Campus Network Security Intrusion Detection Based on Feature Segmentation and Deep Learning

Zhe Chen

Dancing school, Shandong University of Arts, Jinan 250000, China
E-mail: ZheChen2024@outlook.com

Received 19 January 2024; Accepted 26 March 2024

Abstract

At present, the secure campus network strategy adopts technical means such as distinguishing applications and limiting them separately, but they have triggered other new problems, greatly affecting the unity of network resources and data. The relatively dispersed network architecture will inevitably limit the further development and expansion of the campus network. Therefore, when universities plan their networks, they must consider whether the network is safe, complete, smooth, and sustainable for smooth upgrading and development. In order to improve the effect of campus network security intrusion detection, this paper combines feature segmentation and deep learning technology to construct a campus network security intrusion detection model. To reduce the transmission time of query requirements in the grid, this paper improves the replica management mechanism and requires the information server to cache the Bloom Filter structure of nodes in its successor node list. Moreover, this paper uses the Compressed Bloom Filter algorithm to compress the Bloom Filter structure that needs to be transmitted, therefore reducing the network traffic generated during the update process of the Bloom

Journal of Cyber Security and Mobility, Vol. 13_4, 775–802.
doi: 10.13052/jcsm2245-1439.1349

Filter structure copy and avoiding network congestion. It also constructs a campus network security intrusion detection model based on feature segmentation and deep learning. Through experimental verification, the effectiveness of the system in intrusion detection, user evaluation, information processing, and other aspects is verified, and it has certain advantages compared to traditional algorithms. Through experimental research, it can be seen that the campus network security intrusion detection model based on feature segmentation and deep learning proposed in the paper can effectively improve the effect of campus network security monitoring. The method proposed in this article can not only be applied to campus network security, but also to the network security management of enterprises and other units, with certain scalability

Keywords: Feature segmentation, deep learning, campus network, security intrusion.

1 Introduction

In a fundamental sense, absolutely secure networks are impossible. As long as it is used, there are more or less security issues. When discussing security issues, it is actually referring to a certain level of network security. Generally speaking, network security usually comes at the expense of network openness, convenience, and flexibility. Computer network information security is a complex system engineering [1]. It is generally believed in the international community that it not only involves technology, equipment, personnel management, etc., but also should be guaranteed in accordance with legal norms. Only when all aspects are combined, complement each other, and continuously improve, can network information security be effectively achieved. The countermeasures of network information security are only discussed from the technical point of view [2].

Computer network security refers to the smooth operation of the system by using different security technologies and management methods to ensure the authenticity, integrity, and confidentiality communication data. The definition of cybersecurity changes with different perspectives. From the perspective of network users, they are concerned with protecting privacy and security when personal or business information is transmitted in communication. From the perspective of network maintenance and management, the transmission, reading, and writing of local communications need to be protected and restricted to avoid dangers such as backdoor programs,

computer viruses, illegal intrusions, DDoS(Distributed Denial of Service), illegal acquisition of resources, and illegal control, to organize and control the illegal invasion of attackers [3]. From the perspective of information security protection, various behaviors that invade national security information networks need to be effectively identified and promptly stopped, and important information should be prevented from being stolen through intelligent methods to ensure national information security From the perspective of school education and knowledge, bad information on the Internet will have a negative impact on the development of minors, and we must find ways to stop it [4]. Computer network security is actually the security of communication data on the Internet. In a broad sense, the related technologies and methods that usually touch on the reliability, practicality, dissemination, circulation, and control of information on the Internet are all considered areas of computer network security. Computer network security includes the security of physical devices and communication data [5]. Physical equipment includes network communication, Layer 2 switching equipment, routers, computers, etc. It is necessary to build a secure physical network in order to transmit communication data quickly and reliably. Digital information includes network operating systems and system software that maintain the operation of computer networks, as well as stored network user communication data [6].

Computer networks are often threatened by security because the Internet is open, free, decentralized, timely, and interconnected. Computer users can access various network resources at will without being restricted by regions and countries. The frequency of data communication is very fast, so harmful information, can spread rapidly on the Internet [7]. The Internet can break the restrictions of national borders, which means that Internet virus Trojans may not only come from local computer users, but also from computers in any area on the Internet. The Internet is an open system, so it's impossible to predict where the person on the other end of the network will come from. When the computer network is established, it is only designed for convenience and practicability, but it does not consider the security of the entire network system, so everyone, every unit, and company can use the Internet, so the security threats to the computer network exist in many aspects [8].

