ISSN 0146-4116, Automatic Control and Computer Sciences, 2019, Vol. 53, No. 8, pp. 1–8. © Allerton Press, Inc., 2019.
Russian Text © The Author(s), 2019, published in Problemy Informatsionnoi Bezopasnosti, Komp'yuternye Sistemy.

Model for Identifying Cyber Threats to Internet Information Resources

I. V. Luskatov[a, *] **and S. V. Pil'kevich**[a, **]

[a] *Mozhaisky Military Space Academy, St. Petersburg, 197198 Russia*

**e-mail: ambers@list.ru*

***e-mail: mrfa@rambler.ru*

Received April 18, 2018; revised August 6, 2018; accepted September 22, 2018

Abstract—In this paper, we discuss the construction of a model for protecting information resources based on new approaches to active search, unified investigation, and response to cyber threats. The process of identifying cyber threats is based on search prediction methods interconnected with the cyclic nature of the behavioral activity of users of network resources.

Keywords: cyber threats, anomalies, prediction, user behavior, intelligent system

DOI: 10.3103/S0146411619080170

INTRODUCTION

The exponential growth of various cyber threats in the modern world makes the problem of information security extremely important. There are numerous examples of application software that allow solving problems of information protection using universal methods. Nevertheless, the most organic solution would be to provide the information resources themselves with appropriate functionality, which will ensure their protection while taking into account the specifics of concrete operating conditions. The solution of this problem involves consideration of cyber threats that the protection system should withstand and specific features of the information resource being protected.

In this paper, the Internet information resources will be understood as online information platforms (OIPs), i.e., websites whose services have a distinct social orientation and are characterized by close interaction with users providing them with wide opportunities for publishing and sharing various content and storing personalized data. The combination of these properties makes OIPs an appealing target for attacks by information security violators, which, as a rule, disguise themselves as legitimate OIP users and gain profit from spam, theft of personal data and confidential information [1, 2].

CYBER THREAT DETECTION SUBSYSTEM FOR ONLINE INFORMATION PLATFORMS

It appears reasonable to provide OIPs with functions of cyber threat detection and elimination implemented within a similarly named OIP subsystem (Fig. 1).

The intended purpose of the subsystem is to identify anomalous network activity of users of an information resource, evaluate it, and make decisions regarding the subsequent blocking of users with suspicious behavior or temporarily limiting their access to the network resource services.

An approach based on the detection of anomalies is used to identify cyber threats in various systems for which the characteristics of functioning over time were obtained [3]. In the context of this study, an anomaly is understood as any deviation from the model of normal user behavior, e.g., spamming, automatic like boosting, etc. The main controlled parameters of OIPs that have the potential to detect anomalies by analyzing the user network activity are the average number of users, intensity of the message exchange, and time of appearance of reposts, likes, etc.

Analysis of current studies in the field of implementation of cyber threat detection methods [4–6] (Table 1) showed that, based on such parameters as authenticity, efficiency, and reliability, a method based on prognostic evaluation of user behavior is a priority.

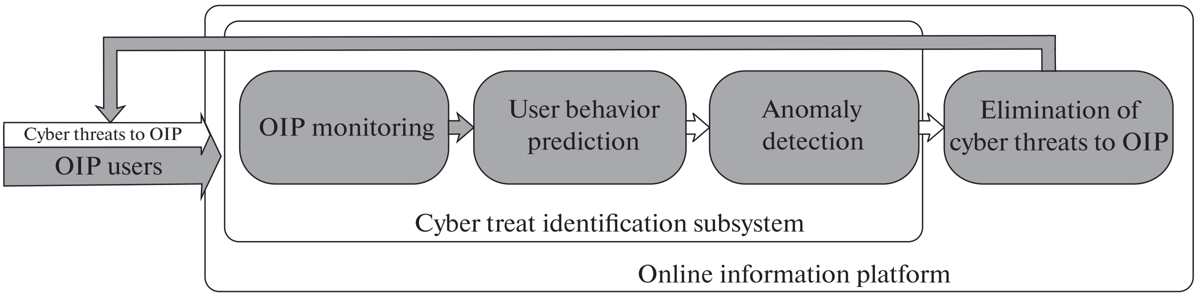
Cyber threats to OIP
OIP users
OIP monitoring
User behavior prediction
Anomaly detection
Elimination of cyber threats to OIP
Cyber treat identification subsystem
Online information platform

Fig. 1. Functional diagram of the interaction between the main elements of the online information platform protection system.

