ТЕХНІЧНІ НАУКИ

UDC 004.056.53

DETECTION OF COMPUTER ATTACKS USING OUTLINER METHOD

Balakin S.V., Zhukov I.A.
National Aviation University

Results of detection of computer attacks by the method of variations (outliers) have been analyzed. Identified errors in the calculation of the deviations and made adjustment that increases performance of the system. Discussed detection of computer attacks by deviations, as well as its capabilities in comparison with other traditional approaches. Developed model of intrusion detection based on information about the behavior of outliers in the network. The techniques of intrusion detection are proposed.
Keywords: attack, computer system, deviation, intrusion, informational system, object.

Formulation of the problem. The intensive use of the Internet network security has become a key foundation for all web applications. Intrusion Detection and identification of attacks by analyzing the information in the records of network processes – all of this can be seen as one of the important ways for effective solutions in the field of network security problems [1].

Invasion could jeopardize the security of both data and the system itself. With the development of information networks and the increase in the data rate, there is a risk of malicious use of the Internet. A need for more reliable and effective control system, crucial network security problem without human interaction.

Using the tools of detection of anomalies and attacks depends on areas of implementation. The narrower scope, the easier it is to apply to it or that research tools.

The possibility of working with neural networks (NN). A distinctive feature of NN is that they cannot be programmed and trained. This is one of the main NN advantages over traditional algorithms. The training consists of connections between neurons that determine the ratio of input and output signals of the neuron. NN based on "learning" and does not allow analytically to calculate the level of errors. The disadvantage is that the topology of the network nodes and weight are determined only after a sufficiently large number of trials and errors. The main disadvantages of IDS (Intrusion Detection Systems) NN:

- lack of precision detection U2R and R2L attacks;
- low reliability of the detection.

Analysis of recent research and publications. To solve these problems, we propose a new approach based IDS deviation method. Detection of deviations is made in order to improve the accuracy and stability of detection. The proposed approach consists of two phases: training to normal data sets and test data sets with samples of intrusion. Different data set is used to prepare the initial stage of IDS in a distributed storage environment. Normal data sets improve the performance of intrusion detection system. With the invasion of the data set, which is used to calculate the value of the error with the training data sets. If the number of errors is incremented by a certain threshold, then the test data set should be regarded as an anomaly.

Various methods can be used for intrusion detection, but each is specific for a particular method [2]. The main purpose of the intrusion detection system – effectively detect attacks. It is important to identify the attack at an early stage, in order to reduce its impact. An approach deviations sculpt the values at which the data set anomaly measured deviation factors (DF).

learning model consists of data sets distributed storage environment to enhance the intrusion detection system performance. Experimental results have shown that this approach detects anomalies efficient than known methods.

The first work is conceptual, as they attempted to use is not defined filters or methods, and theory of probability and mathematical approaches to solving problems.

Known methods of machine learning (temptation-governmental neural network) for intrusion detection.

Development of methods for detection of attacks based on the information about the behavior of outliers in a network devoted to the work.

Bold of unsolved aspects of the problem. Classification of problems contributes to the selection of the best approaches to solving problems. Intrusion detection systems are divided into three categories:

- host-based IDS;
- networking IDS;
- IDS vulnerability assessment.

There are basic models used to analyze the events and intrusion detection:

- misuse detection model – an intrusion detection system by searching for known vulnerabilities or intrusion signatures.

Model of anomaly detection – intrusion detection system by using the search function abnormalities network traffic.

Some IDS can detect signs of intrusion without specifying types of attacks, but they are sensitive to false alarms. We used the proposed IDS approach based on anomaly detection model [3].

Fig. 1. The proposed architecture of the system

The main goal is to develop an IDS based on anomaly detection model, which would have a low threshold of false alarms was adaptive and worked in real time. Fig. 1 shows architecture of the system in which packets are received from the Internet, and SNORT is used to collect data. Original features extracted from the data packets are transmitted to the IDS, then the proposed IDS calculate the distance between the extracted features and the training model. learning model consists of large amounts of distributed data storage environment to improve the intrusion detection system performance. Thus, if the abnormal value exceeds a predetermined threshold, it generates a false alarms [4].

The aim of the study is to developed-thief IDS based on anomaly detection model, which would have a low threshold for false alarms and to la adaptive and worked in real time.

The main material. These samples have normal dense neighborhood, whereas they are dispersed at deviations away from each other. Dropping deviations are the objects of the outer layers [5].

The basic idea of the approach is to assign the data with a factor of the deviation, and to find the data, whose behavior is different compared with more normal flow of information.

The steps of the algorithm used to calculate deviations for these examples are as follows:

- calculate the maximum deviation (D) for each sample data (D);

- calculated distance reachable for each of the examples of D relative to each other (n): the maximum distance (D, n) = max {O (n), d (D, n)} – is the distance between the data of Example D and data Example n;

- calculated by the local density reachability for each example D, based on the return on availability medium distances reachable via MinPts (the minimum number of objects), and sample D with its immediate neighbors;

- are determined to reject all examples of D with respect to the averaged data with coefficients density reachable MinPts nearest neighbors.

