Mathematical Aspects of Application of Neural Networks to Processes with Delays

Zlatinka Kovacheva
Department of Mathematics and Applied Sciences
Middle East College
Muscat, Oman
e-mail: zlatinka@mec.edu.om

Valéry Covachev
Institute of Mathematics and Informatics
Bulgarian Academy of Sciences
Sofia, Bulgaria
e-mail: vcovachev@hotmail.com

Abstract—**Neural networks are used to solve different kinds of problems from a wide range of disciplines. A brief overview of the history and performance of neural networks is given. Some neural network models are presented. Additionally, we summarize our results concerning the existence and global exponential stability of an equilibrium point or periodic solution of these models.**

Keywords-neuron; artificial neural network; processes with delay.

I. INTRODUCTION

Artificial Neural Networks (ANN) are computational paradigms, which implement simplified models of their biological counterparts, biological neural networks.

Although the initial intent of ANN was to explore and reproduce human information processing tasks, such as speech, vision, and knowledge processing, ANN also demonstrated their superior capability for classification and function approximation problems. This has great potential for solving complex problems, such as systems control, data compression, optimization problems, pattern recognition, and system identification.

Neural networks have wide applicability to real world business problems. In fact, they have already been successfully applied in many industries. Since neural networks are best at identifying patterns or trends in data, they are well suited for prediction or forecasting needs including: sales forecasting, industrial process control, customer research, data validation, risk management, target marketing and so on.

ANN are also used in the following specific paradigms: recognition of speakers in communications; diagnosis of hepatitis; recovery of telecommunications from faulty software; interpretation of multi-meaning Chinese words; undersea mine detection; texture analysis; three-dimensional object recognition; hand-written word recognition; and facial recognition [1], [2], [3].

In this paper, we are focusing on the application of neural networks to processes with delay. The rest of this paper is organized as follows: In Section II, we give some information on the history and action of artificial neurons. In Section III, we consider some neural network models, namely, continuous-time neural networks of Hopfield- and Cohen-Grossberg-type and their discrete-time counterparts. We conclude the paper in Section IV.

II. ARTIFICIAL NEURON

An artificial neuron is a device with many inputs and one output. The neuron has two modes of operation; the training mode and the using mode. In the training mode, the neuron can be trained to fire (or not), for particular input patterns. In the using mode, when a taught input pattern is detected at the input, its associated output becomes the current output. If the input pattern does not belong in the taught list of input patterns, the firing rule is used to determine whether to fire or not.

The first artificial neuron was produced in 1943 by the neurophysiologist Warren McCulloch and the logician Walter Pitts [4]. But the technology available at that time did not allow them to do too much. Neural networks process information in a similar way the human brain does. The network is composed of a large number of highly interconnected processing elements (neurons) working in parallel to solve a specific problem. Neural networks learn by example.

In the human brain, a typical neuron collects signals from others through a host of fine structures called dendrites. The neuron sends out spikes of electrical activity through a long, thin stand known as an axon, which splits into thousands of branches. At the end of each branch, a structure called a synapse converts the activity from the axon into electrical effects that inhibit or excite activity from the axon into electrical effects that inhibit or excite activity in the connected neurons.

A more sophisticated neuron is the McCulloch and Pitts model (MCP) [4]. The difference from the previous model is that the inputs are "weighted", the effect that each input has at decision making is dependent on the weight of the particular input. The weight of an input is a number which, when multiplied by the input, gives the weighted input. These weighted inputs are then added together and, if they exceed a pre-set threshold value, the neuron fires. In any other case, the neuron does not fire. In mathematical terms, the neuron fires if and only if

$$\sum_{i=1}^{m} X_i W_i > T,$$

where W_i, $i = \overline{1,m}$, are weights, X_i, $i = \overline{1,m}$, inputs, and T a threshold. The addition of input weights and of the

 ISBN: 978-1-61208-567-8

threshold makes this neuron a very flexible and powerful one. The MCP neuron has the ability to adapt to a particular situation by changing its weights and/or threshold. Various algorithms exist that cause the neuron to “adapt”; the most used ones are the Delta rule and the back error propagation [2],[3]. The former is used in feed-forward networks and the latter in feedback networks.

An important application of neural networks is pattern recognition. Pattern recognition can be implemented by using a feed-forward neural network that has been trained accordingly. During training, the network is trained to associate outputs with input patterns. When the network is used, it identifies the input pattern and tries to output the associated output pattern. The power of neural networks comes to life when a pattern that has no output associated with it, is given as an input. In this case, the network gives the output that corresponds to a taught input pattern that is least different from the given pattern.