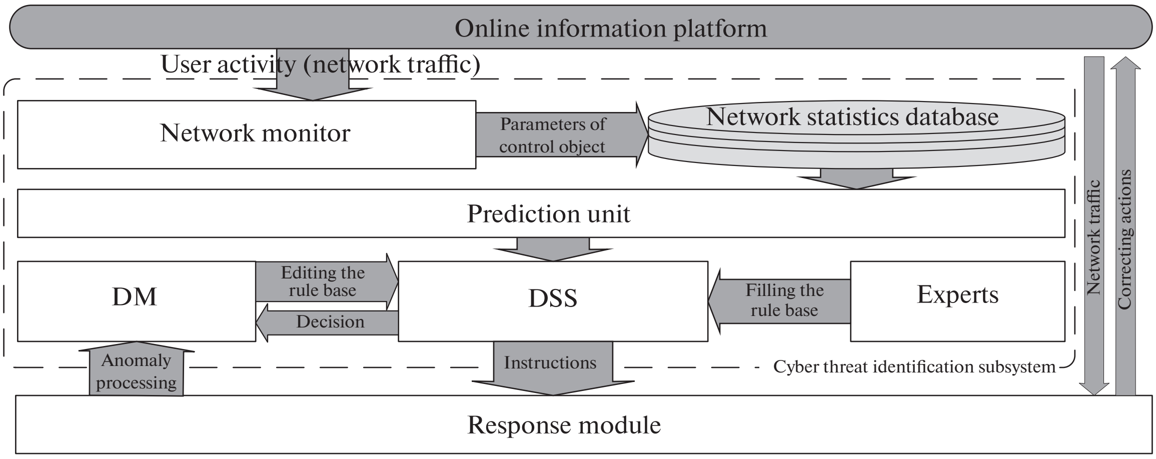

Fig. 2. Generalized scheme of the intellectual system for identifying and responding to OIP anomalies.

Intelligent system for identifying and responding to anomalies of online information PLATFORMS

The implementation of the prognostic method in the cyber threat detection subsystem for information resources makes it necessary to create an intelligent system capable of accumulating and summarizing monitoring data with subsequent automatic construction of rules for responding to identified cyber threats. A general diagram of such a system is shown in Fig. 2.

The source data is the information on the network activity of OIP users extracted from network traffic, as well as information on current threats to OIPs obtained from external sources.

A network monitor collects statistical information on the behavior of OIP users. The accumulated information is stored in the network statistics database, which serves as the basis for a forecast of user network activity. In the decision support system (DSS), the real activity is compared with the predicted one, and the magnitude of the discrepancy is estimated; if this parameter exceeds the specified limit, a decision is made that an anomaly is present and the search for its sources (involved users) is commenced. The DSS rule base is filled by an expert and edited by a decision maker (DM). Based on the results of the decision,

Table 1. Qualitative comparison of the characteristics of cyber threat detection methods

Method name	Estimated parameter		
	authenticity	efficiency	reliability
Expert assessments	Medium	Low	Medium
Criterial	Medium	Low	Medium
Statistical expert	High	Low	High
Prognostic evaluation	High	High	High

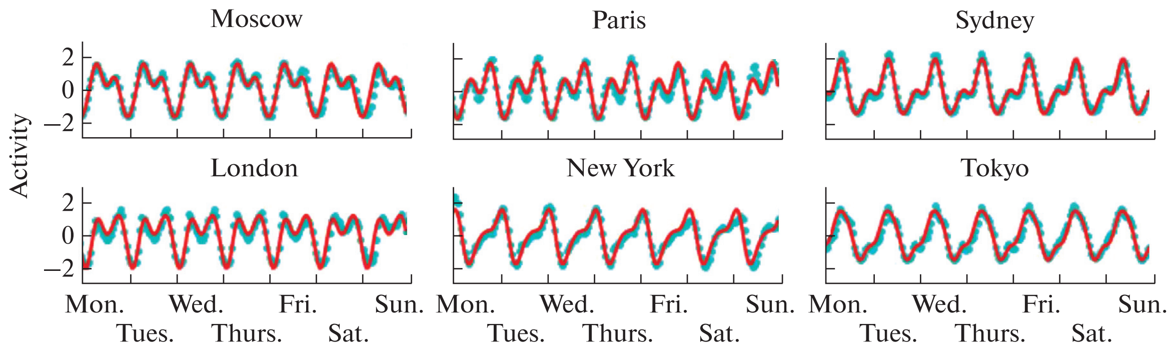

Fig. 3. Twitter user activity and its characteristic periodicity.

specific target designations (instructions) are formed, which are transmitted to the response module. Further, corrective actions are taken, such as warning users, blocking accounts, deleting messages, etc.

