Hindawi
Scientific Programming
Volume 2022, Article ID 7300977, 13 pages
https://doi.org/10.1155/2022/7300977

Research Article

The Construction Method of Computer Network Security Defense System Based on Multisource Big Data

Jie Ma[1] **and Shuanbao Li**[2]

[1]*School of Computer and Information Engineering, Henan Finance University, Zhengzhou 450045, China*
[2]*School of Artificial Intelligence, Henan Finance University, Zhengzhou 450045, China*

Correspondence should be addressed to Jie Ma; majie@hafu.edu.cn

Received 4 March 2022; Revised 1 April 2022; Accepted 6 April 2022; Published 12 May 2022

Academic Editor: Sheng Bin

This article proposes a computer network security defense system model based on multisource big data, which makes the security of computer networks more complete because of the shortage of current computer network security defense systems. The current network security situation is analyzed, and the problems faced in the field of network security are pointed out; then, multisource big data is introduced, and the characteristics of the multisource big data model are analyzed. This article proposes an information system network security model, which formally describes the relationship between network behavior among nodes, security threats, network attacks, and defense capability of security devices in an information system. After that, a network security defense system measurement scheme is constructed based on hierarchical analysis, and quantitative indicators of defense effectiveness are defined in two dimensions: risk severity of security threats and defense response actions of security devices, which enables quantitative evaluation of the security defense system, and then identifies omissions and defects of the network security defense system. A defense system measurement and optimization system is designed based on the proposed network security system measurement and optimization scheme. The application analysis of the project is combined with the actual scenarios. Through the evaluation and optimization of the security defense system in the natural methods, it is proved that the network security defense system evaluation and optimization scheme proposed in this article has reasonableness and effectiveness in the actual application.

1. Introduction

Since the development of computer networks, the Internet has rapidly spread to all corners of life; e-commerce, social networking, and other services have become part of life. In the era of cloud computing and mobile Internet, we have long been inseparable from the convenience brought by computer networks. By analyzing and integrating data from different sources, multisource big data can yield more comprehensive and accurate conclusions than traditional data analysis, and thus has been effectively applied in many fields such as intelligence analysis, image fusion, cyber security, and information retrieval. Still, at the same time, many network security problems also came. In this context, the demand for people who can master computer networks increases day by day. We hope to use computer networks to bring us the most excellent convenience and avoid the security risks it gets as much as possible [1]. In the current network environment, hacking behavior is more hidden and systematic. Network attacks have developed in the direction of complexity, customization, multistep, etc., and targeted cyber-attacks (Targeted Cyber Attack) represented by APT have brought substantial economic losses and data leakage, and other network security threats to governments, enterprises, and institutions, especially with the smart grid, industrial Internet, etc. With the development of critical infrastructure applications, various traditional single detection technologies are highly systematic in the overall discovery, and the targeted malicious behavior is no longer ideal and practical. Targeted cyber-attacks have become the most critical factor affecting current network security [2].

With the continuous development of social informatization, many new industries such as Internet finance, online shopping, and electronic banking have emerged. This change

has changed people's traditional way of life and made them more and more dependent on the Internet [3]. Computer networks bring great convenience to people's lives and improve office efficiency; at the same time, because of their inherent characteristics of sharing and openness. Computer networks are exposed to many security threats, including malicious codes, hacker attacks, software vulnerabilities, and protocol vulnerabilities,, which may cause different degrees of damage to computers. Computer network security vulnerabilities are time-sensitive, and as long as the computer is running, new vulnerabilities will be created continuously. A symmetric encryption algorithm uses the same key for encryption and decryption operations, which is fast and suitable for encrypting a large amount of data, but there is a risk of being cracked; an asymmetric encryption algorithm uses a public key for encryption and a private key for decryption, which is slow and highly secure. In this context, this article focuses on studying active defense-based network security technology; that is, honeypot technology is an emerging dynamic network security defense technology, mainly by deploying some resources in the network to the attacker [4]. These arranged resources seem to have some value but are monitored by the system. When hackers interact with these "valuable" resources, the system collects attack information from the attacker and analyzes and processes the attacker's behavior to understand the attacker's attack tools and methods and promptly deduce the attacker's motive and purpose. on the results provided by the honeypot system, network administrators can effectively prevent and avoid attacks and avoid further attacks on the natural method by the attackers.

Multisource big data can obtain more comprehensive and accurate conclusions than traditional data analysis by analyzing and integrating different sources [5]. It has been effectively applied in many fields such as intelligence analysis, image fusion, network security, and information retrieval. In image retrieval, the fusion of image-related information can obtain more image annotation information and increase the richness of image annotation [6]. In contrast, the fusion of image annotation information and the image itself can improve the accuracy of image annotation. The development of artificial intelligence has driven the rapid growth of image recognition. From the tentative appearance of image recognition using deep learning in the early ImageNet competition to the most advanced image recognition models using such technology, deep learning technology continues to refresh the accuracy of machine recognition of images. High-level semantic annotation of images using deep learning can further improve the effectiveness of image retrieval [7]. It is of great significance and application value how to integrate deep knowledge, data fusion, digital image processing, and other techniques to perform rich high-level semantic annotation on images and use rich image annotation to make image retrieval results more complete to enhance the user's retrieval experience.

2. Related Works

Since 2013, the number of security vulnerabilities included in the information security vulnerability sharing platform has increased each year, with an average growth rate of 21.6% [8]. Computer network security vulnerabilities are time-sensitive, and if computers are running, new vulnerabilities will be created continuously. The overall effectiveness of IP randomization is being studied, but no previous research has thoroughly examined its usefulness and limitations in terms of unpredictability, adaptability, and deployment. In it, researchers propose a technique that allows host IP addresses to be randomized, thereby disrupting both external and internal scanning. A symmetric encryption algorithm uses the same key to encrypt and decrypt operations, which is fast and suitable for encrypting a large amount of data, but there is a risk of being cracked; an asymmetric encryption algorithm uses the public key to encrypt and the private key to decrypt, which is slow and highly secure. The use of a single encryption algorithm always has certain limitations [9]. Hu et al. improved the Vigenere encryption algorithm and combined it with the basic algorithm to form a hybrid encryption algorithm, which significantly improved the security of the encryption algorithm. Intrusion detection technology is real-time monitoring of network behavior. The data collected are analyzed by misuse detection technology and anomaly detection technology to see whether the network behavior violates the predefined security policy without affecting the network performance [10]. Snort is a lightweight network intrusion detection system for detecting various attacks and responding to untrusted events. Harsha et al. propose a design idea of combining Snort with NTOP and implementing and verifying the system [11]. Finally, experiments show that by combining Snort and NTOP, detection can be performed more effectively. In the face of the changing work environment, researchers also focus on intelligent intrusion detection techniques. Dong et al. use machine learning in network intrusion detection and genetic algorithms and decision trees to generate rules for network behavior detection [12].

