Prevention of cyber attacks in smart manufacturing applying modern neural network methods

V Krundyshev* and M Kalinin
Peter the Great St. Petersburg Polytechnic University, Saint Petersburg, Russia
E-mail: vmk@ibks.spbstu.ru

Abstract. Digital transformation is a driver of a modern approach to providing actual cyber security. In the context of the globalization of the economy, digital technologies are actively being introduced into production, factories and plants. Due to the increased mobility of topology and the growing amount of data undergoing the processing, traditional methods of protection becomes ineffective, so the researchers are faced with the task of creating new methods for ensuring cyber security that meet new challenges. The paper analyses new artificial neural network (ANN) architectures corresponding security tasks in cyber physical systems, identifies their major advantages, and evaluates the possibility of their application for solving the problem of attacks prevention in case of machine-to-machine smart manufacturing. A meta-neural system of comprehensive protection against cyber attacks on dynamic routing in intelligent production has been developed. Using the resources of a supercomputer, an experimental study of the developed neuroframework was carried out. Test results showed that the proposed solution provides high accuracy in detecting cyber attacks.

1. Introduction
Currently, the world is on the verge of a new, fourth industrial revolution (Industry 4.0), which will lead to the complete automation of most production processes, and, as a result, increase labor productivity, economic growth and competitiveness of its leading countries. The concept of Industry 4.0 was formulated in 2011 by the president of the World Economic Forum in Davos, Klaus Schwab [1]. The main idea of Industry 4.0 is to accelerate the integration of cyberphysical systems into factory processes, as a result of which a significant part of the production will take place without human intervention. Industry 4.0 is associated with such concepts as Industrial Internet of Things, digital enterprise and smart factories of the future. According to World Bank estimates, Industry 4.0 could bring the world economy up to $ 30 trillion [2]. In fact, the flexible interaction of various physical systems through digital technology is changing the look not only of industry, but of the economy as a whole.

The globalization of the economy has led to the formation of a single market and information space, most of the goods and services are produced by transnational corporations [3]. The unification of geographically distributed participants in the design and production processes into a single network will certainly lead to cyber threats. Due to the peculiarities of modern digital infrastructures and the rapidly growing volume of data being processed, traditional security methods are becoming ineffective, so researchers are faced with the task of creating new cybersecurity methods that meet the current challenges of the time - access control methods between systems in a dynamic infrastructure, global regulation information and control flows.
According to experts from the analytical company Frost & Sullivan, cyber attacks in the energy sector alone cost $13.2 million annually [4]. The results of the Kaspersky Lab study show that incidents involving IIoT devices are among the three threats with the greatest financial damage to companies, and the lack of uniform standards remains one of the main problems in the field of cybersecurity in digital production [5]. Leading experts agree that it is necessary to develop new methods for detecting cyberthreats [6-8]. Artificial intelligence technologies will allow for the efficient processing of intensively arriving unstructured ultra-high volume data and extracting knowledge from it [9, 10]. Figure 1 shows the proposed neural network intrusion detection system in smart manufacturing.

![Figure 1. Cybersecurity of smart manufacturing.](image)

To detect security threats in m2m networks, we have proposed artificial neural networks (ANN) of new generation to be applied for intrusion prevention. To train intrusion prevention system (IPS), there is needed a dataset of samples of intrusion signs that would reflect to the insecure behaviour in the m2m network under different scenarios. Our task is to develop an intrusion prevention system that is based on artificial intelligence, using neural networks and synthetic datasets to maximize the detection of cyber attacks. This paper is organized as follows: in Section 2 an analysis of the related research works devoted to various ANN architectures, Section 3 presents implementation of a neural network system for detecting a complex of attacks, and finally a conclusion is provided in Section 4.

2. Methods

There are a lot of ANN architectures, and before creating a system for prevention of the cyber attacks, we conducted an analysis of numerous scientific papers to determine the applicability of these architectures to solve the problem.

