UDC 004.457 **doi: 10.26906/SUNZ.2022.1.075**

V. Martovytskyi[1], Y. Koltun[1], D. Holubnychyi[2], V. Sukhoteplyi[3]

[1]Kharkiv National University of Radio Electronics, Kharkiv, Ukraine
[2]Simon Kuznets Kharkiv National University of Economics, Kharkiv, Ukraine
3Ivan Kozhedub Kharkiv National Air Force University, Kharkiv, Ukraine

TECHNOLOGY FOR MONITORING THE FUNCTIONING STATE OF DISTRIBUTED COMPUTER SYSTEMS

Abstract. Understanding the state of the system is important to ensure the reliability and stable operation of applications and services. The article presents the functioning process of technology of the multi-agent anomaly monitoring system, that is based on the IDEF0 methodology, describes the DCS state monitoring process and allows to integrate the monitoring system with other components of the information security system. The results of the software package of the research prototype of the DCS anomaly monitoring system are presented. The effectiveness of the developed monitoring system is confirmed by the method of simulation of attacks on the segment of the DCS network. The prototype provides a probability of making the right decision about the presence of abnormal traffic at the level of 97%, errors of the type I - 2% and errors of the type II - 1%

Keywords: IPS, IDS, networks, monitoring, distributed computer systems, multi-agent system.

Introduction

Understanding the state of infrastructure and systems is important for the stable operation of services. Deployment and performance information not only helps the team respond to problems in a timely manner, but also enables them to confidently make all necessary changes. One of the best ways to obtain this information is with a robust monitoring system that collects system metrics, visualizes data, and alerts operators to cyber influences on a computer system.

The network infrastructure monitoring systems are undergoing radical changes caused by the escalation of competition in the market, increasing requirements for the quality of security, technical re-equipment of communication networks, changes in the nature of traffic distribution [1]. All this leads to the need to control a large number of operation parameters of different technologies' networks [2].

The monitoring system not only changes the perception of the operating system, moving from collecting data parameters of individual stations to the operating parameters of the entire network, but also automates many routine processes for collecting and processing the parameters of a distributed computer system. The analysis of this information makes it possible to identify various cases of threats and violations, such as [3]:

- unauthorized connection to the network, missed by the classic means of perimeter protection (IPS / IDS);
- spread of viruses and spyware not detected by regular antivirus tools;
- incorrect actions when using the resources of distributed computer systems. For example, large-scale downloads from torrent trackers, access to network segments that are not accessible, attempting to access confidential information, etc.;
- new devices connecting to network and their behavior;
- errors in the operation of equipment;
- the emergence of "bottlenecks" in the network and other possible violations.

The architecture of the monitoring system of parameters obtained from sensors is characterized not only by their target functions, but also by functionalities that ensure the implementation of target functions, hierarchy and level of parallelism, homogeneity or heterogeneity of modular structure, organization of real-time information collection, data processing and network exchange of information with subscribers [4].

At the same time the following should be provided:

- non-interference of network equipment in the operation;
- constant collection of statistical information, which allows creating full-scale databases needed to analyze network parameters in real time [5];
- ensuring high speed processing of requests for the necessary information resources and services;
- collection, processing, storage of complete information about the state of all components of the telecommunications and information infrastructure of the network in real time, regardless of network architecture, type of switch and supplier;
- creation of a single standardized information center for storing data on the state of systems and networks.

Given the large number of events that accompany the process of diagnostic monitoring, the variety of types of events and devices in the open system being diagnosed, and the need for real-time operation given the high variability of the environment, the task of building diagnostic network monitoring should be considered big data processing. The solution to this problem is associated with the implementation of new paradigms for the development of software systems that support the possibility of distributed interaction of autonomous active devices in the process of solving a specific operational task.

The relevance of the research topic is that the current level of development of information technology and technologies allows creating devices and systems of different purposes and scales for a wide range of monitoring tasks. Monitoring of spatially distributed parameters involves significant use of network technologies, where impressive progress has also been made recently. This progress is

characterized by the following main trends: network specialization, changing the nature of network applications and shared use of different types of networks.