Let us consider the sequence of operations that are applied to the source data and ultimately allow identifying and evaluating anomalies.

Network Monitoring

At the first stage, it is necessary to collect statistical information on the behavior of OIP users. Since the goal of forecasting is to obtain the values of the activity parameters for a certain period of time in the future based on the available data on the OIP payload, it is necessary to extract information on keywords from the content of each message and store the date and time the message appears on the site. A response to the identified anomaly requires information on its source (the account of the OIP user) and place of publication. Thus, to predict the user activity parameters, the following data are extracted from the metadata of the message: author, keywords (markers), destination page, and publication date.

Prediction of the Behavioral Activity of Online Information Platforms Users

The next stage consists in predicting the average values of the integral parameter of the network activity of OIP users. Prediction is implemented by the method of cyclical analysis and includes the following stages. The basis of this is the idea that many real processes have a certain periodicity and repeatability [7, 8]. It was suggested that the user activity is also characterized by a certain periodicity [9]. For example, a global analysis of Twitter showed that its chart of communication intensity throughout the day is similar to a cardiogram (Fig. 3) [10].

The user activity parameters of other social networks can also form peculiar "behavioral patterns" that are repeated in time. Based on this feature, it turned out to be possible to form the predicted values of the OIP user activity and subsequently use these values for detection of anomalies.

The data-selection stage is characterized by the fact that the collected statistics should be represented as a series of data. Let there be statistics on the OIP user activity collected over a period of time T. To obtain a data series, we divide the time period T by Q equal intervals Δt : $Q = T/\Delta t$. The values of T and Δt are selected in such a way that Q is an integer.

Let us introduce the integral parameter of the network activity of OIP users: a kind of "volume" of the jth message (V_j), which is calculated according to the following expression:

$$V_j = k_l N_l + k_r N_r + k_u N_u + k_c N_c,$$

where N_l is the number of likes; N_r is the number of reposts; N_c is the number of comments; N_u is the number of links to the media; and k_l, k_r, k_u, k_c are the normalization coefficients.

Further, for each interval Δt, we sum the volumes of R messages that fall into this time interval: $X_q(\Delta t) = \sum_{j=1}^{R} V_j$, where R is the number of messages that fall into the interval Δt, q is the number of the interval $q \in [1,\ldots,Q]$, and X_q is the data series that describes the time variations in the volume of messages with a sampling rate of Δt.

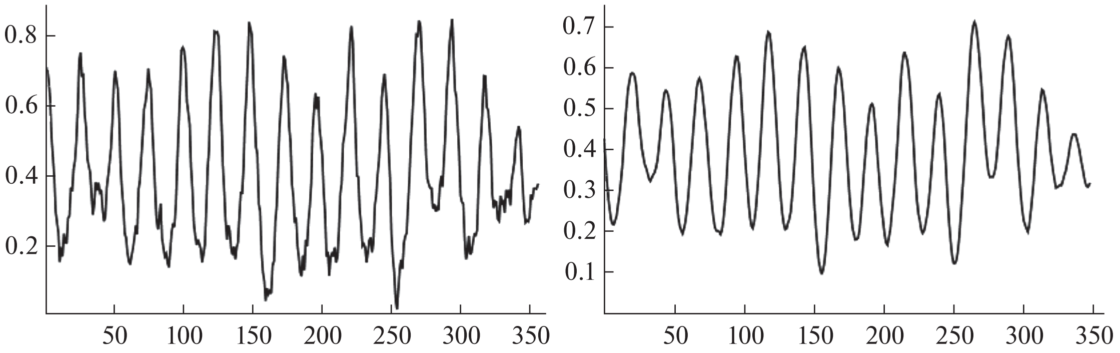

Fig. 4. Initial data on user activity with random fluctuations and the result of their smoothing.

