Searching for Robust Chaos in Discrete Time Neural Networks using Weight Space Exploration[1]

Radu Dogaru, A.T. Murgan, S. Ortmann*, M. Glesner*
Applied Electronics Department, University "Politehnica" of Bucharest,
Bd. Armata Poporului Nr.1, Sect.6, Bucharest, Romania,
* Institute of Microelectronic Systems, Darmstadt University of Technology,
Karlstrasse 15, 64283 Darmstadt, Germany
E-mail: radu_d@lmn.pub.ro

ABSTRACT

A method for analysis and synthesis of recurrent discrete time neural networks with robust chaotic behavior is presented. Using this method, a new nonlinear activation function called saturated - modulus was found to be the most efficient in order to get chaos even in small sizes neural networks. Based on weight space exploring using a searching strategy closely related with genetic algorithms, some new concepts were introduced, namely the ***generic structure*** *of a neural networks population and the concept of* ***descriptor map****. Instead of trying to learn a specified (chaotic) trajectory, entropic and sensitivity maps are computed and displayed for a population of neural networks sharing the same generic structure. The sensitivity map is the particular case of a descriptor map based on Liapunov exponents. These maps can be used to select from a population of neural networks those individuals which best fits with some desired behavior. Moreover, using sensitivity maps,* ***robust chaos*** *was emphasized meaning that for some compact set included in the weight space the chaotic behavior of the network remains unchanged. Experimental results also proved that the activation function used by neurons is strong related to the global aspect of the descriptor maps and they can be efficiently used for synthesis.*

1. Introduction

Chaotic dynamics can occur in recurrent neural networks as long as they are particular cases of nonlinear dynamic systems. Chaos is often defined as a dynamic behavior which is highly sensitive to perturbations. Interesting considerations about the relationships between stochastic and chaotic signals can be found in [11]. According to [18] chaotic attractors are essential for the mechanism of learning and retrieval of information in the olfactory system of rabbits. Although the role of the chaos observed in biological systems is not yet clear, it could be profoundly significant and there is an increasing interest [2],[3],[5],[12],[18],[19] to this problem.

Chaos synchronization and control [6],[14],[15] were recently emphasized as interesting behaviors in nonlinear systems which could also explain storage and retrieval mechanisms in neural networks. In fact , the main difference between chaotic and random signals is that the first one can be obtained using well-defined structures of nonlinear systems and thus they can be controlled.

Interesting perspective for exploiting chaotic dynamic is offered by the synergy between the fields of neural networks and nonlinear dynamics. Different models for chaotic neural networks were proposed in the last years [1],[2],[3],[18] most of them based on biological models. An interesting point of view is presented in [3] regarding the potential of „computation with chaotic attractors" instead of „computation with fixed points" used by actual digital systems. Other researchers [2] consider chaotic systems as "infinite reservoirs of periodic orbits" and thus assigning to each periodic orbit a pattern one hopes to store efficiently high amounts of information. While the field of neurodynamics [10] is mainly concerned with fixed points dynamics, and more recently with the problem of oscillatory neural networks synthesis using specific learning algorithms [13], the problem of chaotic neural network synthesis it is still at the beginning. Most of the chaotic neural networks proposed in literature are based on continuos-time dynamics and standard sigmoidal activation function. Usually they are described by a large number of differential equations and thus they are computational intensive. Instead, we propose a different approach in order to build systems based on computation with chaotic attractors;

---

[1] Accepted to be presented and published in Proceedings ICNN'96, Washington D.C. 2-6 June 1996

