Journal of Computers Vol. 35 No. 3, June 2024, pp. 275-289
doi: 10.53106/199115992024063503019

Construction a Portrait of Network Security Defense Capabilities Driven by Risk Monitoring Data

Jingling Xiao[1], Shuai Zhang[2*], Yong Peng[2], Liang Tong[2],
Wenhong Xu[2], Huan Wang[2], Qiumi Qin[1]

[1] Liuzhou Railway Vocational Technical College, Liuzhou Guangxi 545616, China
{81792402, 276832464}@qq.com

[2] Guangxi University of Science and Technology, Liuzhou Guangxi 545006, China
{221077427, 20230702029}@stdmail.gxust.edu.cn, 15277627934@163.com,
{pengyong, wanghuan}@gxust.edu.cn

Received 1 May 2024; Revised 1 June 2024; Accepted 19 June 2024

Abstract. The construction of the defense capability portrait can accurately locate the weak points of university defenses, locate the risks of the university's network information system, and provide security suggestions and countermeasures for university risk defense. Aiming at the problems in the construction method of network security defense capability portrait, such as large evaluation indicator system, strong subjectivity of weight and lack of theoretical support of profiling technology. The method integrates the hierarchical network security defense capability assessment with profiling technology organically, to simplify the evaluation indicator system. Relying on network security knowledge and statistical methods, it realizes the quantification of portrait labels in different dimensions and uses Analytic Hierarchy Process (AHP) to determine the weight of different indicator items, avoiding the subjectivity and arbitrariness in determining weight factors. Experimental results show that the proposed method can effectively solve the aforementioned problems, enhance the credibility and accuracy of network security defense capability portrait, and provide a theoretical basis for profiling technology. The experimental results show that the method in this paper effectively solves the shortcomings of the traditional network security capability portrait construction. In the real university network data, the evaluation results of this method are consistent with the expected effects, effectively evaluating the network security defense capabilities of universities.

Keywords: network security defense, capability portrait, AHP, statistical method

1 Introduction

As the depth and breadth of Internet application continue to expand, network security incidents occur frequently, and the threat to user information security in cyberspace is constantly rising, Network security defense is attracting more and more attention. [1]. Network security defense capability profiling is based on monitored network data, through the establishment of a corresponding indicator system, to realize the quantification of network security dimension indicators, to reflect the situation of network security defense capability [2]. Attackers hope to invade the system as long as they find a weakness, while defenders need to cope with every possible intrusion, even unimplemented attacks. Traditional network defense decision-making methods rely more on empirical subjective judgments, which are difficult to provide effective and credible suggestions for network security administrators to select defense strategies [3].

At the same time, existing quantified indicators may not fully reflect the true situation of network security defense capabilities. Some indicators may only focus on specific aspects and neglect other important factors. In addition, the selection and weight allocation of indicators may also be subjective and biased. Therefore, the construction of network security defense capability portraits has become one of the important challenges faced by various organizations and enterprises. It is crucial to ensure that important asset networks are effectively protected from threats [4]. The Analytic Hierarchy Process (AHP) can be applied to the hierarchical model of network security assessment, integrating profiling technology to simplify the problem while improving the accuracy of the assessment.

Therefore, in order to timely and accurately assess the network security defense capabilities, a large number

* Corresponding Author

of capability assessment models have been applied [5]. Common evaluation methods include qualitative analysis and quantitative analysis. Qualitative analysis makes security situation judgments based on existing knowledge and experience, while quantitative analysis calculates security evaluation results through mathematical methods [6]. Gafni and Levy assigned importance to 26 elements found by predecessors based on 27 subject experts and calculated relative weights [7]. Graf, Skopik and Whitebloom [8] used expert experience to analyze network abnormal data, thereby analyzing the network security situation, but the method depends on subjectivity and the model structure is not rigorous. Li integrated the subjective and objective security threat levels by using the Bayesian inference method, analyzed and quantified the network security risks caused by various threat sources, achieved the continuity and accumulation of security assessments, and could handle probabilistic uncertainty information, but this method is difficult to train, and repeated training is required to obtain more accurate probability values [9]. Zhang, Chen, Yan and Bian measured the coverage of data sources, visibility, and detection through the ATT&CK framework to evaluate the protection effectiveness of information systems, but the evaluation process of this method strongly relies on human subjectivity, which reduces the accuracy of the evaluation results [10]. Fang, Fu, Gu, Hu, Song, Jaeger and Mphapatra evaluated network security based on network security technical standards, but this method relies heavily on various security standards and lacks flexibility [11]. They also believe that existing IoT security analysis only focuses on a subset of all basic components, and propose a framework called IOTA, which conducts system-level security analysis by constructing attack graphs through two indicators [12]. Chen conduct security analysis through three attack graph indicators. These attack graphs not only provide a comprehensive view of system vulnerabilities, but also help identify and defend against potential attacks [13]. The D-S (Dempster-Shafer) evidence theory is an effective means of multi-source data fusion, which fuses incomplete information through mathematical reasoning methods [14]. Zhu, Wang, Luo, Cai, Peng and Zhang assigned the basic probability distribution function in the D-S evidence theory fusion rule through expert experience, but this method will cause inaccurate data fusion and extremely high subjective dependence on the final network security capability [15]. In terms of data fusion, Hu, Liu, Li, Hu, Xiang and Han proposed an efficient and privacy-preserving data aggregation and trust management scheme for IoT smart grids based on smart contracts, which possesses better results in terms of storage cost, computational complexity, and utility of differential privacy [16].

The above analysis can obtain network security defense capabilities, but its ability characterization still needs further research and analysis. While portrait technology can clearly characterize problems from multiple dimensions, therefore, portrait technology is used to reflect the network security defense capabilities of universities. Benaida [17] proposed that user portraits are a set of tag systems constructed for target users using objective reality data to depict their overall characteristics, and then provide them with personalized services or product supplies. Liu, Sun, Su and Zhang believe that user portraits are dedicated to depicting users' characteristics from multiple aspects to help people better understand users, dig out user needs, and improve user services [18]. Xu and Ying Fang believe that the user portrait method can describe user behavior and predict user behavior, personalized recommendation and service, covering e-commerce, libraries, health care, tourism and other fields [19]. Averyanova, Sushchenko and Ostroumov et al. characterized the portraits of intruders by proposing a network threat analysis and evaluation algorithm by analyzing the vulnerabilities of communication, navigation, control and surveillance equipment of modern drone systems [20].

