Using an ensemble of neural networks trained on an unbalanced sample to classify the state of Internet of Things devices

M E Sukhoparov¹, I A Sikarev¹, T M Tatarnikova¹, I S Lebedev² and V V Karetnikov³

¹ Russian State Hydrometeorological University
² St. Petersburg Institute for Informatics and Automation of the Russian Academy of Sciences
³ Admiral Makarov State University of Maritime and Inland Shipping, St. Petersburg, Russia

E-mail: karetnikovvv@gumrf.ru

Abstract. The paper considers an approach to identifying anomalous situations in network segments of the Internet of Things (IoT) based on an ensemble of classifiers and proposes an ensemble of classifying algorithms for detecting an anomalous situation. Classification objects are represented by multiple parameter tuples. Classifying algorithms are tuned for different types of events and anomalies using training samples of different composition. The use of an ensemble of algorithms allows increasing the accuracy of the results due to collective voting. The experiment performed using three neural networks identical in architecture is described. The variety of classifiers in the ensemble when analyzing the state of IoT devices was formed on the basis of a training sample. The results of the assessment were obtained both for each classifier separately and with the use of the ensemble. Despite the fact that the training sample had an imbalance in relation to classes, the test results using averaging of values by the ensemble of classifiers showed an accuracy value of more than 99%.

1. Introduction

The rapid development of network solutions based on the concept of the Internet of Things necessitates solving a number of issues related to ensuring information security of devices and nodes. In this regard, the tasks of detecting network attacks, recognizing abnormal traffic, analyzing the state of information security of the IoT devices themselves are being solved. The elements of the Internet of Things are mainly located outside the controlled area, providing the solution of a large number of domestic and industrial tasks in an automatic mode without human intervention. The growth in the number of nodes and devices raises the issues of supporting their life cycle, ensuring information and functional security, associated with the complexity of tracking their state.

There are problems of ensuring secure network access, receiving and transmitting messages, machine-to-machine communication, routing, intelligent data processing in the face of constant changes in the network structure, software versions, and integration of devices from different manufacturers into network segments.
Most of the solutions to such problems are based on the accumulated statistical data that appear during operation. We consider network segments, corporate networks and search for traffic anomalies caused by any destructive influences. Machine learning methods are used, for which approaches based on various classifying algorithms are used. For intrusion detection, decision trees, deep neural networks are used, and genetic algorithms for detecting anomalies can be built [1–3]. The incidents are prevented on the basis of models that assess the state of information security using tuples of features [3, 4]. The information security events are detected using approaches aimed at analyzing the cause-and-effect relationships of changes in states occurring in the system. The processes occurring in the system are considered, transition graphs are constructed, which allows predicting the development of the situation and prevent the events preceding negative consequences [5, 6].

The huge variety of elements of the Internet of Things, constant changes in the network architecture, the arbitrary appearance and disconnection of devices by end users and the simultaneous use of different generations of machine-to-machine communication protocols necessitate constant improvement of information security systems, where one of the main areas is the use and adaptation of machine learning methods.

The study of the qualitative indicators of various classifying algorithms depends on the parameters of the training sample and the structure of the information it contains. Typically, the issues of the formation of training samples are associated with their balance, separation and correct interpretation of background and significant patterns, the absence of training objects of a certain type, elements of the feature system, inaccurate ranges of values of variables, imbalance, the appearance of external patterns associated with the conditions for the formation of a training set. However, if it contains information of dissimilar events, a number of problematic aspects of interpreting the data structure arise [7]. In the problems of classifying states, examples are associated with a set of labels \( \{ l_1, l_2, ..., l_n \} \subseteq L \). There are several events that are grouped as related, for example, to safe or dangerous states.

Classification algorithms identify events, which records are labeled with label \( l_i \) from a set of disjoint labels \( L, |L| > 1 \). If \( |L| = 2 \), then binary classification is carried out; and if \( |L| > 2 \), a multi-class classification is implemented [8].

Thus, when constructing classifiers, methods are used that transform a classification problem with several labels either into single, or several classification problems with one label, and methods that expand the possibilities for processing data with multiple labels.

The work considers a finite number of labels \( l_i \) of set \( L \). Each label \( l_i \) belongs to subsets \( L_1 \) or \( L_2 \) of set \( L \), defining dangerous \( C_1 \) and safe \( C_2 \) classes of the set of states \( C \).

Classification objects are a tuple of values \( X = (x_1, ..., x_n) \) of length \( n \geq 2 \).

There is some initial labeled sample \( X \times L \). Known pairs \( (X_i, l_j) \) are defined on the training sample.

It is necessary to construct a classification algorithm for the input vector of values \( X \), which, according to the input tuple \( X \), \( X_i \), maps state \( Z \) into set \( C \) of event classes: \( a : X \times L \rightarrow C \).

