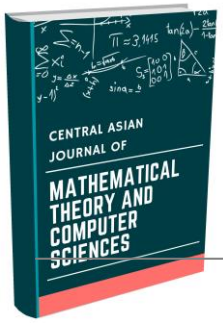

CENTRAL ASIAN JOURNAL OF MATHEMATICAL THEORY AND COMPUTER SCIENCES

https://cajmtcs.centralasianstudies.org

Volume: 04 Issue: 9 | Sep 2023 ISSN: 2660-5309

APPLICATION OF INTELLECTUAL ANALYSIS TO PROTECT INFORMATION IN CORPORATE SYSTEMS

Ibrokhimali Normatov, Inomjon Yarashov, Otabek Tangriberdiyev

National university of Uzbekistan named after Mirzo Ulugbek

Abstract

The article proposes a methodology for constructing an adaptive self-developing information security system for corporate systems. When building the system, methods and basic approaches of intelligent data analysis were used. The system is designed to solve a single problem of protecting computer networks, databases and automatic information processing systems. The elements of using intelligent data analysis methodology in the water supply industry are presented.

ARTICLEINFO

Article history:
Received 3 Jul 2023
Revised form 5 Aug 2023
Accepted 26 Sep 2023

Keywords: *intelligent data analysis, classification, corporate information systems, protection system, database.*

Introduction

Intelligent data analysis (IDA) is one of the progressive methods for analyzing large volumes of data. It is the process of discovering and further applying knowledge or previously unknown information from existing sets[1], the main objectives of which are classification; association; clustering; forecasting; subsequence.

Tools for creating intelligent applications are represented by developments from Cognos, G2 from Gensym Corp, MineSet from Silicon Graphics, Intelligent Miner from IBM, and IDIS from Information Discovery.

Universal IDA tools are quite complex and expensive, so they are not always used in integrated end-user-oriented systems. Intelligent data analysis is used in many areas of modern society, helping to solve a wide variety of problems. These are, for example, insurance, banking, marketing, financial risk analysis, monitoring of equipment and technological processes[9-16], telecommunications, computer security[17-29], etc.

In the field of computer security, IDA methods are closely related to the creation of promising information security systems (IPS). It is the IDA methodology that helps to implement in information security the evolutionary properties of adaptation, self-organization, learning, the possibility of inheritance and representation of the experience of information[30-36] security experts in the form of a system of fuzzy rules accessible for analysis.

Intellectual analysis in corporate information systems

The most complex modern information structures[35-38] aimed at large companies are corporate systems. They are characterized by the use of multiple computers, client-server architecture, specialization of servers, the presence of a single information space, and an extensive network of data reception and transmission. CIS

databases contain huge volumes of data and have all the features of a complex system organization. CIS information spaces include relational and object DBMSs, transactional databases, time series and large-volume numerical data, multidimensional OLAP storages.

Examples of commercial corporate systems are R/3 systems from SAP, Oracle Application, Microsoft Business Solution Navision, the Parus system, an application solution for the 1C: Enterprise 8.0 system "Manufacturing Enterprise Management", corporate information systems from Atlas, corporate information projects based on Lotus Notes/Domino technology. In [2], hybrid intelligent systems are considered that make it possible to effectively combine formalized and informal knowledge through the integration of traditional artificial intelligence tools, examples of the merging of corporate and intelligent systems.

Let's supplement the earlier review with information about the use of IDA in commercial corporate systems.

A comprehensive business intelligence software solution that provides quick access to information and its use in making strategically important decisions from SAP is the SAP Business Intelligence (SAP BI) subsystem.

The core of the solution is a data warehouse designed to store internal and external information, including documentation, video and audio clips. It integrates information across the entire SAP Business Suite platform and provides the ability to quickly respond to market changes, monitor indicators of key success factors, analyze and optimize enterprise performance based on a single business model.

Oracle provides a full range of data mining products - from various tools to ready-made applications - and tailors their use according to the user's problems.

The most popular tools are Oracle OLAP and Oracle Data Mining. OLAP (online analytical data processing) tools are useful when it comes to multidimensional indicators, their hierarchical aggregation and detailing, and modeling. Microsoft SQL Server 2022 provides an integrated environment for creating and working with data mining models. This environment is called Microsoft SQL Server Analysis Services and consists of a set of special tools (Business Intelligence Development Studio, SQL Server Management Studio, Microsoft SQL Server 2022 Integration Services, BI Development Studio). This environment includes data mining algorithms and tools that facilitate the development of a comprehensive solution applicable to a wide variety of projects.

The 1C Enterprise 8.0 system has a special tool – “Data Analysis Subsystem”, which can be built into any platform configuration.

