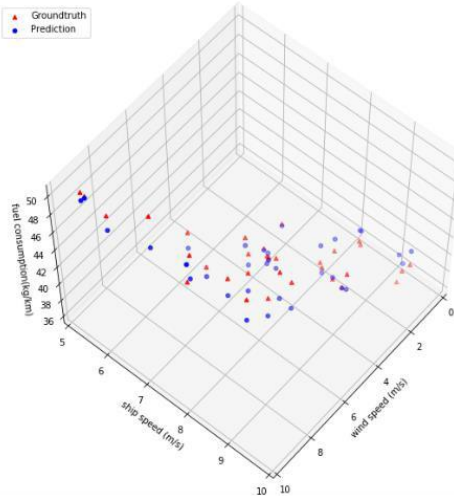


Figure 4. Prediction results of fuel consumption.

5. Conclusions

Prototype design and practice of a ship intelligent platform is proposed. The network architecture and data storage architecture are the key designs of the platform. A hybrid network architecture of fieldbus and real-time ethernet for sensor data is used, and NoSQL database and hadoop distributed file system is employed to obtain real-time performance. An energy efficiency model is established with a machine learning method, and it can be used to predict fuel consumption. The example shows the supporting role of the platform for intelligent applications.

The limitations of this study mainly include two aspects. First, more data is needed to test the network performance. Second, more intelligent applications are needed to

verify the rationality of the platform design. We will apply the platform design to more ships in the future. And more network and database testing will be done. At the same time, more intelligent applications based on the integrated platform will be developed.

References

- [1] Li G, Deng X, Zhou M, et al. Research on data monitoring system for intelligent ship. *Lecture Notes in Electrical Engineering*. 2020 Nov;634(1):234-241.
- [2] Jia S, Ma L, Zhang S. Big data prototype practice for unmanned surface vehicle. In: *Proceedings of the 4th International Conference on Communication and Information Processing*; 2018 Nov 2-4; Qingdao, China: ACM Press; p. 43-47.
- [3] Zhang C, Guo C, Guo MZ. Information fusion based on artificial intelligence method for SINS/GPS integrated navigation of marine vessel. *Journal of Electrical Engineering and Technology*. 2020 May;15(3):1345-1356.
- [4] Xu K. A self-selective correlation ship tracking method for smart ocean systems. *Sensors*. 2019 Feb;19(4):8211-8219.
- [5] Ma L, Xie W, Huang H. Convolutional neural network based obstacle detection for unmanned surface vehicle. *Mathematical Biosciences and Engineering*. 2020 Jan;17(1):845-861.
- [6] Malde K, Handegard NO, Eikvil L, Salberg A. Machine intelligence and the data-driven future of marine science. *ICES Journal of Marine Science*. 2019 Jul;77(4): 1274-1285.
- [7] Prera LP, Mo BS. Machine intelligence based data handling framework for ship energy efficiency. *IEEE Transactions on Vehicular Technology*. 2017 Oct;66(10):8658-8666.
- [8] Park JH, Storch RL. Overview of ship-design expert systems. *Expert Systems*. 2002 Jul;19(3):136-141.
- [9] Bharati S, Podder P, Mondal M. Visualization and prediction of energy consumption in smart homes. *International Journal of Hybrid Intelligent Systems*. 2020 Feb;16(2):81-97.
- [10] Bharati S, Rahman MA, Mondal R, Podder P, et al. Prediction of energy consumed by home appliances with the visualization of plot analysis applying different classification algorithm. *Advances in Intelligent System and Computing*. 2020 Nov;1014(1):246-257.
- [11] Farag YB, Ölçer AI. The development of a ship performance model in varying operating conditions based on ANN and regression techniques. *Ocean Engineering*. 2020 Feb;198(1): Article No. 106972.
- [12] Le LT, Lee G, Park KS, Kim H. Neural network-based fuel consumption estimation for container ships. *Korea. Maritime Policy & Management*. 2020 Jul;47(5): 615-632.

The Bounded Traveling Wave Solutions of a (3+1) Dimensions mKdv-ZK Equation

Yuzhong ZHANG¹

School of Mathematics and Information Technology, Yuxi Normal University, Yuxi, Yunnan, China

Abstract. The bounded traveling waves solutions of a (3+1) dimensions mKdv-ZK equation be investigated by using method of dynamical systems. The exact expressions of bounded periodic waves, solitary waves and kink waves are given. Under fixed parameter condition, the planar simulation graphs of the bounded periodic waves, solitary waves and kinks are obtained by using the software Mathematica 7.

Keywords. mKdv-VZK equation, singular point, periodic waves, solitary waves, kinks

1. Introduction

The well-known Kdv equation has perfect dynamic properties. Since Kdv equation was proposed, many scholars have done a lot of research and obtained a lot of results. Some researchers have extended the Kdv equation, and proposed Kdv-B, mKdv, mKdv-ZK equation, etc. In reference [1], the mKdv-ZK equation is studied with homogeneous balance method, and some traveling solution is given. In this paper, the Vries – Zakharov – Kuznetsov (mKdV-ZK) equation is investigated by using the bifurcation method [2, 3, 4] of dynamical systems:

$$u_t + a_1 u^2 u_x + a_2 u_{xxx} + a_3 (u_{yy} + u_{zz})_x = 0, \quad (1)$$

where a_1 , a_2 and a_3 are real constants, the properties of singular point of mKdv-ZK plane are obtained, and the bifurcation of phase portraits are given. The exact solutions of bounded periodic waves, solitary waves and kinks of mKdv-ZK are obtained by using phase portraits. Under fixed parameter conditions, the plane simulation diagrams of the bounded periodic waves, solitary waves and kinks were obtained by using the software Mathematica 7.

¹ Corresponding Author: Yuzhong Zhang, associate professor, mainly engaged in study of differential equations, School of Mathematics and Information Technology, Yuxi Normal University, Yunnan, China; E-mail: zh1011@yxnu.net.

2. The Bifurcation of Phase Portraits

Letting $\xi = x + y + z - ct$, $u(x, y, z, t) = \varphi(\xi)$, where c is wave velocity, then Eq. (1) can be transformed into:

$$(a_1\varphi^2 - c)\varphi' + (a_2 + 2a_3)\varphi''' = 0. \tag{2}$$

Integrating straightforwardly Eq. (2), and taking integral constant as 0, we get

$$(a_2 + 2a_3)\varphi'' + \frac{1}{3}a_1\varphi^3 - c\varphi = 0. \tag{3}$$

Letting $\varphi' = y$, then Eq. (3) can be transformed into a plane system:

$$\begin{cases} \frac{d\varphi}{d\xi} = y \\ \frac{dy}{d\xi} = \frac{1}{a_2+2a_3}(-\frac{1}{3}a_1\varphi^3 + c\varphi) \end{cases}. \tag{4}$$

Obviously, plane system (4) is a Hamiltonian system, where Hamitonian function as follow Eq. (5):

$$H(\varphi, y) = \frac{1}{2}y^2 + \frac{1}{a_2 + 2a_3}(\frac{1}{12}a_1\varphi^4 - \frac{1}{2}c\varphi^2) = h. \tag{5}$$

Letting $\alpha = \frac{-a_1}{a_2+2a_3}$, obviously, the plane system (4) has the properties of singular points as follow:

- when $a_1c > 0$, the plane system(4) has three singular points $(-\sqrt{\frac{3c}{a_1}}, 0)$, $(0, 0)$, $(\sqrt{\frac{3c}{a_1}}, 0)$ and $H(\sqrt{\frac{3c}{a_1}}, 0) = H(-\sqrt{\frac{3c}{a_1}}, 0)$
 - If $\alpha < 0$, then $(0, 0)$ is a saddle point, $(\sqrt{\frac{3c}{a_1}}, 0)$ and $(-\sqrt{\frac{3c}{a_1}}, 0)$ are two center points.
 - If $\alpha > 0$, then $(0, 0)$ is a center point, $(\sqrt{\frac{3c}{a_1}}, 0)$ and $(-\sqrt{\frac{3c}{a_1}}, 0)$ are two saddle points.
- When $a_1c < 0$, then plane system (4) has only one singular point $(0, 0)$,
 - If $\alpha > 0$, then $(0, 0)$ is a saddle point.
 - If $\alpha < 0$, then $(0, 0)$ is a center point.

From the above analysis, the phase portraits of plane system (4) can be drawn by Eq.(5).

