INVESTIGATION OF THE METHOD OF EVALUATING THE EFFECTIVENESS OF THE INFORMATION SECURITY SYSTEM BASED ON FUZZY INFERENCE

Abstract: As organizations increasingly rely on digital technology to operate, protecting their information and data has become a critical concern. Information security systems are designed to safeguard digital assets against unauthorized access, use, disclosure, disruption, modification, or destruction. However, evaluating the effectiveness of an information security system can be challenging due to the complexity of the system and the diversity of threats it faces. In recent years, researchers have proposed using fuzzy inference to evaluate the effectiveness of information security systems. Fuzzy inference is a mathematical approach that can handle uncertain and imprecise information, making it well-suited for evaluating the effectiveness of information security systems. This research aims to develop a method for evaluating the effectiveness of an information security system based on fuzzy inference. The proposed method uses a set of performance indicators to measure the effectiveness of the system, such as the number of security incidents detected, the response time to security incidents, and the number of false positives and false negatives [1]. These indicators are then combined using fuzzy inference to generate an overall effectiveness score for the system. The proposed method will be evaluated using a real-world case study of an information security system deployed in an organization. The effectiveness score generated by the fuzzy inference method will be compared to the results obtained using traditional evaluation methods, such as the cost-benefit analysis or the return-on-investment analysis. The results of the study will demonstrate the effectiveness and usefulness of the proposed method for evaluating information security systems.

Keywords: information security, audit, fuzzy modeling, cybersecurity, penetration testing
Introduction

Information security has become a vital aspect of our modern society due to the widespread use of technology and the internet. As organizations become more reliant on technology, they must ensure that their information security systems are effective in protecting their sensitive data from cyberattacks. The evaluation of the effectiveness of an information security system is crucial to identifying vulnerabilities and mitigating risks. Traditional evaluation methods rely on numerical measures, but these methods may not provide an accurate representation of the system's effectiveness.

The research topic of this study is the investigation of the method of evaluating the effectiveness of an information security system based on fuzzy inference. Fuzzy logic is a mathematical method that deals with uncertainty and imprecision, making it suitable for evaluating complex systems such as information security. This study aims to investigate the effectiveness of fuzzy inference in evaluating the information security system's performance and identifying potential vulnerabilities.

Hypothesis

The hypothesis of this study is that using fuzzy inference to evaluate the effectiveness of an information security system will provide a more accurate representation of the system's performance and identify potential vulnerabilities that traditional evaluation methods may miss.

Goals

The primary goal of this study is to investigate the effectiveness of fuzzy inference in evaluating the information security system's performance. The study aims to compare the results obtained using fuzzy inference with traditional evaluation methods and determine whether fuzzy inference provides a more accurate representation of the system's performance.

Objectives:

To achieve the goal of this study, the following objectives will be pursued:

1. To review the existing literature on the evaluation of information security systems and fuzzy inference.
2. To develop a methodology for evaluating the effectiveness of an information security system based on fuzzy inference.
3. To apply the developed methodology to a case study and compare the results with traditional evaluation methods.
4. To analyze the results and draw conclusions on the effectiveness of fuzzy inference in evaluating the information security system's performance.

The choice of the fuzzy modeling method in the information security audit in comparison with other traditional methods has several advantages. Firstly, the fuzzy method allows you to represent complex and uncertain data. It is capable of processing vague and uncertain information that traditional methods may not be able to quantify, such as the possibility of a threat, the severity of the threat and the effectiveness of security measures. This allows the auditor to make more informed and accurate decisions [2]. Secondly, the fuzzy method provides a more comprehensive assessment of the information security system. It examines numerous factors that contribute to the overall effectiveness of the system, including the likelihood of a threat, the severity of the threat, the effectiveness of security measures and the potential impact of the threat on information assets. This leads to a more holistic assessment of the system, which allows you to make more effective decisions. Thirdly, the fuzzy method provides greater flexibility and adaptability. Input parameters and output results can be adjusted as needed to reflect changes in the security environment, such as the emergence
of new threats or changes in the effectiveness of security measures. This makes the method more dynamic and responsive to changing circumstances. In general, the fuzzy method in information security audit offers a more accurate, comprehensive, and flexible approach to evaluating the effectiveness of an information security system, which makes it a valuable tool for security professionals and auditors. In the “Tab.1” below you can see its advantages and disadvantages.

