About control of a neuron as an elementary unit of multilayer artificial neural networks with variable signal conductivity

To cite this article: I A Dovgerd et al 2020 IOP Conf. Ser.: Mater. Sci. Eng. 760 012017

View the article online for updates and enhancements.
About control of a neuron as an elementary unit of multilayer artificial neural networks with variable signal conductivity

I A Dovgerd, G A Dovgerd, A V Ignatenkov, A M Olshansky, M G Lysikov and E N Rozenberg

1Design Bureau JSC “Railway Signalling Institute”, 27, b, 1 Nizhegorodskaya St., 109029, Moscow, Russia
2JSC “VTB Capital”, 10, Presnenskaya Embankment, 123112, Moscow, Russia
E-mail: lexolshans@gmail.com

Abstract. The report is devoted to a possibility of a control of each Hopfield’s neuron as an elementary unit of the neural network with variable signal conductivity used for transport scheduling tasks. It is set that direct control of the neuron is impossible in terms of initial data available for initializing the neural network. Also the result of some control algorithms based on statistical computations and signals dynamic is described below.

1. Introduction
Artificial neural networks (ANN) are famous popular algorithms for solving rather difficult and variable tasks at present. Every neural network is a complicated system which could be investigated as a dynamic system. The problem of decreasing the training time of the ANN is very urgent for science and practice and it may be solved using control theory and dynamic representation. It allows us to use computational capacities more efficiently.

At the turn of the XX and XXI centuries in Russia and abroad there were some attempts to apply control theory to ANNs [5, 6, 8]. The characteristic feature of such papers was applying the control theory to multilayer perceptrons with creating optimal time sequences of ANN weight matrices. The sequences were obtained as a result of solving the two-point nonlinear boundary value problem. It led to optimal training rules for ANN weights.

An optimal weight matrix looks like the sum of the products between desired output vectors $R^n$ and concrete output vector calculated permanently during the ANN’s training. Finally we get an optimal sequence of weight matrices for ANN.

Paper [7] analyzes the optimality of hyperparameters of ANN using genetic algorithms.

According to the classical rules of the control theory every task of optimal-controlled ANN should consist of the next elements:
1. A model of the controlled object’s evolution. The ANN may be described as the set of rules or using differential, integral or discrete equations. The state of ANN as a system can be described as a set of activation functions for each layer or each neuron.
2. Initial and final conditions are being set implicitly via weight functions, values of weights, initial trainspeed, sensitivities to output error etc.
3. Outer impacts the ANN usually can be realized implicitly. For example one kind of outer impacts
is even feed an initial vector on ANN’s input.

4. Time period of ANN’s functioning can be set directly (it leads to the control problem with fixed ends) or implicitly (control problem with open right end).

5. In majority of cases constraints of ANN’s state and constraints of control are not determined for ANN’s control problem.

6. In the ANN’s control problems a model of the measurements is often absent because the researcher may observe the ANN’s output directly.

7. Measurement errors and noises are coupled with some peculiarities of PC computation but sometimes are generated specially using functions of temperature fluctuations.

8. The aims of control are usually formulated as requirements for expected solutions, such as minimal output error, accuracy of classification etc.

9. A control algorithm. At present there are some general algorithms for each kind of ANN [4], such as error back propagation, BFGS, LM, gradient descent, annealing etc. But sometimes ANN requires special methods and algorithms. Its feature is absence of strict requirements which can guarantee us convergence of the ANN.

Hence an ANN is a dynamic system contains all elements of the control task. Applying the control theory methods allows us to create new algorithms of ANN’s training and functioning for different tasks, distinctly determine boundaries of using ANNs, investigate ANN’s dependence on outer conditions, mode of initialization, ANN’s hyperparameters etc.

Optimal control of ANN is not well-known till nowadays especially for non-standard kinds of ANN with other topology. One of them is a neural network with variable signal conductivity and it serves for solving traveler salesman problem, scheduling problem etc. [1].

There were 2 attempts to create optimal control of multilayer ANN with variable signal conductivity based on error signal analysis [1] and total weight analysis using special database. The main result is that this kind of ANN is controllable by output, but there is no solution how to realize this control strategy using weights and layers of the ANN. The ANN with variable signal conductivity has about $10^6$-$10^7$ links and weights; because of this computational complexity is more than $10^{10}$.

The goal of the article is to synthesize a control strategy for an elementary neuron of multilayer ANN with variable signal conductivity and to improve existing solutions.

2. Control of a Hopfield’s neuron

Let us describe evolution of an elementary neuron as an electric circuit in terms of J.J. Hopfield’s paper [2].