The relevance of the research topic is that the current level of development of information technology and technologies allows creating devices and systems of different purposes and scales for a wide range of monitoring tasks. Monitoring of spatially distributed parameters involves significant use of network technologies, where impressive progress has also been made recently. This progress is characterized by the following main trends: network specialization, changing the nature of network applications and shared use of different types of networks.

Technology for monitoring the functioning state of distributed computer systems

Understanding the state of the system is important to ensure the reliability and stable operation of applications and services. Deployment performance and productivity information not only helps the team respond to problems in a timely manner, but also enables developers to confidently make configuration changes. One of the best ways to get this information is a robust monitoring system that collects metrics, visualizes data, and alerts operators to failures.

Monitoring technology was developed to assess the state of the DCS using the multi-agent monitoring system presented in [3-4, 7].

This technology determines the conditions and procedure for assessing the state of the DCS using the developed multi-agent monitoring system. Assessment of the state of the DCS by this technology is the analysis of information about the state of the DCS to identify events that do not correspond to the normal functioning of the DCS - cyber influences.

When conducting a study on the state of the DCS, it is necessary to use monitoring agents that analyze the components of the DCS for four groups of parameters: delay, traffic, error rate, saturation, and DCS events from several monitor the functioning of the DCS according to the stages presented in Fig. 1.

Based on the structure of the DCS and the list of available services within the DCS, monitoring agents are placed on the relevant elements of the DCS. Moreover, the administrator must decide on the placement of switch agents and network agent for communication equipment, they can be placed on work stations operating within the DCS, or on dedicated hosts.

After setting up the monitoring system, the agents start collecting data from different sources, converting them to a unified data format and storing them in an intermediate database. At set intervals, the agents transmit data from the intermediate database to the central database, which can act as a distributed database and contain configured replication to increase reliability. This in turn allows you to store copies of the same data on different nodes of the network to speed up search and increase resilience to failures. Next, the training module deployed on the dedicated host of the PC provides training for PC component state assessing model and PC state assessing models.

After training the relevant models, the data analysis models of each monitoring agent are set up, which allows to assess the state of individual PC components and generate events to further start the process of assessing the state of the system as a whole.