At the stage of smoothing the results of monitoring the network activity of OIP users, random oscillations in the analyzed sample are smoothed using the method of centered sliding average. The number of points for smoothing the data is taken equal to L (for each iteration being performed, an odd number of points should be chosen, $L \geq 3$). When calculating the sliding average for L points, $L - 1$ points will be discarded from the initial data series: $(L - 1)/2$ at the beginning and the same number at the end of the series. An example of applying the described approach is illustrated in Fig. 4.

The periodic nature of the analyzed data is manifested with a certain cyclicity. This circumstance predetermines the next stage of processing: the search for possible cycles.

To determine the frequency components of the series, we use the spectral analysis method mathematically based on the Fourier transform.

Using the direct discrete Fourier transform we find the complex amplitudes of the data series X_k obtained as a result of smoothing X_q:

$$Y_n = \sum_{k=1}^{N} X_k e^{\left(-\frac{2\pi i}{N} nk\right)},$$

where N is the number of signal values (length of the data fragment under study) measured over the period, as well as the number of decomposition components; $k = 1,\dots,N$ is the number of discrete time points where the X_k values were measured; i is the imaginary unit; and $n = 1,\dots,N$ is the frequency index.

The power spectrum is calculated based on the complex amplitude values Y_n: $\mathrm{R}_n = |Y_n|^2 = \mathrm{Re}^2(Y_n) + \mathrm{Im}^2(Y_n)$,

Figure 5 shows that the local maxima of the analyzed function can be found both empirically and numerically. The values of the local maxima obtained in such a way indicate the possible presence of cycles. The index n, at which a high value of the power spectrum R_n is observed, will be the value of the cycle frequency.

After determining the possible cycles and their frequencies, we calculate the real amplitude A and phase φ. Let there be b possible cycles whose frequencies make up the set S. The amplitudes and phases of the found cycles can then be calculated by the formulas

$$A_h = \frac{|S_h|}{N} = \frac{1}{N}\sqrt{\mathrm{Re}^2(S_h) + \mathrm{Im}^2(S_h)}, \quad \varphi_h = Arg(S_h) = \arctan\left(\frac{\mathrm{Im}(S_h)}{\mathrm{Re}(S_h)}\right),$$

where $h = 1,\dots,b$, $Arg(S_h)$ is the complex number function, i.e. the angle of the complex number (in radians) corresponding to S_h.

In this case, the function describing the cycle looks as follows: $f_h(t) = A_h \cos(S_h t + \varphi_h)$.

Removing trend components in the analyzed data. The quality of validating the cycles for statistical reliability strongly depends on the existence of directivity in the data. To remove a trend in the data, it is nec-

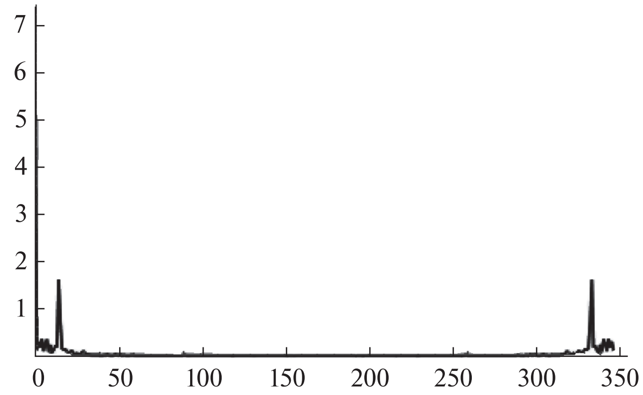

Fig. 5. The result of applying the discrete Fourier transform (power spectrum R_n).

essary for each frequency to calculate a sliding average $\overline{X}_k'$ for a data series X_k with a number of smoothing points $L = \begin{cases} S_h, & \text{if } L = 2p+1 \\ S_h + 1, & \text{if } L = 2p \end{cases}$, $p \in \mathbb{N}$, in this case, we obtain $\overline{X}_k' = \frac{1}{L}\sum_{j=k}^{k+L-1} X_j$.

Further, we subtract the resulting sliding average $\overline{X}_k' : X_k'' = X_k - \overline{X}_k'$ from the original data series X_k. Having thus smoothed the short-term fluctuations of the initial data, we can proceed to validate the possible cycles for statistical significance.

Validation of the cycles in terms of statistical significance. The cycles are assessed using the Fisher criterion and χ^2. The former measures the reliability of the amplitude of the cycle (its shape); the latter measures the reliability of the cycle's phase (its time).