However, due to the complexity and variability of the network environment, the current application of user portrait technology in the field of network security is relatively limited, and the traditional qualitative evaluation method cannot accurately and comprehensively quantify network security capabilities [21]. The network security defense capability evaluation model still has the following problems: (1) The network security data is single and the indicators are not comprehensive: Network security evaluation requires data from all aspects, and the single source of data in the process of network security evaluation cannot effectively reflect the generalization ability of the model, which is a major challenge [22]. (2) The model index system is too large, the quantification method lacks theory, and the weight is subjective: In the hierarchical network security evaluation process, the selection of index weight factors is too arbitrary and the evaluation index system is too complex [23]. (3) The network security defense capability portrait technology lacks theory: The existing portrait technologies are overly dependent on behavioral features, and their depiction ability is limited by the designer's prior knowledge of the analysis scenario, and they are less used in network security defense capabilities [24].

Based on the above research, the construction method of the university network security defense capability portrait based on AHP proposed in this article solves these problems and makes the following contributions:

The proposed method organically combines hierarchical network security defense capability evaluation with portrait technology, taking risk, defense measures, and technical capabilities as the main portrait dimensions, and simplifying the evaluation indicator system.

Based on network security knowledge and statistical methods, it realizes the quantification of portrait labels in different dimensions, uses AHP to determine the weights of different indicator items, and avoids the subjectivity and arbitrariness of weight factor determination.

By formally defining the network security defense capability evaluation model, it enhances the credibility and accuracy of the network security defense capability portrait, and provides a theoretical basis for portrait technology.

To verify the effectiveness of the model in the construction method of the network security defense capability portrait, the problems of the large evaluation indicator system, the strong subjectivity of weight, and the lack of theory in portrait technology, this article conducts research on the "Guangxi University Network Security Data" provided by the Guangxi Education System Network Security Detection Center.

2 Network Security Defense Capability Portrait Construction Model

2.1 Network Security Defense Capability Portrait Construction Framework

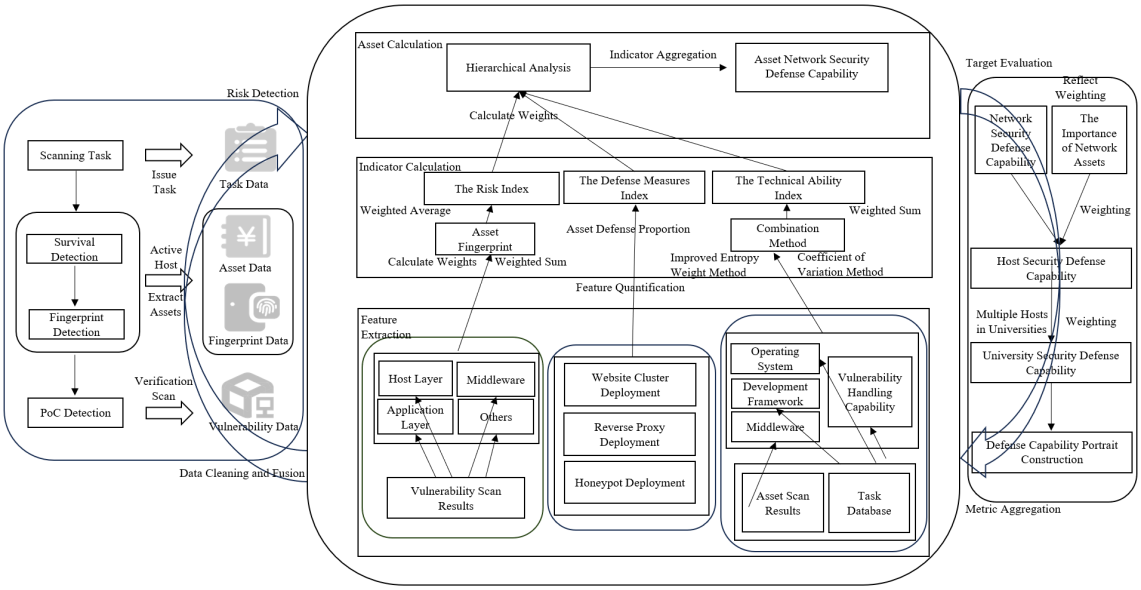

Fig. 1. Framework for constructing network security defense capability portrait

The network security defense capability portrait construction method proposed in this paper is an evaluation model with risk index, defense measures index, and technical ability index as the label system. The overall model structure is shown in Fig. 1, which includes data cleaning and fusion, network security label system quantification calculation, and label aggregation and portrait construction.

Data cleaning and fusion layer: Mainly includes cleaning of original data and fusion of multi-source data. Among them, the original data mainly includes scanned asset data, published task data, detected fingerprint data, and discovered vulnerability data. These datasets contain a large amount of noise, missing values, and outliers. Therefore, we need to clean them up, count and locate the "NaN" characters in the dataset, use the column average interpolation method to fill in the missing values, making the data more accurate and complete; by analyzing each column of the original data, delete all "0", empty and duplicate columns to reduce the feature dimension. To improve data quality and completeness, the data from different sources are fused according to the method of associating IP address ports after cleaning, thereby constructing multiple dimensions.

Network security defense capability label system quantitative calculation layer: mainly completes the extraction of asset elements, the quantification of dimensions and the aggregation of asset network security defense

capabilities. The quantification of the indicator system can intuitively reflect the actual situation of network security defense capabilities and improve decision-making efficiency. First, the risk index is mainly obtained through vulnerability scanning results, mainly including the scores and vulnerability risk levels of fingerprints and vulnerabilities in the host layer, middleware, application layer and other fingerprints; the features of the defense measure index are obtained from the data in the website group, reverse proxy and honeypot deployment; the features of the technical capability index are first obtained from the task database and asset task scanning results, and then mathematical equations are used for calculation and crawler data for correlation analysis, and finally the feature content is obtained. Secondly, complete the quantitative calculation of the risk index, defense measure index, and technical ability index, and obtain specific quantitative values through mathematical calculation methods and statistical models. Finally, complete the comprehensive calculation of network security defense capabilities. According to the three quantitative indicators, calculate the weight factor of each indicator item through the weight adjustment idea of AHP, and then aggregate the indicator items according to the weight factor to obtain the network security defense capability of the assets.