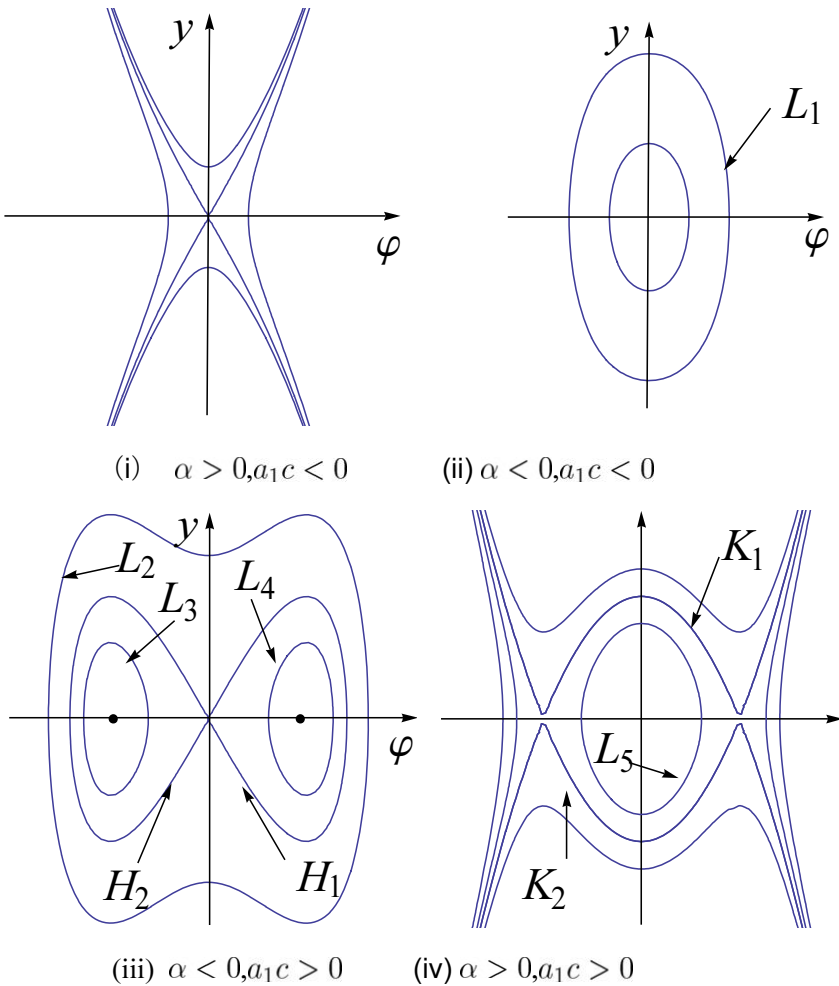


Figure 1 The bifurcation of phase portraits of the system (4)

3. The Traveling Wave Solutions

Setting $(\varphi_0, 0)$ as a initial point, substituting it into Eq.(5), then we get $H(\varphi_0, 0) = \frac{1}{a_2+2a_3}(\frac{1}{12}a_1\varphi_0^4 - \frac{1}{2}c\varphi_0^2) = h_0$.

Letting $H(\varphi, y) = h_0, H(\varphi, 0) = h_0$, then we get

$$\frac{1}{2}y^2 = \frac{1}{a_2 + 2a_2}(\frac{1}{12}a_1\varphi^4 - \frac{1}{2}c\varphi^2) - \frac{1}{a_2 + 2a_3}(\frac{1}{12}a_1\varphi^4 - \frac{1}{2}c\varphi^2), \tag{6}$$

$$y = \pm \sqrt{\frac{-a_1}{6(a_2 + 2a_3)}(\varphi - \varphi_0)(\varphi + \varphi_0)(\varphi^2 + \varphi_0^2 - \frac{6c}{a_1})}, \tag{7}$$

$$\frac{d\varphi}{d\xi} = \pm \sqrt{\frac{\alpha}{6}(\varphi - \varphi_0)(\varphi + \varphi_0)(\varphi^2 + \varphi_0^2 - \frac{6c}{a_1})}, \tag{8}$$

where $\alpha = \frac{-a_1}{a_2+2a_3}$. The traveling wave solutions can be obtained by integrating Eq. (1).

3.1. The Periodic Wave Solutions

Since the traveling wave corresponding to the smooth closed orbit of the phase portrait is a bounded periodic wave, the solution of the bounded smooth periodic wave can be obtained by integrating Eq.(8) along the periodic orbit.

Property 1. Under the following conditions, Eq. (1) has a bounded smooth periodic wave.

(i). If $\alpha < 0, a_1c < 0, \varphi_0 < 0$, then the periodic orbit L_1 (see Figure 1(ii)) determines a periodic wave and its solution is:

$$\varphi = \varphi_0 \operatorname{cn}\left(\left(\varphi_0^2 - \frac{3c}{a_1}\right)\sqrt{\frac{-2\alpha}{3}}\xi, \sqrt{\frac{\varphi_0^2}{2(\varphi_0^2 - \frac{3c}{a_1})}}\right). \tag{9}$$

(ii). If $\alpha > 0, a_1c > 0, -\sqrt{\frac{3c}{a_1}} < \varphi_0 < 0$, then the periodic orbit L_5 (See Figure 1(vi)) determines a periodic wave and its solution is:

$$\varphi = \varphi_0^* - \frac{\varphi_0^{*2} - \varphi_0^2}{\varphi_0^* - \varphi_0 + 2\varphi_0 \operatorname{sn}^2\left((\varphi_0^* - \varphi_0)\sqrt{\frac{\alpha}{12}}\xi, \frac{2\sqrt{-\varphi_0^*\varphi_0}}{\varphi_0^* - \varphi_0}\right)}, \tag{10}$$

where $\varphi_0^* = \sqrt{\frac{6c}{a_1} - \varphi_0^2}$.

(iii). If $\alpha < 0, a_1c > 0, \varphi_0 < -\sqrt{\frac{6c}{a_1}}$, then the periodic orbit L_2 (See Figure 1(iii)) determines a periodic wave and its solution is Eq.(9)

(vi). If $\alpha < 0, a_1c > 0, -\sqrt{\frac{6c}{a_1}} < \varphi_0 < -\sqrt{\frac{3c}{a_1}}$, then the periodic orbit L_3 (See Figure 1(iii)) determines a periodic wave and its solution is:

$$\varphi = \varphi_0^* - \frac{2\varphi_0^*(\varphi_0^* - \varphi_0)}{\varphi_0^* - \varphi_0 + (\varphi_0^* + \varphi_0)\operatorname{sn}^2\left((\varphi_0^* - \varphi_0)\sqrt{\frac{-\alpha}{12}}\xi, \frac{\varphi_0^* + \varphi_0}{\varphi_0^* - \varphi_0}\right)}, \tag{11}$$

where $\varphi_0^* = \sqrt{\frac{6c}{a_1} - \varphi_0^2}$.

(v). If $\alpha < 0, a_1c > 0, 0 < \varphi_0 < \sqrt{\frac{3c}{a_1}}$, then the periodic orbit L_4 (See Figure 1(iii)) determines a periodic wave and its solution is:

$$\varphi = -\varphi_0 + \frac{2\varphi_0(\varphi_0^* + \varphi_0)}{\varphi_0^* + \varphi_0 + (\varphi_0 - \varphi_0^*)\text{sn}^2\left(\left(\varphi_0^* + \varphi_0\right)\sqrt{\frac{-\alpha}{12}}\xi, \frac{\varphi_0^* - \varphi_0}{\varphi_0^* + \varphi_0}\right)}, \tag{12}$$

where $\varphi_0^* = \sqrt{\frac{6c}{a_1} - \varphi_0^2}$.

Proof: (i). when the conditions $\alpha < 0, a_1c < 0$ are satisfied, taking $\varphi_0 < 0$, then the orbit L_1 (See Figure 1(ii)) passing through initial point $(\varphi_0, 0)$ is a smooth closed orbit, and on $\xi - \varphi$ plane, the traveling wave corresponding to it is bounded periodic wave.

Eq.(8) can be transformed into:

$$\frac{d\varphi}{\sqrt{(\varphi - \varphi_0)(-\varphi - \varphi_0)(\varphi^2 + \varphi_0^2 - \frac{6c}{a_1})}} = \pm \sqrt{\frac{-\alpha}{6}} d\xi. \tag{13}$$

Integral Eq.(13) along L_1

$$\int_{\varphi_0}^{\varphi} \frac{d\varphi}{\sqrt{(-\varphi_0 - \varphi)(\varphi - \varphi_0)(\varphi^2 + \varphi_0^2 - \frac{6c}{a_1})}} = \pm \sqrt{\frac{-\alpha}{6}} \int_0^{\xi} d\xi, \tag{14}$$

Using the elliptic integral formula 259. 00 in reference [5] to calculate (20), the bounded smooth periodic wave solution (9) can be obtained.

(ii). when the conditions $\alpha > 0, a_1c > 0$ are satisfied, taking φ_0 to satisfy $-\sqrt{\frac{3c}{a_1}} < \varphi_0 < 0$, then the orbit L_5 (See Figure 1(vi)) passing through initial point $(\varphi_0, 0)$ is a smooth closed orbit, and on $\xi - \varphi$ plane, the traveling wave corresponding to it is bounded periodic wave.

Eq.(8) can be transformed into:

$$\frac{d\varphi}{\sqrt{(\varphi - \varphi_0)(\varphi + \varphi_0)\left(\varphi - \sqrt{\frac{6c}{a_1} - \varphi_0^2}\right)\left(\varphi + \sqrt{\frac{6c}{a_1} + \varphi_0^2}\right)}} = \pm \sqrt{\frac{\alpha}{6}} d\xi. \tag{15}$$

letting $\varphi_0^* = \sqrt{\frac{6c}{a_1} - \varphi_0^2}$, Integral Eq.(15) along L_1

$$\int_{\varphi_0}^{\varphi} \frac{d\varphi}{\sqrt{(\varphi_0^* - \varphi)(-\varphi_0 - \varphi)(\varphi - \varphi_0)(\varphi + \varphi_0^*)}} = \pm \sqrt{\frac{\alpha}{6}} \int_0^{\xi} d\xi, \tag{16}$$

Using the elliptic integral formula 255.00 in reference [5] to calculate Eq.(18), the bounded periodic wave solution Eq.(10) can be obtained.