<table>
<thead>
<tr>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Can model imprecise and uncertain data</td>
<td>Can be computationally intensive</td>
</tr>
<tr>
<td>Can be used to combine different types of data</td>
<td>Requires domain expertise in fuzzy logic</td>
</tr>
<tr>
<td>Can be used to create rule-based systems</td>
<td>May not be suitable for all types of data</td>
</tr>
<tr>
<td>Can provide more detailed information than binary systems</td>
<td>Interpreting results can be difficult</td>
</tr>
<tr>
<td>Can improve decision-making processes</td>
<td>Results may not be as precise as with other methods</td>
</tr>
</tbody>
</table>

It’s worth noting that while there are some disadvantages to using fuzzy inference in information security audit, the advantages often outweigh them, particularly when dealing with imprecise and uncertain data.

**Evaluating Security System Effectiveness through Fuzzy Modeling Techniques**

Fuzzy modeling can be used to assess the effectiveness of a security system by considering multiple input variables and their degrees of membership in different linguistic terms. The fuzzy model can be developed based on the available data on security incidents, the security systems feature, and the feedback from security personnel. The output of the fuzzy model can provide a quantitative assessment of the security system’s effectiveness, which can be used to identify the system’s strengths and weaknesses and to optimize the security measures [3].

The input data for fuzzy inference in security system evaluation can include various types of data related to the state of the control object, such as data on security incidents, system features, and feedback from security personnel. The input data can also include data on impacts outside the control object, such as data on the environment and potential threats. These input variables are fuzzified, which means they are transformed into linguistic terms that represent their degree of membership in a particular category. The next stages involve forming a rule base that describes how the input variables are related to the output variable, aggregating sub-conditions based on the fuzzy logic rules, activating sub-conclusions, and accumulating conclusions. Finally, the output variable is defuzzified to obtain a crisp value that can be used to evaluate the security system’s effectiveness [4]. In addition to the input data, other factors such as control variables, effects and control mechanisms, controls, and other security-related factors may also be considered in developing the fuzzy model. In “Fig. 1”, you can see the process of fuzzy inference.
The main element in this process are fuzzy rules. As a rule, there can be a lot of large fuzzy rules for the output/input, for example, "IF-THEN":

- rule 1: IF \( y_1 = a_{11}, \ AND \ y_2 = a_{12} \ AND \ ... \ AND \ y_i = a_{1j} \) THEN \( x = B_1 \);
- rule 2: IF \( y_1 = a_{21}, \ AND \ y_2 = a_{22} \ AND \ ... \ AND \ y_n = a_{i2} \) THEN \( x = B_1 \);
- rule i: IF \( y_1 = a_{1i}, \ AND \ y_2 = a_{2i} \ AND \ ... \ AND \ y_n = a_{ni} \) THEN \( x = B_i \);

where \( y_1, y_2, ..., y_n \) are input variables;
\( x \) is the output variable;
\( a_{ij} \) – fuzzy areas of definition of input variables defined on universal sets \( y_1, y_2, ..., y_n \);
\( B_i \) – output linguistic variable.

Each fuzzy set corresponds to a membership function \( \mu(a_{ij})x_j \).

Each rule consists of conditional and final parts.

The antecedent of a conditional statement contains the input variables, while the consequent provides the output variable values. To represent expert knowledge in a formal way, a set of rules has been developed for fuzzy inference [5]. These rules are called fuzzy production rules and contain formal knowledge about how to manage objects and their characteristics in various conditions. The rules can be combined with simple operators "AND" and "OR" to form compound and simple statements. Each statement must have redefined membership functions for each linguistic variable's term set. Fuzziness is introduced by correlating the values of term functions and numerical values of input variables. Fuzzification is the process of mapping input values to corresponding linguistic terms for fuzzy inference systems. Fuzzy kernel conditions are also used to determine the degrees of truth for all logic statements in the antecedents of fuzzy rules. The probability function used for this purpose may be arbitrary and based on various assumptions about the system's properties, considering any existing uncertainty.

Aggregation is the process that determines the degree of truth of the fuzzy inference system's position. It uses the function values obtained at the fuzzification stage. For a simple fuzzy statement, the degree of truth corresponds to the validity function's values. When the consequent of a fuzzy production rule is a fuzzy statement, the degree of truth is equal to the algebraic product. When concluding a compound statement, the degree of truth is equal to the algebraic product of the weighting coefficient and the degree of truth of the antecedent of the fuzzy production rule. Weights and values are equal to one by default at the rule base formation stage [6].

