https://www.proconference.org/index.php/gec/article/view/gec26-01-008
DOI: 10.30888/2709-1783.2023-26-01-008

UDC 004.056

POTENTIAL USE OF NEURAL NETWORKS TO DETECT ANOMALIES INTRUSIONS IN NETWORK TRAFFIC

H. Haidur
d.t.s., prof.
ORCID: 0000-0003-0591-3290
S. Gakhov
c.m.s., as.prof.
ORCID: 0000-0001-9011-8210
A. Bryhynets
student
ORCID: 0009-0008-7631-415X
State University of Telecommunications, 03110, Kyiv, Solomenska, 7

Abstract. *The rapid digitalization of the world has led to various attacks on computer systems and networks, so cybersecurity of networks is an extremely important and relevant component of information security today. In our study, we compare two deep learning models, specifically a recurrent neural network and a convolutional neural network, for detecting anomalies in network traffic. Both neural networks have proven to be useful in a wide range of applications. The aforementioned technologies are currently not widespread in intrusion detection and network anomaly detection systems due to their novelty, so they require more thorough research. Conventional machine learning algorithms will eventually become insufficient, as they do not have as good learning capability as deep learning neural networks. We will provide a detailed analysis of the capabilities of recurrent and convolutional neural networks along with long short-term memory layers, which may be useful in further research and applications.*

Key words: *cybersecurity of an information system, intrusion into an information system, intrusion detection, machine learning, neural network.*

Introduction.

The rapid development of information and communication technologies currently offers the necessary means of processing and exchanging information and tries to deal with the growing needs of individuals and society. Unfortunately, the development of computerization was accompanied by the development of methods and means of malicious actions, the motives of which are as numerous as they are dangerous and evolve over time. All this makes the problem of ensuring cybersecurity of information systems of organizations acutely relevant.

A promising area of ensuring the cybersecurity of information systems of organizations in the face of harmful influences is the creation and application of automatic protection systems based on artificial intelligence that detect and respond to intrusions in real time.

The development of an effective anomaly detection system involves extracting relevant information from a large amount of "contaminated" data of high dimensionality. For this purpose, network monitoring devices are used to collect statistical data at high speed. Different anomalies manifest themselves in network statistics in different ways, so it is difficult to develop general models of normal network behavior and anomalies.

Deep learning is best represented by two algorithms: convolutional neural networks (CNNs) for image recognition and recurrent neural networks (RNNs),

which are mainly used for natural language processing and speech recognition [2]. CNNs have a local receptive field and a shared weighting kernel that can represent spatial features by extracting basic visual characteristics such as oriented edges, endpoints, and angles [3]. RNN has a very deep structure that connects the basic neural units in a chronological order and is usually effective for modeling sequential data by training using gate nodes such as long short-term memory (LSTM) units [4]. In addition, the combination of CNN and LSTM layers is being studied to extract temporal and spatial characteristics.

Main text.

Deep learning consists of a variety of machine learning techniques that use streams of nonlinear nodes arranged in multiple layers that extract and transform the values of entity variables from an input vector. The individual layers of such a system have the output of the previous layers as input, except for the initial input layer, which receives signals or input vectors from the external environment.

In addition, unsupervised or supervised methods can be used to train systems. This leads to the possible application of these models to supervised learning tasks such as classification and unsupervised tasks such as pattern analysis. Deep learning models also rely on extracting higher-level entities from lower-level entities to obtain a stratified representation of the input data using an unsupervised learning approach at different levels of entities.

Deep learning uses multiple levels of nonlinear processing nodes to extract and transform features. Successive layers use the output of previous layers as input.

In practice, all deep learning algorithms are neural networks that share some basic properties. They all consist of interconnected neurons organized into layers. They are distinguished by the network architecture (or the way neurons are organized in the network), and sometimes by the way they are formed.

Recurrent Neural Networks (RNNs) are networks that contain feedback and allow storing information.

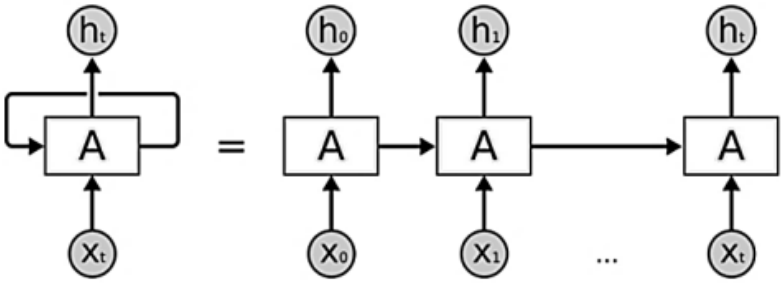

Figure 1 - The recurrent network in deployment

In Figure 1, a fragment of the neural network A takes an input value x_t and returns the value h_t. The presence of feedback allows you to transfer information from one-step of network training to another one.

