Estimation of applicability of modern neural network methods for preventing cyberthreats to self-organizing network infrastructures of digital economy platforms [a,b]

Maxim Kalinin[1*], *Vasiliy* Krundyshev[1], and *Evgeny* Zubkov[1]

[1] Peter the Great St.Petersburg Polytechnic University, Department of Information security of computer systems, 195251, Politechnicheskaya st., 29, Russian Federation

Abstract. The problems of applying neural network methods for solving problems of preventing cyberthreats to flexible self-organizing network infrastructures of digital economy platforms: vehicle adhoc networks, wireless sensor networks, industrial IoT, "smart buildings" and "smart cities" are considered. The applicability of the classic perceptron neural network, recurrent, deep, LSTM neural networks and neural networks ensembles in the restricting conditions of fast training and big data processing are estimated. The use of neural networks with a complex architecture– recurrent and LSTM neural networks – is experimentally justified for building a system of intrusion detection for self-organizing network infrastructures.

1 Introduction

Digitalization and Internetization of economy have specified a new problem – the problem of cyberthreats having active and destructive impact on automated and connected cyberphysical systems and, through them, on entire network infrastructure. The object of protection previously understood as a combination of data with limited access obtains new representation as an element of cyberspace where traditional operations of reading/writing have physical consequences. Cyberthreats impacting the modern infrastructure of digital economy platforms – communications of unmanned transport, wireless sensor networks of digital industry, industrial IoT, "smart buildings", and "smart cities" – are dangerous because

[a] With financial support from the Ministry of Education and Science of the Russian Federation within the framework of the Federal Special Purpose Program "Studies and projects in the priority fields of development of the scientific technological complex of Russia for 2014-2020" (Agreement 14.575.21.0131 of September 26, 2017, 09.2017, unique identifier RFMEFI57517X0131).
[b] The results of the work were obtained with the use of computing resources of the super-computer center of Peter the Great St. Petersburg Polytechnic University – Super-Computer Center "Politekhnichesky" (http://www.spbstu.ru).
[*] Corresponding author: sci@ibks.spbstu.ru

they carry a threat not so much for data as for the work of devices, executive subsystems, physical processes and, further, for resources, finances, ecology. and people life.

Self-organizing network infrastructures (for example, VANET – a network among connected cars, FANET – a network between flighting facilities, MARINET – a network between floating crafts, WSN – wireless sensor networks, IIoT – industrial Internet of Things) have peer-to-peer architecture, support multiple links between hosts and dynamic control of routing at each network node that determines their advantages over traditional networks, that is the possibility of multi-line transfer of data to large distances without stationary retranslators, immunity to spacial variations in network topology, dynamic reconfiguration of a network under the conditions of distorting actions and faults [1]. The revealed opportunities have created new conditions for implementing new cyberthreats including bringing equipment out of operation, traps to cyberphysical systems, cyberterrorism on transport and in industry, organizing large-scaled megabotnets of sensor networks, organizing massive failures of unmanned vehicles, digital fabrics, intellectual industrial and smart city systems.

Distribution of cyberthreats to vehicular and industrial network infrastructures is temporarily restrained by minor penetration of inter-machine communications but there are already incidents with industrial networks and autonomous cars which demonstrate their complete unsafety against occured threats (for example, [2]). As introduction of security means into inter-computer network is a complicated procedure due to the peculiarities of networks of such a type (decentralization, dynamics, peer-to-peer character, etc.) and the limitations of computational resources on hosts, active demand for methods of aprior protection is formed. These methods should provide for timely determining of cyberthreats and, therefore, increasing level of protection without interfering network infrastructure. The basis for solving the specified problem is made of different types of artificial neural networks (ANN) which have appeared last decade; including deep ANN, recurrent networks, LSTM networks, and their ensembles.

2 The related works in the field of preventing cyberthreats to self-organizing network infrastructures

The task of any intrusion detection systems (IDS) is detecting a cyberattack according to some indicators [3]. In contemporary self-organizing network infrastructures, the volume of data to be processed rapidly increases; this information should be processed by IDS in real-time. For example, a network consisting of 1000 cyberobjects of digital factory (there is a control bus and about 20 controllers in each cyberobject) generates about 4 million of data sets per minute, which can influent the undercontrolled processes. Big data significantly complicates the task of developing new IDS – so the task of providing their effectiveness and, primarily, their productivity becomes the essential one.

