Improving Security Incidents Detection for Networked Multilevel Intelligent Control Systems in Railway Transport

Andrey V. Chernov, *Member, IEEE,* Maria A. Butakova,
Ekaterina V. Karpenko, and Oleg O. Kartashov

***Abstract*— Security monitoring and incident management systems have become the main research focus in the area of intelligent railway control systems. In this work, we discuss a system architecture of multilevel intelligent control system in Russian Railway transport and security incident classification and the handling of theprocess. We make a detailed explanation of problems and tasks of security information and event management system as an important part of a multilevel intelligent control system. We use a rough sets theory to detect an abnormal activity in the considered system. Our main result consists in the development of simple and fast detection techniques that are based on rough sets theory and allow investigating a new type of incidents.**

***Keywords* — intelligent transport systems, railway control systems, rough set theory, security information and event management.**

I. Introduction

RAILWAY transport is the main part of the transportation system in Russia. *JSC* "*Russian Railways*" is one of the largest transportation companies in the world, which owns the world's third-longest rail network (about 85 300 km) and related infrastructure in Russia. Due to the enormous amount of cargo and passengers traffic *JSC* "*Russian Railways*" is always dealing with modernizing the automation and telecommunication systems concerned with transportation processes. The information infrastructure of *JSC* "*Russian Railways*" is a new territorially distributed network system of data storage and processing with decision support systems based on optical, mobile and satellite communications. The largest fiber-optic communication network in Russia, with a length of more than 76 000 km, network capacity of 1.6 *Tbps* and more than 1 000 access nodes in all regions of Russia is laid along the railways and provided by *JSC* "*Company TransTeleCom*" [1]. The modern corporate network facilities of *JSC* "*Russian Railways*" are presented in Fig. 1.

Obviously, ensuring the comprehensive integrated security of railway automation, telecommunication and control infrastructure is an important State task. Moreover, in the rail industry the need for computer control systems to perform safety-critical tasks is constantly increasing due to the development of intelligent transport systems(*ITS*) [2]. The intensive development of *ITS* leads to the necessity for significant efforts to provide the required level of integrated security of mission critical objects on railways. *ITS* have intensive network exchange regarding "*Big Data*", therefore, traffic analysis and prediction systems have attracted considerable attention to both scientific research [3] and product development [4].

Information security attacks targeting industrial systems, such as *ITS*, have become more sophisticated, more adaptable and harder to find. Due to these circumstances, it is harder to prevent attacks and unwanted activity in *ITS* traffic. One component of security assurance of mission-critical objects, as well as *ITS*, is security monitoring and security information and event management (*SIEM*) [5]. The *SIEM* consists in continuous control of the parameters of the protected object state, their evaluation, forecasting and timely identification of facts and prerequisites of security violation. In modern computer security research, this direction is referred to as computer awareness. By "computer awareness" we mean knowledge of duties and actions (first of all, actions that violate the established security policy) of both the internal users and external subjects on the use of physical and information resources of the organization.

Despite the fact that some advances have been made in this direction, the scientific and technical issues of *SIEM* of *ITS* remain unsolved. Among various approaches, we concentrate in this work our attention on the incidents handling processes and the technique of rules discovering for *SIEM*. For this purpose, we make use of rough set theory. The rest of the paper is organized as follows. Section II contains the system architecture with a more detailed discussion of our security incidents classification. Section III presents the diagram of security incidents handling process in considering *SIEM* and classification of incidents types. Section IV shows the concept of soft computing solutions based on rough set theory, discusses decision-making tasks and presents an improved technique for a new type of security incident investigation.

Paper received May 11, 2016; revised June 9, 2016; accepted June 11, 2016. Date of publication July 20, 2016. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Vlado Delić.

This paper is a revised and expanded version of the paper presented at the 23rd Telecommunications Forum TELFOR 2015 [10].

The work was financially supported by Russian Foundation for Basic Research (projects 15-01-3067-a; 15-01-4995-a, 16-07-00888-a, 16-01-00597-a).

The authors are with the Information Technologies of Controlling Faculty of Rostov State Transport University, Rostovskogo Strelkovogo Polka Narodnogo Opolchenya 2, 344038 Rostov-na-Donu, Russia (phone: 7-8632-726595; e mail: A.V.Chernov@ieee.org).