Figure 1. Process of fuzzy inference
In the fuzzy inference system, accumulation is the process of finding the membership function for each output linguistic variable. The purpose is to combine all degrees of truth to identify the membership function of each output variable. This process connects fuzzy sets of all the conclusions in the fuzzy rule base. Defuzzification is the transition from the membership function to a clear value, which uses the results of the accumulation of all the output linguistic variables to obtain quantitative values [7].

The theoretical procedures for fuzzy sets are similar to finding the characteristics of the position of random variables in probability theory. To simplify the procedure, a clear number corresponding to the maximum function can be selected, but this method is limited to single-extremal functions. For multiextremal, various defuzzification methods exist, such as the center of gravity, the center of the maxima, the highest of the maxima, the name of the highest, and the median. To obtain the output quantitative parameters, the Mamdani algorithm can be used in defuzzification, which gives interpretable results and can be used with numerical data to apply flexible capabilities [8].

The creation of rules for fuzzy inference follows a structured and coordinated process that involves developing a list of fuzzy production rules in the format of "If...Then...". The antecedents of the fuzzy production rules are formed by combining logical sets with "And", while the consequents of the rules are kept simple [9].

- Fuzzification. Determination of the degree of activation (truth) of each premise of each rule for the given inputs \( \mu_{A_{ij}}(x'_j) \).
- Aggregation of the degrees of truth of the premises for each of the rules using the min operator min.

\[
\begin{align*}
\alpha_1 &= \min\{\mu_{A_{11}}(x'_1), \mu_{A_{12}}(x'_2), \ldots, \mu_{A_{1n}}(x'_j)\}, \\
\alpha_2 &= \min\{\mu_{A_{21}}(x'_1), \mu_{A_{22}}(x'_2), \ldots, \mu_{A_{2n}}(x'_j)\}, \\
& \vdots \\
\alpha_i &= \min\{\mu_{A_{ij}}(x'_1), \mu_{A_{il}}(x'_2), \ldots, \mu_{A_{ij}}(x'_j)\}.
\end{align*}
\]  

The activation of the sub-conclusions of fuzzy production rules is carried out using the method of min-activation.

\[
\begin{align*}
\mu_{B_{1'}}(y) &= \min(\alpha_1, \mu_{B_1}(y)), \\
\mu_{B_{2'}}(y) &= \min(\alpha_2, \mu_{B_2}(y)), \\
& \vdots \\
\mu_{B_{l'}}(y) &= \min(\alpha_l, \mu_{B_l}(y)).
\end{align*}
\]

The process of gathering sub-conclusions of fuzzy production rules involves using the traditional method of max-union of membership functions, which is a common practice in fuzzy logic \( \mu_{B'}(y) = \max\{\mu_{B_{1'}}(y), \mu_{B_{2'}}(y), \ldots, \mu_{B_{l'}}(y)\} \).

Defuzzification is carried out by the method of center of gravity according to the formula.

\[
Y' = \frac{\int_{Y_{min}}^{Y_{max}} y \mu_{B'}(y) dy}{\int_{Y_{min}}^{Y_{max}} \mu_{B'}(y) du}
\]  

The formula calculates the center of gravity of a flat figure with the boundaries defined by the axes of coordinates and the membership function graph of the fuzzy set. It uses the boundaries of the interval support of the output variable \( y \), represented by \( Y_{max} \) and \( Y_{min} \).
A fuzzy inference method for gauging effectiveness

The developed method is based on the approach presented in [9]. By calculating the probability of threat occurrence and implementing countermeasures to address threats to information security, a quantitative assessment of the security system’s effectiveness can be determined. However, due to the subjective nature of the evaluation, fuzzy modeling and fuzzy logic are utilized to provide a quantitative assessment based on qualitative linguistic variables. To present the information, a system of reference fuzzy statements is used to establish a relationship between the values of fuzzy input and output parameters [10], [11], [12]. Linguistic variables are used to describe the input and output parameters, such as “Probability of threat”, “Correlation of measures” and “Effectiveness of the security system”. The assessment of the security system’s effectiveness requires evaluating its productivity in relation to each of the current threats and the quality assessment is based on the adequacy of the measures taken to compensate for security threats. Both input and output variables are described in a formalized form as linguistic variables.

In the form of a linguistic variable, the input and output variables “Compliance of measures”, “Probability of threat”, “Effectiveness of the security system” will be described.