The grid network management is efficient. It combines the geographic information system, which can separate and locate the management work content in detail. The huge work tasks can be classified into specific and detailed small pieces. The reference standard of the classification is mainly

based on the nature, scope, and content of the work; in this way, the powers and responsibilities can be clearly divided, and efficient management can be achieved [9].

The implementation of network security management in gridded campuses can not only reasonably apply the spatial positioning technology but also can use the PSO algorithm to lock the possible flaws or errors in the management and inform the managers of the relevant information. In this way, the managers can After receiving the information fed back from the system, it is possible to deal with the mistakes in the shortest time, which can greatly improve the management efficiency and ensure that the management work can be carried out smoothly [10].

The campus network security management system is based on various modern technologies and concepts to create a platform for the campus, which can classify various tasks in detail so as to realize their systematization and process, so as to realize their systematization and process and ignore the middle part. It can improve management efficiency and achieve a certain degree of standardization and fairness [11].

Compared with traditional campus network security management, network security management is more advanced. It will be divided into different grid units according to the jurisdiction where the campus is located, and then the staff will be allocated, and the regional responsibility system will be implemented to collect, organize, and update. The information that each grid needs to organize and organize includes basic information, security information, related material, etc. [12]. Applying advanced information technology in campus network security management can greatly improve the efficiency and quality of management. This is also the advantage of grid network security management. At the same time, it can also ensure that there is no dead angle in campus services, avoid missing information due to geographical reasons, etc., and have a good guarantee for the integrity and independence of information. It is a more advanced one in the current campus network security management mode [13].

The first is informatization, which is established on the basis of a huge information resource base and database. It is more convenient for information integration, query, and access, and can better provide campus services; the second is standardization. Campus network security management is not only for the people's legitimate rights and interests to be protected, and users are encouraged to participate more in campus life; the third is efficiency. Campus network security management optimizes the management mode and improves the management level with the help of high technology, which can provide

faster and better services. Users solve the difficulties in life; finally, comprehensive, grid-based network security management further refines different campus units, broadens the service area, and makes campus services more comprehensive and complete [14].

Gridded campuses have gradually emerged, so colleges and universities have begun to practice and use the gridded campus network security management model, and continue to absorb experience and combine with local actual conditions to make the gridded campus network security management model more distinctive and effective. Targeted. By analyzing and understanding the grid network security management mode of these campuses, we can better improve the management level of campus network security [15].

The purpose of data encryption is to protect the data, files, passwords, and control information transmitted on the Internet from being eavesdropped or tampered with before being transmitted to the other party, resulting in data being stolen and the authenticity and integrity of the data not being guaranteed. Before the data is transmitted on the Internet, the plaintext of the data is converted into ciphertext by encryption technology, and then transmitted. In this way, even if it is intercepted during the transmission process, the real information cannot be obtained. At the same time, the encryption mechanism can improve the integrity of the data transmission process., authenticity to identify, so as to ensure the integrity and reliability of the data. Therefore, the confidential data transmitted on the campus network must be encrypted before transmission [16].

The backup of the network system refers to the backup of the core equipment and data information in the network so that when the network is damaged, the operation of the network system can be quickly and comprehensively restored. The backup of core equipment mainly includes core switches, core routers, and important servers. The data information includes backup of configuration of network devices, server data, and the like. Backup not only protects the network system from hardware failure or human error but also protects it from unauthorized access by intruders or network attacks and damage to data integrity [17].

The construction of the network security system is actually a lasting confrontation process between the intruder and the anti-intruder. A network security system is not a perfect system that can prevent any attack once and for all, and such a system does not exist. What people are trying to establish is a dynamic network security protection system, a dynamic plus static defense, a passive plus active defense, or even a resistance, a complete security concept of management plus technology. The network security

problem is not a problem that can be solved by adding a firewall to the network [18].

Without authorization, network information data cannot be modified, and during transmission or storage, network information will not be intentionally or accidentally deleted, lost, forged, damaged, modified, replayed, or reordered (integrity requires correct storage and generation of network information, correct transmission), which is a characteristic of network security integrity. The availability feature refers to the network information that authorized entities can access on demand. Network information services allow authorized users or entities in need to access it, or the network can still be used when it needs to be downgraded or partially damaged (usually measured by the ratio of the normal usage time of the system to the entire working time). The function that a network information system can accomplish under specified time and conditions is the reliability characteristics of network security. The confidentiality feature means that information cannot be disclosed to anyone or organization, nor can it be stolen by anyone or organization. Controllability refers to the ability to control the dissemination and content of network information, and not allow the transmission of harmful information through public networks. The non repudiation feature refers to the authenticity of the interaction between participants in the network information system. In other words, all participants cannot deny the completed commitments or operations (digital signature technology is a way to solve non repudiation) [19].