It is designed to help users of a corporate information system find answers to non-trivial questions. Provides automated transformation of data accumulated in the corporate information system into practically useful and well-interpreted patterns, implements grouping of relatively similar objects; search for stable combinations of events and objects (associations); provides the construction of a cause-and-effect hierarchy of conditions leading to certain decisions (decision tree).

A feature of many commercial corporate systems is that security systems are not initially included in their composition, and, despite the availability of tools, must be selected and purchased separately.

This leads to additional costs (financial, time, labor, material, etc.) when purchasing, setting up and operating systems; it requires the development of approvals for the integration of two dissimilar systems.

Intelligent data analysis in ensuring information security

Modern computer systems and networks are in a state of constant development and modification, and the volume of analyzed data in the world doubles every year. Therefore, to ensure the required level of information protection, it is necessary to respond flexibly and quickly to changing conditions, provide reliable protection taking into account constant changes in input influences, and prevent the actions of intruders, i.e. have adaptive and self-developing information security systems.

The purpose of this work is to develop a methodology for using data mining to build an adaptive, self-developing information security system in corporate systems.

The need to use data mining tools in the information security system of corporate systems stems from the heterogeneity of the structures of the information spaces of these systems; difficulties in obtaining analytical information from large databases; a large number of users simultaneously working in the system; requirements for constant monitoring of functioning and making informed management decisions, depending on many factors.

The prerequisites for using IDA in a CIS are client-server technology, distributed databases, the availability of information storage facilities, the use of modern network technologies and a variety of tools used for collecting, processing, visualizing and analyzing data.

A feature of information security systems in corporate systems is a combination of at least three problems: information security in computer networks; ensuring database security; ensuring the safe operation of automatic information processing systems [3].

Intelligent tools often used in computer networks include knowledge bases as part of expert systems, systems based on the Bayesian method, fuzzy logic systems, neural networks, evolutionary methods and hybrid intelligent systems. The main tasks solved by intelligent means of ensuring information security of a computer network are classification and clustering. Intelligent database security tools can be found in [4]. It is indicated that the database information security system must use the tools and objects of the applied database management system (DBMS), database objects and tools, a set of rules and events characterizing user actions. In [5] it is stated that it is the recording of events that allows one to get an idea of what each user is interested in; a list of the main recorded events has been compiled.

The means of ensuring the safe operation of information processing systems include intrusion prevention mechanisms, authorization, delimitation of access rights, cryptographic protection (on storage media, in networks, password protection), and management of user rights. In order to monitor the state of the system, they use signature databases of known attacks, and use system logs and files as the main sources of information, and analyze the contents of network traffic and files.

The traditional approach to building a security system using IDA tools uses artificial neural networks, decision trees and classification algorithms, fuzzy clustering methods, association rules, limited search algorithms, and cluster analysis.

Neural networks are used to monitor the traffic of a protected local network, search for hidden patterns in primary data arrays, and detect intrusions. To predict the value of the target indicator, sets of input variables, mathematical activation functions and weighting coefficients of the input parameters are used. An iterative training loop is performed, the neural network modifies the weighting coefficients until the predicted output parameter matches the actual value. After training, the neural network becomes a model that is used for prediction.

Classification mechanisms are used at the initial level, for example, to systematize protection methods (fuzzy conclusions) according to a vector of fuzzy threats. If the reliability of the classification for known threats is less than a certain level, then if there are signs of an attack, the classification is expanded by introducing a new gradation into the classification - the problem of threat clustering is solved. Associations reveal cause-and-effect relationships and determine probabilities or confidence coefficients, allowing appropriate conclusions to be drawn.

Most publications on the use of intelligent information security systems are devoted to attack detection systems based on the model proposed by Denning. The model contains a set of profiles for legitimate users, compares current activities with the corresponding profile, updates the profile, and reports any anomalies detected.

The disadvantages of the traditional approach are:

1. Knowledge bases are formed by experts, i.e. the principle of including situations in them is subjective.
2. Knowledge bases must be periodically updated, organized, and systematized, which is a labor-intensive and expensive procedure.
3. With the traditional approach, there is a time delay between the appearance of a new attack and means of protection against it (lagged counteraction).
4. Attacks are constantly modified, improved, "disguised" as standard procedures, which requires constant improvement and complication of security measures.

Taking into account the above, the problem of the evolutionary development of information security systems (ISS) is relevant.

Construction of an adaptive self-developing protection system

Building an intelligent data analysis model is part of a larger process that involves everything from formulating data selection and storage issues to model creation to deploying the model to production. Let's move on to describe the features of building an adaptive self-developing system.