While in digital systems with few basic building blocks one can built very complex systems we hope that designing appropriate basic building blocks behaving chaotically they can be also used in building larger systems. These building blocks must be robust to parameter changes and thus we are searching for robust chaotic behavior. Another important feature for these blocks is that they must be easy to be implemented or simulated on computers and thus the discrete-time recurrent neural network structure was chosen as the most appropriate paradigm to embed them. Discrete-time offers the possibility to obtain chaos more easily than continuos time and the simulated models match better with their hardware implementations. While the main feature to be exploited in computation with chaos is the wide spreading of the state space, appropriate tools for measuring the „degree of chaos" were introduced, allowing one to search for the most chaotic and robust networks. A complete analysis and synthesis method called "weight space exploration" (WSE) is proposed in this paper. Using a simple non-monotone activation function (saturated rectifier), we proved that robust chaotic behavior can be easily achieved in very small networks and further applications of them for building larger adaptive systems [8] proved also to be very efficient. Other simulation results proved that sigmoidal neurons are not well suited in order to get chaos even in large neural networks, our method giving a systematic design procedure for discrete time chaotic neural networks which have a simpler model than the one described in [1].

2. Chaotic Neural networks analysis and synthesis

2.1 The method of weight space exploration

The method proposed here is essentially based on two concepts. The first one is the concept of ***running ,*** which is defined as a mapping from the high dimensional space of the neural net ***structure*** (defined by the interconnection weights, the initial status and the nonlinear activation function) to one or two scalar values (called ***dynamic descriptors***) giving the essential information about the dynamic behavior of the network as it was obtained by running it on a manageable enough large time period. The second concept is called ***descriptor map*** . According with this concept, some of the interconnection weights are selected as variable parameters by splitting them in to two groups corresponding with two parameters (p_0 and p_1). Each parameter is ranging in a specified domain, using a finite specified resolution (*res*). We have often used in our simulations *res=32* which corresponds to a population of $res^2 = 1024$ neural networks being under investigation. All individuals in this population share the same ***generic structure*** the differences between them being only given by the values of the parameter weights. For each population, a ***descriptor map*** is assigned and computed. According with the two types of dynamic descriptors described in section 2.3. it is possible to generate both ***entropic maps*** and ***sensitivity maps .*** For each member of the population, a pixel in the descriptor map is assigned, having a color or brightness proportional with the value of the dynamic descriptor obtained by applying to it the ***running*** procedure. The spatial position of each pixel in the map is given by the coordinates $x, y \in \{0,.. \; res-1\}$, the 0 values corresponding with the upper left corner of the map. For each pixel (i.e. neural network individual) a well defined structure is established by taking $: p_0 = {p_0}^{\min} + y \cdot \left({p_0}^{\max} - {p_0}^{\min} \right) / res$. A similar relation is established between the horizontal position *x* and the other parameter. Individuals that are neighbors in the descriptor map have no significant structure differences. This map topology, allow one to observe ***robustness*** of dynamic descriptors as large compact zones having the same brightness (color), or bifurcations as crisp borders between compact zones in the descriptors maps. For each map, a ***calibration pattern*** is down displayed having the left sided color (or gray level) assigned to the minimum value of the dynamic descriptor and the right side color to the maximum value of the same dynamic descriptor.

2.2. Neural network dynamics

For each individual in the neural network population, a discrete time full-connected recurrent neural network model was considered, the dynamics being described by the following equation (the number of neurons is *N*):

$$X(t) \leftarrow W \cdot Y(t) + I, \quad t = \overline{T_i \; \ldots \; T_i + T_a - 1} \tag{1}$$

X, Y, and *I* are *N*-dimensional vectors corresponding with the state, output and respectively the input of the network. The component "s" of the vector *Y* is the activation function of the "source" *"s"* neuron applied to it's state at the previous discrete time moment: $y_s(t) = f(x_s(t-1))$. The dynamics given by (1) is performed from *t=0* until *t=*T_i without computing the dynamic descriptors in order to avoid transitory effects. These descriptors are evaluated only during the ***analysis period*** T_a during a ***running procedure***.

The N x N matrix W is composed by the weights which connects the "source" neurons with the "target" ones : $W = \{w_{t,s}\}$, $s,t \in \{0,..N-1\}$, some of them being variable parameters as it was described in the pervious section.

2.3. Entropic and Sensitivity Descriptors

The ***entropic descriptor*** is essentially based on [9] and it gives information about the disorder in the neural network. While the dynamics described by (1) is deterministic, the number of distinct states (*nrds*) of the system during the analysis period was considered [7] as a entropic descriptor. Fixed points dynamics are thus characterized by *nrds=1,* while large period limit cycles are characterized by large *nrds* values .