Let’s introduce linguistic variables.

1. $\beta_x$ – Probability of a threat the probability of a threat with a definition area $X = [0, 100]$, and a set of base values $T_x = \{a_{x1}, a_{x2}, a_{x3}, a_{x4}, a_{x5}\}$ (very low, low, medium, high, very high).

2. $\beta_y$ – “compliance of measures” (compliance of information security threat compensation measures) with the definition area $Y = [0, 100]$, and the set of basic values $T_y = \{a_1, a_2, a_3, a_4, a_5\}$ (practically absent, small, moderate, high, very high).

3. $\beta_z$ – the effectiveness of an information security system (Evaluation of the effectiveness of information security) with a definition area of $Z = [0, 100]$, and a set of basic values of $T_z = \{a_{z1}, a_{z2}, a_{z3}, a_{z4}, a_{z5}\}$ (not at all effective, insufficiently effective, moderately effective, effective, very effective).

This passage discusses the process of using input parameters $\beta_x$ and $\beta_y$ to generate an output parameter $\beta_z$, and the need to create fuzzy statements to describe the relationship between them. These fuzzy statements will be used to generate fuzzy conclusions, which will be used to form fuzzy rules. To create effective fuzzy rules, certain requirements must be met [13] [14].

The presented evaluation of security system effectiveness involves fuzzy rules which must follow certain criteria. One important criterion is that every linguistic term of the output variable should have at least one rule assigned to it. Additionally, every term of the input variable should also have at least one corresponding rule where it’s used as a prerequisite.

1. If the likelihood of a threat is extremely low and there is a high degree of compliance with measures.
2. Alternatively, if the likelihood of a threat is extremely low and there is a moderate degree of compliance with measures.
3. Or, if the likelihood of a threat is low and there is a high degree of compliance with measures.
4. Or, if the likelihood of a threat is low and there is a moderate degree of compliance with measures.
5. Or, if the likelihood of a threat is average and there is a high degree of compliance with measures.
6. Or, if the likelihood of a threat is high and there is a high degree of compliance with measures.
7. The effectiveness of the IB is very high.
8. If the likelihood of a threat is extremely low and there is a moderate degree of compliance with measures, or the likelihood of a threat is extremely low and there is a low degree of compliance with measures, or the likelihood of a threat is low and there is a moderate degree of compliance with measures.
9. Alternatively, if the likelihood of a threat is average and there is a high degree of compliance with measures.
10. Or, if the likelihood of a threat is high and there is a high degree of compliance with measures.
11. Or, if the likelihood of a threat is very high and there is a very high degree of compliance with measures.
12. Or, if the likelihood of a threat is very high and there is a high degree of compliance with measures.
13. The IB is effective.
14. If the likelihood of a threat is extremely small and measures are not being observed.
15. Alternatively, if the likelihood of a threat is low and there is a low degree of compliance with measures.
16. Either the likelihood of a threat is average and there is a moderate degree of compliance with measures, or the likelihood of a threat is average and there is a low degree of compliance with measures.
17. If the likelihood of a threat is high and there is a moderate degree of compliance with measures, or the likelihood of a threat is very high and there is a moderate degree of compliance with measures, then the IB is moderately effective.
18. If the likelihood of a threat is low and measures are practically nonexistent.
19. Or, if the likelihood of a threat is average and compliance with measures is practically absent.
20. Or, if the likelihood of a threat is high and there is a low degree of compliance with measures.
21. Or, if the likelihood of a threat is very high and there is a low degree of compliance with measures.
22. The effectiveness of the IB is inadequate.
23. If the likelihood of a threat is high and measures are practically nonexistent.
24. Or, if the likelihood of a threat is very high and measures are practically not observed, then the IB is entirely ineffective.

The linguistic values of variables are set along the X and Y axes, the intersection is the values of output variables.

Rationale for selecting the research methodology and instruments.

The efficiency positioning matrix in the table below demonstrates how the input parameter relates to the output. Linguistic variable input values are shown both horizontally and vertically, while the values of the output variable are located at the intersections of the inputs [15] [16].
During the study, it was determined that trapezoidal functions would be utilized as membership functions. These functions are utilized to characterize uncertainties such as “approximately equal,” “average value,” “similar to an object,” and “similar to an object.” The trapezoidal membership function is defined by four parameters \((u_0, u_1, u_2, u_3)\), which are determined through the expert method. Graphs of membership functions are then created based on these parameters [17].