III. NEURAL NETWORK MODELS

A. *Hopfield-Type Neural Networks*

Hopfield-type (additive) networks have been studied intensively during the last two decades and have been applied to optimization problems. The original model [5] used two-state threshold “neurons” that followed a stochastic algorithm: each model neuron i had two states, characterized by the values V_i^0 or V_i^1 (which may often be taken as 0 and 1, respectively). The input of each neuron came from two sources, external inputs I_i and inputs from other neurons. The total input to neuron i is then

$$\text{Input to } i = H_i = \sum_{i\neq j} T_{ij}V_j + I_i,$$

where T_{ij} can be viewed as a description of the synaptic interconnection strength from neuron j to neuron i. The motion of the state of a system of N neurons in the state space describes the computation that the set of neurons is performing. A model, therefore, must describe how the state evolves in time, and the original model describes this in terms of a stochastic evolution. Each neuron samples its input at random times. It changes the value of its output or leaves it fixed according to a threshold rule with thresholds U_i:

$$V_i \to V_i^0 \quad \text{if} \quad \sum_{i\neq j} T_{ij}V_j + I_i < U_i,$$

$$V_i \to V_i^1 \quad \text{if} \quad \sum_{i\neq j} T_{ij}V_j + I_i > U_i.$$

A simple Hopfield-type neural network is the following one:

$$C_i\frac{dx_i(t)}{dt} = -\frac{x_i(t)}{R_i} + \sum_{j=1}^{m} a_{ij} f_j\left(x_j(t)\right) + I_i, \quad i = \overline{1,m},$$

where m denotes the number of units (neurons) in the network, $x_i(t)$ denotes the state of the i-th unit at time t, the positive constants C_i and R_i are the neuron amplifier input capacitance and resistance, respectively, I_i is the constant input from outside the network, $f_j\left(x_j(t)\right)$ denotes the output of the j-th unit on the i-th unit at time t, a_{ij} is the weight (strength) of the synaptic connection between the j-th unit and the i-th unit.

In the formulation of the above system, it is implicitly assumed that the neurons process input, produce output and communicate with each other instantaneously. But this is usually not true and there can be significant time delays both in neural processing and axonal transmission. Such delays can be concentrated (discrete), or continuously distributed over a certain duration of time, finite or infinite.

Most widely studied and used neural networks can be classified as either continuous or discrete. Recently, there has been a somewhat new category of neural networks which are neither purely continuous-time nor purely discrete-time. This third category of neural networks, called impulsive neural networks, displays a combination of characteristics of both the continuous and discrete systems. To the best of our knowledge, impulsive neural networks first appeared in 1999 [6], yet we would mention that after the publication of our paper [7] in 2004 hundreds or maybe thousands of papers devoted to impulsive neural networks appear each year.

In order to solve problems in the fields of optimization, neural control and signal processing, neural networks have to be designed such that there is only one equilibrium point and this equilibrium point is globally asymptotically stable so as to avoid the risk of having spurious equilibria and local minima. In the case of global stability, there is no need to be specific about the initial conditions for the neural circuits since all trajectories starting from anywhere settle down at the same unique equilibrium. If the equilibrium is exponentially asymptotically stable, the convergence is fast for real-time computations.

In our paper [7], we considered several Hopfield-type systems incorporating the aforementioned features. All of them can be put together as the following:

$$\begin{aligned}\frac{dx_i(t)}{dt} &= -a_i x_i(t) + \sum_{j=1}^{m} b_{ij} f_j\left(x_j(t)\right) \qquad (1)\\ &+ \sum_{j=1}^{m} c_{ij} g_j\left(x_j\left(t-\tau_{ij}\right)\right)\\ &+ \sum_{j=1}^{m} d_{ij} h_j\left(\int_0^{\infty} K_{ij}(s) x_j(t-s)\, ds\right) + I_i, \quad t>0, \quad t\neq t_k,\end{aligned}$$

$$\begin{aligned}\Delta x_i(t_k) &= -B_{ik} x_i(t_k) \qquad (2)\\ &+ \int_{t_{k-1}}^{t_k} \psi_{ik}(s) x_i(s)\, ds + \gamma_{ik}, \quad i=\overline{1,m}, \quad k\in\mathbb{N},\end{aligned}$$

with initial values prescribed by piecewise-continuous functions $x_i(s) = \phi_i(s)$ which are bounded for $s \in$