The advantages of the proposed approach the deviations are shown in Figure 2. Clusters are defined as dense sets of related objects. A simple two-dimensional array of data taken from a much larger number of samples in cluster R1 and then in R2. Thus, the density higher than the density K1 K2. For each example, be considered an object within the cluster K1, where the distance between him and his nearest neighbor is greater than the distance between P2 and its nearest neighbor cluster of K2. Therefore, P2 is not regarded as abnormal values.

Therefore, detection of outliers is in the field of statistics. However, in the P1 deviation can be detected using only the nearest neighbor distances. Alternatively, the deviation can capture both values (P1 and P2) in connection with what they consider grouping all points (Figure 2).

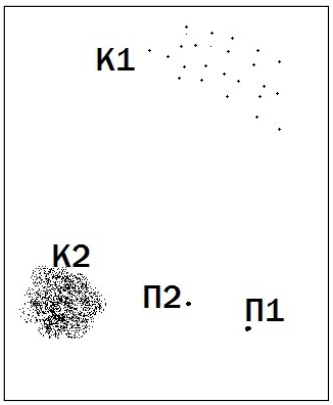

Fig. 2. Accuracy of detection of abnormalities

In the study of the number of experiments it was based on the extracted data to measure the performance of IDS systems. Experiments were carried out on the basis of configuration: Windows 8, Core2Duo (R), CPU T7300 2,90 GHz.

The extracted data set included over two thousand entries compounds. Test data included five thousand records connections. The data set includes the features derived from each compound, and the group of labels identifying a recording status connections for deviations.

The distance and the deviation of the data values are calculated by a method of detection of deviations. These calculations show that the deviation values increase when the distance between lessons and test data increases. The results are shown in the table.

Table 1

The dependence of the distance and the sample variance

№	Distance	Variance
1	3.5	5
2	1.2	2
3	4.6	8
4	2.7	3
5	3.6	7

Conclusions and suggestions. In the approach developed for the detection of deviations to detect intrusion into a computer network. learning model consists of data sets distributed environment, which increases the performance of the system when it detects intruders. The proposed approach is applied using data sets KDD. Machine learning approaches of intrusion detection in computer networks are promising, since it is possible

to fully automate the detection of intrusions in the networks. In the study evaluated the performance of the proposed method, which can detect anomalies in most computer network. The results of using the approach of intrusion detection in computer networks.

Further research is planned to use the results to samples of learning models or test data.

References:

1. Gamayunov, D. Y., Smelianskiy, R. L. Model povedeniya setevih elementov v raspredilitel'nikh vicheslitel'nikh sistemakh [Model of behavior of network objects in distributed computing systems]. Programming, 2007, pp. 20–31.
2. Cramer, H. Mathematical Methods of Statistics. Prinseton University Press, 1962, 590 p.
3. Korchenko, O. G. Postroenie system zashiti informacii na nechetkikh mnozestvakh [Building security systems on fuzzy sets]. Programming, 2006, 320 p.
4. Masayoshi, M. The Design and Implementation of Session Based NIDS. IEICO, 200, pp. 551-562.
5. Sheyner, O. Scenario Graphs and Attack Graphs. Carnegie Mellon University, 2004, 141 p.

Балакін С.В., Жуков І.А.
Національний авіаційний університет

ВИЯВЛЕННЯ КОМП'ЮТЕРНИХ АТАК ЗА ДОПОМОГОЮ МЕТОДУ ВІДХИЛЕНЬ

Анотація
Представлений аналіз результатів виявлення комп'ютерних атак за допомогою методу відхилень (значень відхилення). Обговорюється виявлення комп'ютерних атак методом відхилень, а також його можливості в порівнянні з традиційними підходами. За результатами аналізу визначено похибки в розрахунках відхилень процесів та внесені коректування, які значно підвищують показники продуктивності отримані раніше. Розроблено модель виявлення атак на основі інформації про поведінку значень відхилень в мережі і їх виявлення.
Ключові слова: атака, комп'ютерна система, відхилення, вторгнення, інформаційна система, стан об'єкта.

Балакин С.В., Жуков И.А.
Национальный авиационный университет

ОБНАРУЖЕНИЕ КОМПЬЮТЕРНЫХ АТАК С ПОМОЩЬЮ МЕТОДА ОТКЛОНЕНИЙ

Аннотация
Проведен анализ результатов обнаружения компьютерных атак с помощью метода отклонений (отклоняющихся значений). Обсуждается обнаружение компьютерных атак методом отклонений, а также его возможности в сравнении с традиционными подходами. По результатам анализа определены погрешности в расчетах отклонений процессов и внесены корректировки, которые значительно повышают показатели производительности. Разработана модель обнаружения атак на основе информации о поведении отклоняющихся значений в сети и их выявление.
Ключевые слова: атака, компьютерная система, отклонение, вторжение, информационная система, состояние объекта.