Recent work has provided a general theory of moving target defense using IP address randomization. Among them, the SDNA architecture inserts a hypervisor between each network node so that the host appearance is dynamic for observers and transparent to the operating system, applications, and end-users. The main drawback of the SDNA architecture is the overhead of reconfiguring the hypervisor for each network host [13]. Some scholars have also investigated the use of IP randomization to resist honeypot mapping attacks, which aim to "live" and monitor IPs and block the IPs detected to be under attack. Mahdianpari et al. apply network layer moving target defense to IPv6 and demonstrate that finding a specific host with a change in IPv6 address is a complex problem using standard network scanning procedures [14]. However, none of these techniques provide a transparent scheme that can be carried out for addressing randomization. Do not exploit the full potential of IP address randomization to achieve high unpredictability, do not provide adversary awareness, minimize the overhead required reconfiguration, and have low address translation rates. In addition, while the overall effectiveness of IP randomization is studied, no previous

study has thoroughly examined its usefulness and limitations in terms of unpredictability, adaptability, and deployment. Researchers propose a technique in which host IP addresses can be randomized, thereby disrupting both external and internal scanning. The method is transparent to end hosts, but the randomization is uniform, does not consider adversary behavior, and has no practical implementation.

The purpose of building a security defense system is to reduce or eliminate the security threats that exist in information systems and analyzing threats can help improve system security and reduce the risks faced by information systems. Some scholars have carried out research work related to threat modeling. Sanden and Neideck divided the threat model into four parts: component, threat, impact, and security control [15]. The piece is the target to which the threat model is oriented; for example, to build a threat model for a server, component means server; the threat is the type of threat to which the component is exposed; impact is the hazard to the component from a specific threat, and security control is the mitigation measure for a particular type of threat. The model extends the pure threat model by introducing the concept of defense. Most studies also focus on decomposing targets in specific application scenarios to determine security threat types and define them as threat models. Zhang et al. divide intelligent home networks into sensing, network, and application and service layers and step by step determine the possible risks faced by each layer as a risk model in intelligent home network scenarios. Su et al. also use scenario decomposition to divide the IoT network into an end-node, network, service, and application layer for threat modeling [16]. This threat modeling approach requires different decomposition levels of the target in different scenarios, and the threat model is no longer valid and general after the scenario transformation.

3. Design of a Computer Network Security Defense System Based on Multisource Big Data

3.1. Multisource Big Data Model Building. The main objective of the data preparation platform for multisource big data is to give users knowledge about the structure of data and other aspects through a unified view and then improve the data preparation process flexibility by increasing the interactivity of the data preparation process multisource big data. To achieve the goal of establishing a unified view, it needs to obtain information about the structure of data sources and target views; to achieve the purpose of improving the flexibility of the data preparation process, it needs to provide an editing page where users can select data preparation process components and edit component attributes in the data preparation process editing interface, and then build a data preparation process for multisource big data, and provide process file upload and generation functions [17]. After the procedure is made, the process needs to be mapped and modeled, and then, it can be used to execute the user-defined strategy. When running the user-defined data preparation process, the implementation of each step in the process should be efficient and flexible. It is necessary to optimize the missing data processing algorithm and method. The platform functional analysis system process is shown in Figure 1.

The data source may be a particular or multisource big data in the data preparation phase. When the data source is multisource big data, it is necessary to connect it to the system and establish a unified view of multisource big data. There are three main functions to develop the unified theory of multisource big data. The first is to select the data source, the second is to establish the unified view, and the third is to update the unified theory.

(1) Data source selection: by selecting different data sources to obtain the corresponding data, the system will record the corresponding data table name and table metadata information. The table in the same data source lights a specific correlation, to ensure that there may also be different data sources between tables with a specific correlation. According to the table fields and other information to analyze the relationship between the tables, with the help of the unified view algorithm, the relationship between the tables contained in the data source is displayed so that users have a more intuitive understanding of the data.

(2) Unified view establishment: according to the relationship between the data source tables, establish the unified view of the accessed data sources. When the selection of data sources is completed, you can analyze these data sources and show a unified view of these significant data sources.

(3) Unified view update: the data source in the access system to the next time before re-execution may change, resulting in the initial establishment of a unified view of the data, and the actual situation is different, so the need for active or passive view update before re-execution. A database view is a virtual table formed by a series of association, selection, and other operations in the target database of interest to the user. The database does not store the corresponding data, only the definition of the view. The database system will automatically query the corresponding basic table when operating on the data in the view, thus improving the flexibility of the database application and generating a data model that better fits the user's needs than the primary table. Most of the mainstream database systems support the creation of views based on this database engine. Still, different database systems do not have the same degree of support for view creation. Each database system only supports the creation of views based on this system, which is not universal, so when the need to create views of basic tables in different databases, or even with nondatabase table data, the current database system cannot help. Therefore, a new way of thinking is needed to build a unified view based on multiple sources of big data [18]. The main

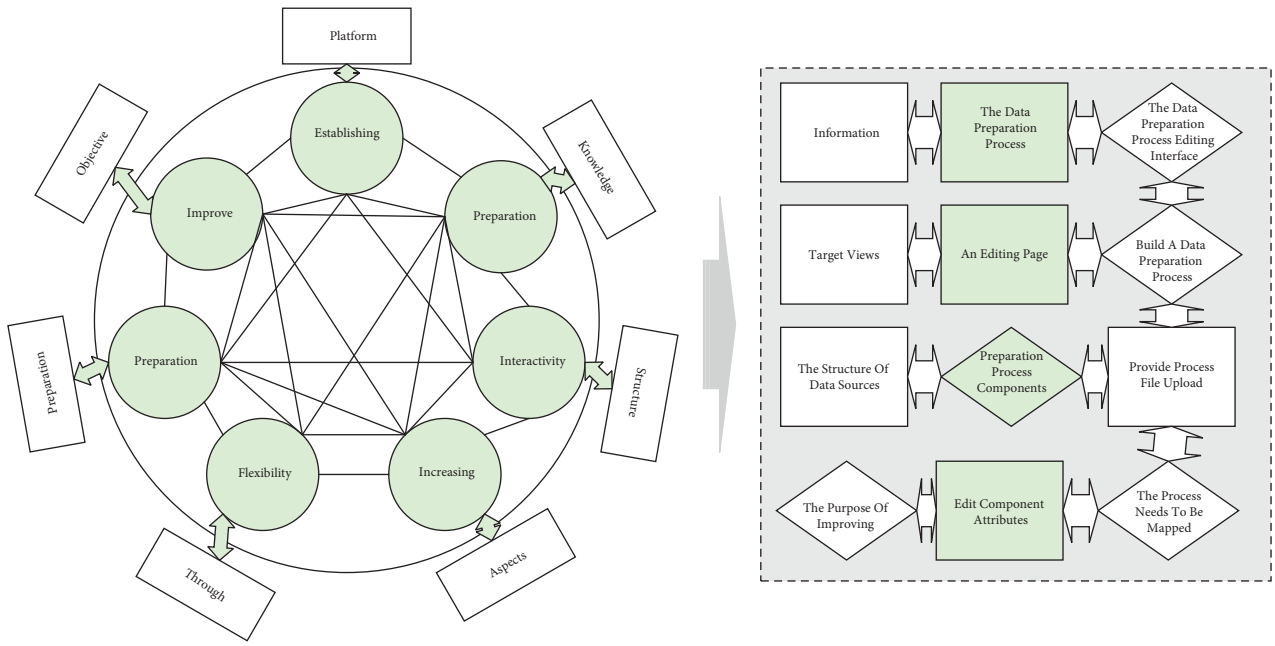

Figure 1: Platform functional analysis system process.

problem of establishing a unified view is that data preparation data may come from the same database or different heterogeneous databases. There are many data redundancies in these databases, which describe the same things but lead to significant differences in the design of the data tables due to different business requirements, causing inconvenience to users who want to obtain unified data. The flowchart of multisource big data is shown in Figure 2. Further aggregated analysis of multisource security events generates anomaly alerts with security semantics; anomaly alerts contain only low-level security information, and there are false alarms that will mislead the attack analysis process; other semantic security data generated by security devices (e.g., IDS, antivirus software, and firewalls) also provide attack-related information, and correlation analysis will connect these dots into a more significant attack scenario.