(iii) when the conditions $\alpha < 0, a_1c > 0$ are satisfied, taking $\varphi_0 < -\sqrt{\frac{6c}{a_1}}$, then the orbit L_2 (See Figure 1(iii)) passing through the initial point $(\varphi_0, 0)$ is a smooth closed orbit, and on $\xi - \varphi$ plane, the traveling wave corresponding to it is a bounded periodic wave. In the same way as (i), the periodic wave solution is Eq.(9).

(vi) when $\alpha < 0, a_1c > 0$, taking φ_0 to satisfy $-\sqrt{\frac{6c}{a_1}} < \varphi_0 < -\sqrt{\frac{3c}{a_1}}$, then the orbit L_3 (See Figure 1(iii)) passing through the initial point $(\varphi_0, 0)$ is a smooth closed orbit, and on $\xi - \varphi$ plane, the traveling wave corresponding to it is a bounded periodic wave.

(8) can be transformed into:

$$\frac{d\varphi}{\sqrt{(\varphi - \varphi_0)(-\varphi - \varphi_0)(\varphi - \sqrt{\frac{6c}{a_1} - \varphi_0^2})(\varphi + \sqrt{\frac{6c}{a_1} + \varphi_0^2})}} = \pm \sqrt{\frac{-\alpha}{6}} d\xi. \tag{17}$$

letting $\varphi_0^* = \sqrt{\frac{6c}{a_1} - \varphi_0^2}$, integral Eq.(17) along L_3

$$\int_{\varphi_0}^{\varphi} \frac{d\varphi}{\sqrt{(-\varphi_0 - \varphi)(\varphi_0^* - \varphi)(-\varphi_0^* - \varphi)(\varphi - \varphi_0^*)}} = \pm \sqrt{\frac{-\alpha}{6}} \int_0^{\xi} d\xi, \tag{18}$$

Using the elliptic integral formula 253.00 in reference [5] to calculate Eq.(18), the bounded periodic wave solution Eq.(11) can be obtained.

(v). when $\alpha < 0, a_1c > 0$, taking φ_0 to satisfy $0 < \varphi_0 < \sqrt{\frac{3c}{a_1}}$, then the orbit L_4 (See Figure 1(iii)) passing through the initial point $(\varphi_0, 0)$ is a smooth closed orbit, and on $\xi - \varphi$ plane, the traveling wave corresponding to it is a bounded periodic wave.

Eq.(8) can be transformed into:

$$\frac{d\varphi}{\sqrt{(\varphi - \varphi_0)(\varphi + \varphi_0)(-\varphi + \sqrt{\frac{6c}{a_1} - \varphi_0^2})(\varphi + \sqrt{\frac{6c}{a_1} + \varphi_0^2})}} = \pm \sqrt{\frac{-\alpha}{6}} d\xi. \tag{19}$$

Letting $\varphi_0^* = \sqrt{\frac{6c}{a_1} - \varphi_0^2}$, integral Eq.(19) along L_4

$$\int_{\varphi_0}^{\varphi} \frac{d\varphi}{\sqrt{(\varphi_0^* - \varphi)(\varphi - \varphi_0)(\varphi + \varphi_0)(\varphi + \varphi_0^*)}} = \pm \sqrt{\frac{-\alpha}{6}} \int_0^{\xi} d\xi, \tag{20}$$

Using the elliptic integral formula 256.00 in reference [5] to calculate Eq.(20), the bounded periodic wave solution Eq.(12) can be obtained.

3.2. Solitary Wave Solutions

Since the traveling wave corresponding to the homoclinic orbit of the phase portraits is a bounded smooth solitary wave, the solution of the bounded smooth solitary wave can be obtained by integrating Eq.(8) along the homoclinic orbit.

Property 3 when $\alpha < 0, a_1c > 0$, there is a solitary wave in Eq. (1).

(i). The homoclinic orbit H_2 (See Figure 1(iii)) determines a downward bounded smooth solitary wave, and the solitary wave solution is:

$$\varphi = -\sqrt{\frac{6c}{a_1}} (\cosh \sqrt{\frac{-\alpha c}{a_1}} \xi)^{-1}. \tag{21}$$

(ii). The homoclinic orbit H_1 (See Figure 1(iii)) determines an upward bounded smooth solitary wave, and the solitary wave solution is:

$$\varphi = \sqrt{\frac{6c}{a_1}} (\cosh \sqrt{\frac{-\alpha c}{a_1}} \xi)^{-1}. \tag{22}$$

Proof: (i). when the conditions $\alpha < 0, a_1c > 0$ are satisfied, then the orbit H_2 passing through point $(-\sqrt{\frac{6c}{a_1}}, 0)$ is a homoclinic orbit, and on the $\xi - \varphi$ plane, the traveling wave corresponding to it is a bounded smooth solitary wave.

Eq.(8) can be transformed to:

$$\frac{d\varphi}{\sqrt{\varphi^2(\frac{6c}{a_1} - \varphi^2)}} = \pm \sqrt{\frac{-\alpha}{6}} d\xi. \tag{23}$$

Integrate Eq. (23) along H_2

$$\int_{-\sqrt{\frac{6c}{a_1}}}^{\varphi} \frac{d\varphi}{\sqrt{\varphi^2(\frac{6c}{a_1} - \varphi^2)}} = \pm \sqrt{\frac{-\alpha}{6}} \int_0^{\xi} d\xi. \tag{24}$$

A bounded smooth solitary wave Eq.(21) can be obtained by calculating Eq.(24).

(ii). the orbit H_2 passing through point $(\sqrt{\frac{6c}{a_1}}, 0)$ is a homoclinic orbit, then on the $\xi - \varphi$ plane, the traveling wave corresponding to it is a bounded smooth solitary wave.

Eq.(8) can be transformed to:

$$\frac{d\varphi}{\sqrt{\varphi^2(\frac{6c}{a_1} - \varphi^2)}} = \pm \sqrt{\frac{-\alpha}{6}} d\xi. \tag{25}$$

Integrate Eq.(25) along H_1

$$\int_{\varphi}^{\sqrt{\frac{6c}{a_1}}} \frac{d\varphi}{\sqrt{\varphi^2(\frac{6c}{a_1} - \varphi^2)}} = \pm \sqrt{\frac{-\alpha}{6}} \int_{\xi}^0 d\xi. \tag{26}$$

The bounded smooth solitary wave Eq.(24) can be obtained by calculating Eq.(26).

3.3. Kink Solutions

Since the traveling wave corresponding to the heteroclinic orbit of the phase portraits is a bounded kink, the solution of the bounded kink can be obtained by integrating Eq.(8) along the heteroclinic orbit.

Property 4 when $\alpha > 0, a_1c > 0$, there are two kink in the Eq. (1).

(i). the heteroclinic orbit K_1 (See Figure 1(iv)) determines a kink, and the solution of the kink is:

$$\varphi = -\sqrt{\frac{3c}{a_1}} + \frac{2\sqrt{\frac{3c}{a_1}}(\sqrt{\frac{3c}{a_1}} + \varphi_0)}{\sqrt{\frac{3c}{a_1}} + \varphi_0 + (\sqrt{\frac{3c}{a_1}} - \varphi_0)e^{-\sqrt{\frac{2c}{a_1}}\xi}}, \tag{27}$$

where $-\sqrt{\frac{3c}{a_1}} < \varphi_0 < \sqrt{\frac{3c}{a_1}}$.

(ii). The heteroclinic orbit K_2 (See Figure 1(iv)) determines a kink, and the solution of the kink is:

$$\varphi = -\sqrt{\frac{3c}{a_1}} + \frac{2\sqrt{\frac{3c}{a_1}}(\sqrt{\frac{3c}{a_1}} + \varphi_0)}{\sqrt{\frac{3c}{a_1}} + \varphi_0 + (\sqrt{\frac{3c}{a_1}} - \varphi_0)e^{\sqrt{\frac{2c}{a_1}}\xi}}, \tag{28}$$

where $-\sqrt{\frac{3c}{a_1}} < \varphi_0 < \sqrt{\frac{3c}{a_1}}$.

Proof: (i)when the conditions $\alpha > 0, a_1c > 0$ are satisfied, taking φ_0 to satisfy $-\sqrt{\frac{3c}{a_1}} < \varphi_0 < \sqrt{\frac{3c}{a_1}}$, the orbit K_1, K_2 passing through saddle point $(\frac{3c}{a_1}, 0)$ are two heteroclinic trajectories, on the $\xi - \varphi$ plane, the traveling waves corresponding to it are two bounded kinks.

Eq.(8) can be transformed to:

$$\frac{d\varphi}{\sqrt{(\sqrt{\frac{3c}{a_1}} + \varphi)^2(\sqrt{\frac{3c}{a_1}} - \varphi)^2}} = \pm\sqrt{\frac{\alpha}{6}}d\xi. \tag{29}$$

Integrate Eq.(29) along K_1, K_2

$$\int_{\varphi_0}^{\varphi} \frac{d\varphi}{(\sqrt{\frac{3c}{a_1}} + \varphi)(\sqrt{\frac{3c}{a_1}} - \varphi)} = \pm\sqrt{\frac{\alpha}{6}} \int_0^{\varphi} d\xi. \tag{30}$$

The two bounded kinks Eq.(27) and Eq.(28) can be obtained by calculating Eq.(30).