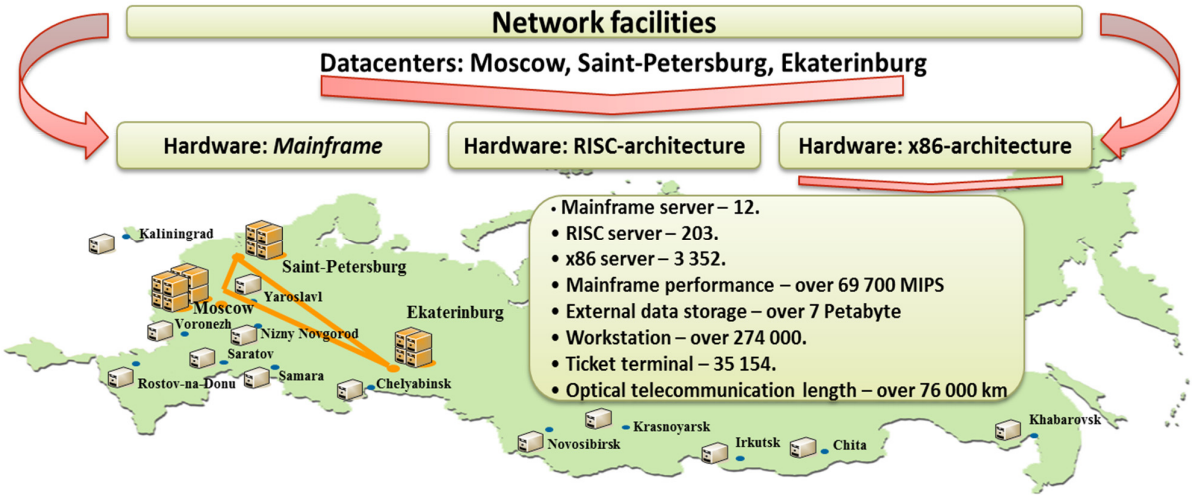

Fig. 1. Corporate network facilities of *JSC* "*Russian Railways*".

II. System Architecture

The above-mentioned Russian railway information infrastructure allows carrying out development and deployment of a multilevel intelligent control system (*MICS*) within *ITS*. In accordance with international standard (*ISA*-95) for developing an automated interface between enterprise and control systems, the hierarchy of automated control systems of industrial enterprise is divided into four levels, forming a functional pyramid: *Business Intelligence*(*BI*) – *Enterprise Resource Planning*(*ERP*)– *Manufacturing Execution System* (*MES*) – *Supervisory Control and Data Acquisition* (*SCADA*).Functional tasks and problems define *MICS* as *MES* with the following multilevel structure, Fig. 2.

The component multilevel architecture of *MICS* is based on the domain-oriented software platform. This platform provides *MES*-level components, multi-agent software for sensor and actuators, other software, and, also, the realization of *SIEM*. The application-oriented part of *MICS* begins with domain ontology. An approach based on ontology is applied in cases of creation of a system with the continuous adaptation, extension and scaling. The *MICS* must support uses, which are not known yet. For the creation of *MICS* with various and complex relations between elements, they shall be appropriately registered and provided in an information model. Because of ontology and with the use of the basic opportunities of a platform, the