The research on data integration and unified logical view in the heterogeneous environment has produced many architectures and technical solutions, such as the earlier Mediator/Wrapper architecture, XML (Extensible Markup Language)-based middleware method, improved ontology and XML-based middleware method, and other solutions. In the Mediator/Wrapper architecture by deploying a wrapper Wrapper for the database, a mediator on the wrapper is deployed to build a star structure, and the mediator is used as an interface to access different databases. And playing a great help in solving the interaction of heterogeneous data resources, but providing no relevant solution to the problem of semantic differences; while the middleware method based on ontology and XML, using ontology as a data source conceptual model explicit the ontology-based and XML-based middleware approaches use ontologies as a precise formal specification of the conceptual model of data sources, which helps to solve the problem of semantic differences between heterogeneous data sources. The single ontology approach uses a global ontology. All data sources are associated and mapped, which can visually reflect how data are obtained from data sources in the mapping relationship. The advantage of the multi-ontology approach, where each data source has a local ontology that describes the information of this data source and does not need to be related to other data sources, is that it has a high degree of modularity and good scalability. Still, it is difficult to interact with other data sources due to the absence of a global ontology. In the ontology-based data integration approach, ontologies are used in three ways: single, multi-ontology, and hybrid. All data sources are associated and mapped, which can intuitively reflect how to obtain data from data sources in the mapping relationship. The advantage of the multi-ontology approach, where each data source has a local ontology that describes the information of this data source and does not need to be related to other data sources, is that it has a high degree of modularity and good scalability. Still, it is difficult to interact with other data sources due to the absence of a global ontology.

3.2. Computer Network Security Defines System Model Design. The quantitative assessment of defense systems in the design phase is to help security staff understand the defense coverage of security threats at the time of design of the defense system, that is, which threats the defense system is designed to defend against, and to identify the shortcomings of the defense system, that is,, which threats the defense system is designed to lack the ability to protect against.

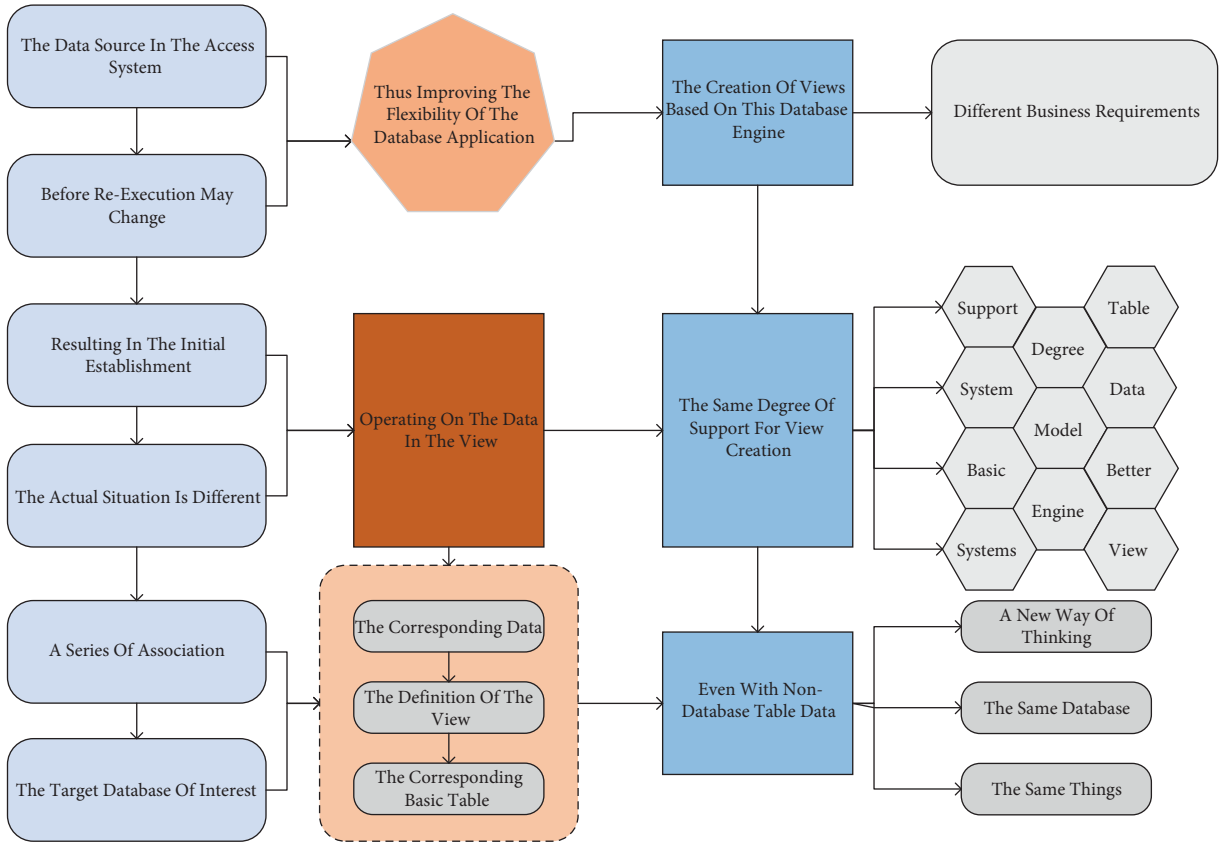

Figure 2: Multisource big data flowchart.

(1) Establish a network security model for the measurement target and determine the security threats. According to the proposed information system network security model, the network topology information of the information system is modeled. The security threats each device node faces in the information system that must be protected are determined.

(2) The risk severity of security threats is used as an evaluation factor in the design phase. The hierarchical analysis method is applied to quantify the risk severity of security threats. Security threats are potential factors that may cause damage to information systems. When the threats are transformed into attacks that pose risks to information systems, security staff defends against attacks by deploying security devices to reduce or eliminate the risks. Therefore, the severity of the risk faced by the information system can be used as a quantitative indicator to measure the defense capability of the defense system [19]. The impact of threat type on risk severity can be viewed as a decision problem, so this article applies hierarchical analysis to quantify the risk severity of threats. Attackers translate threats into specific attacks and achieve attack goals through attack behaviors to bring risks to target assets. Security risks include privilege elevation, data leakage, information collection, data tampering, and service unavailability. Different threats generate different security risks and possess different risk severity for a security threat T at a device node in an information system. The SDNA architecture inserts a hypervisor between each network node, making the host appearance dynamic for observers while transparent to the operating system, applications, and end-users:

$$\sum_{i=1} tw(n-1) = \frac{\sqrt{n-1}}{n+1}. \quad (1)$$

(3) Use the network security model to determine the defense of security devices in the information system against threats to the device nodes. Assuming that the device node is device_1 there is a security threat T, the network threat of device_1 is addr_1, the attacker's network address is, addr_2, the defense result of the security device in the defense system, device_j against the threat T is denoted by $\text{DT}_j = (d_{j1}, d_{j2}, \ldots d_{jn})$, d_{ji} with t_i one-to-one correspondence. While using computer information systems, users will operate them for a long time and frequently. In computer systems, managers often start with an exemplary configuration of the system's security, which protects the computer information system:

$$S = \sum_{i=1}^{n} (r_i + tw_{i-1}) - \mathrm{TW}^T. \quad (2)$$

(4) Calculate the overall defense effectiveness of the defense system. Assuming that there are m assets in the information system, the actual economic value of the assets is used to measure their importance $\mathrm{AW} = (aw_1, \ldots aw_2)$ as the asset importance weight, where $aw_1 : aw_2 : \ldots : aw_m = v_1 : v_2 : \ldots : v_m$, v is the true monetary value of each purchase, where $\sum_{j=1} aw_j$, and finally, the formula for calculating the defense effect of the defense system on the whole information system is obtained:

$$S_1 = \sum_{j=1}^{m} aw_j + \sum_{i=1} tw_i. \quad (3)$$

For the defense results R of the defense system against the threats faced by a single asset, if $r = 0$, it means that the defense system is designed to lack the defense range to reach the support and have the security equipment to defend against the threats, and the security staff needs to add the related security equipment at the appropriate location in the network to protect the asset from the risks t_i brought by the threats. By analogy, by analyzing the defense results of each purchase in the information system, the security shortcomings of the entire network security defense system can be obtained, guiding the security staff to improve the defense system in a targeted manner. Building a security defense system aims to reduce or eliminate security threats to information systems. Analyzing the threats helps improve system security and reduce the risk to information systems. This article categorizes security risks into permission elevation, data leakage, information collection, data tampering, and service unavailability 5 categories. With the development of network security technology, differences in specific application scenarios, and different criteria for defining security risks, the security risk classification in this scheme can be adjusted as needed to suit actual needs. After the data flow reaches the destination address addr_1, the defense results of all the entire defense systems against threats $t_i \mathrm{device}_1$ on the assets can be obtained, and the calculation formula is as follows:

$$S_{t_1} = \sum_{k=1} re_k \cdot rw. \quad (4)$$

After there is a threat to get the defense system on the T defense results $R = (r_1, \ldots, r_n)$, the defense system's effectiveness is calculated by the following formula:

$$S_{t-1} = \frac{\sqrt{r_i \cdot tw_i}}{R \cdot tw^t}. \quad (5)$$

The security defense system can identify data sources, detect them, and make appropriate processing based on data matching results. The design of the security defense system is shown in Figure 3. The defense mechanism consists of a training module, a detection module, a verification module, and a database storage system. In the training module, the system will first collect a large amount of data $\{T_1, T_2, \ldots, T_N\}$ and preprocess the data, put the preprocessed data into the training set database, then use the neural net to train the preprocessed training set, and put the introduced features into the feature library.

Finally, two variables in the scenario are analyzed for their practical significance. The first one is RE the complete response result of the defense system against the threats t_i that device_1 exists in the assets; $re_1 = 0$ indicates that the defense range in the defense system can cover device_1. None of the security devices can detect the danger t_i corresponding attacks; it means that none of the security devices that the defense system covered device_1 can block the attacks corresponding to the threats; it means that none of t_i, $re_3 = 0$ the security devices that the defense system can the attacks t_i corresponding to the threats. If all four values device_1 are present, it means that no security device can defend against the threat $t_i, re_4 = 0$. The second one is $B_{ij}, B_{ij,}$ the response result of the security devices to protect against the dangers device_1 in the assets. The meaning of each value device_1 is RE like the other, but the difference is $B_{ij,}$ the actual response result of the individual security devices to the threats.

4. Analysis of Results

4.1. Multisource Extensive Data Model Analysis. Starting from the demand of data preparation in the process of multisource extensive data analysis and the order of data preparation in the task of data centralization management, the system optimizes and expands the standard data preparation process and unifies management, solving the problems of data preparation in the process of multisource significant data unification [20]. It solves the problems of establishing the view, defining, and organizing the data preparation process components. Computer networks bring great convenience to people's lives and improve office efficiency; at the same time, because of their inherent characteristics of sharing and openness, computer networks face many security threats, including malicious codes, hacker attacks, software vulnerabilities, and protocol vulnerabilities, which may cause different degrees of damage to computers. It saves the data preparation results from realizing the platform for multisource human data and supports data analysis and centralized data management tasks. The platform includes a multisource human data access function, editing process management, and process execution, which unifies the management of multisource big data from access and processing to output and displays. It saves it in the form of a data preparation process to improve the flexibility of multisource big data, and facilitate subsequent tasks.

The weakest attack link poses a more significant obstacle to the attacker, and the attack perpetrator must generate events throughout the attack chain to accomplish the task. However, attack independence shows that the probability of attack chain-related events is low under nonattack conditions,

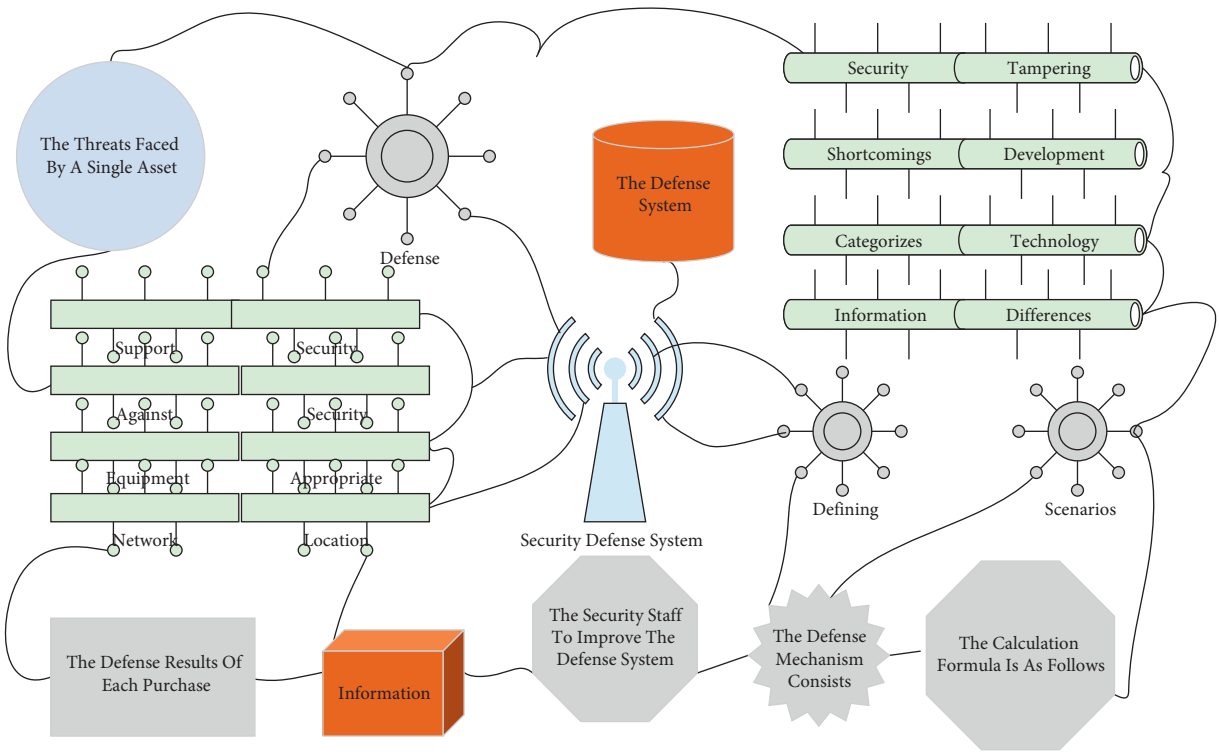

Figure 3: Design of the security defense system.

