Mimic Defense Structured Information System Threat Identification and Centralized Control

Bo Zhang1,a, Weichao Li2,b, Xin Sun3,c, Yufeng Zhao4,d
1State Grid Key Laboratory of Information & Network Security Nanjing, China
2State Key Laboratory of Mathematical Engineering and Advanced Computing Zhengzhou, China
3State Grid Zhejiang Electric Power Research Institute Hangzhou, China
4State Key Laboratory of Mathematical Engineering and Advanced Computing Zhengzhou, China
aemail: zhangbo@geiri.sgcc.com.cn bemail: 010liweichao@163.com cemail: 16526452@qq.com demail: 2298869023@qq.com

Abstract: With the speeding up of the informatization process of our national economy and social development, the degree of informatization in our country is getting higher and higher. Information system has become the key infrastructure of our country. The security of these basic information networks and important information systems has been seriously related to national security and social stability, to the collective interests of enterprises, and to the vital interests of the broad masses of the people. Increasing application in improving work efficiency and economic benefits at the same time, the increasingly complex information system has also brought greater challenges to operation and maintenance managers. The use of known or unknown vulnerabilities, causing damage to information systems, stealing sensitive information, endangering the normal operation of business systems and other behaviors, has become one of the main causes of various security incidents. How to effectively deal with the vulnerability threat and various vulnerability problems of enterprise information systems, concentrate on the implementation of information security vulnerability warning related risks, and provide solutions to related vulnerability threats, have become an urgent problem to be solved. Therefore, we construct the mimic structured enterprise information system by establishing the threat identification system through unified management to form a comprehensive and secure cyber threat management system.

1. Introduction

With the rapid development of the Internet, the network and web application have brought a fundamental change in people's living and thinking. However, as a key security issue, cyberspace security has become more prominent [1]. The traditional defense, mostly using signature detection, configuration access control label and other static, passive protection strategy. The traditional defense has hysteresis effect on the aspect of coping with unknown vulnerability attacks. To enhance the protection capability of various kinds of vulnerabilities, in recent years, active defense has gradually become the central research issue in cyberspace defense technology [2].
Our related work is using dynamic, heterogeneous to achieve structural diversity on active defense model [2]. Zhang et al. introduced a basic construction method and performance index of dynamic and heterogeneous web server [3]. Tong et al. introduced a method of constructing system heterogeneity by using the diversity of software and hardware [4]; Zhang et al. introduced a method of software diversified compilation, which is proposed to resist the threat of attack [5]. Li et al. introduced a heterogeneous tagging method against SQL injection attacks [6]. Ma et al. introduced a similarity solving method, which is proposed to enhance the discriminating and decision-making capabilities of DHR architecture systems [7]. And in another paper, Ma et al. proposed a dynamic scheduling model and method to resist persistent threats [8]. However, the current related work has no effect on Threat Identification and Centralized Control, which may make the identification of attack activities deficient.

2. DHR ARCHITECTURE AND INFORMATION SYSTEM

The DHR architecture is a principle method for implementing mimic defense. In the practical application of the information system based on DHR architecture, in order to ensure the normal security of the enterprise information system, it is usually necessary to configure various functional modules in the mimic defense architecture. Once the configuration is wrong, it will result in Multiple heterogeneous executables generate a collaborative attack vulnerability, which is easily attacked by an attacker to obtain some important data in the information system, causing irreparable damage to the user. Fig. 1 shows the DHR structure with a formalization research [2,3,4].

- The request dispatching and balancing module (RDB) which dispatches and balances user input requests to the appropriate executors.
- The dissimilar virtual server pool (DVSP) that consists of numerous virtual servers to construct heterogeneous.
- The heterogeneous element component (HEC) which is generated in DVSP and waits for scheduling.
- The dynamically executing scheduler (DES) which dynamically chooses heterogeneous elements for the HEA according to the principle of the heterogeneity maximization.
- The heterogeneous executor aggregate (HEA) which receives requests from the RDB, parses them, and passes the executed result to the voter.
- The dissimilar redundant response voter (DRRV) which responses and process requests across the executor aggregate and vote out the inaccurate responses.