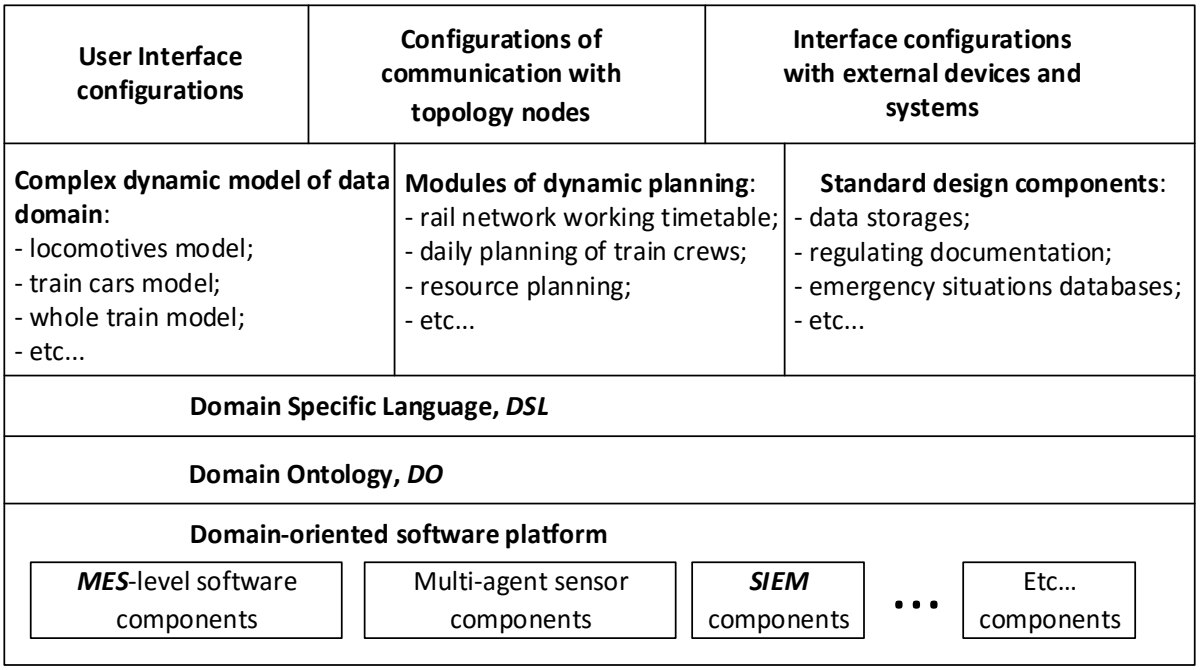

Fig. 2. System architecture of *MICS*.

domain specific language is created. The *DSL* provides a basis for the solution of application-oriented tasks of *MICS*. A solution of application-oriented tasks is realized through the creation of complex dynamic models and modules of the adaptive planning of railway operations. Much attention in case of a system design is paid to components refactoring and reuse. Finally, *MICS* consists of various user interface configurations and interfaces with networked external systems and devices.

Let's consider an outline of the incident management process in our *SIEM*. Taking into account the complex information infrastructure of the "*Russian Railways*", our *SIEM* implements a risk-based approach to data security. The approach consists of: 1) splitting of information resources and systems of *JSC* "*Russian Railways*" into the access categories; 2) determination of potential threats; 3) risk and loss assessment because of threats execution; 4) monitoring of security during functioning of information systems; 5) identifying and responding to security incidents; 6) risk-oriented planning based on prediction of security threats.

As for the specific incidents processed by our *SIEM*, their classification is very extensive and partially will be presented further in Section III.For example, we can designate the most common security incidents concerning corporate *JSC* "*Russian Railways*" telecommunication network. They are malicious network resource scanning; using of malware programs; sharing commercial, corporate information resources; unauthorized connection of notebooks and other types of mobile computers via *Ethernet* and *Wi-Fi* connection; unauthorized connection of mobile devices and smartphones to corporate computers via *USB*, *Bluetooth,* and other methods.

It is evident that we describe only the elementary security incidents, which can be implemented. Most acuteness of this problem appears when trying to implement the targeted cyber and physical attacks on resources of mission critical objects of railway infrastructure from both internal intruders (insiders) and external ones. The category of cyber and physical attacks includes various types of cyber attacks, virus infection, overcoming the access system of information resources, and so forth. The physical attack includes unauthorized disabling of railway infrastructure elements, unauthorized penetration to a protected zone, unauthorized connections to the network, and other. Targeted attacks, also called advanced persistent threats pose the greatest danger to the computerized railway objects.They are quite diverse and are implemented for instance in the form of both directed distributed denial of service attacks (*DDoS*-attacks) and (or) in "low-and-slow" mode (slowly and imperceptibly), i.e. they can be performed during an extended period until the attack object "knows" on the invasion. Examples of such attacks are *Stuxnet*, *Flame*, *Duqu*, *Wiper*, *Gauss* and others.

The greatest danger is a complex application of these types of attacks. Therefore, security monitoring and incident management of mission critical objects should be comprehensive, cover all possible types of attacks against mission critical objects infrastructure and consist in the comprehensive security monitoring of mission critical objects.

Now *SIEM* systems define a new and quite promising direction in the field of information security. These systems implement the principles of so-called a posteriori information security, performed under the assumption that the offender already has (or can embed/upload) necessary malicious hardware and software in the mission critical objects, and which provides the performance of actions realized after implementation of computer attack (or attacks) to ensure the required level of data protection in the system.Comprehensive research on the further improvements of *SIEM* systems was conducted within the European project *MASSIF*[6], which the participants of this project have been actively involved in. As part of the *MASSIF* project, some recommendations and common decisions on the construction of *SIEM* systems of a new generation were developed. In particular, the following components were integrated into *SIEM*. These are an ontological repository module, an element of security analysis, an attack modeling module and a component for visual data analysis component.