According the topology of the considered ANN output voltage equation can be described as (based on [2]):

$$\frac{dx}{dt} = -\frac{x(t)}{t} + \frac{w}{1+e^{-Power}} + I,$$

where $x(t)$ is output voltage of the neuron, $T = RC$ – the product of neuron’s resistance and neuron’s capacity, $I$ – the bias value, $w$ – weight of the income link of the neuron, $Power$ – the value of neurons’ output from previous layer.

For formulation the control problem of the selected neuron let’s transform equation (1) as:

$$\frac{dx}{dt} = -x(t) + u(t) + D,$$

where $u(t)$ is desired control function, $D$ is the sum of the second and the third components of the right part in equation (1).

We also have time period $[t_1, t_2]$ and the functional of control’s quality (3):

$$I = \int_{t_1}^{t_2} u^2(t)dt \rightarrow \min.$$
We arrange that there are no constraints of control. Initial conditions $x(t_1) = x_1$ are also set. As we don’t need the powers of neuron at the finishing of control $x(t_2)$ (we need only neuron’s number to calculate error function, see [3]) so we don’t set the value of neuron’s output at the final control moment. So we faced the classical isoperimetric control Lagrange problem. Let’s solve it using the Pontryagin’s maximum principle.

The Hamiltonian of the system is:

$$H = \Psi(-x(t) + u(t) + D) - u^2(t) = \Psi u(t) - \Psi x(t) + \Psi D - u^2(t).$$  \tag{4}

To find the structure of optimal control, let’s describe the first derivative of the Hamiltonian by the control function:

$$\frac{dH}{du} = \Psi(t) - 2u(t).$$  \tag{5}

The second derivative of it is equal to $-2<0$, thus equation (5) is the maximum of the control function, and the structure of the optimal control is:

$$u_{opt} = \frac{\Psi(t)}{2}.$$  \tag{6}

Let’s create the canonical system of equations to find the optimal control:

$$\begin{cases}
\frac{dx(t)}{dt} = \frac{dH}{d\Psi} = -x(t) + \frac{\Psi(t)}{2}, \\
\frac{d\Psi(t)}{dt} = - \frac{dH}{dx} = \Psi(t).
\end{cases}$$  \tag{7}

The condition of transversality is:

$$\delta F(t_2, x(t_2)) - H(t_2)\delta t_2 + \Psi(t_2)\delta x = 0, t = t_2,$$  \tag{8}

where $F$ is a terminal summand of the equation (3).

Let’s consider formula (8). As we are solving the Lagrange’s problem so the terminal summand is absent and $\delta F(t_2, x(t_2))=0$. The right end of trajectory is fixed thus $\delta t_2=0$ and $H(t_2)\delta t_2=0$. So we see that $\Psi(t_2)\delta x=0$. As the variation $\delta x$ is rather arbitrary, hence $\Psi(t_2) = 0$.

Thus, $x(t_1) = x_1$ and $\Psi(t_2) = 0$ may be considered as necessary conditions for the system (7).

The general solution of the second equation of the system (7) is:

$$\Psi(t) = C \Psi_0 e^{t-t_1}.$$  \tag{9}

Finding the value of $C$ we get the incompatible equation:

$$\Psi_0 e^{t-t_2}.$$  \tag{10}

As a result of the task (1)-(3) we may observe that:

1. The optimal control strategy we have found is a monotonic growing function. It is permitted only in the virtually world. For every neuron of the ANN it means the fixation of the control value at the final moment of control.

2. No such control function can be realized because of equation (10).

We may conclude that every neuron of the ANN with variable signal conductivity in the variant of Lagrange problem is impossible and it requires the set of the neuron’s output function at $t_2$. 


3. Analysis of ANN phase portraits
Also authors analyze the ANN’s behavior using the phase portraits in coordinates \((dE(t)/dt,E(t))\) for the converged and unconverged neural network.

![Phase portraits of the ANN](image)

**Figure 1.** Phase portraits of the ANN: converged – a; unconverged – b.

We may see the similarity. Both of portraits (figure 1) illustrate us sets of quasi-stable cycles and points with a low error level which don’t consist of the cycles. Meanwhile the speed of the error is changing during habitation at the cycles with growth of the total error.

Figure 2 presents a phase portrait in coordinates (acceleration of error, speed of error). Quasi-stable cycles are detected too.

![Phase portrait of the 2nd ordered system](image)

**Figure 2.** A phase portrait of the 2\(^{nd}\) ordered system

Behavior of multilayer ANN with variable signal conductivity [1] can be characterized with sudden jumps of the first and the second derivatives of the error signal. The reason for it is in structure of the network and in the principle of maximal link election during calculation of output.