A special form of a recurrent neural network is the LSTM, which is widely used for processing time series data. In a standard RNN, the output of any layer depends not only on the current input, but is also based on the previous output.

Most modern classical LSTMs now include adjustments, including skipping the forgetting filter layer and cell associations. Other variations also include a less complex Gated Recurrent Unit (GRU) design.

Recently, RNN-LSTM has been widely used to analyze log data based on the similarity of LSTM methods used in natural language processing [5, 6]. The clustering method is used for multiple log entries that are fed into the LSTM network to detect and predict system failure. Generalized LSTM-based detection and diagnosis is used when raw data is processed and then analyzed for detection.

A stacked LSTM is a deep architecture used in log data, where the output of each LSTM layer is the input to the next LSTM layer, and the repeated layer over time can be deployed as a feedforward network. Compared to a conventional RNN, an LSTM requires minimal or no data preprocessing; in addition, it does not require features trained by experts because it works on raw data, and it does not require prior annotation for anomaly functioning.

Convolutional Neural Network (CNN) is a variant of a neural network which aims to learn appropriate feature representations of the input data. A CNN network has two main differences from MLP-derived networks (in our case, RNNs), namely weight distribution and association. Each level of a CNN network can consist of many convolutional kernels that are used to create different feature maps. Each neighboring area of a neuron is connected to a neuron in the feature map of the next layer. In addition, to create a feature map, all spatial layers of the input share a kernel. After several convolution and clustering layers, one or more fully connected layers are used for classification [1].

By using common weights in the CNN, the model can learn the same pattern that occurs in different input positions without having to train separate detectors for each position.

Convolution layers reduce the computational load because they reduce the number of connections between convolution layers. In addition, clustering layers increase the translation invariance properties and improve the reception field of subsequent convolutional layers. Typically, one or more fully connected layers are added at the end of the convolutional network flow, and a loss function is used to measure errors for training purposes.

We have analyzed a number of characteristics of recurrent and convolutional neural networks, as well as their modifications when paired with LSTM. If we consider them from the point of view of the effectiveness of detecting network anomalies, the most effective models are those that additionally use long short-term memory layers.

Adding LSTM layers significantly increases the accuracy of network anomaly detection. For example, in the case of recurrent neural networks, the accuracy rate can increase by almost 30%.

If one chooses between CNN-LSTM and RNN-LSTM, convolutional neural networks will show better results due to the peculiarity of their structure. In addition, recurrent neural networks will take many times longer to detect an anomaly due to the recursion of each element within the network, and when additional layers of long short-term memory are added, it will lead to a serious payload on the resources of the

computer system.

On the other hand, using only layers of long short-term memory will not help solve the problem of time resource utilization. The fact is that LSTM itself has a very low level of detection accuracy, which in some cases is slightly higher than 80%. In general, this solution could be used on small computing systems that do not have a lot of resources, but if we are talking about the highest accuracy, it is better to combine LSTM with one of the existing neural networks.

Summary and conclusions.

Thus, research has proven that the most optimal neural network for detecting network anomalies is a convolutional neural network with long short-term memory. It is much faster than a recurrent neural network and more reliable than conventional long short-term memory layers. Moreover, it is more convenient to use and can detect a wide range of anomalies in the network. This point can be used to build automatic systems for detecting intrusions in the network traffic of an organization's information system.

References:

1. A Convolutional Neural Network for Network Intrusion Detection System / L. Mohammadpour et al. Barcelona, 24–26 October 2018. 2018. P. 50–55.

2. LSTM learning with bayesian and gaussian processing for anomaly detection in industrial IoT/ D. Wu et al. IEEE transactions on industrial informatics. 2020. Vol. 16, no. 8. P. 5244–5253. URL: https://doi.org/10.1109/tii.2019.2952917 (date of access: 15.04.2023).

3. Gradient-based learning applied to document recognition / Y. Lecun et al. Proceedings of the IEEE. 1998. Vol. 86, no. 11. P. 2278–2324. URL: https://doi.org/10.1109/5.726791 (date of access: 15.03.2023).

4. I. Cvitić, D. Perakovic, B. B. Gupta and K. -K. R. Choo, "Boosting-Based DDoS Detection in Internet of Things Systems," in IEEE Internet of Things Journal, vol. 9, no. 3, pp. 2109-2123, 1 Feb.1, 2022, doi: 10.1109/JIOT.2021.3090909.

5. Staudemeyer R. C. Applying long short-term memory recurrent neural networks to intrusion detection. South African Computer Journal. 2015. Vol. 56. URL: https://doi.org/10.18489/sacj.v56i1.248 (date of access: 15.04.2023).

6. Young T., Nammous M. K., Saeed K. Advanced Computing and Systems for Security. Berlin, Germany: Springer; 2019. Natural language processing: speaker, language, and gender identification with LSTM; pp. 143–156.