After examining the available approaches for assessing information security system effectiveness, a method using Mamdani fuzzy inference systems was developed. Due to the subjective nature of the assessment, fuzzy logic was employed, which provides a quantitative evaluation based on qualitative linguistic variables. Trapezoidal functions were chosen as membership functions during the study. The methodology proposed based on input parameters and results enables the auditor to make informed decisions regarding security system upgrades and the implementation of necessary protective measures [18].

Now we will proceed directly to the analysis of the results of the experiment. During the research, trapezoidal functions will be utilized as membership functions to define uncertainties such as “approximately equal,” “average value,” “similar to an object,” etc. These membership functions are described by four parameters \((u_0, u_1, u_2, u_3)\), which are established through an expert method. The membership function graphs are created using the formula below.

\[
\mu(x) = \begin{cases} 
\frac{x-u_0}{u_1-u_0}, & \text{if } u_0 \leq x \leq u_1 \\
1, & \text{if } u_1 \leq x \leq u_2 \\
\frac{u_3-x}{u_3-u_2}, & \text{if } u_2 \leq x \leq u_3 \\
0, & \text{in other cases}
\end{cases}
\]  

(4)

Figures 2, 3, 4 show graphs of membership functions for the linguistic variables “Compliance of measures”, “Probability of threat”, “Effectiveness of the security system”.

<table>
<thead>
<tr>
<th>(\alpha_{x5})</th>
<th>(\alpha_{y5})</th>
<th>(\alpha_{y4})</th>
<th>(\alpha_{y3})</th>
<th>(\alpha_{y2})</th>
<th>(\alpha_{y1})</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\alpha_{x4})</td>
<td>(\alpha_{z4})</td>
<td>(\alpha_{z4})</td>
<td>(\alpha_{z3})</td>
<td>(\alpha_{z2})</td>
<td>(\alpha_{z1})</td>
</tr>
<tr>
<td>(\alpha_{x3})</td>
<td>(\alpha_{z5})</td>
<td>(\alpha_{z4})</td>
<td>(\alpha_{z3})</td>
<td>(\alpha_{z2})</td>
<td>(\alpha_{z1})</td>
</tr>
<tr>
<td>(\alpha_{x2})</td>
<td>(\alpha_{z5})</td>
<td>(\alpha_{z5})</td>
<td>(\alpha_{z4})</td>
<td>(\alpha_{z3})</td>
<td>(\alpha_{z2})</td>
</tr>
<tr>
<td>(\alpha_{x1})</td>
<td>(\alpha_{z5})</td>
<td>(\alpha_{z5})</td>
<td>(\alpha_{z4})</td>
<td>(\alpha_{z3})</td>
<td>(\alpha_{z2})</td>
</tr>
</tbody>
</table>
The shape of the graphs of membership functions is chosen trapezoidal, as indicated earlier. The graphs are based on the opinions of experts.

Graphs of membership functions together with the rules form a knowledge base based on the opinion of experts.
By utilizing the fuzzy Mom-tribute output method, we can generate numerical values for output parameters based on input parameters and a knowledge base. The proposed methodology was practically executed using fuzzy logic fuzzy modeling in the MATLAB software environment, as described in detail in [19]. To execute the method, we established fuzzy rules and created trapezoidal membership functions for input and output linguistic variables. We also configured fuzzy output using the Mamdani algorithm and defuzzification through the center of gravity method. The main outcomes of the collection efficiency assessment method implementation are illustrated in Figures 5 and 6.
A model was created in the MATLAB environment to implement the proposed method, enabling the determination of a numerical evaluation of the security system’s effectiveness.

**Conclusion**

During the theoretical studies conducted, a more robust method of auditing information security has been justified and selected. This method enables the system to be modernized by quantifying the effectiveness of the security measures in countering security threats. The current methods for assessing the likelihood of security threats do not consider several factors, including destructive actions from the implementation of threats against information assets. Additionally, there is insufficient attention during audits paid to quantifying the effectiveness of the security system against current security threats.

To address these shortcomings, a methodology for information security audit has been developed based on a model of the audit process that considers the quantitative assessment of the security system’s effectiveness using Mamdani fuzzy inference systems. The implementation of this methodology is carried out in the FUZZY LOGIC package in the MATLAB environment. Based on input parameters and obtained results, the proposed methodology enables the auditor to make informed decisions regarding the need to modernize the security system and implement the necessary protection measures.

**References**