In order to ensure the security and validity of system configuration, we provide an information processing method and threat identification and control equipment. After obtaining the data access request, the identity information carried by the system will be verified to be qualified, and then the independent processing results of multiple heterogeneous executives on the same input data can be obtained. The voting result and the processing logs corresponding to the plurality of heterogeneous executors can avoid the illegal operation of illegal users. In addition, by combining the voting results
with multiple processing logs, the existence of abnormal heterogeneous executors in multiple heterogeneous executors is analyzed, and abnormal heterogeneous executors are replaced in time, so as to ensure the security, reliability and stability of the system and equipment configuration.

3. CENTRALIZED CONTROL PLATFORM

a) Platform Design

The mimic defense structured enterprise information system and its threat identification and centralized control p-

![Diagram of Platform Design](image)

Figure 2 Platform design

The platform mainly includes: request dispatcher, response voter, multiple heterogeneous executors, data transmission devices and information processing equipment. Among them, request dispatcher, response voter and multiple heterogeneous executives belong to the application of mimetic defense technology, which is not discussed. The application scope of surgery is not mentioned in this article. The threat identification and centralized control platform includes data transmission devices and information processing devices. Fig. 2 shows the design method of the mimic information system management platform.

According to the functional composition of each module, the centralized management and control platform for threat identification of enterprise information system under the framework of mimic defense can be divided into four functional modules: security audit, security management, operation and maintenance and system upgrade. Through independent information processing interface, data communication with dynamic heterogeneous redundant architecture can be realized. Among them, security audit can be used to manage the processing logs and authenticate the requesting users, that is, the above authentication module, information acquisition module and exception determination module; security management is mainly used to ensure that the system does not provide redundant network servers, and in the event of abnormal circumstances, to ensure that the module of this function can belong to the part of security management, and the part of operation and maintenance is mainly used to monitor the running state of the system, and automatically generate specific recommendation policies, such as the above authentication loss. The system upgrade part is mainly used to realize the system upgrade, and the upgrade related modules belong to this part.

b) Information Processing

Information processing devices generate corresponding processing logs for each heterogeneous executor to record the processing of input data by the heterogeneous executor. Combining with multiple processing logs and voting results, we can accurately analyze the existence of abnormal heterogeneous executors in multiple heterogeneous executors, that is, to detect the running state of computer equipment system. The abnormal heterogeneous executor can be switched and reconstructed by the authorized administrator.
Figure 3 Information processing

promptly and reliably, which ensures the stability of the system. After the heterogeneous executor is attacked, the voter can not accurately judge the abnormal heterogeneous executor, and the system can not switch the abnormal heterogeneous executor in time. The information processing flow is shown in Fig. 3.

In the actual application scenario, it is usually divided into internal network and external network. This implementation can isolate the internal network and external network, and use a relatively secure internal network to achieve data transmission and computer equipment access. The internal network is usually a computer communication network, which is composed of computers, external devices and databases connected to each other in a local geographical area. It is a local area network. It can be connected with other local area networks, databases or processing centers through data communication networks or special data circuits. A wide range of information processing systems, therefore, this embodiment uses a local area network to achieve data transmission, the specific transmission mode is not limited.

c) Data Transmission

The data transmission device mainly includes the information processing interface of the business interface, which can realize the synchronous transmission of the processing logs of multiple heterogeneous executives to improve the work efficiency. For the management of information system by authorized administrator, it can be implemented by independent management interface, which is the information processing interface. It should be noted that the information processing interface is different from the bu-
business interface of the computer equipment system, that is, the interface used by ordinary users to access the computer equipment. By analyzing the data transmitted by this information processing interface, we set up an independent information processing interface for the management of computer equipment, so as to avoid the attacker posing as authorized administrator to log on computer equipment, steal important data from computer equipment, or destroy the operation of computer equipment system.