Different approaches to solve the problem of cyberthreats detection are known:

a) statistical approach [3, 4] (a pattern profile of the system is created, then it is being averaged in the process of training, any deviation from the pattern profile is considered to be an intrusion). The disadvantage of this approach is that the intruder may train the system gradually putting the system into the condition that IDS doesn't detect any anomalies;

b) forecasting of the patterns [3, 5] (IDS approximates the profile and predicts the future events basing on the previous ones). However, many of the modern cyberattacks cannot be described by those rules and they will be missed by the detector;

c) artificial neural networks (ANN) [3, 6-8]. The ANN apparatus gives opportunities for parallel processing of data and revealing hidden dependencies and deviations. The process of searching for anomalies consists in selecting significant properties, preparing parameters and training the neural network, at that, a profile of normal activity is formed. During the work,

real data is conveyed to the ANN input and the ANN determines their belonging to this or that class. For the ANN ability for training and generalizing, there is an opportunity for predicting and searching for anomalies information about which did not take part in training.

Though neural network methods have been already actively developed for many years, the interest in this field has been lately increasing and this is coused by progress in IT and demand for fast processing of big data which has appeared. ANN provide significant advantages against other methods of adaptive data processing: accounting hidden dependencies in datasets, automatic selection and optimization of significant attributes for object classification that reduces time for developing and implementing methods of data analysis, solving intellectual tasks by approximation of arbitrary functions. Alternative methods of machine learning and data mining don not work with the dependencies directly and have to use heuristic control. The ANN apparatus solves this problem initially by means of parallel distribution of signals. Only ANNs possess the two last features which enable focusing specifically on ANN methods and that makes developments in this field scientifically and practically significant.

3 Investigation of modern neural networks for intrusion detection in self-organizing network infrastructures

Uncertainty of security perimeter of network infrastructure, absence of security means for protecting self-organizing networks, unsuitability of the traditional methods of network protection created for stationary computer networks confirm the necessity of developing special methods and means for objects of the given type on the basis of the approach of detecting cyberthreats.

Traditional ANNs are multilayer perceptrons with logical transferring functions [7], module ANNs of direct and back signal propagation [7]. Being universal and flexible, they still do not satisfy the increased requirements to the volume of training data sets (the problem of saturation on "big data"), have a low coefficient of effectiveness (not higher than 95% in most cases), the time of training is very long, and they are not able to build dependencies of great complexity. Classic ANNs are characterized by extreme redundancy (width and depth of an ANN) and resource intensity (growth of requirements to the volume of the memory and CPU capacity) [8]. For operative control of security of self-organizing network infrastructures compact and fast ANNs are needed.

The classic perceptron of direct signal distribution and actively developing ANNs have been selected for estimation of applicability and correlation of modern ANNs in comparison with traditional ones for solving the problem of detecting cyberthreats aimed at flexible network infrastructures: recurrent ANN [9], deep ANN [10], LSTM neural network [11], and ensemble of neural networks [12]. Program modeling has been carried out for them with the use of the computing capacities of the supercomputer center of Peter the Great St.Petersburg Polytechnic University – SCC "Polytechnichesky". A typical cyberattack was chosen as a test one: "Black Hole" [12] which is aimed at dynamic routing in self-organizing networks, exploits their specific features and there are no solutions for its detection with traditional IDS. The results obtained in the course of training the investigated ANNs are listed in table 1.

Table 1. The results of training the investigated ANNs.

Neural network model	Error (for the training set)	Size of the training/ working set	Training time, s
Perceptron	0.002	22/11	3.9
Recurrent ANN	0.04	4500/1500	44.5
Deep ANN	0.001	22/11	2.2

Neural network model	Error (for the training set)	Size of the training/ working set	Training time, s
LSTM-network	0.04	4500/1500	50.6