4. The Plane Simulation Graphs of Traveling Wave

Taking fixed parameters, according to the solution of the traveling wave, then the plane simulation graphs of bounded periodic wave, solitary wave and kink can be obtained by using Mathematica7.

Example 1. Letting $\alpha = -1; a_1 = 2, c = 3$, then $-\sqrt{\frac{3c}{a_1}} \approx -2.12132$, $-\sqrt{\frac{6c}{a_1}} = -3$. Letting $\varphi_0 = -2.5$, then $\varphi_0^* = \sqrt{\frac{6c}{a_1} - \varphi_0^2} \approx 1.65831$. Substituting these data into Eq.(11), we draw a plane simulation graph of a bounded periodic wave, as shown in Figure 2(i).

Example 2. Letting $\alpha = -1; a_1 = 2, c = 3$, then $\sqrt{\frac{3c}{a_1}} \approx 2.12132$. Letting $\varphi_0 = 0.5$, then $\varphi_0^* = \sqrt{\frac{6c}{a_1} - \varphi_0} \approx 2.95804$. Substituting these data into Eq.(12), we draw a plane simulation graph of a bounded periodic wave, as shown in Figure 2(ii).

Example 3. Letting $\alpha = -1; a_1 = 2, c = 3$, Substituting these data into Eq.(21) and Eq.(22), we draw respectively two plane simulation graphs of a bounded solitary wave, as shown in Figure 3(iii) and (iv).

Example 4. Letting $\alpha = 1; a_1 = 2, c = 3$, then $\sqrt{\frac{3c}{a_1}} \approx 2.12132$. Letting $\varphi_0 = 0$, substituting these data into Eq.(27) and Eq.(28), we draw respectively two plane simulation graphs of a bounded solitary wave, as shown in Fig. 4(iii) and (vi).

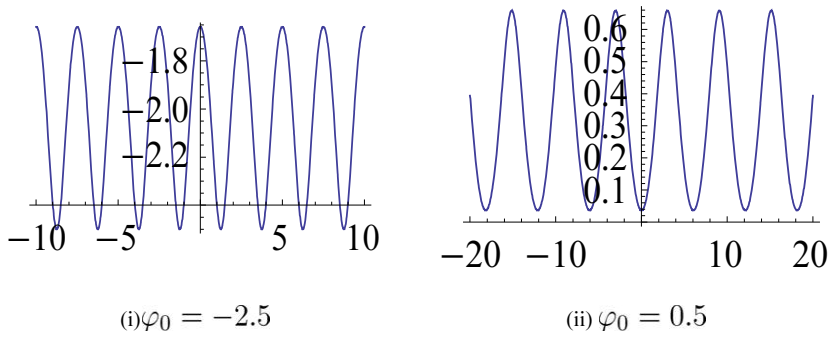


Figure 2. The periodic waves of Eq (1) when $\alpha = -1; a_1 = 2, c = 3$

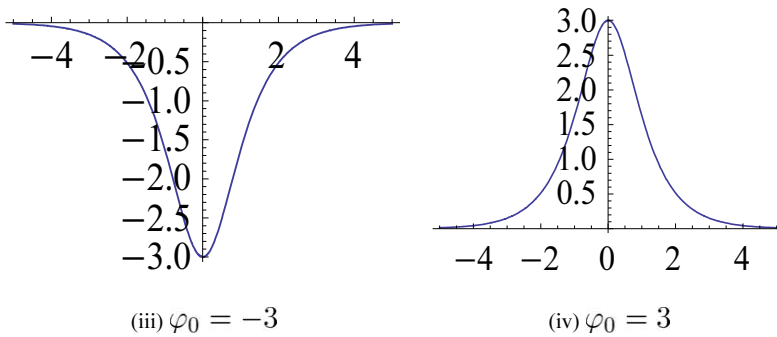


Figure 3. The solitary waves of Eq (1) when $\alpha = -1; a_1 = 2, c = 3$

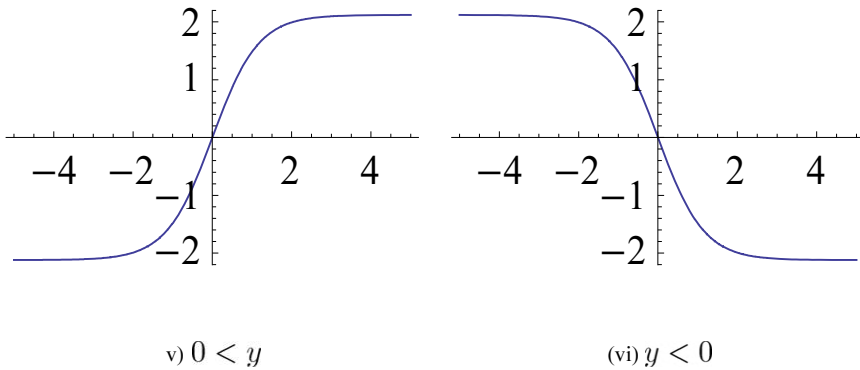


Figure 4. The kinks of Eq (1) when $\alpha = 1; a_1 = 2, c = 3, \varphi_0 = 0$

5. Conclusion

In the study of traveling wave solutions of wave equations, the bifurcation theory of differential dynamic systems has been widely applied. In a Hamiltonian system, the phase graphs bifurcation of a plane system can be drawn by using the bifurcation theory of differential dynamic systems. The orbit on the phase graphs correspond to different traveling wave bifurcation. The smooth closed orbit determines the periodic wave solution of the wave equations. The homoclinic orbit determines the solitary wave solutions of the wave equation. The heteroclinic orbit determines the kink wave solutions of the wave equation. In this paper, the periodic wave solutions, the solitary wave solutions, and kink solutions of mKdv-ZK equation are derived by using the bifurcation theory of differential dynamic systems, and their plane simulation graphs are given.

References

- [1] Tariq KU, Seadawy AR. Soliton solutions of (3+1)-dimensional Korteweg-de Vries Benjamin–Bona–Mahony, Kadomtsev–Petviashvili Benjamin–Bona–Mahony and modified Korteweg de Vries–Zakharov–Kuznetsov equations and their applications in water waves. *Journal of King Saud University – Science* [http://dx. doi. org/10. 1016/j. jksus. 2017. 02. 004](http://dx.doi.org/10.1016/j.jksus.2017.02.004)
- [2] Xie S, Hong X, Lu J. The bifurcation traveling waves of a generalized Broer-Kaup equation. *Appl. Anal. Comput.*, 2020, Online First.
- [3] Xie S, Hong X, Jiang T. Planar bifurcation method of dynamical system for investigating different kinks of bounded traveling wave solutions of a generalized Camassa-Holm equation. *J. Appl. Anal. Comput.*, 2017, 7(1), 78--290.
- [4] Xie S, Zhang Y, He J. Two types of bounded traveling-wave solutions of a two-component Camassa-Holm equation. *Appl. Math. Comput.*, 2013, 219(20), 10271--10282.
- [5] Byrd PF, Friedman MD. *Handbook of elliptic integrals for engineers and scientists*. Berlin: Springer, 1971.

A Survey on Ship Intelligent Cabin

Shuli JIA^a, Fuxin WANG^{a,1}, Mengchao DONG^b and Liyong MA^b

^aAutomation Engineering Department, Shanghai Marine Diesel Engine Research Institute, Shanghai, China

^bSchool of Information Science and Engineering, Harbin Institute of Technology, Weihai, China

Abstract. At present, intelligent ship has become a new hot spot of international maritime research and development. In order to achieve the purpose of safety, reliability, energy conservation, environmental protection, economy and efficiency, Intelligent ship integrates modern information technology, artificial intelligence technology and other new technologies with traditional ship technology. In this essay, the rapid development of intelligent cabin technology in recent years is surveyed. The intelligent cabin technology and the development trend of the current technology is analyzed, and the possible development direction in the future is pointed out. Collection, transmission and storage of sensor data are summarized. The development of status perception of cabin equipment and environment is discussed. Data-driven intelligent applications are summarized, and the development in the future is discussed. Intelligent cabin system is a data driven information system, which involves data acquisition, communication, storage, analysis, visualization and other rich content. Using the data provided by intelligent ship, more and more intelligent applications will be developed for ships.

Keywords. Intelligent ship, ship cabin, survey, data-driven application

1. Introduction

At present, intelligent ship has become a new hot spot of international maritime research and development. In order to achieve the purpose of safety, reliability, energy conservation, environmental protection, economy and efficiency, Intelligent ship integrates modern information technology, artificial intelligence technology and other new technologies with traditional ship technology [1-3]. According to the Intelligent Ship Specification of China Classification Society, the intelligent cabin is one of the important modules in the six intelligent modules, which provides vital support for the normal operation of the intelligent ship.