However, currently existing solutions in the field of commercial *SIEM* systems (*ArcSight*, *EMC*, *Q1 Labs*, *IBM*, *Symantec*, *LogLogic*, *Novell*, *AlienVault*, *Prelude* and others) and new generation systems have significant drawbacks that make it difficult to use them in the *MICS* in particular.

The main drawbacks are as follows: 1) insufficient performance when operating in networks with a large number of sources of security events (the majority of mission critical objects and railway network automation systems can be classified into this class of networks); 2) inability to correlate security events appeared at various levels to secure the infrastructure (network level, level of physical sensors and level of business processes); 3) low reliability and trustworthiness of the information exchange subsystem when mounting modern computer attacks; 4) focus on cyber security monitoring only and not on a comprehensive monitoring covering all elements of railway critical objects as a complex cyber-physical system.

Therefore, new theoretical research in security incident management, including novel security incident detection techniques, becomes necessary.

III. INCIDENT HANDLING AND CLASSIFICATION

Due to the huge telecommunication infrastructure of our *MICS*, we cannot detail the whole underlying network facilities. In Fig. 3 we selected and presented one telecommunication segment, designed to ensure secure communications between two network segments: *MICSLAN* and remote networks segment. A set of network interconnected devices represents the secured telecommunication architecture (*STA*) of *MICS*. As it is

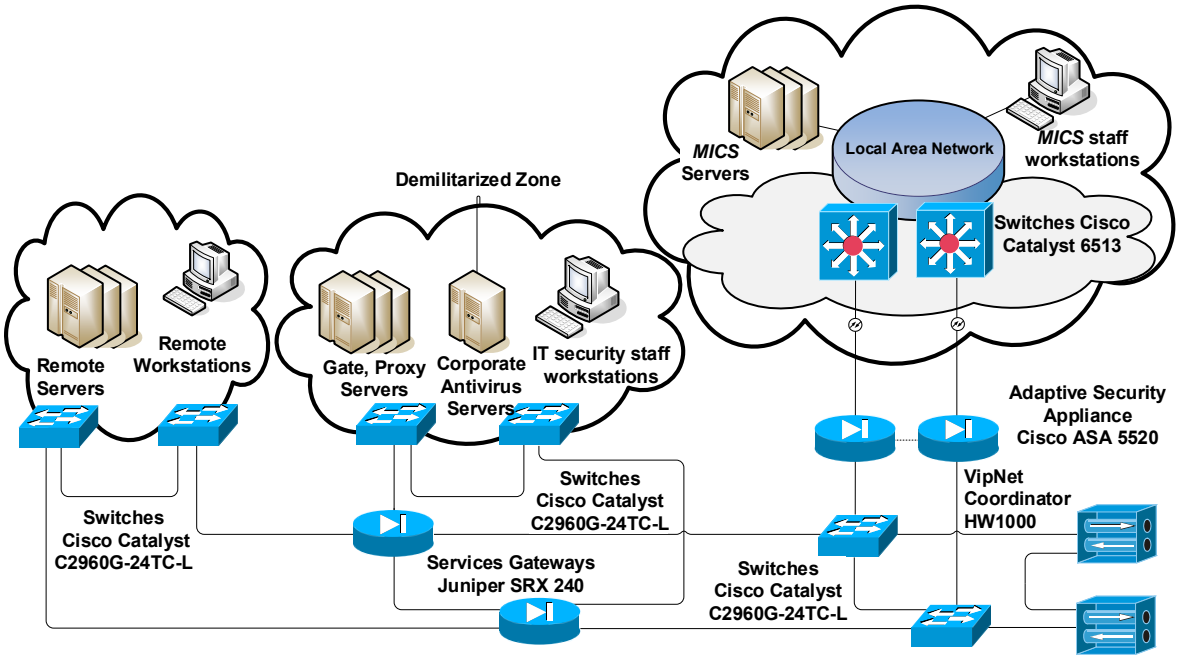

Fig. 3. Secured telecommunication architecture of *MICS*.

shown in Fig.3, a demilitarized zone separates *MICSLAN* and remote servers of various *JSC* "*Russian Railways*" enterprises. *STA* hardware devices provide the security access rules, antivirus protection, firewall and intrusion detection functions. *STA* ensures several protection perimeters implemented as software security services. In addition to hardware facilities, *STA* software services include network traffic monitoring, auditing network events, and system events registering.