When an even signal propagates through the connections between neurons, then at the time of activation of the neuron according to the conditions of the problem there is an insensitive zone. An odd signal propagating through the connections finds the maximum odd connection that can lead to the switched-off neuron. In this case the signal propagates through the next link with largest weight. And this connection can be separated from the point of the desired output by a significant distance (network
connections are initialized randomly). Thus, we observe a sharp increase in the output error signal.

To mitigate this effect, the model of determining the width of trained links for each neuron was changed according to the formula:

$$s = \left( E(t) + (E(t) - E(t-1)) \right)^{1/2},$$

where $s$ is the magnitude of the beam of neuron’s links for training, $E(t)$ is the error, $t$ is the epoch number.

Thus, the behavior of the ANN is controlled not only by the error, but also by its change in the previous step. If the ANN tries to increase the error, the width of the link beam increases.

As a result the phase portraits of the error have changed significantly (Figure 3).

![The phase portrait of even error signal of the neural network](image-a)
![The phase portrait of odd error signal of the neural network](image-b)

**Figure 3.** The phase portrait of the ANN considering the current and previous errors (even the error is in a, odd is in b)

Despite the fact that the fully oscillatory behavior of the ANN could not be eliminated, there are significantly lower error values and the ANN is near them for a long time. This allows us to stop the training of the ANN and get a solution for a smaller number of epochs (in 1.7–2 times) than is achieved in the usual network [3].

Another technique to reduce fluctuations in the output of the ANN during the training process is introduction of so-called control errors and post-training.

The essence of post-training is as follows: at the first epoch ANN yields primary errors and writes them as a control. At following epochs the current errors change and are taken to comparison with the results of the control errors of the previous step. If the new error is greater than the previous control error, the procedure of post-training is started; thereby the ANN tries to reduce the error to the minimum value, simultaneously striving to a quasi-stable position. In case when the new value of error is less than control errors of the previous one, the control errors are overwritten. After a certain number of epochs ANN comes to a quasi-stable state, but even there are jumps, on average, once per 275 epochs. Post-training is carried out 25 times, after which the already changed error results are being displayed.

The results are shown in the graphs below (Figure 4).

4. Conclusions
The ANN may be characterized by complex non-linear behavior. As a result of the research authors found two best modes recommended for the control of error signals. For these modes phase portraits are very similar, but post-training gives fewer oscillations within a given area.

There are no classic rigor techniques based on optimal control theory for the considered artificial neural network with variable signal conductivity. These authors suggest a combination of control methods including traditional ANN training techniques and specific approaches.
Figure 4. The phase portrait of the post-trained ANN (even error is in a, odd is in b)

Architecture of the post-training method is universal and suitable for every other control technique of the ANN (e.g. PID-controller, neuro controller, etc.). The post-training method provides a solution for a smaller number of epochs (in 1.7–2 times) than others.

Comparison between phase portraits of post-training method and ANN training with previous dynamic (with flexible beam) represents that post-training (in average 1 gap per 200 epochs) reduces oscillations of the error signal to a greater extent than the second method (in average 1 gap per 176 epochs).

5. Acknowledgments
The reported study was funded by the RFBR (project number 17-20-01065 “A theory of railway transport systems’ neural network control”).

6. References
[1] Olshansky A and Ignatenkov A 2016 About the task of running an artificial neural network with quasiperiodic nature of the behavior of the error signal Information Technologies 635–8
[2] Hopfield J J and Tank D W 1985 “Neural” computation of decisions in optimization problems Biological cybernetics 52(3) 141–52
[3] Olshansky A and Ignatenkov A 2016 Development of an artificial neural network for the construction of train schedule Bull. of the Ryazan state radio engineer. University 55 73–80
[4] Nazarov A V and Loskutov A 1 2003 Neural network algorithms of forecasting and optimization of systems (St. Peterburg: Science and Technology) 384 p ISBN 5-94387-076-8.
[5] Fahotimi O, Dembo A and Kailath T 1989 Neural network weight matrix synthesis using optimal control techniques Advances in Neural Information Processing Systems-2 (NIPS-2) (USA, Stanford, Denver, Colorado).
[6] Becerikli Y, Konar A-F and Samad T 2003 Intelligent optimal control with dynamic neural networks Neural Networks 16 251–9 Retrieved from: www.elsevier.com/locate/neunet
[7] Choi Y 2007 Neuroevolutionary algorithm and software for image processing: thesis for the degree of candidate of technical Sciences (Tomsk: Tomsk Polytechnic University) 209 p
[8] Andreeva E A and Pustarnakova Yu A 2001 Mathematical model of artificial neural network with Delay Software products and systems 3 6–9