Combining and projecting cycles into the future. Let us suppose that D cycles passed the validation tests of the previous stage. Confirmed cycles allow us to combine the corresponding functions $f(t)$ (obtained at the stage of search for possible cycles) into a general curve that describes the periodicity revealed in the data series: $\overline{V}(t) = \sum_{j=1}^{D} f_j(t)$ (Fig. 6).

The resulting function $\overline{V}(t)$ can be extrapolated and allows us to obtain the predicted value of the activity of OIP users for the upcoming period of time $t' \in (t_{\text{pr}}, t_{\text{fut}}]$.

Anomaly Detection

Based on the data on the predicted and current activity of OIP users, as well as a set of decision rules, it is possible to implement procedures for detection and assessment of potential hazard of anomalies in relation to the OIP security.

Anomalies are searched for based on the comparison of the input user activity parameter (the values of which are calculated in real time based on the results of monitoring the network user activity as described above) with the predicted value. For this, per unit time t, two values are compared: V_{real}, the current value of the network activity parameter, and V_{pred}, the predicted activity value. The parameter deviation is considered anomalous if it exceeds or equals a given threshold value $\alpha : |V_{\text{pred}} - V_{\text{real}}| \geq \alpha$. If an anomaly was detected, the search for its sources is performed. The sources (personal and/or collective author of the message, network information resource) are determined based on the information extracted from the metadata of current OIP traffic.

Further, the magnitude of the anomaly is estimated and a decision is made regarding its processing. There are two main ways to handle anomalies: warning (filtering abnormal activity) and blocking (disabling malicious accounts).

The magnitude of the anomaly is estimated based on the production base of rules. The variety of OIPs, intensity of message exchange on various subjects, and the number of likes and reposts depend not only on the OIP user, but also on the presence or absence of a relevant informational cause and the state of the information space as a whole; this extremely complicates the actions of experts, who, as a rule, operate

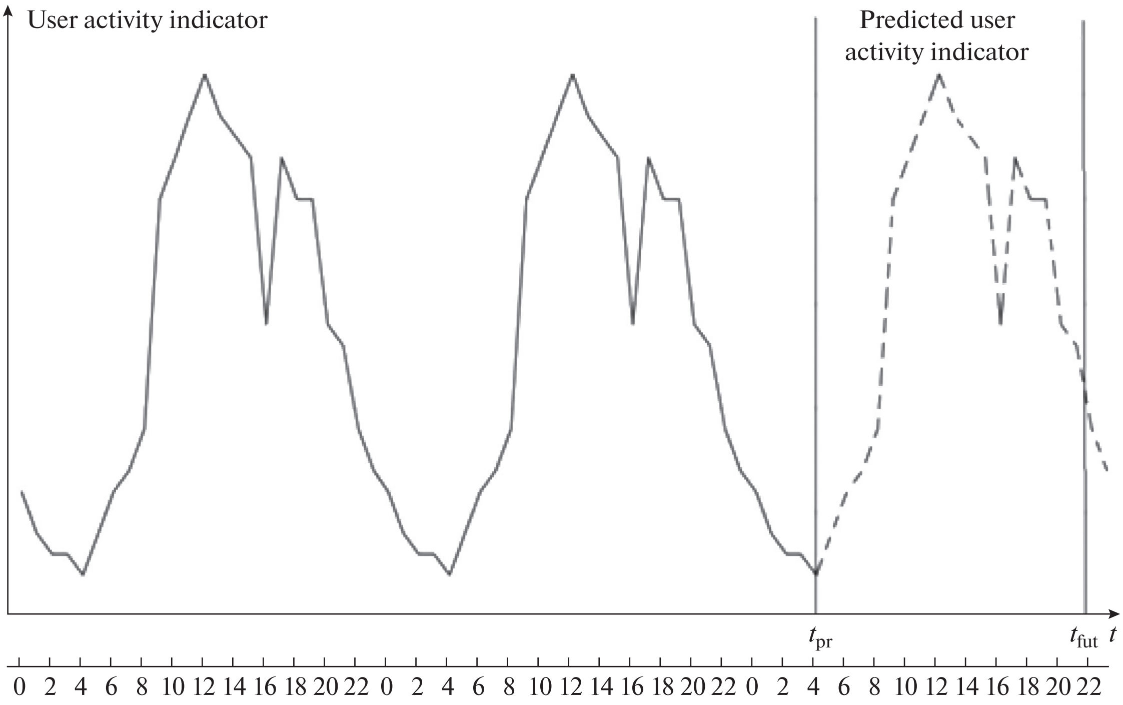

Fig. 6. User activity prediction. $[0, t_{\text{pr}}]$ is the period of monitoring; $(t_{\text{pr}}, t_{\text{fut}}]$ is the predicted period.

with qualitative assessments and face difficulties in specifying precise ranges of quantitative parameters describing the anomaly. These circumstances make it necessary to use the theory of fuzzy information granulation by L. Zade [11].