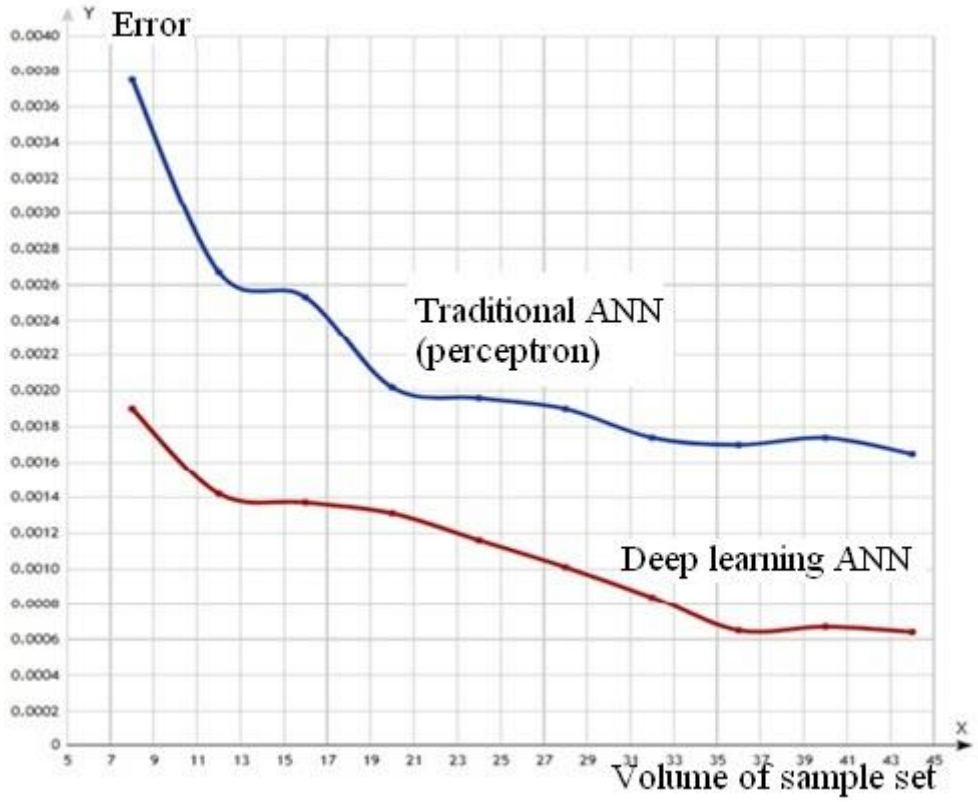

Fig. 1. Dependencies of the error coefficient on the volume of a sample set for perceptron and deep ANN.

ANNs with a simple architecture (perceptron and deep ANN) get a win against modern ANNs with a complex architecture. Parameters of an ANN ensemble were defined as the maximum value of the training time, the maximum value of the error coefficient and the largest set of data which is necessary for training the components of an ANN. Perceptron is not adapted to working with "big data", a recurrent ANN and a deep ANN are not sensitive to "big data" while training. According to the results of modeling, a similarity between the behavior of perceptron and a deep ANN is noted (Fig. 1). When using a set of an analogous size, a deep ANN is trained with a smaller error coefficient, at this, a sharp drop when increasing the set is not observed. Comparing the characteristics of a recurrent ANN and a LSTM network (Fig. 2) demonstrates their similarity, however, a recurrent ANN has a series of drops in the process of training which are missing in the analogous process in an ANN with memory because the latter uses the optimization mechanism – the memory enables locking the result through a feed-back and remembering the conclusions made.

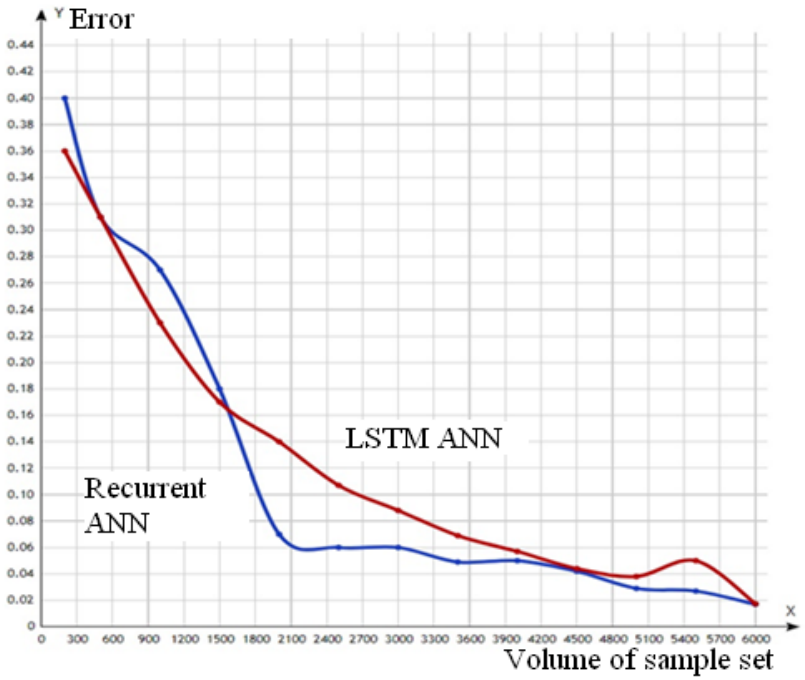

Fig. 2. Dependencies of the error coefficient on the volume of a sample set for recurrent and LSTMnetworks.