The ***sensitivity descriptor*** was introduced in order to evaluate the „degree of chaos“ and it is based on evaluating the Liapunov (or characteristic) exponents λ_i , $i \in \{0 .. N-1\}$ of the dynamic system (1) for all N directions of the state space. When the status of a nonlinear dynamic system is perturbed with a very small amount Δx on direction *i,* after a period *delt* this disturbance can be amplified, the system evolving on different attractors, or it can be diminished, the system being structurally stable and thus evolving on the same attractor as in the not perturbed case. The first case corresponds with $\lambda_i > 0$ while the second one is characterized by $\lambda_i < 0$. According with the theory of chaos [14], a system has chaotic dynamics if it has at least one positive characteristic exponent. Starting from the definition of the characteristic exponents for continuos time nonlinear systems, the following formula was derived in order to compute them for discrete time neural networks:

$$\lambda_i = \frac{1}{delt} \lim_{\substack{T_a \to \infty \\ \Delta x \to 0}} \frac{1}{T_a} \sum_{t=T_i}^{T_i+T_a-1} \ln \left| \frac{x_i(t+delt) - xp_i(t+delt)}{\Delta x} \right| \qquad (2)$$

where $xp_i(t+delt)$ denotes the perturbed state of the neuron i at the moment $t+delt,$ when the status at the moment t was perturbed with Δx and $x_i(t+delt)$ is the not perturbed state of the neuron i. In order to characterize the sensitivity for all directions we have chosen the following sensitivity descriptor (*sd*):

$$sd = \sum_{i=0}^{N-1} \sigma(\lambda_i) \quad \text{where } \sigma(x) = \begin{cases} 0 \text{ if } x \le 0 \\ 1 \text{ if } x > 0 \end{cases} \qquad (3)$$

For chaotic dynamics, $sd>0$ and it reaches the maximum value N when all the characteristic exponents are positive. Networks having large sd values may be especially useful because their attractors are wide distributed in the whole state space. It must be noticed that according with Kaplan-Yorke's conjecture there is a proportionality relationship between the fractal dimension of the chaotic attractor and the value of *sd.*

3. Simulation results

A strategy for finding chaotic neural networks matching some desired behavior can be summarized in the following steps: 1) Choose the generic structure (neural network dimension, nonlinear activation function, the groups of parameter weights); 2) Specify the resolution (and thus the size of the population of neural networks) and the variation domains for each parameter; 3) Evaluate the entropic and the sensitivity maps ; 4) Select a specific network as a pixel in the descriptor maps, according with the desired behavior (e.g. the network must be chaotic with two positive Liapunov exponents) ; 5) (Synthesis) Compute the parameter weights based on the coordinates of the pixel associated with the network. 6) Evaluate the dynamics or/and use the chaotic neural network individual synthesized in step 5).

Steps 1 to 5 can be repeated after "freezing" the last group of parameter weights to those values which ensures the best fitness with the desired behavior. The repeated strategy was successfully applied , the effect being an increasing number of the neural networks individuals having the desired behavior. Using different nonlinear activation functions in generic structures with different numbers of neurons, many descriptor maps were generated and observed. While monotone increasing functions proved to be not suitable in order to get chaos, a new nonlinear activation function was found, which ensures chaotic behavior even for very small number of neurons within large variation domains for weights.

This function corresponds with the rectifying operation often used in electronic instrumentation and it was called saturated modulus (sat-mod) or saturated rectifier:

$$f(x) = \begin{cases} |x| \text{ if } |x| < 1 \\ 1 \text{ if } |x| \geq 1 \end{cases}$$

The most important qualitative feature which has been observed for all sensitivity and entropic maps is the strong relationship between the nonlinear activation function shape and the global aspect of the map. To be more specific, some global parameters such as the percentage of chaotic individuals from a given population with a given *sd* value are almost the same for the same nonlinearity instead of choosing different groups of parameter weights. Some simulation results will be now considered (see Fig.1.):