Network security defense capability portrait construction layer: mainly completes the aggregation of host and university network security defense capabilities. Among them, in the process of host network security defense capability aggregation, the weight of the corresponding network under each host is calculated according to the importance of each network, then the security defense capability of each obtained network is taken into account, and finally integrated into the host's network security; the aggregation of target layer is through weighted synthesis of the number of hosts contained under each target, thereby constructing a hierarchical university network security defense capability portrait according to the hierarchy.

In order to better understand the network security defense capability portrait assessment model, the definitions of each portrait system are given below:

Definition 1: Risk index, refers to the possible risk conditions in the network. Through known vulnerabilities and fingerprint information, these risks may be exploited to cause damage to network assets.

Definition 2: Defense measure index, refers to the main means of network vulnerability protection, these means mainly include the deployment of website clusters, honeypot deployment and reverse proxy deployment, etc.

Definition 3: Technical ability index, refers to the management and protection capabilities of the network's own administrators for assets. Through their technical capabilities, the generation of vulnerabilities can be prevented.

2.2 Formal Expression of Evaluation Model

(1) F refers to the set of university information. The set of university information includes multiple host nodes H, F= $\{F_{h1}, F_{h2}, \ldots, F_{hm}\}$. For any host node $F_h \in F$, it is aggregated by the network assets it contains and the value of the assets, H = (H_A, Assert). For any network asset H_A in the host, it can be represented by a six-tuple (ip+port, L, D, T, w, re), where H_A is the unique identifier of the network asset represented by ip+port; L represents all types of vulnerabilities existing in the network assets, including host layer, application layer, middleware and other categories; D is the defense measures possessed by the network assets; T is the technicality of the network assets; w is the weight of the network assets in the host; re is the value of the network assets.

(2) RE refers to the set of network asset values. The value of network assets refers to the services running in the network and the network systems running, which are assigned values according to their importance in the network. Any asset on the host has a value $re \in RE$ all have a certain value, represented by reV.

(3) Fp refers to the set of fingerprint information. The fingerprint $fp \in FP$ of any network asset can be represented as a quaternion (id, finger, c, score), where id is the unique identifier, finger is the fingerprint name, c is the fingerprint category, and score is the fingerprint score.

(4) A refers to the set of vulnerability information. For any vulnerability $a \in A$ in the network, a four-tuple (id, poc, c, severity) can be used to represent it, where id is the unique identifier, poc is the vulnerability name, c is the vulnerability category, and severity is the level of danger of the vulnerability.

(5) D refers to the combination of defense capability information. The defense measures $d \in D$ for any network asset can be represented by a binary group (id, dV), where id is a unique identifier, dV is the defense capability of the defense device.

(6) T refers to the set of technical capability information. The technical capabilities $t \in T$ for any network asset can all be represented by a quintuple (id, OS, MW, WF, VCM), where id is a unique identifier, OS is the technical capabilities of the operating system, MW is the technical capabilities of the middleware, WF is the

technical capabilities of the development framework, and VCM is the technical capabilities of the vulnerability handling.

SA denotes the network security defense capability, which consists of host information H, asset value RE, fingerprint information Fp, vulnerability information A, defense capability D, technical capability T. It is denoted as SA = (H, RE, Fp, A, D, T).

3 Quantitative Calculation of Evaluation Indicators

The network security defense capability portrait is characterized by the combined effects of asset risk, defense measures, and technical ability, so it is necessary to quantify the risk index, defense measure index, and technical ability index.

3.1 Quantitative Calculation of Risk Index

The risk index of the network SA(S) refers to the potential risk situation of the network, which is mainly composed of four elements: application layer, host layer, middleware, and others. This paper calculates the weights of each element category of risk capability according to the fingerprint category and score in the fingerprint database, and the average danger level of the poc vulnerabilities corresponding to these element categories. The process of calculating the risk index of the network is shown in equation (1)-(2):

$$\mathrm{SA}_j = \sum_{i}^{n} Score_i * Severity_i \ . \tag{1}$$

$$SA(A) = \sum_{i=1}^{n} w_i * SA(A)_i, \ \lambda_i = \frac{n_i}{n}, \ w_i = \frac{\lambda_i \overline{Severity_i}}{\sum_{k=1}^{n} \lambda_k \overline{Severity_k}} \ . \tag{2}$$

Where, SA_j represents the risk index of the j-th class of assets. n_i represents the number of fingerprints of the i-th element category, λ_i represents the proportion of fingerprints in the i-th element category, severity i represents the vulnerability danger level of the fingerprint corresponding to the i-th element category in the poc vulnerability library, w_i represents the weight of the risk index in the i-th element category, $Severity_{ALi}$ represents the danger level of the poc vulnerability corresponding to the fingerprint of the i-th asset under the application layer.

3.2 Quantitative Calculation of Defense Measures Index

The defense measure index of the network SA(D) refers to the protective ability of the network when facing vulnerability risks. The defense measure index reflects the defense ability and reliability level of the current assets by calculating the ratio of the defense ability dV_i of the assets X_i to the overall defense ability of all assets. The specific calculation process is shown in equation (3). The indicators include protective measures such as reverse proxy deployment, website group deployment, and honeypot deployment.

$$SA(D) = \frac{\sum_{p=0}^{k} dV_{ip}}{\sum_{j=0}^{n} dV_j} \ . \tag{3}$$

Where, d_{ip} is the quantified value of the ability of the p-th defense device owned by the i-th asset, k is the number of defense measures a certain asset has, n is the total number of defense measures, and $k \leq n$. The indicators include the deployment of reverse proxies, website clusters, and honeypots, etc.