Table 1 provides a list of the main data sources and the information they contain to be analyzed.

Table 1 – Sources of analyzed data

Data source	Analyzed information
log files of running subsystems	time and type of operations performed, essence of operations, password compliance, failures when establishing communication with a remote machine, emergency stop diagnostics
network traffic	loading of network equipment, use of communication channels, network activity
directories and logs of registration of users and events	User ID codes, correctness of passwords, actions performed
lists of functional tasks	chains of interconnected calls to tasks and processes
access rights information	compliance with the regulations for accessing resources
information about the operation of the mail system	statistics, volumes and targeting of mailings and postal receipts, subject of messages
text files	thematic focus
Excel workbooks	security, presence/absence of macros
tables with attributes of executable files	file types, creation and modification dates, authors of changes and their rights, control of "immutability", addresses of reference modules, checksums

Sources of data for analysis are system event logs, temporary files of servers and workstations, log files of running subsystems, network traffic, directories and user and event logs, lists of functional tasks and information about access rights, information about the operation of the mail system, text files , Excel workbooks, emails, tables with attributes of executable files, etc.

Working hypotheses:

1. User activity, targeted access to system resources and processes occurring in the system can be recorded and an adequate model can be built.

2. An event (or sequence of events) corresponding to a generalized attack model is truly an attack, and the use of anticipatory or simultaneous counteraction algorithms is justified.
3. The system can monitor the operation of the system software and, if it detects damage, restore protection and automatically resume downloading of lost or damaged files.

The mechanism for using IDA for adaptive self-developing information security can be divided into a number of stages.

1. Statement of the problem. At this stage, the requirements are analyzed, the problems that will be solved, the metrics by which the model will be evaluated are determined, and the tasks for the data mining project are defined.

This stage examines the levels of data confidentiality, user needs and rights in relation to available data, identification and authentication methods traditionally used in the enterprise.

In this case, system information security risks can be defined as a function of three variables:

- the likelihood of the existence of threats (potentially possible events, intentional or accidental, that could have an undesirable impact on the corporate system or its parts, or on information assets and, as a result, on the company's business processes);
- the likelihood of the existence of vulnerabilities (deficiencies or shortcomings in the system, due to which it becomes possible for unwanted influence on it from intruders, unqualified personnel or malicious code);
- potential losses, which are potential direct and indirect financial losses resulting from the implementation of threats and the presence of vulnerabilities

At the same time, the decision to expand the classifications of attacks and protection mechanisms is made in accordance with the system of assessing the reliability of neutralizing threats in the context of individual protection mechanisms. The feasibility of using a protection mechanism as part of multi-level information security systems can be justified, for example, using the matrix of reliability of using protection mechanisms to neutralize threats [6]:

$$x_i = \sqrt[n]{\prod_{j=1}^{n} me_{ij}}, \qquad i = 1, \ldots, m$$

where me_{ij} – elements of the credibility matrix "threats - protection mechanisms".

Conclusion

Intelligent data analysis is a necessary and modern addition to such a large information structure as a corporate system. One of its components is the information security system. Security means must be constantly improved and developed, which is why the mechanism proposed in the work for constructing an adaptive self-developing information security system is relevant, and the use of fast algorithms along with IDA will increase the efficiency of the system, which is the topic of a separate study.

References

1. Han J. Data Mining: Concepts and Techniques / J. Han, M. Kamber // Morgan Kaufmann. – 2000.
2. Маслова Н.А. Концептуальные особенности построения интеллектуальных корпоративных систем предприятий водоснабжающей отрясли / Н.А. Маслова // Штучний інтелект. – 2006. – № 4. –С. 443-452.