The Rummler-Brache diagram [7] of the overall process of incidents handling in our *SIEM* is shown in Fig. 4. In the goals of classification, we define network security incidents as a cause-effect chain of events produced by *STA* of *MICS* and transferred to our *SIEM*. Not all registered events are network security incidents, only can have a negative effect. There are several groups of possible incidents. They are the system malfunction incidents depending on abnormal hardware and software activity, user errors incidents resulting from an incorrect user's activity, incidents caused by internal misuse and abuse, incidents of various external network attacks and theft, interruption of service incidents and other unforeseen events resulting from changes of *MICS* hardware and software configuration.

Let's consider a partial classification of network security incidents in our *SIEM*.

System malfunction incidents include:

1) malicious network ports activity;

2) network interface card malfunctioning;

3) *CRC* errors caused by network hardware;

4)*RAM* and data storage failures;

5) malfunction of network equipment (switches, routers, hubs, netbridges);

6) abnormal network subsystem activity;

7) network drivers and other software errors;

8) non-identified network software processes.

Incidents caused by user errors include:

1) mistakes made by authorized non-*IT* staff;

2) mistakes made by *IT* staff responsible for maintaining network infrastructure;

3) actions which aren't regarded by users as potentially dangerous (e.g. connecting external *USB* devices, establishing *Internet* connection without permission by using various wireless devices).

In addition to previous, another group of user-generated incidents exists. They are user actions that bypass or contravene security policy. This group of incidents contains:

1) deliberately gaining access to computer and network systems to which it is not authorized;

2) intentionally sharing access passwords and other keys to internal resources;

3) installing or modifying system software to change user privileges;

4) downloading inappropriate content or sending confidential corporate data;

5) installation of unauthorized communication services (e.g. messengers, proxy and web servers);

6) modifying, changing and insertion of false data into documents, files, databases and their associated transactions.

Moreover, finally, the broadest group of security incidents involves intentional external cyber attacks. We list the most known incidents:

1) password cracking;

2) network scanning and collecting information about network resources;

3) modifying network traffic;