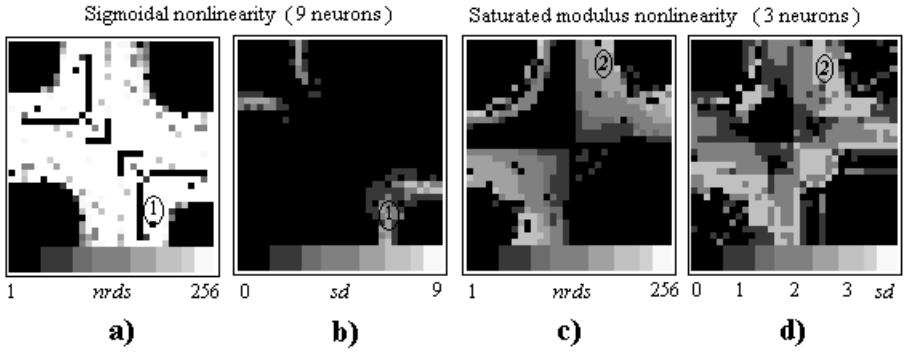

Fig. 1 Entropic and sensitivity maps for two different nonlinear activation functions

First, chaotic individuals were searched according with the sequence of six steps by using the sigmoidal function: $f(x) = \tanh(2x)$. The resulted entropic and sensitivity maps are represented in Fig 1.a. and Fig. 1.b. and it can be easily seen that even for a population of networks with 9 neurons, there are very few individuals with chaotic behavior. In 3 neurons populations chaotic behavior was never found. The interconnection matrix *W* corresponds with a 2D cellular neural network structure, where each neuron is connected with its four topologic neighbors (left, right, up an down) by the same set of synaptic weights. There is also a feedback signal for each neuron. Thus, the matrix *W* has 0 valued elements except the main diagonal (1 valued) , the second and fourth lower diagonals (1 and respectively *p1* valued) and the second and the fourth upper diagonals (-1.75 and respectively *p0* valued). For most of the pairs (*p0,p1)* periodic signals with high entropy are generated. The sensitivity to initial conditions is very low confirming the small values of the *sd* in Fig. 1.b). For the non-monotonic "saturated rectifier" the resulted maps are represented in Fig.1.c. (entropic) and Fig. 1.d. (sensitivity) and they were obtained for a three neurons generic structure characterized by:

$$W = \begin{bmatrix} 1 & p_0 & 0 \\ p_1 & 1 & p_0 \\ 0 & p_1 & 1 \end{bmatrix}, \quad I = \begin{bmatrix} -0.5 \\ -0.5 \\ -0.5 \end{bmatrix}, \quad \begin{cases} p_0^{\min} = p_1^{\min} = -2 \\ p_0^{\max} = p_1^{\max} = 2 \end{cases} \text{ and } res = 32$$

Now, more than 70% of the individuals are chaotic ones, many of them having a high sensitivity (*sd=3).* Comparing with the case of sigmoidal nonlinearity where most of the individuals characterized by high entropy (see white pixels in Fig. 1.a.) are not chaotic ones, the saturated-rectifier activation function ensures that all chaotic individuals have also high entropy.