3.3 Quantitative Calculation of Technical Ability Index

The technical ability index of the network SA(T) refers to the management and protective ability of the network's administrators for assets. The technical ability index is calculated by evaluating the security level OS(g) of the operating system used in the asset, the security level of the middleware MV(g), the security level of the development framework WF(g), and the vulnerability handling ability VCM in the evaluation model. Where, according to the independently collected China National Vulnerability Database (CNVD) vulnerability database, the vulnerabilities are classified as {very dangerous, dangerous, general, safe, and very safe}, which are corresponding to the values of the vulnerabilities of security level OS(g), MW(g), MW(g), are classified as {very dangerous, dangerous, general, safe, very safe}, corresponding to the values {1, 2, 3, 4, 5}. At the same time, a segmentation function VCM(t) is introduced to reflect the relationship between the speed of vulnerability resolution and technical capability. Usually, the vulnerability is solved in a short time, which indicates that the technical capability of the asset is strong. The specific calculation process of an asset's vulnerability handling capability is shown in equation (4).

$$VCM(t)=\begin{cases} 0.9.....0\le t\le 3 \\ 0.9-\dfrac{\arctan((t-3)*\dfrac{2}{\pi})}{5}.....t>3 \end{cases} . \tag{4}$$

The vulnerability handling capability function is VCM(t), t representing the number of days the vulnerability has been unresolved, when the number of days it has been detected is less than 3 days, the vulnerability takes the value of 0.9. The longer the vulnerability has been unresolved, the lower the value is VCM(t), indicating that the vulnerability handling capability of the asset is poor.

Indicators of different nature are treated differently, with positive indicators having a positive impact on the target technical capacity index and negative indicators having a negative impact.

$$\begin{aligned} a_{OS(g)j}&=\frac{x_{OS(g)j}-x_{OS(g)(\min)}}{x_{OS(g)(\max)}-x_{OS(g)(\min)}}. \\ b_{OS(g)j}&=\frac{x_{OS(g)\max}-x_{OS(g)j}}{x_{OS(g)(\max)}-x_{OS(g)(\min)}}. \end{aligned} . \tag{5}$$

Where, $a_{os(g)j}$ is the dimensionless way of the security level of the operating system if it is a positive indicator, the dimensionless way of the security level of the operating system $b_{os(g)j}$ if it is a negative indicator, $x_{os(g)j}$ is the raw indicator value of the j-th evaluation object of OS(g), $x_{os(g)min}$ is the minimum value, and $x_{os(g)max}$ is the maximum value.

Since the elements of the technical ability index of assets are quantitative data, the improved entropy weight-variation coefficient method is used to calculate the weights of various factors and reduce the subjectivity of the evaluation process [25]. The specific calculation process of the technical ability index of the network includes dimensionless, modified entropy weight method weight calculation, variation coefficient method weight calculation and weight fusion calculation, and the specific calculation process of the technical ability index of the network is shown in equations (6)-(11).

$$P_{(OS(g)j}=\frac{a_{(OS(g)j}+10^{(-4)}}{\sum_{j=1}^{n}(a_{(OS(g)j}+10^{(-4)})} . \tag{6}$$

$$H_{OS(g)}=-K*(\sum_{j=1}^{n}P_{OS(g)j}\ln P_{OS(g)j}) . \tag{7}$$

$$w_{OS(g)} = \frac{1 - H_{OS(g)}}{\sum (1 - H_i)} \ . \tag{8}$$

Where, to prevent meaningless logarithmic operations during the calculation, $P_{OS(g)j}$ is the proportion of the optimized indicator value, $k = (\ln n)^{-1}$, $H_{OS(g)}$ is the entropy value of OS(g), and $w_{OS(g)}$ is the weight of the OS(g) element.

Let $\sigma_{OS(g)}$ represent the standard deviation of the security level of the operating system, $x_{OS(g)}$ represent the mean value of the data in the OS(g) column, $v_{os(g)}$ represent the coefficient of variation obtained in the OS(g) column, and $w_{OS(g)}$ represent the weight of the indicator in the OS(g) column, and the coefficient of variation is calculated by the following equation:

$$v_{OS(g)} = \frac{\sigma_{OS(g)}}{x_{OS(g)}} \ . \tag{9}$$

$$w_{OS(g)} = \frac{v_{OS(g)}}{\sum_{i=1}^{m} v_i} \ . \tag{10}$$

Finally, the combined weights are obtained by combining the weights of the two methods according to the preference coefficients.

$$w_{OS(g)} = \lambda \sigma_{OS(g)} + (1 - \lambda) \beta_{OS(g)} \ . \tag{11}$$

Where, $\sigma_{OS(g)}$ is the variation coefficient weight, $\beta_{OS(g)}$ is the entropy weight, λ is the preference coefficient, and $w_{OS(g)}$ is the comprehensive weight.

The weights of each element are calculated using the improved entropy weight-variation coefficient method, and the technical ability index is calculated according to equation (12).

$$\begin{cases} SA(T) = \alpha * OS(g) + \beta * MW(g) + \gamma * WF(g) + \delta * VCM(g) \\ \alpha + \beta + \gamma + \delta = 1 \end{cases} . \tag{12}$$

Where, SA(T) is the quantified value of the technical ability of the network, and α, β, γ, δ are the weights of the security levels of the operating system, middleware, development framework, and vulnerability handling capability, respectively.

4 Determination of Dimension Weighting Factors

The determination of the evaluation index weight factor is mainly based on the idea of the analytic hierarchy process to determine the weight, and the weights of the risk index dimension, defense measure index dimension, and technical ability index dimension are solved. The specific process includes establishing a judgment matrix, calculating index weights, calculating the maximum eigenvalue, and conducting a consistency check in four parts [26]. The following will elaborate on the four parts in detail.

4.1 Constructing a Judgment Matrix

This paper adopts the hierarchical analysis method to determine the weight of each index, from the AHP hierarchical analysis process can be known as the comparison object for the riskiness index, the defense measures index, the technical ability index three parts, the index vector W = [SA(A), SA(D), SA(T)] ^{T}. Compare the im-

portance between them two by two to determine the degree of priority. Based on the satty ratio nine scale system, we can get the pairwise comparison matrix H composed of the ratio of each index as:

$$H = \begin{bmatrix} SA(A)/SA(A) & SA(A)/SA(D) & SA(A)/SA(T) \\ SA(D)/SA(A) & SA(D)/SA(D) & SA(D)/SA(T) \\ SA(T)/SA(A) & SA(T)/SA(D) & SA(T)/SA(T) \end{bmatrix} . \quad (13)$$

