Some approaches of improving the quality of artificial neural network training

Y.N. Rozenberg[1], A.M. Olshansky[1], I.A. Dovgerd[1], G.A. Dovgerd[1], A.V. Ignatenkov[2], P.V. Ignatenkov[3]

[1]JSC Railway Signalling Institute, Malaya Sukharevskaya square 10, Moscow, Russia, 127051
[2]JSC VTB Capital AM, Presnenskaya nab. 10, Moscow, Russia, 123112
[3]Smolensk State University, Przhevalsky str. 4, Smolensk, Russia, 214000

Abstracts. The paper is devoted to a problem of regular improving the quality of artificial neural network (ANN) training. The object of study is a complex neural network which consists of 2-dimensional Kohonen network and Wilshaw and von der Malsburg network. These networks are applied to a timetable problem for transport systems. The main existing results of using optimal control theory for ANN training are analyzed; authors suggest a new technique based on the direct neural control. Authors give comparative values of error during a training process for the traditional methods and a new approach. It is presented that this new technique is better than the traditional one for considered neural networks.

1. About optimal control in neural networks tasks

Last time the area of application for neural networks has been significantly extended. The most popular tasks are synthesis of control systems, identification tasks, data processing, information recovery tasks, scheduling problems and other original activities e.g. creating new pictures and arts.
Despite routine modifications of the structure and topologies of ANN and training methods, ANN is a system which is controllable only by using sets of recommendations, based on heuristic approaches [2], numerical experiments etc. The majority of authors directly note that the quality of ANN training and the development and creation of neural network solutions is a complicated scientific problem. Sometimes we may see attempts of the combined use ANN and optimal control theory as a rigor mathematical method for any tasks.

We should separately note the paper [6] where the author created a genetic algorithm to optimize the vector of hyperparameters for convolutional neural networks. The closest result is in [7] where an asynchrony mover is a control object and two neural networks are suggested. The first network creates a control signal; the second catches the difference between the desired output and the measurable output.

The paper [3] is devoted to constructing the optimal time sequences which consists of weights between neurons of a dynamic ANN. In [3] the two-point boundary value nonlinear problem is solved. It yields optimal rules of the ANN training. The weight matrix of the ANN in every time step (epoch) is set as an optimal time sequence. Authors note that in the best case the weight matrix at the final time step relates to the symmetric matrix constructed by J.J. Hopfield for associated memory [4].

Initial conditions are set as an input vector concatenated several training samples.

The functional (the criterion) of quality minimizes the value which is an opposite value of correlation between the output of the neuron and the desired output of the neuron at the final time step of controlling. During the time interval between the first step and the last step of controlling the functional penalizes miscorrelation level between the desired output and the answer of activation function of each neuron.

In this case an optimal control strategy is founded as Lagrange problem for a task of an optimal program control of the multilayered perceptron with a sigmoid activation function.

Another way of control is applying PID-controllers as a control technique. Authors organized and implicated about 1200 starts of the ANN with different parameters of the proportional, the integral and the differential error components and the disturbations value from 5 till 60 points per every time step. It is not a not very efficient method of control because it provides only 10% stable trajectories. The stability is taken into consideration in a Lyapunov sense [5]. Computational experiments illustrate that the marginal critical value of the disturbations feed to the ANN is no more than 10-15% of the middle error in the stable mode. This result could not be appreciated as convenient for practice.

The goal of the paper is to increase the quality of training Kohonen neural networks and ANN based on such a type. The general model of the considered ANN is described in [1] and looks like 2-dimensional modification of the Kohonen and Wilshaw and von der Malsburg network. This kind of ANN is used to create timetables for various processes including transport applications.

2. Direct Neurocontrol for multilayers artificial neural networks

Except the traditional training algorithm authors suggest a direct neurocontrol mode for training. The controllable object is a multilayered ANN with variable signal conductivity [1], a three layer perceptron with sigmoid activation functions is considered as a controller.

The main scheme is given in the fig.1.