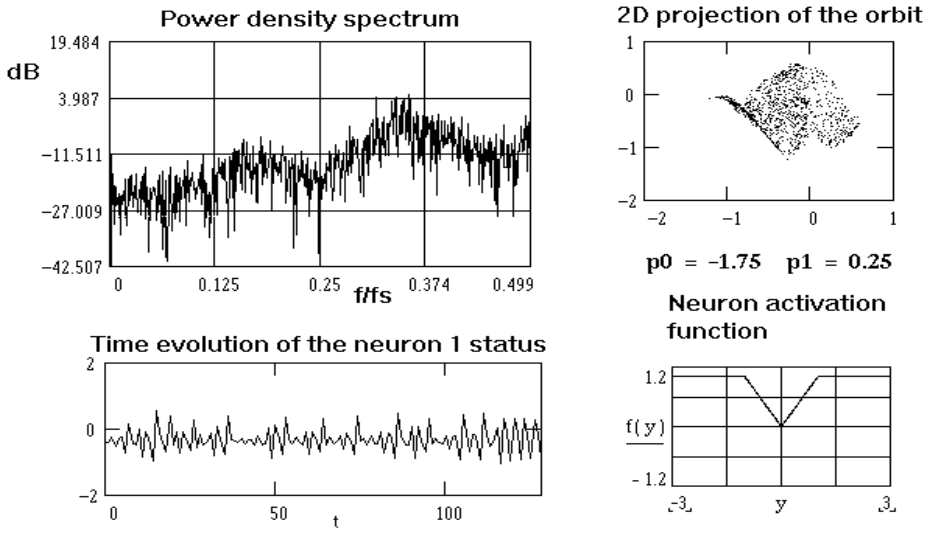

Fig. 2. Dynamics of a particular chaotic neural network obtained by using the "sat-mod" activation function

As a chaotic neural network synthesis example let us consider the network marked with "2" in Fig. 1.d. According with its associated pixel brightness and position, *sd=3* and the parameters are: *p0=-1,75* and *p1=0.25* . Based on the previously mentioned generic structure we have got a well defined chaotic neural network, further simulations presented in Fig. 2. confirming that all state variables behave chaotic, the associated signals having a power spectral density which is exponential proportional with the frequency . Another interesting aspect related to the chaotic behavior in small neural networks using sat-mod activation function is the ***robustness***. The chaotic behavior is robust while it remains unchanged (the same value of the sensitivity descriptor) for large tolerances of the weight parameters. In Fig. 3. this property is emphasized on entropic and sensitivity maps of another saturation-rectifier based generic structure using a rectangle „A" to show a robustness domain projection in the 2D space.

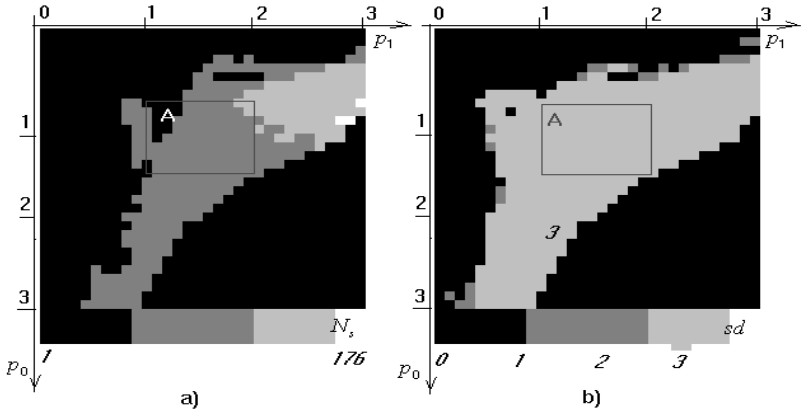

Fig. 3. Robustness domain in the entropic (left) and sensitivity (right) maps of a 3-cell chaotic neural network.

Robustness is an important property while it allows to design a mapping from a higher dimensional space represented by all points within a robustness domain (the chaotic network weights) to the low dimensional chaotic signal space associated with one output of the network. This property allow one to efficiently build analog devices for generating broadband signals.

4. Conclusions

The WSE method based on computing entropic and sensitivity dynamic descriptors has proved to be an efficient tool to investigate oscillatory and chaotic phenomena in neural networks allowing also chaotic neural networks synthesis. Using a new nonlinear activation function (saturated-rectifier), it was proved that robust chaotic behavior can be easily achieved even in small size neural networks. These small networks can be used as basic building blocks for building large adaptive systems based on computation with chaos. The saturated rectifier neuron function may be easily implemented in both analog and digital technologies and thus larger

neural system using computation with chaos may be easily design and implement using small chaotic neural networks as basic building blocks.

References

[1] K. Aihara, T. Takabe, and M. Toyoda. “Chaotic Neural Networks”, *Phys. Lett. A, vol 144* , pp. 333-340, March 1990.