2. Materials and Methods

Due to the presence of different design features, interfaces and communication protocols for IoT devices, individual events that will be monitored on some nodes cannot be identified in other segments. This may be due to different characteristics of devices, different software versions, manufacturing peculiarities, use of incompatible interfaces. One of the approaches to solving this problematic situation may consist in independent training of classifying algorithms, according to the data of certain types of devices and network segments of the Internet of Things, with their subsequent combining into ensembles.

The choice of classification algorithms is determined by the need to achieve the required values of quality indicators. In the case of the appearance of complex structures, a large dimension of the processed data in the problems of classification, forecasting, to achieve a given quality, various compositions of individual algorithms are used [9–11].
The use of compositions of classifying algorithms presupposes the fulfillment of a number of conditions for achieving a diversity (difference) of classifiers in an ensemble. The solutions to the indicated problematic issue are primarily related to the fact that algorithms of "different nature" are used, compositions of "weak" classifiers are formed [11–13]. However, there is no unequivocal opinion on the issue of diversity (difference) of classifiers [14–17].

The proposed approach to achieving a variety of classifiers is associated with the formation of a training sample. Under the conditions of the functioning of the system, the appearance of destructive influences, the emergence of specific conditions of the external and internal environment, it is not always possible to know in advance all the events that may occur. Therefore, the training sample may not fully reflect the information about the analyzed events.

To neutralize this effect, it is proposed to form a training sample in such a way that it contains a limited subset of events for each classifier. Each classifier specializes only in a certain part of the classification objects, marked with appropriate labels. The trained classifier may not "know" about events in one class that are marked with different labels. The effect of under-training is artificially created, which is solved by using the ensemble.

At the first stage, classes of objects and their attribute space are determined. At the second stage, the parameters of objects of each of the classes are set. The set of parameters depends on the algorithm, the limitations of the infrastructure segment environment, where it will identify the state as the main one. The third stage is the construction of a training sample for each classifier. The fourth stage is the formation of the aggregation function and the decision rule.

The difference in the approach is that not just a weak classifier for the ensemble is formed, but a strong one on a limited set of classes. The variety of destructive influences, their uniqueness leads to the fact that in order to achieve diversity, an ensemble is formed not from just weak classifiers, but from strong ones in determining the initially specified events.

3. Results and Discussion

In practice, situations arise when a dataset record about an event can belong to several classes at the same time, for example, a DOS-attack and parasitic traffic, which makes it necessary to combine quite different events into one class. Achieving the difference between the algorithms included in the model, their training occurs independently of each other on randomly selected subsets of the training sample. The initial training set of records of \( X \) tuples is split into subsamples, where it is initially assumed that each classifier "specializes" in a certain class of events.

For the experiment, the Matlab package was used, where the basic algorithms were two-layer neural networks [12].

\[
a_k(X) = \text{sign} \left( \sum_{t=1}^{T} w_{jt} \delta \left( \sum_{j=1}^{n} w_{ji} X^j \right) \right),
\]

where \( T \) is the number of neuron of the hidden layer, \( w_{jt} \) is the bond weight between the \( j \)-th feature and the \( t \)-th neuron of the hidden layer, \( w_{ij} \) is the bond weight between the \( i \)-th neuron of the hidden layer and output neuron, \( \delta \) is the activation function.

The ensemble of basic algorithms is described by the expression [12] (2):

\[
a(X) = \Phi \left( F(a_1(X), \ldots, a_k(X)) \right),
\]

where \( \Phi \) is the decisive rule.

Let us create a training sample of each classifier in such a way that the algorithm "specializes" on certain classes.

To do this, we artificially divide it in such a way that the set of training examples for each classifier contains records about some events, but information on others is absent or minimized.

We consider a finite number of labels \( l_i \) of set \( L \), the information of which is contained in tuples \( X_i \). Known pairs \((X_i, l_i)\) are determined on the training sample. Each label \( l_i \) is included in subsets \( L_1 \) or \( L_2 \).
of set $L$ that define dangerous $C_1$ and safe $C_2$ classes of states. The training sample is divided into subspaces equal in terms of the number of training examples for each classifying algorithm. A single subspace may not contain tuples labeled with specific labels from subset $L_1$ or $L_2$. There is an independent training of classifiers, where each algorithm recognizes the labels of its subset.

The test set is fed to an ensemble of classifiers consisting of networks of the same structure (Figure 1).

Below are the results of the classification of three identical two-layer neural networks after training on different training samples. A vector from the test set was fed to the input, and a binary classification was carried out at the output. The size of the training and test sample was 70:30. The training set was divided into three parts. The training sample for the first and second networks was implemented by random selection. For the third network, 90% of the entries belonged to set $C_1$.

The results obtained after processing the test set for each network and the entire ensemble are shown in Figure 2.

The test results using averaging of values by an ensemble of classifiers show an accuracy value of more than 99%, despite the fact that the training sample had an imbalance in relation to classes.
Figure 2. Results of testing networks separately and the overall result of the classifier

4. Conclusion
The increase in the number of new types of attacks and destructive effects on IoT devices necessitates the analysis of a large number of parameters of the system's functioning. Most of the events that occur in the system under the influence of the external environment are unique and can manifest themselves in different ways in various devices, nodes and elements.