A large number of studies have been devoted to recurrent neural networks (RNN) [11,12]. The authors note the following advantages: contain the mechanism of "memory" of the network, solve complex problems, the results of the work are very accurate. However, there are several disadvantages: imply a long learning curve for the network, require a large dataset for training the network. ANNs of this architecture are a more complex class of ANNs, but this class is able to perform the existing task in view of the network's "memory" mechanism. The architecture is applicable.
Recently gaining popularity spiking neural networks [13]. Advantages of such architecture: work with physical quantities, copy the processes of biological neural networks. However, there is one significant drawback: require special hardware to operate. ANNs of this architecture work with data that are not real numbers. The architecture is not applicable.

Neural networks of direct propagation (perceptrons) are considered in detail in these articles [14,15]. The authors note the following advantages of such neural networks: have a simple architecture, imply a short learning curve, have a short work time. Disadvantage: the accuracy of the results depends on the complexity of the task. ANNs of this architecture are ANN with the simplest architecture, which is why it is necessary to consider how ANN with such architecture will cope with the task. The architecture is applicable.

There are a number of studies on neural networks with fuzzy logic [16,17]. The researchers note advantages and disadvantages. Advantages: use the mechanisms of fuzzy logic and the results of the work are very accurate. Disadvantage: require non-standard data generation mechanisms. ANNs of this architecture refer to fuzzy logic, implying not a standard logical result (0 or 1). The architecture is not applicable.

In [18], T. Bakir et al. examine another architecture of ANN - neural networks with wavelet transform (ANN wavelet), the main advantages are: have a simple architecture, a special activation function allows you to perform the classification task, have a short work time. The disadvantage is the accuracy of the results depends on the complexity of the task. ANNs of this architecture are a modification of the classical perceptron architecture in order to perform data classification and can solve the existing problem. The architecture is applicable.

In addition, the great attention of researchers is focused on recurrent neural networks with long short-term memory (LSTM RNN) [19-21]. Advantages: contain the mechanism of "memory" of the network, solve complex problems, the results of the work are very accurate. However, there are several disadvantages: mean a long time learning network, require a large dataset for training the network. ANNs of this architecture solve the problem of the classical recurrent neural network with the help of the improved "memory" mechanism of the network. The architecture is applicable.

There are also a number of works devoted to generative adversarial neural networks [22, 23]. Among the advantages of the researchers note: work with data of complex nature (sound, image), applicable to solving security problems associated with data of complex nature (providing authentication). The disadvantages are: have a complex architecture, imply a long learning curve for the network. ANNs of this architecture work with composite data. And the main task is simple data (real numbers). The architecture is not applicable.

V. Sze et al. in [24] examined deep neural networks (DNN), noting the advantages: have a deep learning structure and the results of the work are of medium accuracy. Main disadvantage is the accuracy of the results depends on the complexity of the task. ANN of this architecture, being a logical continuation of the architecture of neural networks of direct distribution, can solve the existing problem. The architecture is applicable.

In addition, the authors in [25,26] investigated meta-neural networks and heterogeneous ensembles of neural networks. Advantages: are a collection of neural network models and have average values for all parameters of those ANNs that are part of the ensemble. Disadvantages: the accuracy of the results depends on the ANN included in the ensemble, mean a long time learning network and require a large dataset for training the network. The ANN of this architecture will be considered as a cluster of those ANNs that are applicable to the existing task. The architecture is applicable.

Therefore, in order to determine which ANN would better detect a specific attack, it was decided to conduct research on the following ANN architectures: perceptron, RNN, DNN, ANN wavelet, LSTM RNN and a meta-neural network / heterogeneous ensemble of neural networks. Next, it was necessary to simulate these architectures in order to determine the best ANN architecture suitable for solving the task of detecting an attack on the network.
3. Results and Discussion
Since for each attack from the class of attacks aimed at dynamic routing, a detection model and its associated neural network method were developed, then it was necessary to consider the interaction of neural network methods.