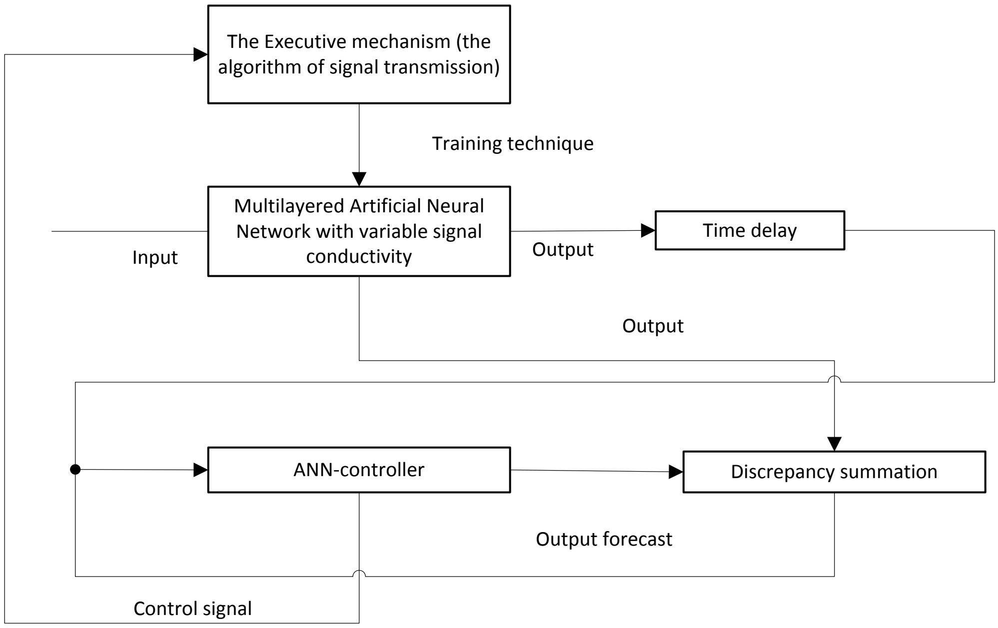

Figure 1. The scheme of a direct neurocontrol mode.

The ANN-controller is trained by the aggregation of triple sets “ The level of error per epoch” – “The level of error at the previous moment” – “The control signal from a previous time step to present time step” or “The previous level of error” – “The current level of error” – “The control signal”.

The current error signal of the ANN and the previous one are gathered and entered the trained and ready multilayer perceptron. An answer signal of the ANN-controller entered the discrepancy summation and actuating mechanism (an algorithm). Hereinafter the value of summated discrepancy is also fed by the ANN-controller.

The control scheme described above was tested for the concrete scheduling problem (the railway branch Arkhara – Volochaevka, 27 railway stations). The task included 185 trains per 24 hours.

The results of testing are given in the table 1.

Table 1. A comparison of different training methods.

Training error (points)	Traditional algorithms	Direct neurocontrol
Min	75	193
Max	134795	57895
Median	5469	210
Average	16548	384
SD	6687	1180
Rate of error overshoot (per 100 epochs)	50	0,4

3. Conclusions

Thus the paper shows us the principal opportunity to control the multilayered artificial neural network with variable signal conductivity. The three layered perceptron with the sigmoidal activation function can be considered as a controller. The quality of solutions achieved by multilayered artificial neural networks with direct neurocontrol shifts the level of solutions with traditional training algorithms.

4. Acknowledgements

This research was provided by Russian Foundation for Basic Research (the project #17-20-01065 “A theory of railway transport system neural network control”).

5. References

[1] Olshansky, A. One approach to control of a neural network with variable signal conductivity / A. Olshansky, A. Ignatenkov // Information Tecnologiesn and Nanotechnologies (ITNT). – 2017. – P. 984-987.

[2] Nazarov, A.V. Neural networks algorithms of forecast and system optimization / A.V. Nazarov, A.I. Loskutov – Saint Petersburg: Science and technic, 2003. – P. 384.

[3] Fahotimi, O. Neural network weight matrix synthesis using optimal control techniques / O. Fahotimi, A. Dembo, T. Kailath // USA, Stanford,1989.

[4] Hopfield, J.J. Neural networks and physical systems with emergent collective computational abilities // Proc. Natl. Acad. Sci. USA. – 1982. – Vol. 79. – P. 2554-2558.

[5] Dorf, R.C. Modern control systems / R.C. Dorf, R.H. Bishop – Pearson, 2011.

[6] Tsoy, Y.R. Neuroevolution algorithm and software for image processing: The dissertation on competition of a scientific degree of Candidate of technical sciences / Y.R. Tsoy – Tomsk: Tomsk Polythechnic University, 2007. – P. 209.

[7] Sagdatullin, A.M. A neural network controller for the velocity value of an asynchronic motor // Theses of Russian Congress for Control Problems – Moscow: Institute of control sciences of Russian Academy of Sciences, 2014. – P. 4485-4498.