It can be seen that the computer equipment which only contain dynamic heterogeneous redundancy architecture have added information processing equipment relative to the transmission computer equipment, which can monitor the running state of the system, alarm and deal with the abnormal situation, set and modify the management strategy, upgrade the system and so on from many aspects, thus ensuring the stability of the operation of computer equipment. In addition, through the security cooperation processing of information processing interface, the vulnerabilities of cooperative attack caused by misconfiguration can be found in time, so that the attacker can not gain access to computer equipment and steal or destroy important data, which can improve the security of computer equipment.

4. THREAT IDENTIFICATION

a) Association Analysis of Alarm Data

Based on the threat of mimic ruling reasoning methods of abnormal alarm information that created by inconsistency degree of the data flow is too single to trigger some problems, such as the low identification of attack behavior and the difficulty of attack traceability. In order to enhance the accuracy and effectiveness of warning information and facilitate and facilitate the analysis of common malicious attacks, the verdict log and the log record in the actuator container are merged to analyze the attack threat more effectively [9]. Fig. 4 is a flow diagram of the generation of abnormal alarm data for the mimic defense system.

Firstly, the system input distributes the request to a set of N actuator containers, which parse and calculate the request, and send their calculation results to a decision module. Then, the ruling module makes a multi-mode ruling on the calculation results of the N actuator containers. According to the degree of inconsistency of the multi-mode ruling, the results with higher degree of inconsistency are marked as abnormal alarm data. When the exception information is recorded in the ruling log, the executor container index and the failure type recorded in the ruling log are used to determine whether the exception occurs in the application log from the log of the specific application, and then the related records of the process running with or without the exception are determined in the system log, and the relevant records of the process running with or without the exception are based on the type of failure recorded in the ruling log. Threat level is divided according to the situation that appears in the application level. Finally, the alarm information is ge-

![Figure 4 Flow diagram of generating abnormal alarm data](image-url)
Specifically, the decision log recodes decision information data, such as the ruling inconsistent degree, executive body index, the decision to visit the url, the enforcement of the ruling time; It is recorded in the PHP application log: response time, execution file, PHP parsing result status, and other program code execution data; It is recorded in the Web service log, such as the field of response time, response url, response status, and application service running data. It is recorded in the SQL application log that response time, SQL execution statement, SQL execution result state, and other SQL statements process the result data. Recorded in the operating system log, system files, port operation, running time, such as system call process data. The threat level is established by identifying threat generation at multiple levels. For example, if a Web service uses three executive body of the container mimic defense architecture, each execution body has four layers: application logic layer, Web service layer, data storage layer and system architecture. The abnormal alarm data generated by heterogeneous execution body container at different levels are obtained. According to the threat level, the threat level is from high to low: system architecture layer, data storage layer, Web service layer and application logic layer. The specific threat level is classified according to the inconsistency of response output recorded in the ruling log.

For example, if all three of the actuators are treated the same, no attack threat is considered and the threat level is 0. If the three execution results obtained by the executors are different, it is deemed that the attack is not effective on all three executors, and threat is only found in the application logic layer, then it is considered as level 1 threat. Threats found in the Web services layer are considered as level 2 threats. Threats found in the data storage layer are considered as level 3 threats. Threats identified at the system architecture level are considered level 4 threats. If two of the three execution results obtained by the processing of the three executors are the same, and the same two execution results are different from the other execution results, the attack is considered to be effective on one executor, and the attack is considered as a 5–8 threat from the application logic layer of the system to the system architecture layer. The classification of threat level can be adjusted according to the different choice of simulation defense architecture and decision strategy.

b) The classification prediction method of threat

When the ruling log is merged with the multi-level system log, there will be a lot of data information in the log record, including contain_id, url, time, IP, port, file and so on. Therefore, for the purpose of classification prediction of attack threat, the historical information knowledge base is built according to the feature fields extracted from the fusion log. The deep data information is mined and learned through deep learning, so as to achieve the classification prediction of threats [10].

