GALAXY INTERNATIONAL INTERDISCIPLINARY RESEARCH JOURNAL (GIIRJ)
ISSN (E): 2347-6915
Vol. 12, Issue 12 December (2024)

DETECTION OF INFORMATION SECURITY VIOLATIONS IN DATA FLOWS IN THE ELECTRONIC GOVERNMENT SYSTEM

Gafurov Sh. R.
Independent Researcher of TUIT named after Muhammad al-Khwarizmi

ABSTRACT

This article presents a proposal for analyzing messages to detect information security violations in the flow of information in the electronic government system, and visualizing messages based on the results of the analysis .

Keywords: E-government, filtering , integrity , analysis , visualization, security monitoring.

INTRODUCTION

In almost all methods used to detect information security breaches in data flows, the efficiency of the method directly depends on the mathematical apparatus chosen in the method. In particular, when considering the data flow in the e-government system, the mathematical expectation and the dispersion of the distribution in adjacent time intervals should not differ significantly, therefore, their estimates should be based on moments specially created for this purpose. can be obtained by interpolation from the values stored in the table. When analyzing scientific works in this area, it was found that the quadratic interpolation method allows obtaining the most accurate values for estimating the mathematical expectation and variance in the data flow in the electronic government system.

Therefore, a time table is used to track the flow of information security messages. In the proposed method, 24 values of mathematical expectation and 24 difference values corresponding to each hour of the day are taken in the time table. Usually, in organizations or in voluntary information systems, the dimensions of the time table are selected based on the flow of information security messages. The values stored in the moment table are dynamically updated using an exponential moving average after receiving the next number of information security messages. The exponential moving average (EMA) method is understood as an exponentially decreasing and moving EM value with the distance of the calculated value from the current observation value. The combination of methods for interpolating moment table values and calculating values using ECHO allows us to assess the nature of the circulation of information security messages and long-term trends [1].

The method for detecting information security violations in the flow of information security messages in the e-government system (proposed method) is implemented in a sequence of four main stages that f_1 are x_t used each time the next number of information security messages is received, which are:

- Interpolate the values in the table of t –moments and estimate the parameters of the normal distribution for the chi time interval ;F_t
- x_t- update the moment table values according to the obtained value;
- - F_t estimate the value by calculating its standard value x_t based on the distribution parameters ;Z_t
- Z_t based on standard price I_x calculate the cost.

- Unlike known approaches, this method allows for dynamic threshold calculations, the number of information security alerts in the adjacent time interval, and the identification of rapid attacks and threats distributed over time at the e-government system level. Therefore, the first step, namely the interpolation of the mathematical expectation and the distribution of the number of information security messages in the table of moments, is carried out in the computational flow. For this, it is necessary to know the mathematical expectation and distribution of the number of messages in order to determine the boundary values of the number of information security messages [2].
- Calculations of the mathematical expectation and distribution of the number of information security messages in an e-government system over a certain period of time can be carried out using several methods. However, the advantage of the proposed approach is that it allows for the identification of deviations based on comparison with reliable data and, therefore, the accuracy of anomaly detection. The main disadvantages of current methods are that, firstly, a large amount of data (estimates of the expected value and variance) must be stored for each time interval, and secondly, this solution cannot be scaled up. From the point of view of the architecture of the e-government system and the monitoring system used in it, changing the time interval requires changing the database structure and the program code that implements similar functions.
- When identifying information security breaches in the information flow of the e-government system, it is necessary to estimate the mathematical expectation and distribution of the number of information security messages in adjacent time intervals based on empirical data, to monitor the changes in these values, and then only to calculate the basic values of the mathematical expectation and the distribution of the number of information security messages for each hour. It is necessary to save. Using mathematical methods, the remaining values of the calculations are calculated. Based on the main task of the monitoring system in the e-government system, the first task is to construct a function that can calculate the mathematical expectation and the distribution of the number of information security messages from certain points with high accuracy. This task is called curve fitting and can be solved using extrapolation or interpolation methods. In the case of e-government data coming from a cyclical cycle, extrapolation methods are not considered to be very suitable for this situation, since to predict an indicator based on them, reliable data on the dynamics of the selected function over at least 5 time intervals is required. The interpolation method is more suitable for solving this problem in the data flow of e-government systems, as it is simpler and more accurate than extrapolation [3].

There are the following possible options for interpolation using polynomials:

- linear;
- quadratic;
- cubic .

The higher the degree of the polynomial , the more accurate the interpolation result will be, but the more difficult it will be to calculate the values. Therefore, it is necessary to choose the simplest interpolation method that provides the required accuracy . For the statistical data collected in the e-government system, the mathematical expectation and distribution of the number of information security messages at each given time interval are calculated, and then

these values are compared with the estimated results obtained by various interpolation methods over the same time interval, taking into account the time intervals. In the proposed method, it is recommended to use the Lagrange interpolation polynomial.

$$L(x) = \sum_{j=0}^{n} y_j l_j(x) \qquad (1)$$

Here the basic polynomials are calculated according to the following formula:

$$l_j(x) = \prod_{i=0, j\neq}^{n} \frac{x - x_i}{x_j - x_i} = \frac{x - x_o}{x_j - x_o} \dots \frac{x - x_{j-1}}{x_j - x_{j-1}} \frac{x - x_{j+1}}{x_j - x_{j+1}} \dots \frac{x - x_n}{x_j - x_n} \qquad (2)$$

The results obtained for estimating the mean and variance are presented in the table below. The tables contain data for 10-minute intervals. Similar results were obtained for the following time intervals: 20 minutes, 5 minutes, 3 minutes, and 1 minute.

Table 1. Comparison of interpolation methods

Naming	Interpolation method	Variation range (message)	Standard deviation (message)	Coefficient of vibration	Linear coefficient of variation
Scorecard. Information security is waiting for the number of messages	Linear	9.2	1.5	0.083	0.013
	Square	5.3	0.7	0.048	0.006
	Cube	5.2	0.7	0.047	0.006
Assessing the change in the number of information security messages.	Linear	1.1	0.2	0.088	0.013
	Square	0.7	0.1	0.052	0.007
	Cube	0.6	0.1	0.047	0.006

CONCLUSION

This method allows to calculate the mathematical expectation and distribution of the number of information security messages in the adjacent time interval by interpolation using key values stored in a specially created table of moments to detect information security violations in the information flow of the electronic government system. As a result, it is possible to optimize the volume of stored information. Interpolation using polynomials is often used today. This is because polynomials are easy to calculate and find their derivatives analytically, and the set of polynomials is densely packed in the space of continuous functions.

REFERENCES

1. Alhitmi HK, Mardiah A, Al-Sulaiti KI, Abbas J (2024) Data security and privacy concerns of AI-driven marketing in the context of economics and business field: an exploration into possible solutions. Cogent Bus Manag 11(1).
2. Fayzullajon, B., Sharifjon, G., Sherzod, S. (2023). Methods for Assessing Information Security Incidents in the Enterprise and Making Decisions. In: Ranganathan, G., Fernando, X., Rocha, Á. (eds) Inventive Communication and Computational Technologies. Lecture

Notes in Networks and Systems, vol 383. Springer, Singapore. https://doi.org/10.1007/978-981-19-4960-9_12.

3. Beckman L., Hultin Rosenberg J., Jebari K. Artificial intelligence and democratic legitimacy. The problem of publicity in public authority //AI & SOCIETY. – 2024. – Т. 39. – №. 3. – С. 975-984.