3. Шаньгин В.Ф. Защита информации в распределенных корпоративных сетях и системах / В.Ф. Шаньгин, А.В. Соколов. – Изд-во: ДМК, 2002. – 134 с.
4. Корнеев В.В. Базы данных: интеллектуальная обработка информации/ В.В. Корнеев, А.Ф. Гареев, С.В. Васютин, В.В. Райх. – М. : Нолидж, 2000. – 352 с.
5. Маслова Н.А. Информационная безопасность систем управления базами данных / Маслова Н.А. // Комп'ютерна математика. Оптимізація обчислень : зб. наук. праць. – Київ : ІК НАН України, 2001. – Т. 1. – С. 271-280.
6. Гончаров М. Модифицированный древовидный алгоритм Байеса для решения задач классификации / Гончаров М. – Spellabs, 2007.
7. Задірака В.К. Т-ефективні алгоритми наближеного розв'язування задач обчислювальної математики / В.К. Задірака, М.Д. Бабич, А.І. Березовський та ін. – К., 2003. – 216 с.
8. Kabulov A. et al. Algorithmic method of security of the Internet of Things based on steganographic coding //2021 IEEE International IOT, Electronics and Mechatronics Conference (IEMTRONICS). – IEEE, 2021. – С. 1-5.
9. Kabulov A., Kalandarov I., Yarashov I. Problems of algorithmization of control of complex systems based on functioning tables in dynamic control systems //2021 International Conference on Information Science and Communications Technologies (ICISCT). – IEEE, 2021. – С. 1-4.
10. A. Kabulov, I. Saymanov, I. Yarashov and A. Karimov, "Using Algorithmic Modeling to Control User Access Based on Functioning Table," 2022 IEEE International IOT, Electronics and Mechatronics Conference (IEMTRONICS), Toronto, ON, Canada, 2022, pp. 1-5, doi: 10.1109/IEMTRONICS55184.2022.9795850.
11. A. Kabulov, I. Normatov, I. Kalandarov and I. Yarashov, "Development of An Algorithmic Model And Methods For Managing Production Systems Based On Algebra Over Functioning Tables," 2021 International Conference on Information Science and Communications Technologies (ICISCT), Tashkent, Uzbekistan, 2021, pp. 1-4, doi: 10.1109/ICISCT52966.2021.9670307.
12. A. Kabulov and I. Yarashov, "Mathematical model of Information Processing in the Ecological Monitoring Information System," 2021 International Conference on Information Science and Communications Technologies (ICISCT), Tashkent, Uzbekistan, 2021, pp. 1-4, doi: 10.1109/ICISCT52966.2021.9670192.
13. A. Kabulov, I. Yarashov and A. Otakhonov, "Algorithmic Analysis of the System Based on the Functioning Table and Information Security," 2022 IEEE International IOT, Electronics and Mechatronics Conference (IEMTRONICS), Toronto, ON, Canada, 2022, pp. 1-5, doi: 10.1109/IEMTRONICS55184.2022.9795746.
14. Kabulov A. V. et al. COMPUTER VIRUSES AND VIRUS PROTECTION PROBLEMS //Science and Education. – 2020. – Т. 1. – №. 9. – С. 179-184.
15. Madrahimova D., Yarashov I. Limited in solving problems of computational mathematics the use of elements //Science and Education. – 2020. – Т. 1. – №. 6. – С. 7-14.
16. Yarashov I. Algorithmic Formalization Of User Access To The Ecological Monitoring Information System //2021 International Conference on Information Science and Communications Technologies (ICISCT). – IEEE, 2021. – С. 1-3.
17. Kabulov A. et al. Algorithmic method of security of the Internet of Things based on steganographic coding. 2021 IEEE International IOT //Electronics and Mechatronics Conference, IEMTRONICS.–2021. – 2021.

18. Kabulov A., Muhammadiyev F., Yarashov I. Analysis of information system threats //Science and Education. – 2020. – Т. 1. – №. 8. – С. 86-91.

19. Kabulov A., Yarashov I., Vasiyeva D. Security Threats and Challenges in Iot Technologies //Science and Education. – 2021. – Т. 2. – №. 1. – С. 170-178.

20. Gaynazarov S. M. et al. Algorithm of mobile application for medicine search //Science and Education. – 2020. – Т. 1. – №. 8. – С. 600-605.

21. Yarashov I., Normatov I., Mamatov A. The structure of the ecological information processing database and its organization //International Conference on Multidimensional Research and Innovative Technological Analyses. – 2022. – С. 114-117.

22. Yarashov I., Normatov I., Mamatov A. Ecological information processing technologies and information security //International Conference on Multidimensional Research and Innovative Technological Analyses. – 2022. – С. 73-76.

23. Kabulov A., Yarashov I., Mirzataev S. Development of the implementation of IoT monitoring system based on Node-Red technology //Karakalpak Scientific Journal. – 2022. – Т. 5. – №. 2. – С. 55-64.

24. Кабулов А. В., Болтаев Ш. Т. АЛГОРИТМИЧЕСКИЕ АВТОМАТНЫЕ МОДЕЛИ И МЕТОДЫ СОЗДАНИЯ РАСПРЕДЕЛЕННЫХ МИКРОПРОЦЕССОРНЫХ СИСТЕМ УПРАВЛЕНИЯ И ИНФОРМАЦИОННОЙ БЕЗОПАСНОСТИ.