The result of research was the neural network system (Figure 1), consisting of the following ANN:
- neural network method based on LSTM RNN with activation function "Rectifier", optimization method "Adam" and size of training dataset is 5000 elements, detecting the attack "Black Hole";
- neural network method based on LSTM RNN with the activation function of sigmoid, optimization method "Adam" and size is training dataset of 5000 elements, detecting attack "Gray hole";
- neural network method based on ANN wavelet with activation function in the form of wavelet transform, optimization method "Adadelta" and the size of training dataset is 400 elements, detecting the attack "Wormhole";
- neural network method based on the LSTM RNN with the activation function "Rectifier", the optimization method "Adam" and the size of the training dataset is 1000 elements that detects the "Denial of Service" attack.

This new neural network framework is a complex of different neural network methods (Fig. 2) each of which is addressed to a specific type of cyber attack.

![Figure 2. Structure of a meta-neural framework.](image)

Table 1 shows a result of testing the neural network framework in case where all ANNs were involved and when each ANN has received datasets of the same size at the input.

| Volume of dataset | Accuracy, % | Data set processing time, s |
|-------------------|-------------|-----------------------------|
| 100 000           | 96.31       | 82                          |
| 500 000           | 96.18       | 204                         |
| 1 000 000         | 96.23       | 412                         |
| 2 000 000         | 96.26       | 778                         |
| Volume of dataset | Accuracy, % | Data set processing time, s |
|--------------------|-------------|----------------------------|
| 3 000 000          | 96.33       | 1124                       |
| 4 000 000          | 96.02       | 1345                       |
| 5 000 000          | 96.13       | 1602                       |
| 10 000 000         | 96.10       | 1872                       |

Based on the test results, it can be concluded that the neural network framework coped with the task of preventing the attacks specific to m2m infrastructures in smart manufacturing, i.e. the attacks on dynamic network routing. In spite of the big data volume (5…10 million), there is a significant increase in the data processing time.

4. Conclusion
An extensive study was conducted to protect m2m networks in smart manufacturing from cyberattacks. Protection has been developed by determining the normal (legitimate) activity of nodes and the characteristics that correspond to it. Violation of signs signals the attack in the target network and allows us to identify it. In determining the better configuration of ANN, it has been decided to focus on 3 components - the functions of activation of the hidden layer, the optimization methods and the dimension of the training dataset. The accuracy of the ANN is a guarantee of detecting an attack. To confirm that the accuracy is observed not only at the training stage, then for each network a testing phase was conducted. This stage showed how each network handles data that is not part of the training dataset, and how long it takes an ANN to process large data. The final stage of the investigation consisted in examining the neural network system, which is capable of detecting the entire class of attacks aimed at dynamic routing. Since this system consists of ANNs that have passed the testing phase for specific attacks, the entire system also goes through this stage and gives the exact answer.

Acknowledgments
The work was funded by the Russian Federation Presidential grants for support of young scientists and postgraduate students (SP-443.2019.5).
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] Schwab K 2016 The Fourth Industrial Revolution
[2] World Economic Forum The Fourth Industrial Revolution Davos 2016 [online], Available at: https://www.weforum.org/agenda/2016/01/the-fourth-industrial-revolution-what-it-means-and-how-to-respond
[3] Kordos M and Vojtovic S 2016 Transnational corporations in the global world economic environment Proc. of the 3rd International Conference on New Challenges in Management and Organization: Organization and Leadership, 2 May 2016, Dubai, UAE pp150–158
[4] Global Industrial Cybersecurity Services Market, Forecast to 2022 2016 [online], Available at: https://go.frost.com/EI_PR_EDebowska_K2C6_IndustrialCyberSecurity_Nov18
[5] Kaspersky Lab From Shaking Their Hands to Paying Off Their Debts: Third party Cybersecurity Failures Cost Businesses the Most 2017 [online], Available at: https://www.kaspersky.com/about/press-releases/2017_third-party-cybersecurity-failures-cost-businesses-the-most
[6] Kalinin M, Krundyshev V and Zubkov E 2018 Estimation of applicability of modern neural network methods for preventing cyberthreats to self-organizing network infrastructures of digital economy platforms The Convergence of Digital and Physical Worlds: Technological Economic and Social Challenges (CC-TESC2018) SHS Web of Conf, vol 44 pp 1–7
[7] Lavrova D, Zegzhda D and A. Yarmak 2019 Using GRU neural network for cyber-attack detection in automated process control systems 2019 IEEE International Black Sea Conference on Communications and Networking (BlackSeaCom), Sochi, Russia, pp 1–3