so it can be assumed that the co-occurrence of attack chain events is considered accidental, and the association between multiple occasions in the attack chain can help analysts reconstruct the attack scenario, and the association analysis based on various source events is also the main idea of this article. From the above analysis, research based on network attack chains has unique advantages in describing attack scenarios and helping to understand them; in addition, through forwarding mapping and reverse reasoning of attack chains, critical steps in the attack implementation process can be discovered and quickly intercepted in the attack defense. Due to the lagging nature of defense means, the correlation analysis and detection of APT often lag the actual attack. From the perspective of the dynamic security model, the time of attack discovery may be longer than the time of attack success: $Et = Dt + Rt - Pt > 0$. Only $Pt > Dt + Rt$. The security of the system can be ensured. Applying the MCKC network attack chain model can shorten the attack discovery time, thus ensuring the system's security. It is a keyword in the relationship description, and $n(t, R)$ is the number of occurrences in R, and the calculation of $P(t|\theta_{ees})$ is given in the following equation. To achieve the goal of establishing a unified view, it needs to obtain information on the structure of the data source and target views; to achieve the purpose of improving the flexibility of the data preparation process, it needs to provide an editing page where users can select data preparation process components and edit component properties in the data preparation process editing interface, and then build a data preparation process for multiple sources of big data, while providing process file upload and generation functions:

$$P\left(\frac{\theta}{t}\right) = \frac{1}{d+1}\sum_{t=1} p\left(\frac{t-1}{\theta}\right). \tag{6}$$

The MCKC model outperforms previous approaches in most metrics. As a lightweight model, the model simplifies the attack process into five stages and retains the representation of the lateral movement of the internal network in a recursive structure, where the bidirectional analysis method supports metadata analysis and cyclic iterative analysis, which is closer to the human analytical cognition. The advantage of the proposed MCKC model is that by fusing data information from multiple sources and offering a human analyst-like reasoning process, it reduces the cognitive burden of analysts in understanding and analyzing complex attack events in large enterprise networks. Subsequently, it allows adding more attack processes and novel probes in typical network environments to address the scalability of the analysis method [21]. However, the MCKC model is deficient in quantitative analysis, an essential issue in APT attack research. Attack activities need to conceal their activities to avoid detection by defenders. Business and probabilistic anomalies are often inconsistent in data performance, making the metric anomaly problem challenging to solve. Our future research uses the CAPEC-CVE model for attack metrics and risk assessment. The results of the CAPEC-CVE correlation are shown in Figure 4.

The purpose of the forward analysis is to discover progressively all possible impacts that depend on the starting point of the attack by evaluating the effects of an APT attack, starting from one attack entry point. The network attack chain represents the penetration process between nodes in

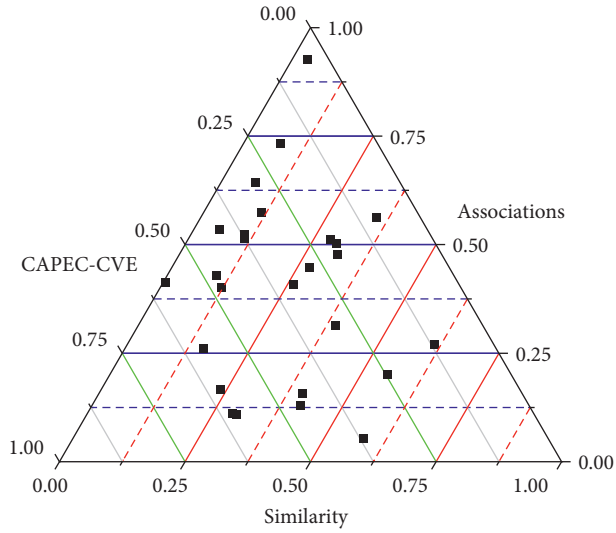

Figure 4: CAPEC-CVE correlation results.

the network. It is possible that an attack did not occur per se or that part of the attack has occurred but was not detected. The attack scenario cascades the chain of attacks between nodes into a new chain, with the intermediate nodes acting as a dabbling link for another network attack chain implementation. The forward analysis process relies heavily on classification mapping algorithms. After preprocessing, the raw data are represented as formatted events and stored centrally, where different data sources correspond to different phases of the attack. Database view is a virtual table formed by a series of association, selection, and other operations of the table of interest to the user in the target database; the database does not store the corresponding data but only the definition of the view, and the database system will automatically query the corresponding basic table when operating on the data in the view, thus improving the flexibility of the database application and producing a data model that better fits the user's needs than the primary table. Previous studies show that the number of security events and the relationship of events in the raw logs is not precise. Therefore, it is necessary to aggregate further and analyze multisource security events to generate anomaly alerts with security semantics; the anomaly alerts contain only low-level security information, and there are false alerts that will mislead the attack analysis process; other semantic security data generated by security devices (e.g., IDS, antivirus software, and firewalls) also provide attack-related information, and correlation analysis will connect these points into a more significant attack scenario [22]. With the guidance and intervention of expert knowledge, security alerts can be mapped to different attack phases, and eventually, attack scenarios can be reconstructed using attack chains. To verify the role of the knowledge graph-based correlation entity recommendation algorithm in entity association, the analysis of attack category recommendation and prediction accuracy is performed. Three dimensions of prediction accuracy (precision), recall (recall), and F1 value are selected in this article to compare the association results of different attack types. The comparison of recommendation algorithm results is shown in Figure 5.

Provide data sources for a data preparation platform oriented to multiple sources of big data. By accessing data from different data sources, the correlation between the algorithm lookup tables is established using a unified view and displayed to give users a more intuitive understanding of the data. This step realizes the function of accessing significant data sources, setting a unified view among the accessed data sources, and updating the idea according to certain conditions. The visualized components represent the atomic steps in the data preparation process. In image retrieval, the fusion of image-related information can be used to obtain more image annotation information and increase the richness of image annotation. The fusion of image annotation information and image information itself can improve the accuracy of image annotation. The data preparation process is established by dragging and dropping elements and creating lines between them to indicate the sequence of nuclear step execution. You can open the established process file and re-edit it or execute it directly to reuse the process file by uploading the process file. The visualized process components are converted into corresponding data structures, the data flow is set according to the relationship between members, and the model is checked for completeness and correctness. This step realizes the functions of generating process files, developing process models, and saving process execution results.