Sensitive information such as access categories and sources in abnormal access traffic is analyzed according to the data record of the simulated defense threat perception log. In the abnormal access traffic recorded by the mimetic device, it is divided into six categories:

Figure 5 The alarm data correlation diagram

nerated from the multiple determined attack threats. The alarm data correlation diagram is shown in Fig. 5.
- Attempt of XSS attack. Include trying javascript function in the use of abnormal access request, such as "publish/main/9/javascript: the history, the back ()", "publish/main/17/javascript: void (0)!", etc.
- Attempt of webshell connection. Include trying .php, .asp, .jsp and other types of abnormal access requests, such as "/plus/mytag_is.php", "/index.asp", etc.
- Attempt of XML utilization. Include exception access requests that try to make use of external XML files, such as "/rss/bullet_2_0.xml", etc.
- Attempt of directory browsing. Include an attempt directory burst exception access request, such as "/yishi/", "/admin/", etc.
- Attempt of variable utilization. Include exception access requests that try to take advantage of variable operations in the source code, such as "public/column/4664041?type = 4 & catId = 4694378 & action = list ", etc.
- Attempt of the other exception access. Include abnormal access requests for scanning backup files that may exist on the site, such as "/robot.txt", "/www_cert_org_cn.rar", "/4652100.html", etc.

Of the 6 types of abnormal access, there were 29 XSS cross-site attack attempts, 237 webshell connection attempts, 749 XML utilization attempts, 2839 directory browsing attempts, 237 variable utilization attempts and 6,798 other exceptions. The threat classification of mimic equipment is shown in Fig. 6.

c) Visualization of Threat Situation

As early as 1977, John elaborated on the profound impact of visualization on data analysis [11]. For the warning data after the classification of threat level, the information needs to be expressed more clearly and effectively by means of graphics and images. Using data visualization technology, therefore, will each data representation for system situation elements, a large number of data collection system show figure, at the same time, in the form of multidimensional data system situation of each attribute value, from the aspect of different dimensional data analysis, thus further observation and analysis of data [12]. The visualization of simulated threat situation adopts the method of generating situation display diagram, which mainly includes general threat trend diagram, average threat trend diagram, overall request trend diagram, threat source distribution diagram and threat rank diagram. By classifying threat data and presenting them according to different display requirements, features are aggregated, and different feature aggregation results are generated according to different contents that need to be counted, and then displayed visually.

Data types can be divided into the following 5 categories in the visualization demonstration of mimic threat situation:

- for data with classification attributes, each classification attribute has a value. Therefore, it is easy to grasp the range of data sets by classifying data. Taking the ranking chart of attacked
websites as an example, cluster the domain name information in the threat perception log to obtain the number of threats generated by all different networks. Through the histogram, the numerical value of each classification and the most value in the classification can be clearly observed.

- for data containing overall and partial relationships, it is necessary to be able to clearly express the hierarchical structure of the data and keep all the groups together for analysis. Taking the attack category proportion map as an example, the attack categories recorded in the threat perception log were classified and summarized, the number of attack categories and total attack amount were calculated, and the proportion of each attack category was displayed through the pie chart.

- for data containing nested relationships of subcategories, multi-dimensional data is usually needed to be introduced for observation, so as to obtain richer connotation information. Taking the execution body state diagram as an example, if we want to know the attack situation of multiple execution body containers, we need to increase the data dimension of the execution body container category in the attack classification data. The proportion of each attack type in the running state of each execution body container is shown through the multiple ring graphs.

- for data that changes with time, the change of the data at what time, the degree of the change, and the reason for the change of the data trend are all contents that need to be paid attention to by the time series data. Take the attack trend diagram as an example to summarize the number of threats over a period of time, and then connect the number of threats at each time point in a chronological order to form a curve, so as to clearly observe the number of attacks and the trend of attacks.