Taking into account the sensibility of an ANN to the volume of a sample set during training and taking into account the fact that self-organizing network infrastructures according to their characteristics are the objects in which "big data" about the security of nods will be constantly generated and it is necessary to process them while solving the tasks of detecting cyberthreats, it is necessary to validate ANNs with respect to their readiness to process large datasets because it is one of the main advantages of modern ANNs over traditional ANNs and other detection models. In order to understand how well networks function with a big datasets, for exact comparison, three formulas of calculating the error coefficient are distinguished [13]:

mean squared error (MSE): $\frac{(R_{O_1}-R_{\Pi_1})^2+(R_{O_2}-R_{\Pi_2})^2+\cdots+(R_{O_N}-R_{\Pi_N})^2}{N}$;

root mean squared error (Root MSE): $\sqrt{\frac{(R_{O_1}-R_{\Pi_1})^2+(R_{O_2}-R_{\Pi_2})^2+\cdots+(R_{O_N}-R_{\Pi_N})^2}{N}}$;

squared arctangent (Arctan): $\frac{\arctan^2(R_{O_1}-R_{\Pi_1})+\arctan^2(R_{O_2}-R_{\Pi_2})+\cdots+\arctan^2(R_{O_N}-R_{\Pi_N})}{N}$.

where R_{O/Π_i} is the relation of expected/obtained result, where *i* is an index, $i = 1, 2, \ldots, N$; N is the volume of data which is processed in the ANN. In Table 2 it is shown how ANNs are able to process "big data" (the volume is 4 mln attributes).

Perceptron demonstrates short time of processing but its error on a big dataset is quite large comparing with other ANNs. It is also unstable at the stage of training. A deep ANN is void of this problem at the training stage but at the stage of processing it demonstrates unsatisfactory results as well because ANNs with such architecture are more suitable for the tasks where many resulting responses are predetermined. It is recommended to use modern ANNs with a complicated architecture, recurrent ANNs and their ensembles, for solving the task.

Table 2. Comparison of ANNs for processing "big data".

Neural network model	Error (for the working set)			Processing time, s
	MSE	*Root MSE*	*Arctan*	
Perceptron	0.37	0.61	0.23	2.1
Recurrent ANN	0.01	0.13	0.01	28.8
LSTM-network	0.02	0.14	0.01	27.2
Deep ANN	0.50	0.71	0.31	3.0

4 Conclusion

The experimental results of investigation demonstrate inapplicability of traditional approaches to solving the task . The traditional neural network becomes useless for big data processing in time and resource limitations.. Validating recurrent ANN and LSTM networks shows that complicated modern ANN are able to cope with the task because they apply the mechanisms which optimize their operation (for example, memorization). Efficiency of solving the task can be raised by simultaneous application of several polytypical ANNs forming heterogeneous ensembles of neural networks.

Using new neural network methods will enable building modern systems for detecting intrusions for the first time; they will make it possible to block unauthorized activities in time in flexible self-organizing network infrastructures of digital economy.

Project is financially supported by Ministry of Education and Science of Russian Federation, Federal Program "Researching and Development in Priority Directions of Scientific and Technological Sphere in Russia within 2014-2020" (Contract No. 14.575.21.0131, September 26, 2017, the unique identifier of the agreement RFMEFI57517X0131).

Project results are achieved using the resources of supercomputer center of Peter the Great St.Petersburg Polytechnic University – SCC “Polytechnichesky” (www.spbstu.ru).

.

References

1. B. Krishna, International Journals of Advanced Research in Computer Science and Software Engineering, **7(7)** (2017)
2. R. Langer. To Kill a Centrifuge [online] Available at: www.langner.com/wp-content/uploads/2017/03/to-kill-a-centrifuge.pdf (2018)
3. M. Erritali, B. El Ouahidi, International Journal of Engineering and Technology, **5(2)** (2013)
4. K. R. Karthikeyan, A. Indra. Intrusion Detection Tools and Techniques – A Survey. International Journal of Computer Theory and Engineering, **2(6)** (2010)
5. X. Li. Probabilistic techniques for intrusion detection based on computer audit data. *IEEE* Transactions on Systems Man and Cybernetics. Part A: Systems and Humans, **31(4)** (2001)
6. S. Sodiya, O. A. Ojesanmi, O. C. Akinola, O. Aborisade, International Journal of Computer Applications, **106 (18)** (2014)
7. M. A. Widrow, Proceedings of the IEEE, **78(9)** (1990)
8. S. Nikolenko, A. Kadurin, E. Arkhangelskaya, *Deep education. Immersion into the world of neural networks (*"Piter" Publishing House, Saint Petersburg, 2018)
9. A. Mallya. *Introduction to RNNs* [online] Available at: *http://slazebni.cs.illinois.edu/spring17/lec02_rnn.pdf* (2018)
10. V. Sze, Y. H. Chen, T. J. Yang, J. S. Emer, Proceedings of the IEEE, **105(12)** (2017)

11. R. Adhikari, R. K. Agrawal, International Journal of Computer Applications, **32 (7)** (2011)
12. N. K. Chaubey, International Journal of Security and Its Applications, **10(5)** (2016)
13. M. V. Shcherbakov, A. Brebels, N. L. Shcherbakova, A. P. Tyukov, T. A. Janovsky, V. A. Kamaev, A Survey of Forecast Error Measures [online] Available at: www.researchgate.net/publication/281718517_A_survey_of_forecast_error_measures (2018)