[2] A. Babloyantz and J.A. Sepulchre, „Information processing by spatio-temporal chaotic networks“, in Proceedings ICANN’93, pp. 670-675.

[3] B. Baird, F. Eeckman, „A neural network computer architecture for computation with oscillatory and chaotic attractors“, in Proceedings of the 2’nd International Conference of Fuzzy Logic & Neural Networks , Iizuka , Japan, 1992, pp. 981-984.

[4] G. H.Bateni, C.D. McGillen, “A Chaotic Direct-Sequence Spread-Spectrum Communication System”, in IEEE Tr. on Communications, Vol. 42, No. 2/3/4, 1994, pp. 524.

[5] D.L. Birx and S.J. Pipenberg, “Chaotic oscillators and complex mapping feed forward networks (CMFFNS) for signal detection in noisy environments”, in Proceedings IJCNN ‘92, vol II, pp. 881-888, IEEE 0-7803-0559-0/92.

[6] L.O. Chua, L Kocarev, K. Eckert and M. Itoh, " Experimental chaos synchronization in Chua's circuit" , in Int. Journal of Bifurcations and Chaos, vol.2 pp. 705-708, 1992.

[7] R. Dogaru, A.T. Murgan, " Robust oscillations and bifurcations in Cellular Neural Networks", in Proceedings CNNA-94 Rome- Italy , IEEE Catalog Number 94TH0693-2 , pp. 303-308., 1994.

[8] R. Dogaru, A.T. Murgan , “Chaotic resonance theory, a new approach for pattern storage and retrieval in neural networks“, in Proceedings ICNN’95 , Vol. 6, pp. 3048, 1995.

[9] A.M. Fraser, " Information and entropy in strange attractors", in IEEE Tr. on Information Theory, Vol. 35, No. 2, March 1989, pp. 245-262.

[10] S. Haykin, " Neural Networks, a comprehensive foundation" , Macmillan College Publishing Company , 1994

[11] G. Kubin, “What is chaotic signal ?”, in Proceedings of the 1995 IEEE Workshop on Nonlinear Signal and Image Processing, Ed. I. Pitas, pp. 141-144.

[12] S. Nara, P. Davis, and H. Totsuji , " Memory search using complex dynamics in a recurrent neural network model ", in Neural Networks, Vol. 6, pp. 963-973, 1993.

[13] O. Nerrand, P. Roussel-Ragot, D. Urbani, L. Personnaz, and G. Dreyfus, " Training recurrent neural networks: why and how ? An illustration in dynamicall process modeling ", in IEEE Tr. on Neural Networks, Vol. 5, No. 2, March 1994, pp. 178-183.

[14] N. Kanou and Y. Horio, “Switched-current chaotic neural network with chaotic simulated annealing”, in Proceedings ICNN’95, Perth, Australia, Vol. 6, pp. 3146-3149, 1995.

[15] M.J. Ogorzalek, "Taming Chaos - Part I; Synchronization", IEEE Tr. on Circuits and Systems - I , Vol. 40, No. 10, October 1993, pp. 693- 699.

[16] D. Ruelle, " Chaotic evolution and strange attractors", Cambridge University Press 1989.

[17] W. Schwartz, K. Kutzer, T. Kilias, “Generation of Broadband Signals by Chaotic Discrete-Time Series”, in Proceedings of the 1995 IEEE Workshop on Nonlinear Signal and Image Processing, Ed. I. Pitas, pp. 129-132.

[18] W. J. Freeman, „Tutorial on neurobiology: from single neurons to brain chaos“ in International Journal of Bifurcation and Chaos, 2, No. 3, pp. 451-482, 1992.

[19] P. Thiran & M. Hasler, „Information processing using stable and unstable oscillations: a tutorial“ , in Proceedings CNNA-94 Rome- Italy , IEEE Catalog Number 94TH0693-2 , pp. 127-136, 1994.

[20] F. Zou, J. Nossek, "Bifurcation and Chaos in Cellular Neural Networks ", *IEEE Tr. on Circuits and Systems - I ,* Vol. 40, Nr. 3, pp. 166-173, March 1993.