4.2 Compute the Weight Vector

The computation of eigenvectors can be obtained by normalizing the column vectors after arithmetic or geometric averaging [27], in this paper, we use arithmetic averaging to compute the specific computation process:

$$H = \begin{bmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} \\ \alpha_{31} & \alpha_{32} & \alpha_{33} \end{bmatrix} \to \begin{bmatrix} \alpha_{11}+\alpha_{12}+\alpha_{13} \\ \alpha_{21}+\alpha_{22}+\alpha_{23} \\ \alpha_{31}+\alpha_{32}+\alpha_{33} \end{bmatrix} \to [\overline{a_1} \quad \overline{a_2} \quad \overline{a_3}]^T \to [\overline{a_1}' \quad \overline{a_2}' \quad \overline{a_3}']^T = W . \quad (14)$$

4.3 Maximum Eigenvalue Calculation

Since the matrix H is consistent and inverse, the eigenvalues of the matrix H are obtained by right-multiplying the matrix H with the weight vector W, and since the matrix H is a positive matrix, there exists a maximum eigenvalue λmax=n that is not zero, as shown in the following equation:

The maximum eigenvalue is calculated by the equation:

$$HW = \lambda W . \quad (15)$$

$$\lambda_{\max} = \frac{1}{n}\sum_{i=1}^{n}\frac{(HW)_i}{W_i} . \quad (16)$$

Where (HW)i is the i-th element of HW.

4.4 Consistency Test

To ensure the accuracy and rigor of the assessment, the results must be tested for consistency. A judgment matrix is considered to have an acceptable level of inconsistency if CR < 0.1 [28].

$$CI = \frac{\lambda_{\max} - n}{n-1} . \quad (17)$$

$$CR = \frac{CI}{RI} . \quad (18)$$

Where, RI is the random consistency index, which is obtained from the table based on the order of the matrix H. When n=3, RI=0.58. CI is the consistency index, which is determined by the largest eigenvalue and the order of matrix H.

5 Portrait Aggregation

Since we have obtained the quantified values and weights of each dimension of the assets, we need to further analyze and aggregate the network assets [29]. According to the hierarchical thinking, this article aggregates the overall network security defense capability portrait of each university by constructing an asset network security defense capability portrait.

5.1 Evaluation Indicator Aggregation

The network security defense capability portrait of assets includes dimensions such as risk index, defense measures index, and technical ability index, and the technical ability index is calculated according to equation (19).

$$\begin{cases} SA(Asset) = \alpha * SA(A) + \beta * SA(D) + \gamma * SA(T) \\ \alpha + \beta + \gamma = 1 \end{cases} . \tag{19}$$

5.2 Target Layer Aggregation

In network assets, the information reflected by different assets of the same host is different, and their asset values are also different, and their importance in the network is also different. Therefore, asset weight is used to reflect the importance of network assets on the host, and the importance of network assets is generally the sum of the values of various resources on the assets. The specific calculation process is shown in equation (20).

$$w_i = \frac{\sum_{k=1}^{n} reV_{ik}}{\sum_{j=1}^{m}\sum_{k=1}^{n} reV_{jk}} . \tag{20}$$

Where, n is the number of resources owned by asset i, reV_{ik} is the value of the k-th resource of host i, $\sum_{k=1}^{n} reV_{ik}$ is the total value of the i-th network asset, m is the number of network assets owned by a host, and w_i is the importance of network asset in all network assets of the host.

The network security defense capability of the host is composed of the network security defense capabilities of various assets under the host, and is obtained by merging according to the weight of each network asset in the host. The specific calculation process is shown in equation (21).

$$SA(H) = \sum_{i=1}^{n} w_i * SA(Asset)_i . \tag{21}$$

Where n is the number of network assets owned by the host, w_i is the weight of the i-th network asset, SA(Asset) is the network security defense capability value of the asset, and SA(H) is the network security defense capability value of the host.

The network security defense capability of the university is our final goal. It is reflected by the network security defense capability of the host nodes owned by the school, and is obtained by the weighted sum and fusion of various hosts. The specific calculation process is shown in equation (22). Where, n is the number of hosts owned by the university, and SA(F) is the overall network security defense capability of the university.

$$SA(F) = \frac{\sum_{i=1}^{n} SA(H)_i}{n} . \tag{22}$$

6 Experimental Results and Analysis

This paper uses python to write the university network security defense capability portrait model and uses SQL to store real data, and uses python and SQL to interact for experimental testing. The experimental environment is a Windows host, and the processor is an AMD Ryzen 5 5600H with Radeon Graphics.

6.1 Experimental Data

To verify the practical utility of the network security defense capability portrait construction model proposed in this article, this article uses the "Guangxi University Network Security Data" provided by the Guangxi Education System Network Security Detection Center for research. The experimental data includes asset data, fingerprint data, vulnerability data, and task data, etc. The asset data features include domain name, IP, port, asset fingerprint label, operating system, vulnerability, defense device, asset service information, update time, etc.; fingerprint data features include fingerprint manufacturer, fingerprint type, fingerprint name, fingerprint year, fingerprint category, fingerprint service, fingerprint label, fingerprint score, etc.; vulnerability data features include vulnerability type, vulnerability name, vulnerability type, vulnerability name, disclosure time, danger level, vulnerability description, solution, etc.; task data features include task creation time, end time, university name, task period, task progress, vulnerability situation, etc.

Since the experimental data involves many assets, a random quantitative analysis and visualization of the data of 5 universities in the Guangxi University Network Security Dataset are conducted. Among the selected 5 universities, University A has 1 host and 26 assets; University B has 1 host and 112 assets; University C has 3 hosts and 61 assets; University D has 7 hosts and 21 assets; University E has 2 hosts and 6 assets. By selecting data, a comprehensive assessment of the quantitative results of the security defense capability of the model in actual universities is conducted. In order to avoid data leakage, this article desensitizes the names of universities and IP addresses in the data.

6.2 Experimental Results

In accordance with the model design requirements of this paper, first, in the data cleaning and fusion layer, the original data is cleaned, and useful data such as asset type, honeypot and reverse proxy are extracted from the json format data in the data column of asset data; replace illegal characters in the data set; delete duplicate asset data; Convert and calculate time data with python's time package; The value of OS column is "NaN", and the average value of different assets under the same host will be used to insert. Secondly, the multi-source data fusion method based on IP address port association is used to merge data from different sources into a unified data source with the same data pattern. Finally, using the different dimension capability evaluation algorithm introduced in Section 3, calculate the different evaluation indicators of various universities.