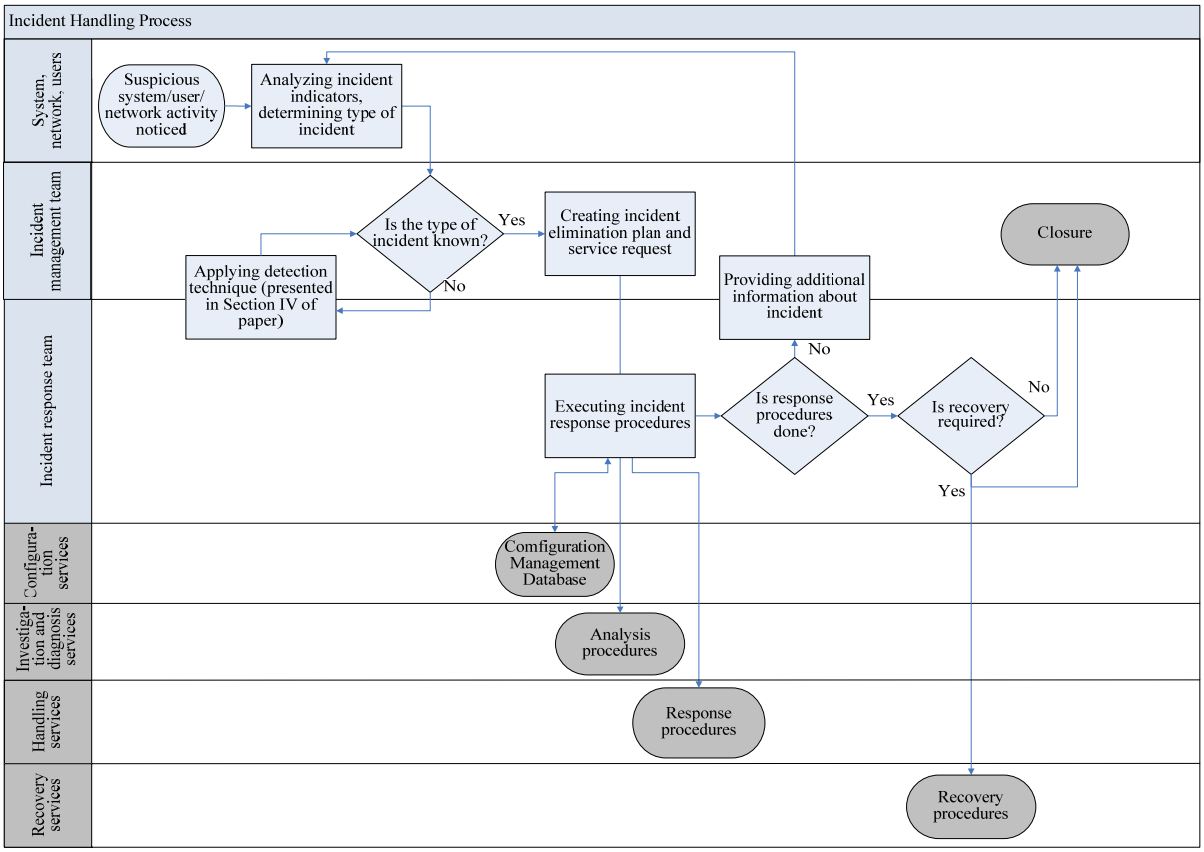

Fig. 4. Incident handling process.

4) introducing malware;

5) unauthorized interception of transmitted data and so forth.

IV. NEW SECURITY INCIDENT DETECTION BASED ON ROUGH SETS THEORY

Nearly every day we witness the appearance of a new malicious software and computer attack. That is why our knowledge about such security incidents cannot be exhaustive. In the process of research of parameters of a new incident, we derive new knowledge for decision making to prevent risks and damage occurrence. In this section, we briefly survey rough set theory as a mathematical framework intended for generating decision rules in the conditions of imprecise and insufficient knowledge.

Rough set theory [8] was invented by Z. Pawlak in early 1980s as new mathematical framework to deal with imprecise knowledge. Many researchers developed later this theory and used it to solve a variety of problems. The chief advantage of this theory consists in the fact, that rough sets theory does not require a priori information about data, such as probability in probability theory or member function in fuzzy sets theory. Further, we will consider how the theory of rough sets can be used for decision-making problems.

Imprecise knowledge or concepts can be defined approximately regarding learning sets, if we use the idea of rough sets which consists of the approximation of a set by a pair of sets, called lower approximation and upper approximation. Lower approximation consists of attributes, which surely define our concept, and upper approximation consists of attributes, which possibly define our concept. The boundary region between lower and upper approximations consists of attributes, which cannot be classified precisely based on our information. Because we have a classification task concerning the detection of an abnormal security incident and our aim is to classify suspicious (normal and abnormal) activity, let us formulate the decision task as follows.

Formally, let the pair $P=\left(O,\widetilde{R}\right)$ be an approximation space, where O is a nonempty finite set of objects called the universe, or regarding classification is our training set, and $\widetilde{R}\subseteq O\times O$ is an indiscernibility relation, which forms a set of equivalence classes on O. Thus, we define the fact, if $\left(x,y\right)\in\widetilde{R}$, then it follows x and y are indiscernible in A with respect to attributes of decision rules $R_i\langle C\rangle\rightarrow\langle D\rangle$, where C is a class of conditions and D is a class of decision, which consists of a conjugate of simple rules has a type $\langle attribute\rangle\rightarrow\langle value\rangle$.Also, we have got the set of equivalence classes defined as

$$R^*\left(D\right)=\left\{e_{1,},e_2,...,e_n\right\}.$$

Next, let us define a set $X \subseteq O$. Low approximation is

$$\underline{A}X = \bigcup_{e_i \subset X} e_i ,$$

and upper approximation is

$$\overline{A}X = \bigcup_{e_i \cap X \neq 0} e_i .$$

The pair $\langle \underline{A}X, \overline{A}X \rangle$ forms a rough set, and boundary region is

$$B_{n_p} = \overline{A}X - \underline{A}X .$$

The accuracy of an approximation of the set X of the set A of attributes is valid only for finite sets and is defined as

$$\mu_A(X) = \frac{|(\underline{A}X)|}{|(\overline{A}X)|},$$

where $|X|$ is a cardinal number of the set X.

We note that if the X is definable in A, then $\mu_A(X) = 1$, otherwise $\mu_A(X) < 1$. Finally, we write the definition for a whole information system, where we make a decision.