4.2. Computer Network Security Defines System Implementation. Security defense systems include access control, identification, authentication, and encryption methods. Management is the process of using a given set of security resources to achieve a goal. Security defense system management includes risk assessment, planning, system and service acquisition, authentication, maintenance, policies, standards, and procedures. People are a critical part of the information assurance program, and security defense systems include security personnel and personnel security. When designing, using, and managing security defenses, people need to be security-aware, educated, and trained in safety. Some attacks are caused internally by computer managers, so it is essential to be aware of security defense systems. Users will operate computer information systems for long periods and frequently use them. In computer systems, managers often start with an exemplary configuration of the system's security, which protects the computer information system to some extent. The computer system cannot effectively identify and stop the user's improper operation, malicious attacks, sending spam, and SQL injection. These malicious operations bring a significant challenge to system management. We do detailed explanation and analysis from four aspects: user behavior analysis, access address, spamming, and SQL injection. The variation in user accuracy is shown in Figure 6.

Comparative tests were carried out on the experimental data in terms of the change in the accuracy of the training set with the number of iterations, the difference in the accuracy

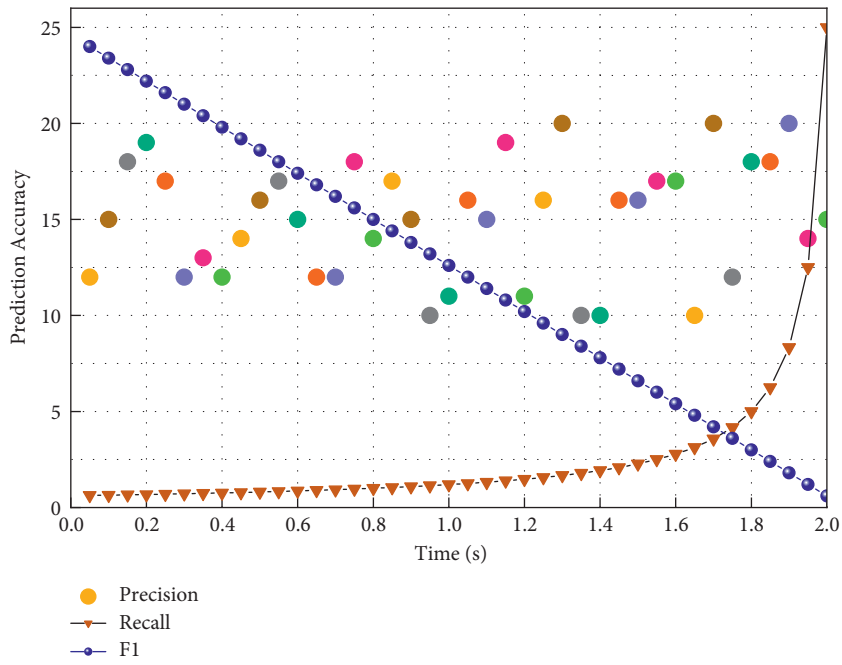

Figure 5: Comparison of recommendation algorithm results.

of the validation set with the number of iterations, the change in the value of the loss function of the training set with the number of iterations, and the change in the value of the loss function of the validation set with the number of iterations. The accuracy, recall, and f1-score of various classifications in each data are also tested and analyzed. Experiments are performed to compare and try the traditional classification method multilayer perceptron, recurrent neural network, and simple Bayesian methods. For the training of the convolutional neural network on user behavior text data, the dimension of the word vector is set as 20, the length of the word sequence is 300 set as, the number of classifications is set as 9 class, the size of each convolutional kernel is selected 5280 as the vocabulary size is 100 set as, the number of fully connected layer neurons is set as 128, the dropout retention ratio is 0.5 set as, the learning rate is specified as 0.001, and training is performed 5 times. After five times of subtraining, the experimental test environment is shown in Table 1.

By measuring the security defense system in the actual scenario, the results are obtained from the calculation, the defense effectiveness score in the design phase is 0.5244, and the actual defense effectiveness score in the deployment phase is 0.4013. For example, the defense system can only detect, alert and record detailed logs of attacks for denial-of-service attacks, and cannot block denial-of-service attacks, so it is necessary to add security devices that can block denial-of-service attacks in the defense system. Intrusion detection technology is real-time monitoring of network behavior. The data collected are analyzed by misuse detection technology and anomaly detection technology to see whether the network behavior violates the predefined security policy without affecting the network performance. In the design-phase evaluation, we analyzed the intermediate variable defense result R. The threats that the current defense system cannot defend include privilege elevation, authentication function defects, weak passwords, network sniffing, privilege misuse, buffer overflow, and process hijacking. In the deployment-phase evaluation, the analysis of the intermediate variables defense result R allows us to obtain the following defects of the existing security devices in the current defense system: (1) unable to block denial-of-service attacks; (2) unable to defend against cross-site request forgery; (3) unable to defend against API exploitation; (4) unable to secure request and response tampering; (5) unable to warn against access control defects; (6) unable to stop brute force cracking; and (7) unable to detect, caution against, and record redirection attacks. The identification statistics of the security defense system are shown in Figure 7.

By traversing the deployment methods of existing security devices through eight improvement rules, the number of intermediate results in the first phase of the algorithm is reduced by 1798, the number of intermediate results in the second phase by 976, and the number of developments in the third phase by 24, which significantly reduces the space explosion of the algorithm results and finally generates a total of 24 effective deployment methods. The evaluation of the 24 deployment methods was conducted by applying for the defense system evaluation program. The evaluation results were compared to obtain two optimal deployment methods. The defense effectiveness rating was 0.6392 in the design phase, and the actual effectiveness rating was 0.5067 in the deployment phase. Compared with the original deployment method, the network antivirus device and web application firewall are moved to the entrance of the information system network so that some threats to the Web server, DNS server, and office system can be defended. The