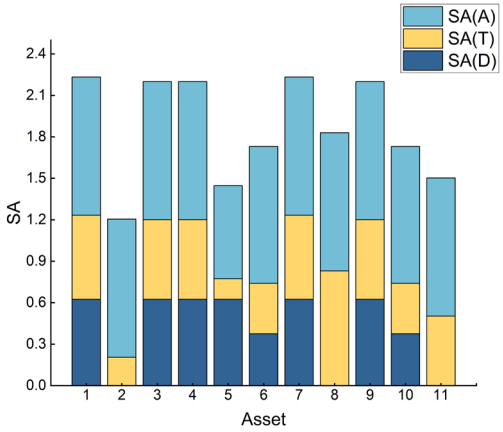

Fig. 2. Different dimension ability values of each asset in university C

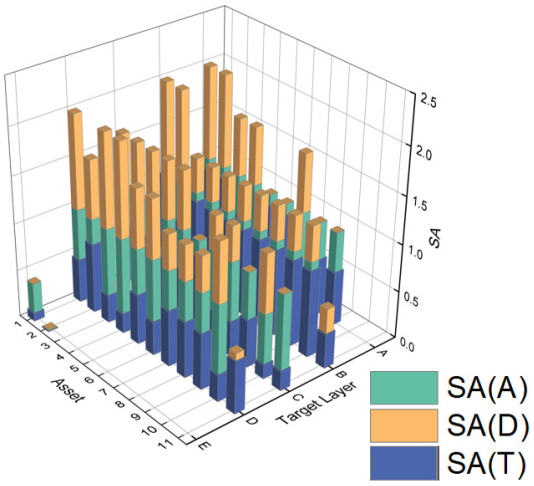

Fig. 3. Distribution diagram of different dimension ability for each asset

Taking an asset of University C as an example, the asset has a vulnerability of remote code execution in the wisdom platform Home Page Config. Ashx, a reverse proxy defense device, a system of Microsoft Windows, a vulnerability handling time of 10 days, and a service of platform login system. Based on the above information, equation (1) is used to get SA(A) = 0.6738. Use equation (2) to get SA(D) = 0.625. Use the improved entropy weight-coefficient of variation method to calculate SA(T) = 0.1478.

Based on the above calculation method, the network security defense capability values of various assets of University C are shown in Fig. 2, and University C has a total of 2 hosts and 11 assets. Through data analysis, it is found that the 2nd, 8th, and 11th assets have no defense equipment, so SA(D) = 0 for these network equipment assets. Because the 5th, 6th, and 10th assets detected vulnerabilities, the network risk indices of these three assets are $SA(A_5) = 0.6738$, $SA(A_6) = 0.9894$, and $SA(A_{10}) = 0.9894$ respectively. At the same time, the 1st and 8th assets have no recent vulnerability behavior and the security levels of the operating system and middleware are high. The network technology capability index of these two are $SA(T_1) = 0.608$, $SA(T_8) = 0.829$ respectively.

After the abilities of each dimension are calculated, it is necessary to aggregate the dimensions. First, this article uses the weight factor calculation method in Section 4 to calculate the weight factors of each dimension. According to the calculation results of the indicators, the weights of different assets in each dimension of University C are calculated. In order to evaluate the reasonableness of the above weight factors, according to equation (18), the consistency ratio at each moment is calculated. If the consistency ratio indicator CR is all below 0.1, it indicates that the calculation of the weight factor is reasonable.

Taking the first asset of University C as an example, according to the method of determining the weight factor, the consistency ratio is calculated to be CR=0.05156, indicating that the weight factor is reasonable. The values of the weight index α, β and γ are 0.196, 0.311, and 0.493, respectively, so the network security defense capability of the first asset is SA(Asset) = 0.196 * 1 + 0.493 * 0.608 + 0.63 * 0.311 = 0.69.

Similarly, the abilities of different dimension indicators of other universities' network assets can be calculated. The specific results are shown in Fig 3. According to the importance weight of network assets, using equations (4)-(7), the network security defense capability values of assets, host network security defense capability values, and the safety defense capability value of University C are calculated to be $SA(H_1) = 0.639$, $SA(F_c) = 0.566$, respectively. The network security defense capabilities of other universities at various levels are shown in Fig 5.

Looking at the overall network security defense capabilities of universities, each university shows unique advantages and areas that need to be improved. Among them, the network defense capability of University A shows a bipolar characteristic, the number of defense devices of some assets is sparse, leading to a lower overall defense capability, while other assets have abundant defense devices and have outstanding network security defense levels. The network defense capabilities of University B's assets are generally stable at around 0.7, but

there are also a few assets with defense capabilities below 0.45, which is mainly due to its insufficient overall defense capabilities and generally low indices in each dimension. For University C and D, the network security defense capabilities of their assets mainly fall below 0.7, which is mainly due to their limited number of defense devices and relatively backward technical capabilities, unable to update tool versions in time. Finally, the overall network security defense capability of University E is extremely low, and the defense capabilities of its assets in each dimension are almost zero, suggesting that University E may not pay enough attention to asset protection.