Given an information system IS we define a quadruple as a decision system as follows

$$IS = \langle O, Q, V, \rho \rangle, \quad (1)$$

where O is a training set; $Q = C \cup D, C \cap D = \varnothing, q \in Q$, i.e. Q is partitioned to two non-intersected sets, namely C - condition attributes, and D - decision attributes; $V = \bigcup_{q \in Q} V_q$ - is a set of domain attributes; $\rho : O \times Q \to V$, $\rho(x,q) \in V_q$, $\forall q \in Q, x \in Q$.

We formulate the security incident detection techniques in the following steps.

Step 1. Let N define a set of normal *SIEM* activity and can be represented as sequence

$$N = \{n_1^l, n_2^l, ..., n_k^l\},$$

where k is a number of activity properties, and l is a number of the sequence.

Step 2. Consider equation (1), where the decision is $D = \{normal, abnormal\}$, and, obviously $C = k$.

Step 3. Let $P = \{p_1^j, p_2^j, ..., p_m^j\}$ define a set of abnormal processes in *SIEM*.

Step 4. In the general case, we have got $m \neq n$, but we must equalize the dimensions of normal and abnormal processes. Such methods are well known (for example, sliding window), and we will not detail them.

Step 5. Obviously O consists of all sequences of normal and abnormal *SIEM* activity, and technically is represented as rows of IS. Also, in this step we obtained two classes, which are denoted Des_{normal} and $Des_{abnormal}$. Then we must calculate lower approximations $\underline{A}Des_{normal}$ and $\underline{A}Des_{abnormal}$.

Step 6. To distinguish normal and abnormal *SIEM* indicators, we need to calculate the positive region as

$$POS_A(D) = \underline{A}Des_{normal} \bigcup \underline{A}Des_{abnormal} . \quad (2)$$

The calculation of equation (2) is not simple, and we use one of the algorithms inducing decision rules. Among them, one can find *LEM1* and*LEM2* [9] algorithms to generate "*IF-THEN*" rules.

V. Conclusion

The main motivation of the proposed security incident detection technique is making the fast indicator for the *SIEM*, which is designed as a part of *MICS*. Our study is based on rough sets theory, because our knowledge about the appearance of new security threats and attack cannot be precise. Our study shows that it is possible to distinguish a malicious activity within *SIEM* looking at the part of the abnormal sequence. Finally, we hope that our proposed technique without complex calculations will be useful to *SIEM* operation in the online modes.

References

[1] *CompanyOverview of Closed Joint-Stock Company TransTeleCom* Available:http://www.bloomberg.com/research/stocks/private/snapshot.asp?privcapId=9767065

[2] *Intelligent transport systems. Innovating for the transport of the future*Available: http://ec.europa.eu/transport/themes/its/index_en.htm

[3] Yisheng Lv; Yanjie Duan; Wenwen Kang; Zhengxi Li; Fei-Yue Wang, "Traffic Flow Prediction With Big Data: A Deep Learning Approach," in *Intelligent Transportation Systems, IEEE Transactions on* , vol.16, no.2, pp.865-873, April 2015.

[4] Omar Santos, *Network Security with NetFlow and IPFIX: Big Data Analytics for Information Security*. 1nd ed. Indianapolis, IN: Cisco Press, 2015.

[5] Miller, D.; Harris, S.; Harper, A.; VanDyke, S.; Blask, C. *Security Information and Event Management (SIEM) Implementation*; McGraw-Hill Companies: Columbus, OH, USA, 2011.

[6] *MAnagement of Security information and events in Service Infrastructures.*Available: http://www.massif-project.eu/

[7] Rummler, G.A., Brache, A.P.*Improving Performance. How to Manage the White Space on the Organization Chart*. Third Edition. Jossey-Bass a Wiley Imprint, John Wiley&Sons, 2013.

[8] Pawlak, Z. Rough sets. *Int.J. Comput. Inf Sci*. 11, 1982.pp. 341-356.

[9] Jerzy W. Grzymala-Busse, Paolo Werbrouck. "On the Best Search Method in the LEM1 and LEM2 Algorithms" in *Incomplete Information: Rough Set Analysis. Studies in Fuzziness and Soft Computing*, v. 13.Physica-Verlag HD, 1998. pp. 75-91.

[10] A. V. Chernov, M. A. Butakova and E. V. Karpenko, "Security incident detection technique for multilevel intelligent control systems on railway transport in Russia," *Telecommunications Forum Telfor (TELFOR), 2015 23rd*, Belgrade, 2015, pp. 1-4.