The article shows that the accuracy of class recognition can be improved by combining a number of classifying algorithms into an ensemble. An approach to the formation of an ensemble of classifiers is presented, where each classifying algorithm included in it, by forming a training sample that "specializes" on certain events. As a result, in order to ensure the difference between the aggregated classifying algorithms, an ensemble is obtained not of weak classifiers, but of those strong in their field.

The approach was evaluated using an ensemble of two-layer neural networks trained on samples containing selected event records, which increased the identification accuracy up to 99.9%.

References
[1] Gao D, Reiter M and Song D 2009 Beyond output voting: Detecting compromised replicas using HMM-based behavioral distance *IEEE Transactions on Dependable and Secure Computing* 6(2) 96–110 DOI: 10.1109/TDSC.2008.39
[2] Devesh M, Kant A K, Suchit Y R, Tanuja P and Kumar S N 2020 Fruition of CPS and IoT in Context of Industry 4.0 *Intelligent Communication, Control and Devices. Advances in Intelligent Systems and Computing* 989 367–375
[3] Sokolov S, Nyrkov A, Knysh T and Shvets A 2021 Countering Cyberattacks During Information Operations *Proceedings of the XIII International Scientific Conference on Architecture and Construction* 2020 (Singapore: Springer) pp 84–100 DOI: 10.1007/978-981-33-6208-6_9
[4] Bevir M K, O'Sullivan V T and Wyatt D G 1981 Computation of electromagnetic flowmeter characteristics from magnetic field data *Journal of Physics D Applied Physics* **14**(3) 373–388 DOI: 10.1088/0022-3727/14/3/007

[5] Semenov V V, Lebedev I S, Sukhoparov M E and Salakhutdinova K I 2019 Application of an Autonomous Object Behavior Model to Classify the Cybersecurity State In *Internet of Things, Smart Spaces, and Next Generation Networks and Systems* (pp 104–112) DOI: 10.1007/978-3-030-30859-9_9

[6] Karetnikov V and Sazonov A 2021 Fuzzy models of the dangerous situations prediction *Transportation research procedia* **54** 12–22 DOI: 10.1016/j.trpro.2021.02.043

[7] Xiao T, Xia T, Yang Y, Huang C and Wang X 2015 Learning from massive noisy labeled data for image classification *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* pp 2691–2699

[8] Tsoumakas G and Katakis I 2007 Multi-Label Classification: An Overview *International Journal of Data Warehousing and Mining (IJDWM)* **3** 1–13

[9] Tatarnikova T, Sikarev I, Karetnikov V and Butsanets A 2021 Statistical research and modeling network traffic *ES Web of Conferences – EDP Sciences* **244** 07002 DOI: 10.1051/e3sconf/202124407002

[10] Sukhoparov M E, Lebedev I S and Semenov V V 2020 Information Security State Analysis of Elements of Industry 4.0 Devices in Information Systems In *Internet of Things, Smart Spaces, and Next Generation Networks and Systems* (Cham: Springer) pp 119–125 DOI: 10.1007/978-3-030-65726-0_11

[11] Sukhoparov M E, Lebedev I S and Garanin A V 2020 Application of classifier sequences in the task of state analysis of Internet of Things devices *Computing, Telecommunications and Control* **13**(3) 44–54 DOI: 10.18721/JCSTCS.13304

[12] Nyrkov A P, Belousov A S and Sokolov S S 2015 Algorithmic support of optimization of multicast data transmission in networks with dynamic routing *Modern Applied Science* **9**(5) 162 DOI: 10.5539/mas.v9n5p162

[13] Ageev S et al. 2020 Intellectual method of operational evaluation of the network element state to ensure the quality of services in corporate multiservice communications networks *ES Web of Conferences* **203** 05017 DOI: 10.1051/e3sconf/202020305017

[14] Zhou Z-H 2012 *Ensemble Methods: Foundations and Algorithms* (New York: Chapman & Hall/CRC) 222 p ISBN: 978-1-4398-3003-1

[15] Yu Y, Zhou Z and Ting K M 2007 Cocktail Ensemble for Regression *Seventh IEEE International Conference on Data Mining (ICDM 2007)* (Omaha: NE) pp 721–726 DOI: 10.1109/ICDM.2007.60

[16] Tatarnikova T M 2019 Restricting data leakage through non-obvious features of Android 5 smartphone *Management information systems* **5** 24–29 DOI: 10.31799/1684-8853-2019-5-30-37/

[17] Sovetov B Y, Tatarnikova T M and Cehanovsky V V 2019 Detection system for threats of the presence of hazardous substance in the environment *Proceedings of 2019 22nd International Conference on Soft Computing and Measurements* pp 121–124 DOI: 10.1109/SCM.2019.8903771