25. I. Yarashov, "Development of a reliable method for grouping users in user access control based on a Functioning table," 2022 International Conference on Information Science and Communications Technologies (ICISCT), Tashkent, Uzbekistan, 2022, pp. 1-5, doi: 10.1109/ICISCT55600.2022.10146787.

26. S. Toshmatov, I. Yarashov, A. Otakhonov and A. Ismatillayev, "Designing an algorithmic formalization of threat actions based on a Functioning table," 2022 International Conference on Information Science and Communications Technologies (ICISCT), Tashkent, Uzbekistan, 2022, pp. 1-5, doi: 10.1109/ICISCT55600.2022.10146987.

27. I. Normatov, I. Yarashov, A. Otakhonov and B. Ergashev, "Construction of reliable well distribution functions based on the principle of invariance for convenient user access control," 2022 International Conference on Information Science and Communications Technologies (ICISCT), Tashkent, Uzbekistan, 2022, pp. 1-5, doi: 10.1109/ICISCT55600.2022.10146952.

28. Бабаджанов А. Ф. и др. Алгоритмический анализ системы защиты информации на основе таблиц функционирования //International Journal of Contemporary Scientific and Technical Research. – 2022. – С. 216-219.

29. Normatov I., Yarashov I., Boboqulov B. Development of models for describing the processing of environmental information in security problems of controlling a protection system based on Petri nets //Central Asian journal of mathematical theory and computer sciences. – 2022. – Т. 3. – №. 12. – С. 229-239.

30. Kabulov A., Yarashov I., Daniyarov B. Systematic analysis of blockchain data storage and sharing technology //Central Asian journal of mathematical theory and computer sciences. – 2022. – Т. 3. – №. 12. – С. 240-247.

31. Jumaniyozov Z. G. et al. Checking the condition of the shutter in the water distribution system using a laser sensor //Science and Education. – 2023. – Т. 4. – №. 6. – С. 430-435.

32. Jumaboyeva A., Yarashov I. Maxsus maktabgacha ta'lim tashkilotlarida nutqida nuqsoni bo'lgan bolalarni axborot texnologiyalari asosida pedagogik metodlar orqali tahlil qilish// O'zbekistonda ilmiy -

amaliy tadqiqotlar mavzusida Respublika 17-ko'p tarmoqli ilmiy masofaviy onlayn konferentsiya.-2020.-C.249-250.

33. Kabulov A.V., Yarashov I.K. Algorithmic model of synthesis and elimination of risks based on Functioning table. Modern problems of applied mathematics and information technologies al-Khwarizmi 2021: abstracts of the international scientific conference. – Fergana. 2021. p.205-206.
34. Kabulov A.V., Yarashov I.K. Algorithmic modeling user access control based on Functioning table. Modern problems of applied mathematics and information technologies al-Khwarizmi 2021: abstracts of the international scientific conference. – Fergana. 2021. p.206-207.
35. Kabulov A.V., Yarashov I.K., Kalandarov I.I., Otakhonov A.A. Algorithmic analysis of a system based on a Functioning table and importance for information security. Modern problems of applied mathematics and information technologies al-Khwarizmi 2021: abstracts of the international scientific conference. – Fergana. 2021. p.207-208.
36. Yarashov I, Normurodov D. “Parol bo‘yicha autentifikasiyalashning asosiy tahdidlari va shaxsiy parolning zaiflik”. Uzliksiz ma'naviy tarbiya kontsepsiyasini amalga oshirishdagi ommaviy axborot vositalarining roli mavzusida Respublika onlayn ilmiy-amaliy konferentsiya, 2020.pp 492-496.
37. Islambek Saymanov, Inomjon Yarashov. “IoT arxitekturasida funksional darajalari tahlili”. Ijtimoiy sohalarni raqamlashtirishda innovasion texnologiyalarning o'rni va ahamiyati Respublika ilmiy-amaliy konferensiya. 2020. Karshi, pp 359-361.
38. Inomjon Yarashov, Normatov Dilmurod. “Kiber fizik tizimlar va Iot tizimlarning qiyosiy tahlili”. Axborot-kommunikasiya texnologiyalari va telekommunikasiyalarning zamonaviy muammolari va yechimlari Respublika ilmiy-texnik konferensiya, 2020 . Fergana, pp 338-340.
39. Sharipova, N. H. (2022). The Ways of Increasing the Efficiency of Cross-Border Payments in Uzbekistan. Central Asian Journal of Innovations on Tourism Management and Finance, 3(6), 40-47.
40. Sharipova, N. (2021). IMPROVING THE ACCESSIBILITY AND QUALITY OF FINANCIAL SERVICES IN UZBEKISTAN. Экономика: анализы и прогнозы, (2), 69-73.