$(-\infty, 0]$. The coefficient $a_i > 0$ is the rate with which the i-th unit self-regulates or resets its potential when isolated from other units and inputs; $f_j(\cdot)$, $g_j(\cdot)$, $h_j(\cdot)$ denote activation functions; the parameters b_{ij}, c_{ij}, d_{ij} represent the weights (or strengths) of the synaptic connections between the j-th unit and the i-th unit; the constant I_i represents an input signal introduced from outside the network to the i-th unit; τ_{ij} are nonnegative numbers whose presence indicates the delayed transmission of signals at time $t - \tau_{ij}$ from the j-th unit to the i-th unit; the delay kernels $K_{ij}(s)$ incorporate the fading past effects (or fading memories) of the j-th unit on the i-th unit; $\Delta x_i(t_k) = x_i(t_k + 0) - x_i(t_k - 0)$ denote impulsive state displacements at fixed instants of time t_k $(k \in \mathbb{N})$ involving integral terms whose kernels $\psi_{ik}(s): [t_{k-1}, t_k] \to \mathbb{R}$ are measurable functions, essentially bounded on the respective interval. Here, it is assumed that the sequence of times $\{t_k\}_{k=1}^{\infty}$ satisfies $0 = t_0 < t_1 < t_2 < \cdots < t_k \to \infty$ as $k \to \infty$; B_{ik} and γ_{ik} are some constants.

Using the Contraction Mapping Principle (Banach's Fixed Point Theorem), we found sufficient conditions for the existence of a unique equilibrium point of the above system. Further, using a suitable Lyapunov functional we found sufficient conditions for the global exponential stability of the equilibrium (that is, each solution of the system tends exponentially to the equilibrium point).

Recently, we considered a class of Hopfield neural networks with integral impulsive conditions and finite distributed delays, formulated in the form of an ω-periodic system of impulsive delay differential equations

$$\frac{dx_i(t)}{dt} = -a_i(t)x_i(t)$$
$$+\sum_{j=1}^{m} b_{ij} f_j\left(\int_0^{\omega} K_{ij}(s)x_j(t-s)\,ds\right) + I_i(t), \quad t \neq t_k,$$

$$\Delta x_i(t_k) = -\alpha_{ik}x_i(t_k)$$
$$+\sum_{j=1}^{m} B_{ijk}\Phi_j\left(\int_{t_{k-1}}^{t_k} c_{ijk}(s)x_j(s)\,ds\right) + \gamma_{ik},$$
$$i = \overline{1,m}, \quad k \in \mathbb{Z}.$$

Using the Contraction Mapping Principle, we found sufficient conditions for the existence of a unique ω-periodic solution. Moreover, if an ω-periodic solution exists, using an appropriate Lyapunov functional we found sufficient conditions for its global exponential stability. We noted that the above-mentioned ω-periodic solution can be found approximately by the method of successive approximations.

B. *Cohen-Grossberg Neural Networks*

We have also studied continuous-time impulsive neural networks more general than the Hopfield-type neural networks, such as the Cohen-Grossberg neural networks. Thus, in [8] we considered the impulsive Cohen-Grossberg neural network with S-type delays

$$\frac{dx_i(t)}{dt} = a_i\big(x_i(t)\big)\left[-b_i\big(x_i(t)\big) + \sum_{j=1}^{m} c_{ij} f_j\left(x_j(t)\right)\right.$$
$$\left.+\sum_{j=1}^{m} d_{ij}\int_{-\tau}^{0} g_j\left(x_j(t+\theta)\right)\,d\eta_{ij}(\theta) + I_i\right], t > 0, t \neq t_k,$$

$$\Delta x_i(t_k) = -B_{ik}x_i(t_k) + \int_{-\tau}^{0} x_i(t_k+\theta)\,d\zeta_k(\theta) + \gamma_{ik},$$

$$i = \overline{1,m}, \quad k \in \mathbb{N},$$

with initial values prescribed by piecewise-continuous functions $x_i(s) = \Phi_i(s)$ with discontinuities of the first kind for $s \in [-\tau, 0]$. Here $a_i(x_i)$ denotes an amplification function; $b_i(x_i)$ denotes an appropriate function which supports the stabilizing (or negative) feedback term $a_i(x_i)b_i(x_i)$ of the unit i; the past effect of the j-th unit on the i-th unit is given by a Lebesgue-Stieltjes integral; the impulsive state displacements at fixed moments of time t_k, $k \in \mathbb{N}$, also involve Lebesgue-Stieltjes integrals. This type of delays in the presence of impulses is more general than the usual types of delays studied in the literature. In fact, concentrated delays correspond to the points of discontinuity of the bounded variation functions.