- for the data affected by spatial changes, the spatial location information expressed in the data is the most important. Taking the display of the attack world map as an example, the longitude and latitude information of the source of the attack is mapped to the two-dimensional world map, and different colors are used to mark the source of the attack in the map. The number of different attack sources can be obviously observed through the differences of colors.

5. Evaluation

To exclude the risk factors and evaluate the Centralized control platform efficiently, we construct a PHP information system with various vulnerabilities, shut down the front-end filtering function and allow users to attack on the internet. All tests were performed on Intel core i7 CPU clocked at 3.1 GHz on machines with 32 GB RAM running PHP.

Fig. 7 shows that from 11:00 to 12:30, the highest number of visits is 210 times per minute, the lowest number of visits is 0 times per minute, the average number of 26.65 times per minute. The information service system runs normally and smoothly.

Fig. 8 shows that the total number of attacks recorded during the day is 1857 and the attacks started at 9:00 and increased significantly after 11:00. The Centralized control platform can effectively record threats.

6. Conclusions

This paper provides a method of mimic information system and threat identification and centralized control, whi-
ch provides a solution for the implementation of information security vulnerability early warning risk and vulnerability threat analysis. The centralized management and control platform is constructed through the information processing and data transmission interface. The anomaly alarm data association, threat classification prediction and threat situation visualization are used to analyze the anomaly response of the system. In this way, it can effectively help security personnel to analyze and evaluate the effect of mimic defense.

ACKNOWLEDGMENTS
This work is Supported by the science and technology project of State Grid Corporation of China: "Research and application of Mimic Defense Gateway applied to web application for power grid" (Grand No. 52110118001F)

REFERENCES
[1] Scarfone, Karen, Benigni, Dan, and Grance, Tim. 2017. Cyber security standards. Smart Business Akron/canton.
[2] Wu JX. 2016. Research on Cyber Mimic Defense. J. Cyber Security. 1, 4 (Sep. 2016), 1-10.
[3] ZHANG Z, Ma BL, Wu JX. 2017. The Test and Analysis of Prototype of Mimic Defense in Web Servers. J. Cyber Security. 2, 1 (Jan. 2017), 13-28.
[4] Tong Q, Zhang Z, Wu JX. 2017. The Active Defense Technology Based on the Software/Hardware Diversity. J. Cyber Security. 1, 1 (Jan. 2017), 1-12.
[5] Zhang YJ, PANG JM, ZHANG Z, WU JX. 2017. Mimic Security Defence Strategy Based on Software Diversity. J. Computer Science. 45, 2 (Mar. 2017), 215-221.
[6] Li WC, Zhang Z, Wang LQ. 2018. Improvement in diversify active defense for web application by using language and database heterogeneity. In Proceedings of the 11th IEEE International Conference on Anti-counterfeiting, Security and Identification (ASID). IEEE, 30-35.
[7] Ma BL, Zhang Z, Liu JX. 2018. A Similarity calculation method applied to dynamic heterogeneous web server system. J. Comput Eng Desig. 1, 282-287.
[8] Ma BL, Zhang Z, Zhu YS. 2017. A formalization research on web server and scheduling strategy for heterogeneity. Advanced Information Management, Communicates, Electronic and Automation Control Conference (IMCEC). IEEE, 1447-1451.
[9] Rinzler K, Cino R, Scully B. Situational awareness system security features[J]. 2018.
[10] Onwubiko C. Security operations centre: Situation awareness, threat intelligence and cybercrime[C]// International Conference on Cyber Security and Protection of Digital Services. IEEE, 2017.
[11] Situational Awareness Using 3D Visualizations[C]// International Conference on Cyber Warfare and Security Icews. 2018.
[12] Lopezcuevas A, Medinaperez M A, Monroy R, et al. FiToViz: A Visualisation Approach for Real-time Risk Situation Awareness[J]. IEEE Transactions on Affective Computing, 2018, PP(99):1-1.