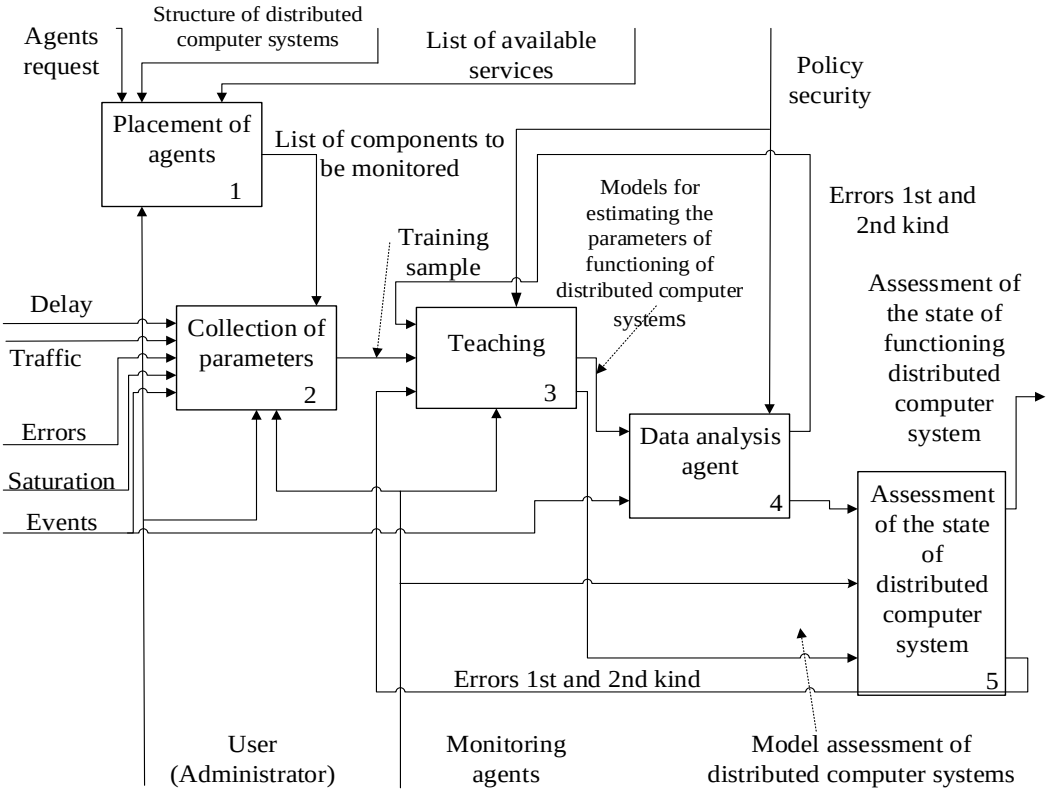

Fig. 1. IDEF1-diagram of the PC state assessing methodology using the developed multi-agent monitoring system

If the agent made a mistake of the 1st or 2nd kind, this information complements the training sample. In the case of a big number of such errors, the agent undergoes the procedure of retraining the PC component state assessing model. The same retraining procedure takes place with the PC state assessing model.

Architecture of multi-agent monitoring system

The role of the monitoring system is the timely provision of relevant information to the operator of the computer system, both as a whole and for individual nodes of the controlled complex. Simple and quick access to this information allows you to respond in a timely manner to all negative changes in the system, to find the causes of problems and failures. The monitoring system can be a means of debugging and monitoring the behavior of running user programs, which allows real-time monitoring of system resources and identify bottlenecks in it.

To date, there are many strategies and tools for monitoring the components of PC, collecting important data and responding to errors and changing conditions in different environments. But as software methods and infrastructure projects develop, monitoring must be adapted to meet new challenges and control parameters relative to unfamiliar territory.

The monitoring system not only changes the perception of the operating system, moving from collecting data parameters of individual stations to the operating parameters of the entire DCS, but also automates many routine processes for collecting and processing parameters.

The architecture of MS designed to solve problems of collecting and storing parameters obtained from sensors, is characterized not only by its target functions, but also functionality that ensures the implementation of target functions, hierarchy and degree of parallelism, homogeneity or heterogeneity of modular structure, organization of information collection real-time, data processing and network exchange of information with subscribers [6]. To solve the above mentioned problems, the architecture of the monitoring system using stand-alone software agents was developed, which is presented in Fig. 2.