Intelligent technology can be divided into three levels, computational intelligence, perceptual intelligence and cognitive intelligence. The intelligent cabin mainly includes the cabin's perceptual intelligence and cognitive intelligence. Perceptual intelligent of ship cabin technology is a series of process. Sensors are used to obtain equipment and environmental information of ship cabin, and then the effective feature information is processed and understood. Cognitive intelligent of ship cabin technology is to further enhance the perceptual information to a cognizable level, such as memory, understanding,

¹ Fuxin WANG, Corresponding author, Automation Engineering Department, Shanghai Marine Diesel Engine Research Institute, Shanghai, China; E-mail: 18801909355@163.com.

planning, decision-making and so on. Data analysis or intelligent diagnosis methods are used in the state assessment and fault diagnosis of ship cabin systems and equipment, which is an important part of ship intelligence.

Marine ship cabin equipment mainly includes three categories as shown in Fig.1, they are power and control equipment, anti-pollution equipment and other auxiliary equipment. Power and control equipment includes main engine, auxiliary engine, boiler and steering gear. Anti-pollution equipment includes oil-water separators, incinerators, and domestic sewage treatment devices. Other auxiliary equipment includes oil separator, water generator, pumping device, etc. It can be seen that ship cabin equipment is the core of ship power and operation, so the use of intelligent technology to realize the perception and cognition of these core equipment of ship cabin will play an irreplaceable role in the development of ship technology.

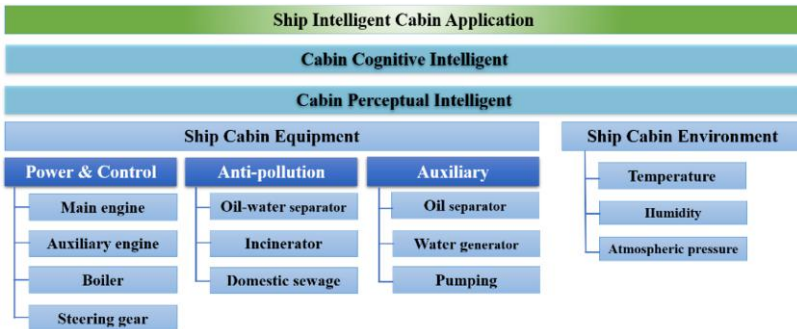


Figure 1. Architecture of intelligent ship cabin.

Although there have been a lot of literature on the research of ship intelligent cabin, but there is not a survey on these researches of intelligent cabin. The main contribution of this paper is as follows. First, the rapid development of intelligent cabin technology in recent five years is surveyed. As far as we know, there is no existing comprehensive study on intelligent cabin technology. Second, the development trend of the current technology on intelligent cabin is analyzed, and the possible development direction in the future is pointed.

2. Collection, Transmission and Storage of Sensor Data

The intelligent cabin mainly uses sensor, data acquisition module and communication equipment to collect data from equipment and system of ship cabin. The specific acquisition equipment and systems include main engine system, gearbox system, pitch control propeller system, fuel system, lubricating oil system, cooling water system, exhaust system, air system, steering gear system, oil separator system, boiler system and auxiliary engine system, as well as video monitoring information [4-5].

For data transmission, communication methods include serial port, USB and ethernet. Fieldbus technology is also used frequently, especially in control networks containing sensors and actuators. CAN bus is commonly used as field bus in ship systems [5].

The data storage development of intelligent ships has gone through two stages. The first stage is centralized storage, generally using relational databases [4-5]. The second

stage mainly adopts distributed storage technology [2,6]. The real-time monitoring of a small number of equipment can be realized through the central database. But when managing a large number of devices, distributed storage is more flexible and efficient. Distributed database engine is the core of the system, which is responsible for SQL parsing, rewriting and execution, and managing many storage nodes at the bottom. In the process of cabin data management, different scale database clusters can be flexibly constructed. By dividing the business data into different database storage nodes, the database has less pressure when facing massive data. By distributing the user's SQL request to the sub-workstations of each node for execution, the computing resources of each node are fully utilized, and the operating efficiency of the server cluster is improved as well.

3. Status Perception of Cabin Equipment and Environment

In summary, for marine ship cabin equipment such as diesel generator sets and propulsion motors, the collected signals can be classified into fast varying signals and slowly varying signals.

For slowly varying signals, such as speed, temperature, pressure and other thermal signals, the sampling frequency of these signals is low. Such data can be obtained through communication with control equipment or directly sampled, and the collected data are directly stored in the database. The feature extraction of slowly varying signal is mainly based on time domain analysis. The operation status of equipment is sensed through peak-peak, root mean square, kurtosis and other methods [7-8].

For fast varying signals, generally, the feature extraction is mainly based on frequency domain analysis and time-frequency domain analysis, and the operating status of the equipment is sensed through changes in the frequency spectrum and related parameters. Vibration signal is the most familiar signal [9], it can be analyzed through time domain, frequency domain, and time-frequency domain feature analysis. And the fusion analysis of feature information can be used to perceive the operating status of the device.

Multi-sensor information fusion plays an important role in device state perception. The complementary or redundant information of multiple sensors in time or space, and comprehensive judgment based on the information of multiple sensors, can effectively improve the accuracy of status perception [10-12]. Engine room equipment, especially the main power equipment, is a highly coupled system of thermodynamics, dynamics, and control. This makes the composite faults of gas circuits and mechanical components show the characteristics of coupling of fault characteristics and complex fault mechanisms, and it is difficult to characterize equipment faults with a single knowledge. In order to improve the fault perception accuracy, it is necessary to study the feature fusion method on the basis of the coupling characteristics of the equipment operation mechanism in the multi-source information fusion stage [10-12]. Using a full range of multiple attributes to redundantly describe faults may be a future research direction.

Fault diagnosis is one of the most studied problems in device state awareness. The problem of fault diagnosis is also representative in perception problems [12-16]. The following is a summary of the research methods and development of state perception, represented by the research of fault diagnosis. Fault diagnosis methods mainly include analytical method, signal processing method and knowledge-based method. The method

based on analytical method can determine whether the model changes in the actual operation process and the degree of change, so it could determine whether the fault occurs and the possible causes of the fault [8]. The signal processing method directly uses the signal model to process and analyze the original signal, extract the corresponding fault characteristic information and judge whether it is faulty [9]. The knowledge-based method does not require a quantitative mathematical model, it will comprehensively use the information of the diagnosis target and use expert knowledge to complete the diagnosis. A lot of research has focused on this method, and the methods used include neural networks [9,13], fuzzy mathematics [12], and other intelligent method including machine learning [14,15] are used here.

4. Data-driven Intelligent Applications

Due to the large number of ship information systems and the huge amount of information transmission between each other, the intelligent cabin can provide data-driven support for applications. These applications include big data [1,2,14] and artificial intelligence methods [16-19]. There are important research advances in the intelligent collision avoidance decision-making application [18-21].

The failure assessment and prediction of ship cabin equipment has been widely studied [22-29]. Generally speaking, fault prediction is divided into state prediction and life prediction. Classical fault prediction methods mainly include function fitting, Kalman filter, particle filter, grey model, neural network, et al. Recently, data-driven intelligent methods have been widely employed for fault diagnosis.

It can be seen that scholars have carried out in-depth research on fault diagnosis and prediction of marine engine [22], and the application of new intelligent algorithm is also constantly innovating. At present, the related research work mainly focuses on the accuracy of fault classification, and there are few research results on the optimization of algorithm cost and calculation efficiency. However, in the actual operation of the ship, the response speed and calculation cost are the key indicators to evaluate the fault diagnosis and prediction system. Therefore, how to accurately diagnose and predict the fault based on low cost and high response speed is a problem that cannot be ignored in the real ship application.

Energy efficacy is another hot topic for intelligent applications of ship [28-32]. Researches on intelligent integrated navigation and obstacle avoidance are also emerging. It is foreseeable that applications in this area will explode rapidly after the wide citation of smart cabin. We expect these intelligent applications could take root and grow.

5. Conclusions

This paper summarizes the development of intelligent cabin technology. Intelligent cabin is the integration of modern information technology, artificial intelligence technology with traditional ship technology. Intelligent cabin system is a data driven information system, which involves data acquisition, communication, storage, analysis, visualization and other rich content. By providing data support, many intelligent applications based on data become possible. And the data and applications can provide strong support for user operation, maintenance and decision-making. With the development of data sensing

technology, 5G communication technology, real-time data storage technology, visualization technology, artificial intelligence and other advanced technologies, intelligent cabin system will play an increasingly important role in the actual ship.