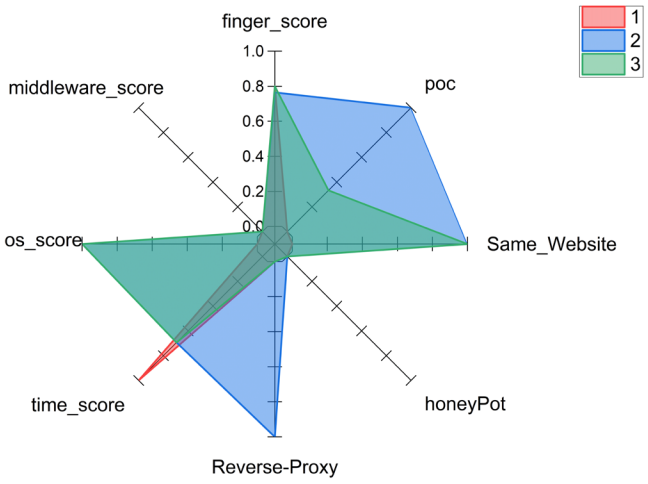

Fig. 4. Network security capability portrait indicator system

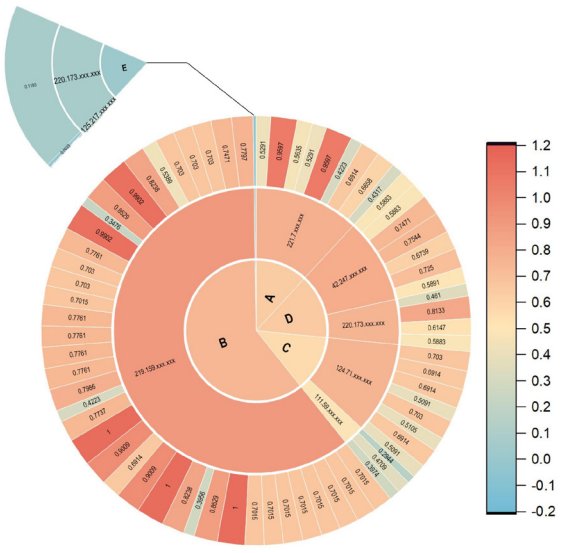

Fig. 5. Network security capability values at different levels

6.3 Experimental Analysis

This article first processes real network security data according to the proposed network security defense capability portrait model, and merges different sources of data into a unified data source with the same data pattern

through the method of associating IP address ports. Secondly, basic elements are selected for the fused data, and effective quantitative evaluation is carried out on the network risk index, defense measure index, and technical ability index of the assets. Finally, the weights of different dimensions are calculated according to statistical methods and aggregated.

Three network security capability portrait indicator systems selected through Fig. 4 show that although there is no vulnerability threat in the first asset, its website cluster, middleware, honeypot, reverse proxy, and system scores are relatively low, resulting in the network security defense situation of this asset is still not ideal. This shows that the evaluation of network security capabilities is multifaceted, and all aspects must be well done in order to get a secure network environment. The middleware score in each asset's network and the honeypot situation are relatively low, indicating that the defense measures level of each university is relatively weak and needs to be further strengthened.

By quantifying the ability indicators of each dimension and comparing them with the actual situation, it can be considered that the constructed risk index, defense measure index, and technical ability index three-dimensional evaluation indicators can effectively solve the problem of large indicator system, where the larger the ability value, the better the security situation. The model's safety ability values for different dimensions of University C are shown in Fig. 2.

Through the analytic hierarchy process to calculate the weights of different dimensions, according to the actual situation of each asset, calculate the weights of different dimensions of each asset, and judge the effectiveness and rationality of the weight factors by calculating the consistency ratio CR of each asset. Finally, the network security defense capability portrait of universities is calculated and constructed according to the importance of different assets in reality. The actual network security defense capability of the assets and the calculation results in Fig. 5 show that vulnerabilities were found in Universities B, C, D, and E in the target layer during the vulnerability scanning process, with fewer defense devices, and it takes about 10 days to handle the vulnerabilities after they are discovered. The risk indices of these targets are poor, and the defense measure indices and technical ability indices are significantly reduced. At the same time, although University A has no vulnerabilities, its overall security defense capability is still not the best, mainly because its defense measure index and technical ability index are poor. These evaluation results are consistent with the actual detection results of the universities. The experimental results show that the construction method that integrates the analytic hierarchy process and the ability portrait technology can reflect the network security defense capabilities of each university under real conditions, proving the accuracy and effectiveness of the network security defense capability evaluation model and the fusion of the ability portrait.

7 Conclusion

In view of the issues such as the overly large multi-question evaluation indicator system in the construction of network security defense capability portraits, the strong subjectivity of weight allocation, and the lack of sufficient theoretical support for profiling technology. This paper proposes an innovative and practical solution, which cleverly combines hierarchical network security defense capability evaluation with feature analysis technology, aiming to reduce redundancy, simplify the evaluation indicator system, and thus more accurately measure network security defense capabilities.

For better quantitative analysis, we rely on network security knowledge and statistical methods to achieve the quantification of portrait labels in different dimensions. This not only helps us to understand the importance of each dimension more deeply, but also provides us with a scientific method to measure and evaluate them.

In terms of weight allocation, we use the AHP to determine the weights of various indicator items. This method constructs a matrix, compares the relative importance of weights, thereby avoiding the subjectivity and arbitrariness in the process of determining the weight coefficient, making the weight allocation more objective and accurate.

Experiments were carried out on the "Guangxi University Network Security Dataset" provided by the Guangxi Education System Network Security Detection Center, and good results were obtained. The experiment shows that the construction method of the university network security defense capability portrait based on AHP effectively solves the problems of a large indicator system, difficulty in quantitative calculation, and strong subjectivity of weights. Therefore, this research has strong practicality and feasibility. However, this research also has certain limitations.

Despite achieving certain results in building cybersecurity defense capability portraits, this study has its lim-

itations. To overcome these limitations, future work will focus on expanding the diversity and quality of the dataset, exploring the use of open-source and crowdsourced data, while ensuring data quality and privacy protection. This will help build a more comprehensive and diverse dataset, providing a solid data foundation for the construction of cybersecurity defense capability portraits. Moreover, with the rapid development of artificial intelligence technologies, future research will explore the application of these advanced algorithms to the AHP and other related assessment methods to enhance the automation and accuracy of the evaluation process. By collecting and integrating cybersecurity datasets from different industries and domains, the model's generalization ability will be enhanced, ensuring adaptability to various network environments and security challenges. Such a comprehensive approach will help improve the credibility and accuracy of cybersecurity defense capability portraits, making a greater contribution to the development of the cybersecurity field.

8 Acknowledgement

Jingling Xiao and Shuai Zhang are the co-first authors of this paper. The authors would like to thank the anonymous reviewers for their valuable comments. This research was supported by the National Natural Science Foundation of China (62106053), the Natural Science Foundation of Guangxi Province of China (2024GXNSFAA010242), the 2023 Guangxi Vocational Education Teaching Reform Research Project (GXGZJG2023A043) Vocational College Teachers' "Micro-competencies, Community-oriented, Individualized" Digital Literacy Enhancement Path and Practice.

References

[1] Z. Wu, L. Tian, Y. Zhang, Y. Wang, Y. Du, Network Attack and Defense Modeling and System Security Analysis: A Novel Approach Using Stochastic Evolutionary Game Petri Net, Security and Communication Networks (2021)(2021) 4005877.