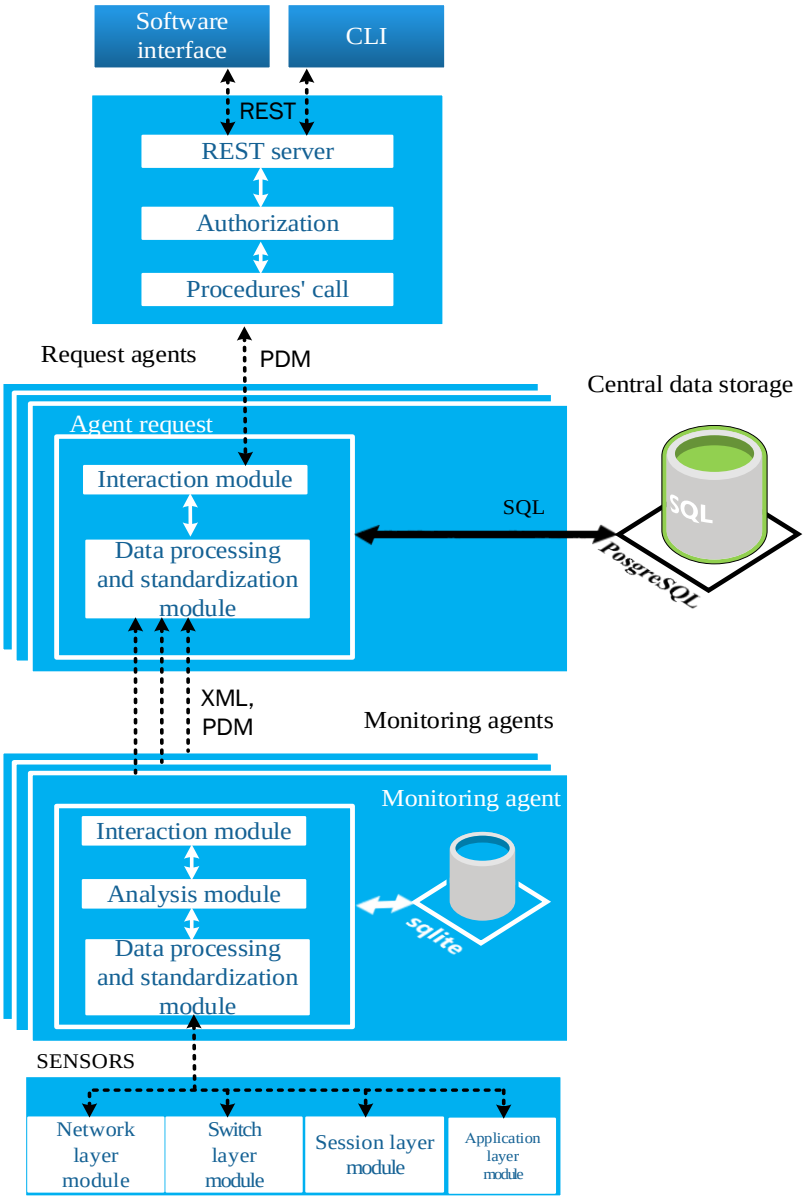

Fig. 2. Architecture system monitoring

The top-most module of the system is responsible for organizing user access to the system and displaying monitoring data in the form of graphs and a list of abnormal conditions to the user. Speed is not so important at this stage, because the software running in this module can run outside the computing complex, referring only to the data provided by the request agents.

The monitoring system stores the parameters of the DCS components, which are also called metrics. A metric is a necessary and sufficient set of characteristics of attributes and operations designed to describe the state of a component. For each specific component, the metric acquires specific set of characteristics' values.

The software package that implements the architecture must monitor the consistency of data over time. In practice, the metric stores only a set of data for a relatively short period of time and organizes effective access to them. The rest of the data collected from RAM in the intermediate data warehouse is transparently accessible to the user upon request.

Architecture involves the dynamic formation of a hierarchical structure, the node of which can be any entity defined by a data source or sensor. Thus, there may be DCS monitoring metrics for grid, clusters, computing nodes, and tasks.

For interaction between all agents, it is proposed to use a group of intelligent request agents whose purpose is to coordinate information collection agents, restructure the information obtained and implement protocols and messaging mechanisms between all agents of the model.

Functions of each information collection agent:

- Collection and accumulation of data in the intermediate storage of small volume;
- Real-time data processing;
- Adjustment of sampling intervals;
- Requesting to agents for additional measurements and a comprehensive analysis of the situation;

Monitoring agents can be divided into several groups.

A switch agent and a network agent that provide data collection from the first two levels described above. Since the operation of the channel and network layers is provided mainly by active network equipment and is usually implemented by the following components: network adapters, repeaters, bridges, hubs, switches, routers, to minimize interference with network equipment, these agents will work based on the SNMP protocol. MIB files will be used as intermediate data storage.

The task of these agents is to standardize data from files for further transmission to request agents. Agents are also tasked with managing the delivery of alarms, as the SNMP protocol works by the means of the unreliable UDP protocol [7].

Session agent that collects information about the user name, terminal line name, astronomical start time of the session, the duration of inactivity of the terminal line since the last exchange, the process ID of the shell command interpreter for each user running the system.

Depending on the operating system, intermediate storages may differ.

For example, for UNIX systems, such storages will be the system files /etc/utmp, /etc/wtmp, /etc/inittab.

The application agent is responsible for collecting data from various applications specific to a particular computer system.

Query agents aim to process requests to sample data from users of the collection system, coordinate other agents to collect the necessary information, and restructure the information obtained to store statistics about the system as a whole.

The introduction of such agents and the software implementation of standardized interfaces between them allow using independent developers' software at different levels. For example, sensors can be data files. However, all sensors form a single structure of metrics, equally accessible to different components of the MS.

Thus, the monitoring system built on this architecture can work in parallel with the already deployed monitoring tools, replacing them at some levels, which allows you to change and expand the set of available functions of these systems.