References

- [1] Cao M, Guo C. Key technologies of big data and its development in intelligent ship. In Proceedings of International Conference on Robotics and Artificial Intelligence (ICRAI); 2017 Dec 29- 31; Shanghai, China: ACM Press; p. 61-65.
- [2] Jia S, Ma L, Zhang S. Big data prototype practice for unmanned surface vehicle. In: Proceedings of the 4th International Conference on Communication and Information Processing; 2018 Nov 2-4; Qingdao, China: ACM Press; p. 43-47.
- [3] Im I, Shin D, and Jeong J. Components for smart autonomous ship architecture based on intelligent information technology. In Proceedings of International Conference on Mobile Systems and Pervasive Computing; 2018 Aug 13-15; Calle Gomera, Spain: Elsevier Press; p. 91-98.
- [4] Li G, Deng X, Zhou M, et al. Research on data monitoring system for intelligent ship. Lecture Notes in Electrical Engineering. 2020 Nov; 634(1):234-241.
- [5] Luo X. Research on communication technology of ship integrated monitoring system based on OPC. In Proceedings of International Conference on Intelligent Transportation, Big Data and Smart City; 2020, Jan 11-12; Vientiane, Laos: IEEE Press; p. 528-531.
- [6] Barsocchi P, Ferro E, Rosa D, et al. E-cabin: A software architecture for passenger comfort and cruise ship management. Sensors. 2019 Nov; 19(22): 49781-19789.
- [7] Chen L, Wang S, Wang H, et al. Model construction and application of machinery fault diagnosis of ships based on technology of resonant demodulation. Acta Technica CSAV. 2017 Feb; 62(2):477-485.
- [8] Kowalski J, Krawczyk B, Woniak M. Fault diagnosis of marine 4-stroke diesel engines using a one-vs-one extreme learning ensemble. Engineering Applications of Artificial Intelligence. 2017 Jan; 57(1):134-141.
- [9] Li Z, Chen J, Zi Y, He S. A sensor-dependent vibration data driven fault identification method via autonomous variational mode decomposition for transmission system of shipborne antenna. Sensors and Actuators, A: Physical. 2018 Jan; 279(1):376-389.
- [10] Zhang C, Guo C, Guo M. Information fusion based on artificial intelligence method for SINS/GPS integrated navigation of marine vessel. Journal of Electrical Engineering and Technology. 2020 May;15(3):1345-1356.
- [11] Xu H, Zhou Z, Han X, Li W. Self-tuning weighted fusion estimation method for intelligent ship. Journal of Huazhong University of Science and Technology (Natural Science Edition). 2019 Mar; 47(3):25-30.
- [12] Zhu Y, Geng L. Research on SDG fault diagnosis of ocean shipping boiler system based on fuzzy granular computing under data fusion. Polish Maritime Research. 2018 Jan; 25(1):92-97.
- [13] Wu, G. Fault detection method for ship equipment based on BP neural network. In Proceedings of International Conference on Robots and Intelligent System; 2018 May 26-27; Changsha, China: IEEE Press; p.556-559.
- [14] Yang M, Shi W. Research on fault diagnosis of ship power system based on improved particle swarm optimization neural network algorithm. In Proceedings of IEEE Advanced Information Technology, Electronic and Automation Control Conference; 2018 Oct 12-14; Chongqing, China: IEEE Press; p.108-113.
- [15] Wang Z, Wang Y, Yang X. Research on multi-agent fault diagnosis system for ship integrated power system. In Proceedings of Chinese Automation Congress; 2018 Nov 30- Dec 2; Xi'an, China: IEEE Press; p.3137-3140.
- [16] Feng D, Li Y. Research on intelligent diagnosis method for large-scale ship engine fault in non-deterministic environment. Polish Maritime Research, 2017 May; 24(S3):200-206.
- [17] Ma L, Xie W, Huang H. Convolutional neural network based obstacle detection for unmanned surface vehicle. Mathematical Biosciences and Engineering. 2020 Jan; 17(1):845-861.
- [18] Arendt R, Kopczyński A, Sychalski P. Centralized and distributed structures of intelligent systems for aided design of ship automation. In Proceedings of International Conference on Information Systems Architecture and Technology; 2017 Sep 17-19; Szklarska Poland: Springer Press; p.310-319.
- [19] Losiewicz Z, Nikoczuk P, Pielka D. Application of artificial intelligence in the process of supporting the ship owner's decision in the management of ship machinery crew in the aspect of shipping safety. In Proceedings of International Conference on Knowledge-Based and Intelligent Information Engineering Systems; 2019 Sep 4-6; Budapest Hungary: Elsevier Press; p.2197-2205.

- [20] Sedova N, Sedov V, Bazhenov R. Intelligent system of automatic ship course operation based on the neural network governor mechanism. In Proceedings of International Conference on Industrial Engineering, Applications and Manufacturing; 2019 Mar 25-29; Sochi Russia: IEEE Press; p.1-5.
- [21] Wang C, Zhang X, Chong L, et al. Research on intelligent collision avoidance decision-making of unmanned ship in unknown environments. *Evolving Systems*. 2019 Apr; 10(4):649-658.
- [22] Du L, Banda O, Kujala P. An intelligent method for real-time ship collision risk assessment and visualization. In Proceedings of International Conference on Collision and Grounding of Ships and Offshore Structures; 2019 Oct 21-23; Lisbon Portugal: CRC Press; p.293-300.
- [23] Borkowski P. Inference engine in an intelligent ship course-keeping system. *Computational Intelligence and Neuroscience*. 2017 Nov; 2017(11): 1-8.
- [24] Gao M, Shi G, Li S. Online prediction of ship behavior with automatic identification system sensor data using bidirectional long short-term memory recurrent neural network. *Sensors*. 2018 Dec; 18(12):1-12.
- [25] Huang B, Zou Z. Online prediction of ship roll motion in irregular waves using a fixed grid wavelet network. *Journal of Ship Mechanics*. 2020 Jun; 24(6):693-705.
- [26] Jeon M, Noh Y, Shin Y, et al. Prediction of ship fuel consumption by using an artificial neural network. *Journal of Mechanical Science and Technology*. 2018 Dec; 32(12):5785-5796.
- [27] Nie Z, Yuan Y, Xu D, et al. Research on support vector regression model based on different kernels for short-term prediction of ship motion. In Proceedings of International Symposium on Computational Intelligence and Design; 2019 Dec 14-15; Hangzhou, China: IEEE Press; p.61-64.
- [28] Ye Y, Cheng Y, Lai J, et al. Fault prediction model of ship unloader based on improved association rules. *China Mechanical Engineering*. 2019 Oct; 30(20):2463-2472.
- [29] Orihara H, Tsujimoto M, Performance prediction of full-scale ship and analysis by means of on-board monitoring. Part2: Validation of full-scale performance predictions in actual seas. *Journal of Marine Science and Technology (Japan)*. 2018 Apr; 23(4):782-801.
- [30] Prera L, Mo B. Machine intelligence based data handling framework for ship energy efficiency. *IEEE Transactions on Vehicular Technology*. 2017 Oct; 66(10):8658-8666.
- [31] Kalikatzarakis M, Geertsma R, Boonen E, et al., Ship energy management for hybrid propulsion and power supply with shore charging. *Control Engineering Practice*. 2018 Jan; 76(1):133-154.
- [32] Tan Z, Chen S, Lin H, et al. Ship intelligent energy efficiency management system design. In Proceedings of International Ocean and Polar Engineering Conference; 2019 Jun 16-21; Honolulu, United states: International Society of Offshore and Polar Engineers; p.4412-4417.

An Identity-Based Directed Proxy Multi-Signature Scheme

Weiping ZUO¹

Department of Mathematics, Tianshui Normal College, Tianshui, China

Abstract. Directed signature is introduced when the signed message contains privacy which is sensitive to the receiver, it is widely used in some special occasions involving signature privacy, such as electronic bidding, electronic voting, personal business activities, etc. This paper presents a new directed proxy multi-signature scheme, which integrated the directed signature and proxy multi-signature. In the proposed scheme, the agent generate a signature on behalf of the delegators, only a receiver specified by the delegator convince whether the truth of signature or not. At last, we analyze the characteristics and discuss the security of the scheme.

Keywords. Identity, Directed signature, Bilinear map

1. Introduction

Shamir[1] proposed identity based cryptosystem, in this type of cryptosystem, the public key of user is obtained directly from his basic information, such as username, ID number, Electronic mailbox, address and so on, while the secret key is obtained from private key generator which is called PKG and it is a trusted third party. The identity based cryptosystem has many advantages by comparison with public key cryptosystem, which make the acquisition of the public key simply, reduce the key management burden greatly and bring great convenience in practical application.

In the real world, people often need to delegate their signing rights to a reliable agent, which ensures the agent can sign on documents on behalf of the delegator, the same as in the electronic information society. In order to solve this problem, In 1996, Mambo[2] introduced the notion of proxy signature, in this type of schemes, an agent obtain the signing right form a delegator, and then the agent generates a signature on behalf of the delegator. Due to the technology of proxy signature is widely used in many fields, the study on proxy signature has attracted more and more attention. But in many electronic scenarios, sometimes it may be necessary to allow an agent to sign on behalf of the delegators at same time, for example, many departments of a company need to appoint an agent to sign on documents jointly at the same time, which is the concept of proxy multi-signature in fact, and the first concrete scheme was given by Yi[3]. Later, more and more concrete schemes [4-8] were introduced.

In the traditional signature, the signer has no restrictions on a verifier of the signature, anyone who obtains the signature convince whether the truth of signature or not. But, the public validity of signature is undesirable in some special occasions

¹ Corresponding Author; E-mail: wpzuo@126.com.

involving signature privacy, such as electronic bidding, electronic voting, personal business activities, etc. In order to protect signature privacy, Lim[9] introduced the concept of directed signature, in this type of schemes, only the receiver specified by the delegator convince whether the truth of signature or not with his secret key, while the others cannot convince it. Later, more and more concrete directed schemes [10-14] were introduced. Motivated with the above, we present a new concrete directed scheme which integrated the directed signature and proxy multi-signature. The new scheme has the characteristic of designated verifiability and it is widely used in many electronic scenarios.

The rest part of our paper has been organized as follows. The second section introduces the bilinear pairings and complexity assumption. In the third section, a new concrete directed proxy multi-signature scheme is proposed. We discuss the security of our scheme in the fourth section. The fifth section concludes remarks.