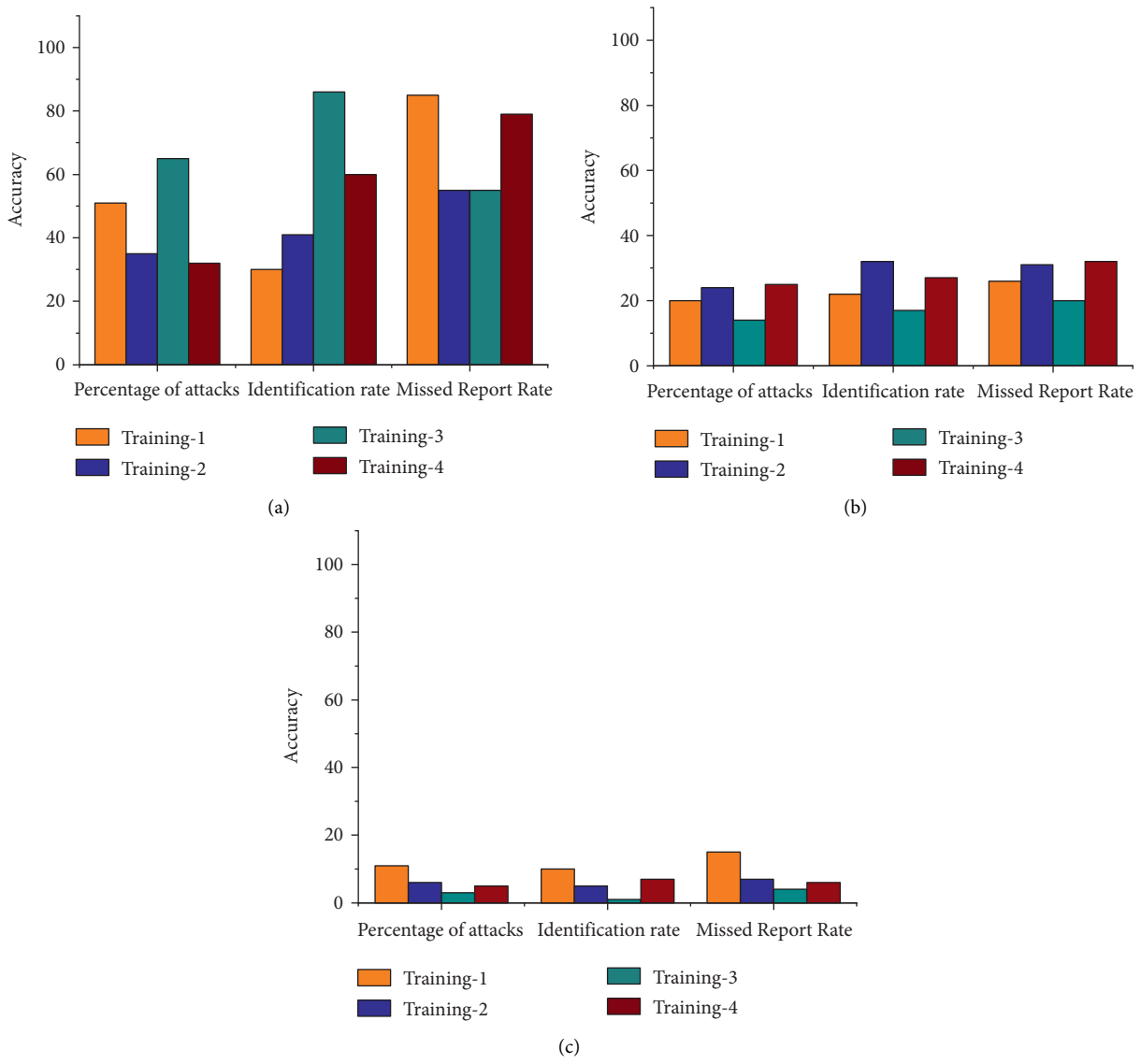

Figure 6: Accuracy of user behavior.

Table 1: Experimental test environment.

Software/ Hardware	Versions	Operability	Safety
TensorFlow	4.00.0-28-generic	The client does not need to install special programs	Support web browser can be used, with simple operation and interaction interface is simple
Python	1.10–4	The interface and operation can be very rich, but usually requires the installation	Can be deployed directly on the WAN to control user access through permission control management the advantage of greater interactivity
Architecture	3.6–6	Safety performance can be easily guaranteed	Speed is often slower than C/S architecture
MEM	x86–128	Wide applicability	Generally, use private protocols for interaction the advantage of faster response time
Cores	16G	Narrow application area, most of them only support some models and require specific network conditions	Simple and easy to develop and maintain, easy to expand business
Operating system	8G	High development and maintenance costs	The advantage of greater interactivity
Linux version	Ubuntu-16	The advantage of faster response time	Defense system against denial-of-service attacks

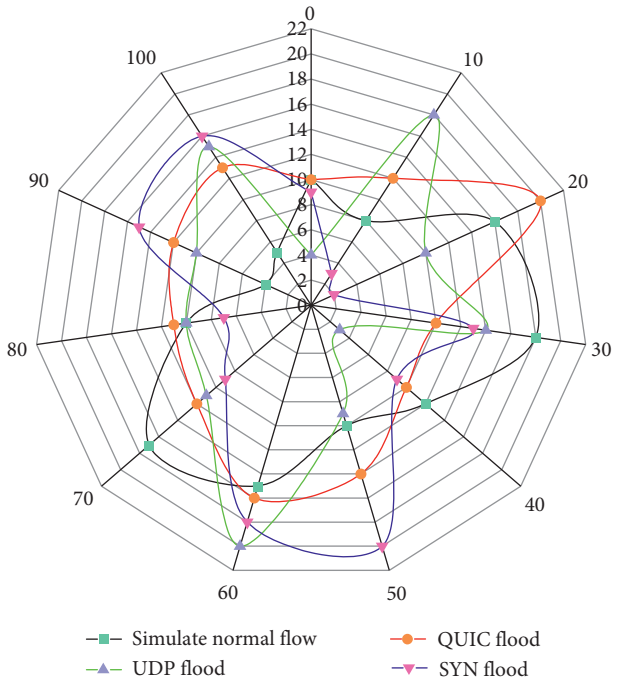

Figure 7: Security defense system identification statistics.

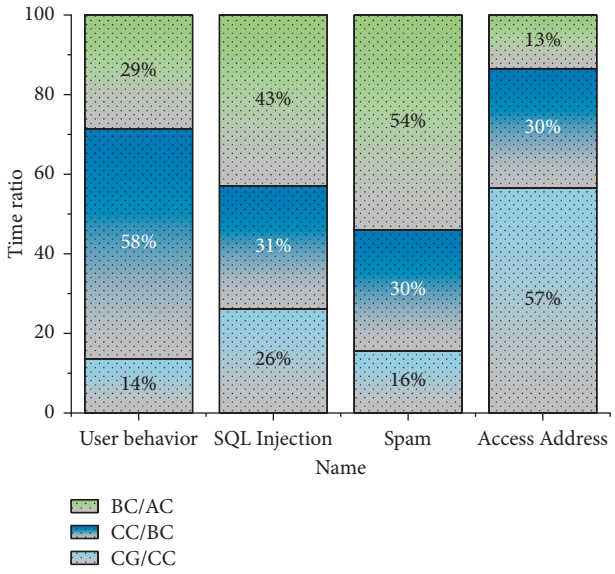

Figure 8: Security defense system results in the analysis comparison chart.

defense effect of the defense system on the above three business devices is improved, thus improving the defense effect of the defense system on the information system. The weakest attack link poses a more significant obstacle to the attacker, and the attack perpetrator must generate events throughout the attack chain to accomplish the task. However, attack independence shows that the probability of attack chain-related events is low under nonattack conditions, so it can be assumed that the co-occurrence of attack chain events is considered accidental, and the correlation between multiple occasions in the attack chain can help analysts reconstruct the attack scenario, and the correlation analysis based on various source events is also the main idea of this article. Finally, by differential optimization, the defense effectiveness of the defense system is evaluated by removing one security device at a time, and the differential evaluation results of seven security devices are compared. Removing the intrusion detection system at location L10 in the deployment does not reduce the quantitative value of the overall defense effectiveness of the defense system, indicating that the defense range of the L9 intrusion detection system includes the defense range of the intrusion detection system at L10, which the intrusion detection system at L10 can be removed to save resources. A comparison of the security defense system analysis results is shown in Figure 8.

A general description of the construction of the virtual machine experimental platform for the defense system is given. The installation of the virtual machine and the configuration of related software are discussed in detail to build the required practical environment for the experiments of the active defense system. The stability and effectiveness of the active defense system were tested in the built LAN environment, and the experimental results showed that the network hosts of several representative operation types protected by the dynamic defense system could generally communicate without any sense of transparency and with minimal impact on the performance of the original network system. When an attacker scans the LAN protected by the active defense system and tries to detect the network topology, he cannot watch the actual IP address information of the LAN hosts. The dynamic defense system can effectively resist the attacker's scanning probes and increase the attacker's attack cost.