Automated workplace of the monitoring server administrator

The software complex for monitoring of the DCS functioning is designed to monitor the condition of the components of the DCS, as well as elements of the network infrastructure. This complex allows monitoring security events and other events occurring at the nodes of the DCS network, as well as timely identifying problems with the nodes and promptly informing users about problems that arise during operation.

The monitoring software package consists of the following components:

- monitoring server - a software server that performs the following functions:

• collects and stores information on the current state of network nodes and other elements of the infrastructure of the DCS network;

• analyzes the values of the components states' parameters of the DCS and generates messages about the detected events;

• notifies system operators and administrators of changes in the status of monitoring objects and detected events, as well as provides API for interaction with external information systems.

- monitoring workstation - the workplace of the operator or administrator of the monitoring server, which allows you to manage monitoring agents through a secure channel. Data access and event notification of monitoring agents are accessed through graphical user interfaces.
- monitoring nodes - network elements, the status of which is monitored by monitoring agents.
- monitoring agent - a component of client software that is located on the monitoring node and provides collection and transmission of data on the status of the node to the monitoring server.

Objectives of the DCS functioning monitoring system:

- monitoring of distributed network nodes (determination of failures, security events and other events);
- monitoring of components, equipment and other parts of computer systems (determination of failures, security events and other events);

- monitoring of mobile devices and nodes of remote network users, regardless of the method of their connection to the communication network;

- monitoring of network infrastructure facilities (routers, switches and other components), peripherals, such as printers, MFPs that support the SNMP protocol.

To deploy the monitoring system, it is necessary to scan the elements of the network structure of the DCS, which sets the range of addresses to scan the active nodes of the network. After scanning the nodes using the "TCP SYN scan" method, the services running on each host are analyzed.

The next step in the deployment of the monitoring system based on the structure of the DCS and the list of available services within the DCS are monitoring agents. The result of the agent placement is shown in Fig. 3.

Query agents are shown in red, network agents in blue, switch agents that are hosted on the appropriate hosts on the network in yellow, session agents in green, and application agents in dark blue.

Then the monitoring system begins to monitor the parameters of the distributed computer system and displays information about suspicious malicious activity in the system.

It is worth noting the important prospects for the development of the proposed monitoring system. The developed prototype has good scalability and functional flexibility.

In addition, the prototype has the ability to embed as a kernel or add-on module in the system of protection against network attacks in systems, which are presented in the articles [8, 9].

The developed MS allows monitoring of DCS not only as a single computer system, but also of all its components separately, which makes it possible to comprehensively assess the state of the system as a whole.

Further research includes the continuation in the direction of studying new types of anomalies not only in the information and communication environment of DCS, but also anomalies in the operation of the system nodes, as well as improving system architecture, improving efficiency and refining monitoring system algorithms using machine learning methods.

Evaluation of the developed monitoring system efficiency

The efficiency of the proposed CM is evaluated only to detect abnormal behavior of network traffic based on the set of parameters of network connections {NW}, which is implemented by analyzing the incoming traffic using an ensemble of classifiers.

Then an experimental study of the effectiveness of the proposed model for detecting abnormal behavior was conducted, the results of which are given in table. 1.

Table 1 – **Experimental study of the proposed model effectiveness in the presence of traffic anomalies**

Number of tests	Number of correct decisions	Number of incorrect decisions	
		Errors of the type I	Errors of the type II
10000	9696	203	101

The tests were performed on a computing cluster deployed on training laboratory's workstation under the influence of a slow DDoS attack and an attack over the entire time interval.

The number of tests was limited to 10,000 due to the fact that the results of the study after reaching this number of tests showed a steady trend. The probability of making the right decision about the presence of abnormal traffic was 97%, errors of the type I - 2% and errors of the type II- 1%.

Thus, the proposed model is effective for detecting anomalies in the functioning of the computing cluster.

These studies can be used to improve existing subsystems for monitoring supercomputer technologies, as well as be the basis for creating a fundamentally new neural network multi-agent anomalous events detection monitoring system.