For the above system, sufficient conditions are found for the existence of a unique equilibrium point and its global exponential stability. Examples of impulsive systems satisfying the sufficient conditions obtained are given, namely, the differential system with S-type delays

$$\dot{x}_1(t) = (2 + \sin x_1(t))[-2x_1(t) + 0.1\arctan x_1(t)$$
$$+0.15\arctan x_2(t) + 0.1\int_{-1}^{0} x_1(t+\theta)\,de^{\theta}$$
$$\left.+0.15\int_{-1}^{0} x_2(t+\theta)\,de^{\theta}\right],$$

$$\dot{x}_2(t) = (3 + \sin x_2(t))[-3x_2(t) + 0.15\arctan x_1(t)$$
$$-0.2\arctan x_2(t) + 0.1\int_{-1}^{0} x_1(t+\theta)\,de^{\theta}$$
$$\left.-0.2\int_{-1}^{0} x_2(t+\theta)\,de^{\theta}\right],$$

provided with one of the following three sets of impulse conditions:

$$\Delta x_1(t_k) = -\frac{1}{2}x_1(t_k) + \frac{1}{4}\int_{-1}^{0} x_1(t_k+\theta)\, de^{\theta},$$
$$\Delta x_2(t_k) = -\frac{1}{4}x_2(t_k) + \frac{1}{4}\int_{-1}^{0} x_2(t_k+\theta)\, de^{\theta};$$

$$\Delta x_1(t_k) = -100x_1(t_k) + \int_{-1}^{0} x_1(t_k+\theta)\, de^{\theta},$$
$$\Delta x_2(t_k) = -50x_2(t_k) + \int_{-1}^{0} x_2(t_k+\theta)\, de^{\theta},$$
$$t_k = 10k, \quad k \in \mathbb{N};$$

$$\Delta x_1(t_k) = -(k+1)x_1(t_k) + k\int_{-1}^{0} x_1(t_k+\theta)\, de^{\theta},$$
$$\Delta x_2(t_k) = -(k^2+1)x_2(t_k) + k^2\int_{-1}^{0} x_2(t_k+\theta)\, de^{\theta},$$
$$t_k = k^2, \quad k \in \mathbb{N}.$$

In all three cases, the equilibrium point $(0,0)^T$ is globally exponentially stable with Lyapunov exponent respectively: 1 in the first case, 0.039 in the second case, and any $\lambda \in (0,0.5)$ in the third case.

C. Discrete-Time Neural Networks

For different Hopfield-type and Cohen-Grossberg neural networks, we have found their discrete-time counterparts and found sufficient conditions for existence and global exponential stability of equilibria and periodic solutions.

Here, we recall just the results of our paper [9], where we obtain a discrete-time counterpart of system (1), (2). Let $h > 0$ denote a uniform discretization step size and $[t/h]$ denote the greatest integer in t/h. For convenience, we denote $[t/h]$, $n \in \{0\} \cup \mathbb{N}$, and, by an abuse of notation, write $x_i(n)$ instead of $x_i(nh)$. Further on, we denote $\kappa_{ij} = [\tau_{ij}/h]$, $i,j = \overline{1,m}$. Finally, we replace the integral terms $\int_0^{\infty} K_{ij}(s)x_j(t-s)\,ds$, $i,j = \overline{1,m}$, by sums of the form $\sum_{p=1}^{\infty}\mathcal{K}_{ij}(p)x_j(n-p)$, where $p = [s/h]$, $\mathcal{K}_{ij}(p)$ stands for $\mathcal{K}_{ij}(ph)$ and $x_j(n-p)$ for $x_j((n-p)h)$.

Now, on the interval $[nh,(n+1)h)$ $(n \in \{0\} \cup \mathbb{N})$ we approximate system (1) by

$$\frac{dx_i(s)}{ds} = -a_i x_i(s) + \sum_{j-1}^{m} b_{ij} f_j\left(x_j(n)\right) + \sum_{j-1}^{m} c_{ij} g_j\left(x_j(n-\kappa_{ij})\right) + \sum_{j-1}^{m} d_{ij} h_j\left(\sum_{p=1}^{\infty}\mathcal{K}_{ij}(p)x_j(n-p)\right) + I_i. \tag{3}$$