2. Preliminaries

Some preliminaries which include bilinear map and complexity assumption are given in this section.

2.1. Bilinear Map

Suppose G_1 be a group with addition, G_2 be a group with multiplication, where G_1 has a generator P and a prime order q , the same as G_2 . bilinear map $e: G_1 \times G_1 \rightarrow G_2$, which has characteristics as follows.

- Bilinearity: $\forall a, b \in Z_q^*, P, Q \in G_1, e(aP, bQ) = e(P, Q)^{ab}$.
- Nondegeneracy: $\exists P \in G_1$, such that $e(P, P) \neq 1$.

2.2. Complexity Assumption

This section revisits the computational Diffie-Hellman problem.

- Computational Diffie-Hellman(CDH) Problem: $\forall a, b \in Z_q^*, P, aP, bP \in G_1$, compute abP , where G_1 is a group with addition and it has a generator P and a prime order q .
- Computational Diffie-Hellman Assumption: Suppose that \mathbf{A} is an adversary, if no such \mathbf{A} can solve CDH problem in polynomial time at most t with probability at least ε .

3. Our Scheme

There are three type of entities which include the delegators $A_1 \dots A_n$, the agent B and the designated receiver C . A_i ($i=1 \dots n$) with identity ID_i , B with identity ID_B , C with identity ID_C . Our scheme is described as follows.

3.1. Setup

Assume G_1 be a group with addition, G_2 be a group with multiplication, where G_1 has a generator P and a prime order q , the same as G_2 . bilinear map $e: G_1 \times G_1 \rightarrow G_2$, k is a system security parameter. $H_i: \{0,1\}^* \rightarrow G_1$ ($i=1..3$), $H_4: \{0,1\}^* \rightarrow Z_q^*$ are cryptographic hash functions. **PKG** randomly chooses $s \in Z_q^*$ and regards it as secret key, then computes $P_{pub} = sP$ and keep s secretly.

3.2. Extraction

According to ID_i of a user, **PKG** generates $Q_{ID_i} = H_1(ID_i)$ and $s_{ID_i} = sQ_{ID_i}$, where Q_{ID_i} is public key, s_{ID_i} is private key. Thus, delegator A_i ($i=1 \dots n$) has the key pair (Q_{ID_i}, s_{ID_i}) respectively, the agent B has the key pair (Q_{ID_B}, s_{ID_B}) , the designated receiver C has the key pair $(Q_{ID_C} = s_{ID_C})$.

3.3. Proxy Key Generation

$A\{A_1 \dots A_n\}$ and B do the following steps to finish delegation under w , where w is the warrant for signed message. Finally, B generates the proxy key s_p .

- For all $i=1 \dots n$, A_i computes $R = \sum_{i=1}^n R_i$, $V_i = s_{ID_i} + r_i H_2(w || R)$, sends (w, R_i, V_i) to B .
- For all $i=1 \dots n$, B confirms (w, R_i, V_i) by an equation such that $e(P, V_i) = e(P_{pub}, Q_{ID_i})e(R_i, H_2(w || R))$. If the equation holds, he accepts signature, otherwise rejects.
- For all $i=1 \dots n$, If B confirms (w, R_i, V_i) , he computes

$$s_p = \sum_{i=1}^n V_i + H_4(w || R || ID_B) s_{ID_B}.$$

3.4. Proxy Multi-Signature Generation

B can sign a message $m \in \{0,1\}^*$ on behalf of $A\{A_1 \dots A_n\}$ by using s_p , he performs the following steps.

- Randomly chooses $u_1, u_2 \in Z_q^*$ and computes $U = u_1P, W = u_2Q_{ID_B}$.
- Computes $V = s_p + u_1H$, where $H = H_3(m \parallel w \parallel R \parallel e(s_{ID_B}, u_2 Q_{ID_C}))$.
- The directed proxy multi-signature is $\sigma = (w, R, U, W, V)$.

3.5. Proxy Multi-Signature Verification

In order to verify $\sigma = (w, R, U, W, V)$, C checks the signed message m and the delegation, if m accords with w and the delegation which $A\{A_1 \dots A_n\}$ authority B is effective, C performs the following computations with his private key s_{ID_C} .

- Computes $H = H_3(m \parallel w \parallel R \parallel e(s_{ID_C}, W))$ and convince whether the equation as follows holds or not. \square
- $e(P, V) = e(P, \sum_{i=1}^n (Q_{ID_i}) e(P, R)^{H_2(w \parallel R)} e(P_{pub}, Q_{ID_B})^{H_4(w \parallel R \parallel ID_B)} e(U, H))$

4. The Analysis of Our Scheme

This section analyze our scheme in detail, the analysis shows that our scheme satisfies the unforgeability, the designated verifiability and the nonrepudiation.

4.1. Correctness

We can put the mathematical formula of proxy multi-signature generation phase (section 3.4) into the equation of proxy multi-signature verification phase (section 3.5), and confirm whether the equation as follows holds or not.

$$\begin{aligned}
 e(P, V) &= e(P, s_p + u_1H) = e(P, \sum_{i=1}^n (V_i + H_4(w \parallel R \parallel ID_B)) + u_1H) \\
 &= e(P, \sum_{i=1}^n (s_{ID_i}) e(P, \sum_{i=1}^n r_i H_2(w \parallel R)) e(P, H_4(w \parallel R \parallel ID_B) s_{ID_B}) e(P, u_1H)) \\
 &= e(P, \sum_{i=1}^n (Q_{ID_i}) e(P, R)^{H_2(w \parallel R)} e(P_{pub}, Q_{ID_B})^{H_4(w \parallel R \parallel ID_B)} e(U, H))
 \end{aligned}$$

where $H = H_3(m \parallel w \parallel R \parallel e(s_{ID_C}, W))$.

4.2. Designated Verifiability

In phase of verification, C must use secret key s_{ID_C} compute H , but $H=H_3(m || w || R || e(s_{ID_C}, W))$ includes the private key s_{ID_C} of C . Only C verify the whether the truth of signature or not, any third party is unable to verify the truth of signature without the private key s_{ID_C} . Therefore, the proposed scheme satisfies the designated verifiability.

4.3. Unforgeability

The others tend to fake a valid directed signature is unfeasible without the proxy key s_p , Since $s_p = \sum_{i=1}^n V_i + H_4(w || R || ID_B) s_{ID_B}$ includes the private key s_{ID_B} of B , however, s_{ID_B} is kept secretly by B , Solving s_{ID_B} from s_p is equivalent to solving discrete logarithm problem. Therefore, the proposed scheme satisfies the unforgeability.

5. Conclusion

Due to the directed signature is applicable where the signed message contains privacy which is sensitive to the receiver and it is used in many electronic scenarios which including electronic transaction and personal business activities. Considering the significance of directed scheme, this paper present a new directed scheme. The new scheme is secure and it satisfies the security requirements of the unforgeability, the designated verifiability and the nonrepudiation.

References

- [1] A. Shamir. Identity-based cryptosystems and signature schemes, Advances in Cryptology: CRYPTO 1984, LNCS 196, Springer-Verlag Press, Berlin,1984, pp. 47-53.
- [2] M. Mambo, K. Usuda and E. Okamoto. Proxy signature for Delegating signing operation, Proc of the 3rd ACM Conference on Computer and Communications Security, ACM Press, New York,1996, pp. 48-57.
- [3] L. J. Yi, G. Q. Bai and G. Z. Xiao. Proxy Multi-Signature Scheme: a New Type of Proxy Signature Scheme, Electronics Letter, vol. 36, Jun. 2000, pp. 527-528.
- [4] F. Cao, Z. F. Cao. A secure identity-based proxy multi-signature scheme, Information Sciences, vol. 179, Mar. 2009, pp.192-302.
- [5] C. Hsu, T. Wu and W. He. New proxy multi-signature scheme, Applied Mathematics and Computation, vol. 162, Mar. 2005, pp.1201-1206.
- [6] J. Ji, D. Li. A new proxy multi-signature scheme, Journal of Computer Research and Development, vol. 41, Apr. 2004, pp. 715-719.
- [7] X. X. Li, K. H. Chen. Multi-proxy signature and proxy multi-signature schemes from bilinear pairings, LNCS 3320, Springer-Verlag Press, Singapore, 2004, pp. 591-595.
- [8] H. Du, J. Wang, Y. N. Liu. Independent verification of proxy multi-signature scheme, International Journal of Computational Science and Engineering Vol. 9, Apr. 2014, PP. 301-311.
- [9] C. H. Lim, P. J. Lee. Modified Maurer-Yacobi's Scheme and its applications, Proceedings of the Australasian Conference on Information Security and Privacy. Berlin: Springer-Verlag, 1992, pp. 308-323.