Let us consider the structure of a rule base (RB) in more detail. To estimate the magnitude of the anomaly, the following linguistic variables with the corresponding term sets are used:

– $V \in$ {low, below average, medium, above average, high} is the deviation value.

– $M \in$ {low, below average, medium, above average, persistent} is the frequency of occurrence of the anomaly.

– $I \in$ {insignificant, below average, medium, above average, high} is the number of sources of anomalies.

– $W \in$ {insignificant, below average, medium, above average, high} is the volume of traffic from one source.

The output parameter $E \in$ {insignificant, below average, medium, above average, high} is the magnitude of the anomaly.

The rules are stored using the methodology of expert systems. Let us give some examples of the rules:

– IF $V \in$ {FROM low TO below average} and $M \in$ {FROM low TO medium} and $I \in$ {FROM insignificant TO medium} and $W \in$ {FROM insignificant TO medium} THEN $E =$ “insignificant”.

– IF $V \in$ {FROM above average TO high} and $M \in$ {FROM average TO persistent} and $I \in$ {FROM above average TO high} and $W \in$ {FROM above average TO high} THEN $E =$ “high”.

Introduction of fuzzy sets to the list of scientific and methodological apparatus requires specifying the boundaries of values for the corresponding membership functions ($\mu_V, \mu_M, \mu_I, \mu_W$ and μ_E). Since these values will be individual for each OIP, let us indicate their common characteristics.

The study uses bell-shaped membership functions, i.e. functions that look as follows: $\mu(x, a, b, c) = \dfrac{1}{1 + \left(\dfrac{x - c}{a}\right)^{2b}}$. The most typical values $(a, b, \text{and}\, c)$ of these membership functions were obtained by direct methods (both single and group).

Table 2. Correlation between the recognition results and true state

		Result of anomaly classification	
		Presence of attack	Regular mode
True state	Presence of attack	True (*A*)	Error kind 1 (missed attack, *C*)
	Regular mode	Error kind 2 (false positive, *B*)	True (*D*)

Response to Identified Anomalies of Online Information Platforms

The final operations implemented by the DSS consist in forming the sets of instructions sent to the response module to neutralize the detected cyber attacks.

The rules that set the parameters for filtering a compromised account (the source of the detected anomaly) are selected depending on the history of its existence. If the user has not previously been compromised and is characterized by a long time of existence on the resource, the account is temporarily blocked until the registration data is confirmed; otherwise, the account is deleted.

The general model for the formation of rules to manage the corrective actions can be represented as follows:

$$G = Func(e, Z, Flt, U),$$

where *Func* is the function of the formation of control rules; *e* is the predicted values of user activity; *Z* is the accounts of the anomaly sources; *Flt* is the parameter the characterizes the response (e.g., filtering time) and is determined depending on the value of *e*; and *U* is the list of exceptions.

Anomaly Detection and Assessment Model

Based on the scheme for the search and assessment of anomalies, we construct a model for detection and assessment of anomalies based on prediction of user activity:

$$e = \begin{cases} Prod(V, M, I, W, E), & \text{if } |V_{\text{pred}} - V_{\text{real}}| \geq \alpha \\ 0, & \text{else} \end{cases},$$

where *Prod* is the procedure for the use of production rules; and α is the threshold value, based on which a decision is made regarding the presence or absence of an anomaly.

The necessity to take actions to eliminate the anomaly is determined based on the information about the anomaly, as well as the DM settings, based on which exceptions $\{Z\}$ are formed. Further, the corrective actions are managed directly and the anomaly report is prepared.

This intelligent system can be classified as recognizing; therefore, the problem of choosing the threshold value α of the decision rule is an optimization problem of finding the relationship between the errors of the first and second kind. Possible combinations of the classification results and initial data are shown in the contingency table (Table 2).