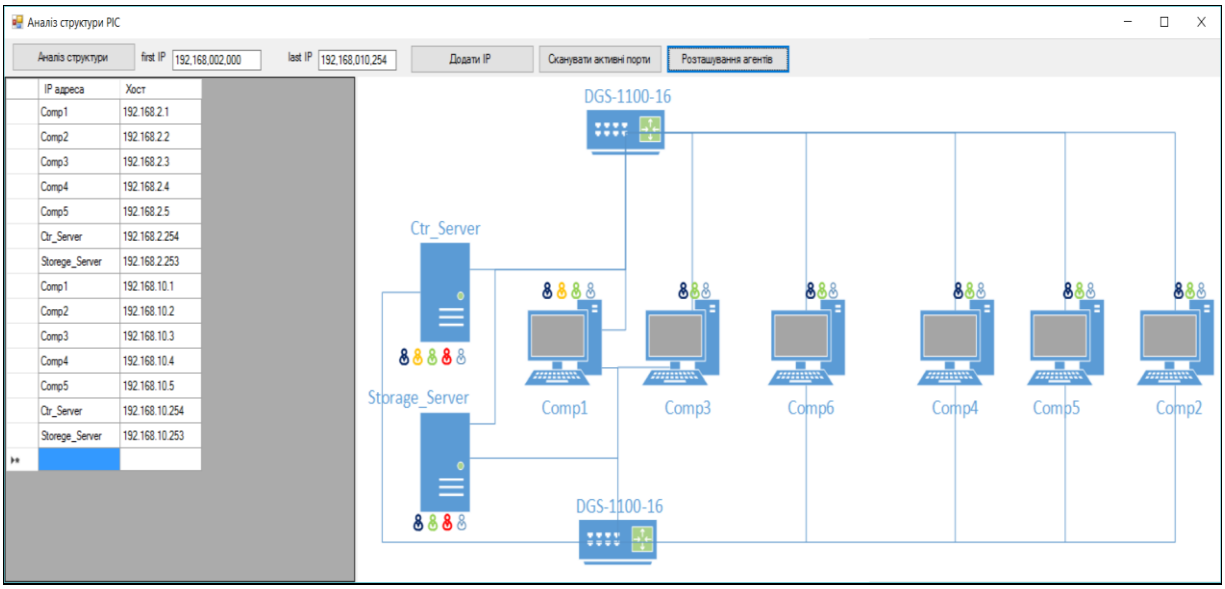

Fig. 3. Agents location by the components of a distributed computer system

Conclusion

In the article is solved the actual scientific and practical problem of improvement of anomalies detection indicators of DCS functioning in the conditions of cybernetic influences of external and internal environment by construction of models and methods on the basis of technologies of data mining.

The structure and architecture of the monitoring system using autonomous software agents have been developed. The architecture of MS is designed to solve the problem of collecting and storing parameters obtained from sensors, characterized not only by their target functions, but also functionality.

The functioning process technology of multi-agent anomalies monitoring system while functioning DCS components based on IDEF0 methodology is developed, detailing process of monitoring of DCS condition and allowing integrating monitoring system with other components of information protection system.

The results of the software package of the research prototype of the DCS anomaly monitoring system are presented. The effectiveness of the developed monitoring system is confirmed by the method of simulation of attacks on the segment of the DCS network. The prototype provides a probability of making the right decision about the presence of abnormal traffic at the level of 97%, errors of the type I - 2% and errors of the type II - 1%.

Further research should focus on improving and enhancing the method of assessing the state of the computing node and the state of running tasks in the computer system.