- [10] J. Zhang, Y. Yang, X. Niu. Efficient Provable Secure ID-Based Directed Signature Scheme without Random Oracle, *Advances in Neural Networks-ISSN 2009*, LNCS 5553, Springer-Verlag Press, Berlin, pp. 318–327.
- [11] Q. Wei, J. He, H. Shao. A directed signature scheme and its application to group key initial distribution, *Proceedings of the International Conference on Interaction Sciences: Information Technology, Culture and Human*. Seoul, Korea: ACM Press, 2009, pp. 265-269.
- [12] D. X. Wang, H.M. Zhu, J.K. Teng. Directed signature scheme based on identity with Fairness, *Journal of Yunnan University: Natural Sciences Edition*, Vol. 33, Nov. 2011, pp. 658-661.
- [13] W. P. Zuo, Y. F. Liu, S. F. Wang. An ID-based Designated-Verifier Proxy Multi-signature Scheme, *Fourth International Conference on Multimedia, Information Networking and Security*, 2012, pp. 576-579.
- [14] N. B. Gayathri, R. V. Rao, P. V. Reddy. Efficient and Provably Secure Pairing Free ID-Based Directed Signature Scheme, *International Symposium on Security in Computing and Communication*, 2017, pp. 28-38.

Subject Index

3D air-ground channel	337	coverage	416
3D modelling	272	cross-based framework	30
5G	346	cryogenic	266
60GHz	346	cutting edge technology	127
6D pose estimation	89	D2D communication	383
accounting information fusion	200	data mining	21
adaptive denoising	30	data-driven application	453
adaptive network	206	decision support	200
AES-128	374	decision tree	152
alpha cut	98	deep learning	40, 71, 115
anisotropic diffusion	40	DeepLog	152
Anomaly detection	152	deformation control	80
AoA	337	density	254
AoD	337	depression	12
artificial intelligence	21, 200	detection	12, 354
assembly drawing	272	digital control process	61
autonomous system	136	digital mock-up	184
bagging	165	digital product definition	184
beamsteering	346	digital twin	184
beat interference	322	direct-detection	322
behavior features	12	directed signature	459
big data	71	domain randomization	89
big data analysis	200	Doppler spread	337
bilinear map	459	drift error reduction	80
binary opcode analysis	146	DropConnect	178
bridge monitoring	416	DSP	266
broadcast collision	241	dual-metal gate	392
broadcast coverage	241	edge computing	374
CAPoNeF	400	edge-preserving smoothing	30
CIR	337	energy disaggregation	54
classification	40	energy efficiency	435
cloud	21	energy reduction	254
cloud computing	248	ensemble	165
Cn band	424	ensemble methods	21
CNN	115	enterprise	248
complex networks	3	event distraction	192
compressive covariance sensing	363	expected interval	98
computational intelligence	3	explainable AI	152
construction	127	Facebook	12, 172
control	288	facial emotion recognition	115
conversation analysis (CA)	172	failure factors	254
covariance estimation methods	363	feedback of states	288
covariance matching	363	filtering	266

fin field-effect transistor	392	multi-data processing	71
fixture design	272	multi-order adaptive	206
formal concepts analysis	21	multi-source heterogeneous data	127
formalism Euler – Lagrange	288	navigation	424
forwarding probability	241	Naïve-Bayes	46, 221
forwarding strategy	241	negative hypergeometric	
fourth dimension	136	distribution	406
FPGA	266	neural network	178
fuzzy arithmetic	98	NFC	374
fuzzy equations	98	NLP	192
fuzzy number	98	Octave	136
genetic algorithm	3	one-to-many mapping encryption	406
GRUs	115	optical OFDM	322
guard band	322	order-preserving searchable	
guided filter	30	encryption	406
hacker	107	oriented FAST and rotated BRIEF	
hacking	107	(ORB)	46
hacking tools	107	orthogonal MSK	424
hardware security module	248	overfitting	178
human gait	288	periodic waves	442
H^∞ state estimation	229	personnel localization	278
identity	459	preprocessing	363
IEEE 802.11ay	346	principal component analysis	
influence maximization	3	(PCA)	54, 354
information security	310	process variation	392
integrated platform	435	public cloud	248
intelligence accounting information		public organization(s)	107, 310
fusion system	200	qualitative analysis	136
intelligent ship	435, 453	RAA	346
intelligent systems	71	radiometric identification	400
invasive ductal carcinoma (IDC)	46	range query	406
Internet of Things (IoT)	374	ranging performance	424
joints	331	Raspberry Pi II	115
K-nearest neighbors	46	reflect-array antenna	346
key management	248	representativeness	54
kinks	442	resource allocation	383
knowledge distillation	165	RF fingerprinting	400
least squares	363	robotic process automation	
linear matrix inequalities (LMIs)	229	platform	115
linear sparse ruler	363	routing algorithms	254
log anomaly detection	152	safety monitoring and	
logistic regression	146	management	278
machine learning	21, 178, 400	SDR	400
malicious program	107	security	248, 400
MANETs	241	security management	310
mixture	354	security standards	310
mKdv-VZK equation	442	self-modeling network	206
mmWave	346	sensor	254, 435
monitoring	127	sentiment analysis	192, 221

seven degrees of freedom	331	survey	453
Shapley value	152	synthetic dataset	89
ship cabin	453	system	127
similar version detection	146	T test	354
singular point	442	thin-walled high rib	80
singular region	61	three-dimensional reconstruction	278
social media research	172	tool path optimization	61
social network activity	12	total variation	40
social networking mining	192	twin objective functions	221
software analysis	146	two time-varying delays	229
solitary waves	442	two-dimensional positioning	302
sparrow search algorithm	416	UG software	272
spatial resolution	302	uncertain neural networks	229
spherical parallel mechanism	331	vehicular communications	383
substation	278	web server	248
super-large integral frame part	80	wireless networks	400
supervised classification	21	wireless sensor networks	416
support vector machines		work-function variation	392
(SVM)	46, 221	ZigBee	374

This page intentionally left blank

Author Index

Abramonte, R.	98, 136	Han, Y.	40
Aekwarangkoon, S.	12	He, C.	383
Aliaga, A.	98, 136	Hemtanon, S.	12
Alwan, N.A.S.	363	Hong, B.-W.	40
Anwar, S.	89	Hu, D.	406
Azibi, H.	21	Hu, X.	229
Bao, Q.	184	Hu, Y.	30
Bao, Y.	374	Huang, S.	80, 331
Becker, M.	288	Ivchenko, A.	400
Bobulski, J.	71	Ji, H.	80
Bolkhovskaya, O.	337, 346	Jia, S.	453
Bolotin, I.	346	Jin, Z.	302
Carvalho, C.B.	54	Kallweit, S.	89
Chen, C.-H.	v	Ke, J.	424
Chen, H.	278	Kim, D.	40
Chen, J.	302	Kim, M.	40
Chen, Q.	383	Kittphattanabawon, N.	12
Chen, X.	322, 424	Kokarev, A.	337
Chen, Z.	331	Kubanek, M.	71
Choi, H.-T.	40	Li, K.	30
Cui, Z.	331	Li, Q.	416
Dai, L.	392	Li, Y.	192, 272
Dai, S.	184	Lim, H.-i.	146, 178
Dantas, P.V.	54	Lin, M.	392
DarkSide Collaboration	266	Liu, C.	221
Das Gupta, H.	165	Liu, Y.	416
Ding, J.	229	Long, T.	406
Dong, M.	453	Lopez, K.M.M.	46
Dorokhin, S.	400	Lou, Q.	241
Du, S.	89	Lu, X.	424
Duan, X.	200	Luo, H.	435
Elokhin, A.	337	Lyu, W.	392
Engemann, H.	89	Ma, L.	435, 453
Escobar, E.	98, 136	Maddouri, M.	21
Espinoza, M.G.T.	107	Magboo, Ma.S.A.	46
Farina, M.	172	Mai, X.	278
Foltz, K.	248	Maltsev, A.	337, 346
Gallegos, L.E.M.	107, 310	Martinez R., A.D.	266
Gao, J.	354	Meddouri, N.	21
González, R.F.P.	107	Mora P, Z.A.	288
Guo, X.	424	Nigmatullin, R.	400
Gutierrez, F.	98, 136	Ning, C.	89
Ham, S.	40	Ospina-Henao, P.A.	288

Park, Y.	40	Xiao, D.	229
Peng, Y.	61, 80, 416	Xiao, Q.	127
Petrosian, O.	152	Xie, J.	354
Qian, G.	383	Xing, D.	254
Ronquillo, M.L.A.	310	Yan, H.	406
Sabino S. Júnior, W.	54	Yan, M.	30
Saxena, S.	115	Yang, F.	278
Seleznev, V.	346	Yang, H.	331
Sergeev, V.	337	Yang, P.	322
Shen, X.	61	Yang, Q.	221
Sheng, V.S.	165	Yang, Y.	278
Simpson, W.R.	248	Yin, A.	406
Sun, N.	302	Yu, S.	127
Sun, Z.	374	Yu, T.	392
Tallón-Ballesteros, A.J.	v	Yu, Y.	184, 435
Tan, S.	127	Zezzatti, A.O.	310
Toapanta, S.M.T.	107, 310	Zhang, K.	165
Topîrceanu, A.	3	Zhang, Y.	442
Treur, J.	206	Zhao, G.	184
Tripathi, S.	115	Zhao, J.	30
Tsb, S.	115	Zhao, L.	61, 80
Valencia, C.H.	288	Zhao, R.	383
Vásquez, S.M.	288	Zhao, Z.	241, 392
Wang, F.	435, 453	Zheng, Y.	383
Wang, X.	424	Zheng, Z.	354
Wang, Y.	61, 80, 229	Zhu, D.	331
Wang, Z.	192	Zou, J.	152
Wei, B.	192	Zuo, W.	459
Wu, H.	278		