[8] Demidov R, Pechenkin A, Zegzhda P and Kalinin M 2018 Application Model of Modern Artificial Neural Network Methods for the Analysis of Information Systems Security Automatic Control and Computer Sciences, December 2018, vol 52 i 8 pp 965-970

[9] Kalinina O, Balchik E and Barykin S 2019 Innovative management neural network modelling based on logistic theory MATEC Web of Conferences, vol 239, 04021

[10] Zegzhda P, Zegzhda D, Pavlenko E and Ignatev G 2018 Applying deep learning techniques for Android malware detection Proceedings of the 11th International Conference on Security of Information and Networks, Cardiff, United Kingdom, pp 1–7

[11] Fernández S 2007 An application of recurrent neural networks to discriminative keyword spotting Proc. of the 17th Int. Conf. on Artificial Neural Networks (Springer-Verlag) pp 220–229

[12] Graves A and Schmidhuber J 2009 Offline handwriting recognition with multidimensional recurrent neural networks Advances in neural information processing systems pp 545–552

[13] Montuschi P 2015 Spiking Neural Network Architecture Computer vol 48 i 10 p 6

[14] Govindarajan M and Chandrasekaran R 2014 A Hybrid Multilayer Perceptron Neural Network for Direct Marketing DOI: 10.4018/ijkb.2012070104

[15] Gallant S 1990 Perceptron-based learning algorithms IEEE Transactions on Neural Networks, vol 1 no 2 pp 179–191

[16] Ramya R, Anandanatarajan R, et al. 2012 Applications of fuzzy logic and artificial neural network for solving real world problem IEEE-Int. Conf. On Artificial Intelligence, Science And Management (ICAESM -2012) pp 443–448

[17] Ghalia M and Alouani A 1995 Artificial neural networks and fuzzy logic for system modeling and control: a comparative study Proc. of the 27th Southeastern Symp. on System Theory pp 258–262

[18] Bakir T, Boussaid B, et al. 2016 Artificial neural network based on wavelet transform and feature extraction for a wind turbine diagnosis system 2016 14th Int. Conf. on Control, Automation, Robotics and Vision (ICARCV) pp 1–6

[19] Chao L, Tao J, et al. 2016 Long short term memory recurrent neural network based encoding method for emotion recognition in video IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP)

[20] Dai J, Liang S, et al. 2016 Long short-term memory recurrent neural network based segment features for music genre classification 10th Int. Symp. on Chinese Spoken Language Processing (ISCSLP) pp 1–5

[21] Youmans M, Spaninou C and Qiu P 2017 Long short-term memory recurrent neural networks for antibacterial peptide identification IEEE Int. Conf. on Bioinformatics and Biomedicine (BIBM) pp 498–502

[22] Radford A, Metz L and Chintala S 2015 Unsupervised representation learning with deep convolutional generative adversarial networks arXiv preprint arXiv:1511.06434

[23] Yin C, Zhu Y, et al. 2018 An enhancing framework for botnet detection using generative adversarial networks Int. Conf. on Artificial Intelligence and Big Data (ICAIBD), pp 228–234

[24] Sze V, Chen Y, et al. 2017 Efficient Processing of Deep Neural Networks: A Tutorial and Survey in Proc. of the IEEE vol 105 no 12 pp 2295–2329

[25] Naik D and Mammone R 1992 Meta-neural networks that learn by learning Int. Joint Conference on Neural Networks Baltimore, USA

[26] Hans A and Udluft S 2010 Ensembles of Neural Networks for Robust Reinforcement Learning 2010 9th Int. Conf. on Machine Learning and Applications pp 401–406