We rewrite equation (3) in the form

$$\frac{d}{ds}\left(x_i(s)e^{a_i s}\right) = e^{a_i s}\left(\sum_{j-1}^{m} b_{ij} f_j\left(x_j(n)\right) + \sum_{j-1}^{m} c_{ij} g_j\left(x_j(n-\kappa_{ij})\right) + \sum_{j-1}^{m} d_{ij} h_j\left(\sum_{p=1}^{\infty}\mathcal{K}_{ij}(p)x_j(n-p)\right) + I_i\right), i = \overline{1,m},$$

and integrate it over the interval $[nh,(n+1)h]$ to obtain

$$x_i(n+1) = e^{-a_i h}x_i(n) + \frac{1-e^{-a_i h}}{a_i}\left(\sum_{j-1}^{m} b_{ij} f_j\left(x_j(n)\right) + \sum_{j-1}^{m} c_{ij} g_j\left(x_j(n-\kappa_{ij})\right) + \sum_{j-1}^{m} d_{ij} h_j\left(\sum_{p=1}^{\infty}\mathcal{K}_{ij}(p)x_j(n-p)\right) + I_i\right), i = \overline{1,m}, \tag{4}$$
$$n \in \{0\} \cup \mathbb{N}, \quad i = \overline{1,m}.$$

This system is the discrete-time analogue of the system without impulses (1). It is provided with initial values of the form $x_i(-\ell) = \varphi_i(-\ell)$ $(\ell \in \{0\} \cup \mathbb{N})$, where the sequences $\{\varphi_i(-\ell)\}_{\ell=0}^{\infty}$ are bounded for all $i = \overline{1,m}$. The method used here is called *semi-discretization* [1]. It is easy to see that systems (1) and (4) have the same equilibria if any.

Further on, denote $n_k = [t_k/h]$ we approximate the impulsive conditions (2) by

$$x_i(n_k^+) - x_i(n_k^-) = \sum_{\ell=n_{k-1}+1}^{n_k} B_{ik\ell}x_i(\ell) + \gamma_{ik}, \tag{5}$$
$$i = \overline{1,m}, \quad k \in \mathbb{N}$$

where, for convenience, $n_0 = -1$ and $B_{ik\ell}$ are suitably chosen constants.

Finally, we find sufficient conditions for the global exponential stability of the unique equilibrium point of the system (4), (5).

IV. CONCLUSION

In the present paper we gave a short overview of the history, performance and applications of neurons and neural networks. We presented several neural network models and our results concerning the existence and global exponential stability of an equilibrium point or periodic solution of these models.

REFERENCES

[1] S. Mohamad and K. Gopalsamy, "Dynamics of a class of discrete-time neural networks and their continuous-time counterparts," Math. Comput. Simulat., vol. 53, pp. 1-39, 2000. ISSN: 03784754

[2] A. I. Galushkin, Neural Networks Theory. Berlin-Heidelberg: Springer-Verlag, 2007. ISBN: 9783540481249

[3] J. Heaton, Introduction to the Math of Neural Networks. Heaton Research. ISBN: 9781604390339

[4] W. McCulloch and W. Pitts, "A logical calculus of the ideas immanent in nervous activity," Bull. Math. Biophys., vol. 9, pp. 127-147, 1943. ISSN: 0007-4985

[5] J. J. Hopfield, "Neurons with graded response have collective computational properties like those of two-state neurons," Proc. Natl. Acad. Sci. USA, vol. 81, pp. 3088-3092, 1984. ISSN: 1091-6490

[6] Z.-H. Guan and G. Chen, "On delayed impulsive Hopfield neural networks," Neural Netw., vol. 12, pp. 273-280, 1999. ISSN: 0893-6080

[7] H. Akça, R. Alassar, V. Covachev, Z. Covacheva, and E. Al-Zahrani, "Continuous-time additive Hopfield-type neural networks with impulses," J. Math. Anal. Appl., vol. 290, pp. 436-451, 2004. ISSN: 0022-247X, 1096-0813

[8] H. Akça and V. Covachev, "Impulsive Cohen-Grossberg neural networks with S-type distributed delays," Tatra Mt. Math. Publ., vol. 48, pp. 1-13, 2011. ISSN: 1210-3195

[9] H. Akça, R. Alassar, V. Covachev, and Z. Covacheva, "Discrete counterparts of continuous-time additive Hopfield-type neural networks with impulses," Dyn. Syst. Appl., vol. 13, pp. 77-92, 2004. ISSN: 1056-2176

 ISBN: 978-1-61208-567-8