5. Conclusion

The advent of the Internet era has made network security more and more critical, and quantitative assessment of network security has gradually received attention from all walks of life. Traditional quantitative assessment techniques for network security have failed to meet the actual needs. Based on the analysis of today's demand for cybersecurity modeling and quantitative assessment and existing methods, this article first outlines the development and research overview of cybersecurity and quantitative evaluation, then introduces the research status of cybersecurity modeling and quantitative assessment, and finally analyzes the shortcomings in existing research. The results of the corresponding work are detailed for information system network security modeling, network security defense system measurement, and optimization techniques, and are validated by applying the solutions proposed in this article to practical scenarios. To solve the problem of optimization and improvement of defense systems in network security, this article first designs a rule-based algorithm for traversing the deployment mode of security devices in defense systems to vary the deployment mode of security devices. On this basis, a differentiated optimization and security device importance analysis method is designed. Then, a network security

defense system optimization scheme is constructed to give the optimal deployment method of the defense system in combination with the measurement scheme to solve the problem of repeated stacking of security device functions and ineffective deployment. This article applies the proposed network security defense system measurement and optimization scheme to design an efficient defense system measurement and optimization system. The system consists of a data entry module, defense system evaluation module, deployment method traversal module, and defense system optimization module four. The system can provide users with defense system evaluation and optimization services. The proposed network security defense system measurement and optimization scheme are also applied with the actual scenario. The effectiveness of the network security defense system scheme proposed in this article is proved by analyzing the measurement and optimization results.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this article.

Acknowledgments

This research was supported by the National Natural Science Foundation of China (Grant nos. U1636107 and 61972297), the Science and Technology Project of Henan Province (China) (Grant nos. 182102210215, 192102210288, and 212102210538), the Soft Science Project of Henan Province (China) (Grant no. 182400410482), and the postgraduate education innovation and quality improvement project of Henan University (Grant no. SYL20040121).

References

[1] X. Chen, D. Zhao, W. Zhong, and Y. Jiufeng, "Research on information sharing technology of mental health alliance based on multi-source heterogeneous data fusion algorithms," *Academic Journal of Computing & Information Science*, vol. 2, no. 1, pp. 74–80, 2019.

[2] A. Ju, Y. Guo, and T. Li, "MCKC: a modified cyber kill chain model for cognitive APTs analysis within Enterprise multimedia network," *Multimedia Tools and Applications*, vol. 79, no. 39-40, pp. 29923–29949, 2020.

[3] L. Wang, M. Peng, and Q. Zhou, "Pre-impact fall detection based on multi-source CNN ensemble," *IEEE Sensors Journal*, vol. 20, no. 10, pp. 5442–5451, 2020.

[4] Y. Xiong and F. Zhang, "Effect of human settlements on urban thermal environment and factor analysis based on multi-source data: a case study of Changsha city," *Journal of Geographical Sciences*, vol. 31, no. 6, pp. 819–838, 2021.

[5] I. Kiaei and S. Lotfifard, "Fault section identification in smart distribution systems using multi-source data based on fuzzy Petri nets," *IEEE Transactions on Smart Grid*, vol. 11, no. 1, pp. 74–83, 2020.

[6] X. Lu, J. Han, Q. Ren, H. Dai, J. Li, and J. Ou, "Network threat detection based on correlation analysis of multi-platform multi-source alert data," *Multimedia Tools and Applications*, vol. 79, no. 45-46, pp. 33349–33363, 2020.

[7] A. Huang and F. Wu, "Two-stage adaptive integration of multi-source heterogeneous data based on an improved random subspace and prediction of default risk of micro-credit," *Neural Computing & Applications*, vol. 33, no. 9, pp. 4065–4075, 2021.

[8] M. S. Tehrany, S. Jones, F. Shabani, F. Martínez-Álvarez, and D. Tien Bui, "A novel ensemble modeling approach for the spatial prediction of tropical forest fire susceptibility using Logit Boost machine learning classifier and multi-source geospatial data," *Theoretical and Applied Climatology*, vol. 137, no. 1-2, pp. 637–653, 2019.

[9] Z. Xiong, H. Xu, W. Li, and Z. Cai, "Multi-source adversarial sample attack on autonomous vehicles," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 3, pp. 2822–2835, 2021.

[10] J. Hu, S. Cai, T. Huang et al., "Vehicle travel destination prediction method based on multi-source data," *Automotive Innovation*, vol. 4, no. 3, pp. 315–327, 2021.

[11] S. S. Harsha, H. Simhadri, and K. Raghu, "Distinctly trained multi-source CNN for multi camera based vehicle tracking system," *International Journal of Recent Technology and Engineering*, vol. 8, no. 2, pp. 624–634, 2019.

[12] Y. Dong, B. Hu, S. Zhang, Y. Huang, G. Nong, and H. Xin, "Research on North Gulf distributed big data submarine 3D terrain computing system based on remote sensing and multi-beam," *Soft Computing*, vol. 24, no. 8, pp. 5847–5857, 2020.

[13] Q. Guo, S. Jin, M. Li et al., "Application of deep learning in ecological resource research: theories, methods, and challenges," *Science China Earth Sciences*, vol. 63, no. 10, pp. 1457–1474, 2020.

[14] M. Mahdianpari, B. Salehi, F. Mohammadimanesh et al., "Big data for a big country: the first generation of Canadian wetland inventory map at a spatial resolution of 10-m using Sentinel-1 and Sentinel-2 data on the Google Earth Engine cloud computing platform," *Canadian Journal of Remote Sensing*, vol. 46, no. 1, pp. 15–33, 2020.

[15] N. Sanden and G. Neideck, "Learnings from the development of public sector multi-source enduring linked data assets," *Australian Journal of Social Issues*, vol. 56, no. 2, pp. 288–300, 2021.

[16] J. Su, R. Zhang, X. Zhao, Z. Hanwen, Li Fei, and K. Le, "A multi-source data based analysis framework for urban greenway safety," *Tehnički Vjesnik*, vol. 28, no. 1, pp. 193–202, 2021.

[17] Y. Lin, H. Ge, S. Chen, and M. Pecht, "Two-level fault diagnosis RBF networks for auto-transformer rectifier units using multi-source features," *Journal of power electronics*, vol. 20, no. 3, pp. 754–763, 2020.

[18] D. Wang, J. Yu, B. Liu, C. Long, P. Chen, and Z. Chong, "Integrated energy efficiency evaluation of a multi-source multi-load desalination micro-energy network," *Global Energy Interconnection*, vol. 3, no. 2, pp. 128–139, 2020.

[19] Y. Tang and M. Elhoseny, "Computer network security evaluation simulation model based on neural network,"

Journal of Intelligent and Fuzzy Systems, vol. 37, no. 3, pp. 3197–3204, 2019.

[20] L. Deng, D. Li, X. Yao, and H. Wang, "Retracted article: mobile network intrusion detection for IoT system based on transfer learning algorithm," *Cluster Computing*, vol. 22, no. S4, pp. 9889–9904, 2019.

[21] C. Zhou, H. Wang, C. Wang et al., "Geoscience knowledge graph in the big data era," *Science China Earth Sciences*, vol. 64, no. 7, pp. 1105–1114, 2021.

[22] S.-P. Wang and D.-M. Zhao, "A hierarchical power grid fault diagnosis method using multi-source information," *IEEE Transactions on Smart Grid*, vol. 11, no. 3, pp. 2067–2079, 2020.