In this case, a number of indicators are distinguished [12], among which we highlight the following:

– *TPR*, sensitivity of the classification algorithm (the proportion of falsely classified situations) $TPR = A/(A + C)$.

– *PPV*, accuracy of the classification algorithm (the proportion of correctly classified situations) $PPV = A/(A + B)$.

As a generalizing characteristic of the indicators, a balanced *F*-measure was chosen: $F = 2\frac{TPR \cdot PPV}{TPR + PPV}$. The analysis of the graph of the *F*-measure allows one to make an informed decision on the choice of the value α.

CONCLUSIONS

Thus, the model presented for cyber threat detection is a central element of the system for protection of Internet information resources; it is based on an intelligent system for identifying and responding to anomalies. In relation to the protection of online information platforms, the proposed model allows us to formalize a mechanism for predicting the behavioral user activity, identifying and responding to anomalies of online information platforms. The above formalisms implemented in the functional modules of the protection subsystem of the information system make it possible to increase the effectiveness of protection against spam attacks, block compromised accounts and remove malicious accounts to ensure a comfortable user experience on a web resource.

REFERENCES

1. Kirichenko, L. Radivilova, T., and Baranovskii, A., Detection of cyber threats through analysis of social networks, *Int. J. Inf. Technol. Knowl.,* 2017, vol. 11, no. 1, pp. 23–48.
2. Zegzhda, P.D., Malyshev, E.V., and Pavlenko, E.Yu., The use of an artificial neural network to detect automatically managed accounts in social networks, *Autom. Control Comput. Sci.,* 2017, vol. 51, no. 8, pp. 874–880.
3. Lavrova, D., Poltavtseva, M., and Shtyrkina, A., Security analysis of cyber-physical systems network infrastructure, *Proceedings*—2018 *IEEE Industrial Cyber-Physical Systems,* 2018, pp. 818–823. https://doi.org/10.1109/ICPHYS.2018.8390812
4. Zegzhda, D., Zegzhda, P., Pechenkin, A., and Poltavtseva, M., Modeling of information systems to their security evaluation, *SIN'17 Proceedings of the 10th International Conference on Security of Information and Networks,* Jaipur, 2017, pp. 295–298.
5. Poltavtseva, M. and Zegzhda, P., Heterogeneous semi-structured objects analysis, *Adv. Intell. Syst. Comput.,* 2018, vol. 868, pp. 1259–1270. https://doi.org/10.1007/978-3-030-01054-6_88
6. Kalinin, M., Krundyshev, V., and Zubkov, E., Estimation of applicability of modern neural network methods for preventing cyberthreats to self-organizing network infrastructures of digital economy platforms, *IV International Scientific Conference "The Convergence of Digital and Physical Worlds: Technological, Economic and Social Challenges" (CC-TESC2018),* St. Petersburg, 2018. https://doi.org/10.1051/shsconf/20184400044
7. Zegzhda, P.D., Lavrova, D.S., and Shtyrkina, A.A., Multifractal analysis of internet backbone traffic for detecting denial of service attacks, *Autom. Control Comput. Sci.,* 2018, vol. 52, no. 8, pp. 936–944.
8. Zegzhda, D.P. and Pavlenko, E.Yu., Cyber-physical system homeostatic security management, *Autom. Control Comput. Sci.,* 2017, vol. 51, no. 8, pp. 805–816.
9. Yang, C.C. and Sageman, M., Analysis of terrorist social networks with fractal views, *J. Inf. Sci.,* 2009, vol. 35, no. 3, pp. 299–320.
10. Morales, A.J., Vavilala, V., Benito, R.M., and Bar-Yam, Y., Global patterns of synchronization in human communications, *J. R. Soc. Interface,* 2017, vol, 14, no. 128. https://figshare.com/collections/Supplementary_material_from_Global_patterns_of_synchronization_in_human_communications_/3694468. Accessed January 29, 2018.
11. Zadeh, L.A., Towards a theory of fuzzy information granulation and its centrality in human reasoning and fuzzy logic, *Fuzzy Sets Syst.,* 1997, no. 4, pp. 103–111.
12. One ROC Curve and Cutoff Analysis, NCSS Statistical Software. https://www.NCSS.com/wp-content/themes/ncss/pdf/Procedures/One_ROC_Curve_and_Cutoff_Analysis.pdf. Accessed October 7, 2017.

Translated by M. Chubarova

SPELL: 1. OK