[2] X. Zang, J. Gong, X. Zhang, G. Li, Attack scenario reconstruction via fusing heterogeneous threat intelligence, Computers & Security 133(2023) 103420.

[3] X. Liu, H. Zhang, J. Ma, Y. Zhang, J. Tan, Research review of network defense decision-making methods based on attack and defense game, Chinese Journal of Network and Information Security 8(1)(2022) 1-14.

[4] R. Leszczyna, Review of cybersecurity assessment methods: Applicability perspective, Computers & Security 108(2021) 102376.

[5] H. Wang, Z. Chen, X. Feng, X. Di, D. Liu, J. Zhao, X. Sui, Research on network security situation assessment and quantification method based on analytic hierarchy process, Wireless Personal Communications 102(2)(2018) 1401-1420.

[6] L. Hu, H. Li, S. Dong, Weapon System Network Security Assessment Based on Combination Weighting and Fuzzy Gray Clustering, Fire Control & Command Control 45(9)(2020) 22-28.

[7] R. Gafni, Y. Levy, Experts' feedback on the cybersecurity footprint elements: in pursuit of a quantifiable measure of SMBs' cybersecurity posture, Information & Computer Security 31(5)(2023) 601-623.

[8] R. Graf, F. Skopik, K. Whitebloom, A decision support model for situational awareness in national cyber operations centers, in: Proc. 2016 International Conference on Cyber Situational Awareness, Data Analytics and Assessment, 2016.

[9] X. Li, Research on network security evaluation model based on Bayesian algorithm, Electronic Design Engineering, 29(5)(2021) 154-158.

[10] F. Zhang, D. Cheng, Q. Yan, J. Bian, Research on network security assessment and detection technology based on ATT&CK framework, Journal of Information Security Research 8(8)(2022) 751-759, http://www.sicris.cn/CN/Y2022/V8/I8/751.

[11] Z. Fang, H. Fu, T. Gu, P. Hu, J. Song, T. Jaeger, P. Mphapatra, Iota: A framework for analyzing system-level security of IoTs, in: Proc. 2022 IEEE/ACM 7th International Conference on Internet-of-Things Design and Implementation, 2022.

[12] Z. Fang, H. Fu, T. Gu, P. Hu, J. Song, T. Jaeger, P. Mphapatra, Towards System-Level Security Analysis of IoT Using Attack Graphs, IEEE Transactions on Mobile Computing 23(2)(2022) 1142-1155.

[13] F. Chen, Synergistic construction of system of cyberspace security technology, standard and system, Digital Technology & Application 39(11)(2021) 219-221, DOI: 10.19695/j.cnki.cn12-1369.2021.11.69.

[14] J. Sun, K. Wang, J. Huang, Z. Li, X. Lin, Z. Li, H. Zhao, N Mei, A multi-criteria fusion method for identifying broken neutral lines in low-voltage distribution networks containing new energy sources based on weight-correction D-S theory, Power System Protection Control 51(11)(2023) 1-14.

[15] L. Zhu, W. Wang, R. Luo, Z. Cai, S Peng, Z. Zhang, Situational awareness of E-learning system based on cyber-attack and vulnerability, in: Proc. 2021 Advances in Web-Based Learning–ICWL, 2021.

[16] C. Hu, Z. Liu, R. Li, P. Hu, T. Xiang, M Han, Smart contract assisted privacy-preserving data aggregation and management scheme for smart grid, IEEE Transactions on Dependable and Secure Computing (2023) 1-17, DOI: 10.1109/TDSC.2023.3300749.
[17] M. Benaida, Developing and extending usability heuristics evaluation for user interface design via AHP, Soft Computing 27(14)(2023) 9693-9707.
[18] H. Liu, J. Sun, Y. Su, Y. Zhang, Literature Review of Persona at Home and Abroad, Information studies: Theory & Application 41(11)(2018) 155.
[19] F. Xu, J. Ying, A Review of User Profile Research at Home and Abroad, Research on Library Science 12(2)(2020) 7-16.
[20] Y. Averyanova, O. Sushchenko, I. Ostroumov, N. Kuzmenko, M. Zaliskyi, O. Solomentsev, B. Kuznetsov, T. Nikitina, O. Havrylenko, A.Popov, V. Volosyuk, O. Shmatko, N. Ruzhentsev, S. Zhyla, V. Pavlikov, K. Dergachov, E. Tserne, UAS cyber security hazards analysis and approach to qualitative assessment, in: Proc. 2021 IDSCS, 2021.
[21] J. Deane, W. Baker, L. Rees, Cybersecurity in Supply Chains: Quantifying Risk, Journal of Computer Information Systems 63(3)(2023) 507-521.
[22] Y. Yu, A network security situation assessment method based on fusion model, Discover Applied Sciences 6(3)(2024) 1-9.
[23] R. Kaur, D. Gabrijelčič, T. Klobučar, Artificial intelligence for cybersecurity: Literature review and future research directions, Information Fusion 97(2023) 101804.
[24] M. Wang, Research on intelligent analysis and profiling methods of Internet end target, [dissertation] Chengdu: University of Electronic Science and Technology of China, 2020. DOI: 10.27005/d.cnki.gdzku.2020.003920
[25] Z. Tang, Y. Shen, An objective empowerment-based approach for assessing the importance of cyber assets, Network Security Technology & Application (4)(2021) 40-42.
[26] J. Xu, Q. Zhang, Z. Lu, Z. Xu, Y Shi, Evaluation index System of Bank Network Security based on Analytic Hierarchy Process, Communications Technology 51(1)(2018) 227-232, DOI: 10.3969/j.issn.1002-0802.2018.01.039.
[27] H. Wang, Z. Chen, X. Feng, X. Di, D. Liu, J. Zhao, X. Sui, Research on network security situation assessment and quantification method based on analytic hierarchy process, Wireless Personal Communications 102(2018) 1401-1420.
[28] J. Yi, L. Guo, AHP-Based network security situation assessment for industrial internet of things, Electronics 12(16)(2023) 3458.
[29] C. Wang, L. Xu, Q. Li, J. Ma, Network security situation awareness mechanism based on behavioral portrait construction, Journal of Computer Applications (2024), DOI: 10.11772/j.issn.1001-9081.2024010141.