REFERENCES

1. Mukhin, V., Kuchuk, N., Kosenko, N., Kuchuk, H. and Kosenko, V. Decomposition Method for Synthesizing the Computer System Architecture, Advances in Intelligent Systems and Computing", AISC, vol. 938, pp 289-300, DOI: 10.1007/978-3-030-16621-2_27
2. Kuchuk, N., Mozhaiev, O., Mozhaiev, M., Kuchuk, H. Method for calculating of R-learning traffic peakedness 2017 4th International Scientific-Practical Conference Problems of Infocommunications Science and Technology, PIC S and T 2017 – Proceedings, pp. 359-362, DOI: 10.1109/INFOCOMMST.2017.8246416
3. I. Ruban, V. Martovytskyi and N Lukova-Chuiko, Approach to Classifying the State of a Network Based on Statistical Parameters for Detecting Anomalies in the Information Structure of a Computing System, Cybernetics and Systems Analysis, vol. 54, no. 2, pp. 302-309, 2018.
4. I. Ruban, V. Martovytskyi and N. Lukova-Chuiko, Designing a monitoring model for cluster super-computer, Eastern-European Journal of Enterprise Technologies, vol. 6, no. 84, pp. 32-37, 2016
5. Tkachov, Vitalii, et al. Overlay Network Infrastructure for Remote Control of Radio Astronomy Observatory, 2019 IEEE International Conference on Advanced Trends in Information Theory (ATIT). IEEE, 2019
6. G. Kuchuk, A. Kovalenko, I.E. Komari, A. Svyrydov, V. Kharchenko Improving big data centers energy efficiency: Traffic based model and method. Studies in Systems, Decision and Control, vol 171. Kharchenko, V., Kondratenko, Y., Kacprzyk, J. (Eds.). Springer Nature Switzerland AG, 2019. Pp. 161-183. DOI: http://doi.org/10.1007/978-3-030-00253-4_8
7. V. Martovytskyi, K. Smelyakov, D. Pribylnov, A. Chupryna Investigation of network infrastructure control parameters for effective intellectual analysis IEEE 14th International Conference on Advanced Trends in Radioelecrtronics, Telecommunications and Computer Engineering (TCSET), 20-24 Feb. 2018. – Pp. 983-986. DOI: 10.1109/TCSET.2018.8336359
8. Liashenko, O., Barkovska, O., Al-Atroshi, C., Datsok, O., Liashenko, S. Model of the work of the neurocontroller to control fuzzy data from the sensors of the climate control subsystem "smart house" International Journal of Advanced Trends in Computer Science and Engineering, 8(1), Pp. 70-74
9. Kovalenko A., Shamraev A., Shamraeva E., Dovbnya A., Ilyunin O. Green Microcontrollers in Control Systems for Magnetic Elements of Linear Electron Accelerators. Green IT Engineering: Concepts, Models, Complex Systems Architectures. Studies in Systems, Decision and Control series. Springer International Publishing Switzerland, 2017. Pp. 283-305. DOI: http://dx.doi.org/10.1007/978-3-319- 44162-7_15

Received (Надійшла) 24.12.2021
Accepted for publication (Прийнята до друку) 02.02.2022

Технологія моніторингу
стану функціонування розподілених комп'ютерних систем

В. О. Мартовицький, Ю. М. Колтун, Д. Ю. Голубничий, В. М. Сухотеплий

Анотація. В статті представлена актуальна науково-практична задача покращення показників виявлення аномалій функціонування розподілених комп'ютерних систем в умовах кібернетичних впливів зовнішнього та внутрішнього середовища. Представлена структура та архітектура системи моніторингу з використанням автономних програмних агентів. Архітектура системи моніторингу призначена для вирішення завдань збору та зберігання параметрів отриманих від датчиків, що характеризуються не тільки їх цільовими функціями, але і функціональними можливостями. Розроблено технологію процесу функціонування мультиагентної системи моніторингу аномалій при функціонуванні компонентів розподілених комп'ютерних систем, що базуються на методології IDEF1, що деталізують процес моніторингу стану розподілених комп'ютерних систем и дозволяють інтегрувати систему моніторингу з іншими компонентами системи захисту інформації. Представлено результати роботи програмного комплексу системи моніторингу аномалій функціонування розподілених комп'ютерних систем. Ефективність розробленої системи моніторингу підтверджена методом імітаційного моделювання атак на сегмент мережі розподілених комп'ютерних систем. Прототип забезпечує ймовірність прийняття вірного рішення про наявність аномального трафіку склало 91%, помилки першого роду - 5% і помилки другого роду - 4%. Подальші дослідження доцільно спрямувати на вдосконалення та покращення методу оцінки стану обчислювального вузла та стану запущених завдань в комп'ютерній системі.

Ключові слова: IPS, IDS, мережі, моніторинг, розподілені комп'ютерні системи, багатоагентна система.