In dynamic security technology, intrusion detection is one of the most important techniques. Attackers are rapidly changing in the network environment and lack proactive response. Strengthening the isolation of traditional operating systems and firewall technology can better ensure system security. If used in conjunction with the "traditional" position of firewall technology, it can greatly enhance and improve the level of system security. In fact, intrusion detection refers to the discovery of intrusion behavior, mainly collecting key point information of computer systems and networks, and then analyzing it to determine whether there are attack phenomena and violations of security policies in the system or network [20].

For event generators, their purpose is to retrieve events from the computing environment and provide them to other parts of the system; For event analyzers, analyze the data and obtain analysis results; For the response unit, the functional unit of the actual response analysis results can change the file attributes, issue alarms, or disconnect. For event databases, it mainly stores all intermediate and final data, which can be simple text files or complex databases [21].

Network Intrusion Detection System (NIDS) refers to the collection and analysis of network packets to determine whether they are abnormal behavior. The information source of network-based IDS is raw network packets, which can be monitored and analyzed for the transmission of all shared networks using network cards operating in hybrid mode. Mainly for real-time monitoring of network critical path information, monitoring all groups on the network, and analyzing available objects. The data source for its attack analysis is raw network packet data. Generally, the role of the network adapter is to monitor and analyze all network transmitted communications in real-time. Once an attack is detected, the system will issue notifications, generate interrupts, and alarms. Host based intrusion detection systems (HIDS) are mainly used to protect servers running critical applications. It detects intrusions by monitoring and analyzing host audit records and log files. This log contains evidence of anomalies and adverse activities in the system to determine whether someone is intruding or successfully intruding into the system [22]. By reviewing the log files, successful intrusions or intrusion attempts can be identified and corresponding emergency response procedures can be quickly initiated. Distributed Intrusion Detection System usually refers to an intrusion detection system deployed in a large-scale network environment, mainly monitoring the security of the entire network environment, including the host system and the network itself. With the increasing reliance on network infrastructure, ensuring the security and prevention of attacks of these facilities has become increasingly urgent. In order to protect network infrastructure, a mechanism is needed to detect in real-time activity patterns that may indicate abnormal or malicious behavior, and respond through automatic response measures. The main function of intrusion detection systems is to constantly monitor computer traffic. If any abnormal code flow is found, the detection mechanism is immediately triggered to avoid damage to internal devices. This also divides intrusion detection into three basic steps: information collection, data analysis, and response [23].

The main research content of this paper is how to quickly locate resources in a large amount of grid information, reduce the overall time consumption in the process of resource query demand from sending to receiving feedback, so as to find hidden unsafe information from a large amount of information. Unlike traditional algorithms, the traditional algorithm finds intrusion objects through identification, but there is inevitably a fish in the net when the amount of data is huge, The method proposed in this article is to effectively identify security risks while ensuring the efficiency of processing massive information, providing a theoretical basis for security

protection systems in the context of massive information in the Internet of Things era.

This paper combines feature segmentation and deep learning technology to construct a campus network security intrusion detection model to improve the security of campus network systems.

Considering the scalability and robustness of information service models, as well as the cost issues of information service queries, updates, and other aspects, a structured P2P network model was borrowed, and the DHT algorithm was used for distributed resource search to design the topology structure of grid information services. Therefore, this paper analyzes the campus network detection algorithm model in the second part, focuses on the P2P network technology and BloomFilter algorithm, builds the algorithm model, then constructs its structure in the third part, and designs experiments to verify its performance. Through research, it verifies the progressiveness and reliability of the algorithm model in this paper, and summarizes the research results and follow-up research directions in the conclusion.

2 Campus Network Detection Algorithm

Distributed P2P technology not only solves the problem of single point failure of centralized directory service nodes and high-level index service nodes in centralized and hierarchical grid information service models, but also eliminates the system bottleneck problem caused by them, improving the robustness and scalability of the system.

This article draws on the structured P2P network model to design the topology of grid information services, using the DHT algorithm for distributed resource search. Firstly, it considers the scalability and robustness of the information service model, and secondly, it takes into account the cost issues of information service queries, updates, and other aspects.

2.1 P2P Network Technology

P2P (Peer to Peer) technology, also known as peer-to-peer technology, no longer relies solely on a few servers, but on the computing power and bandwidth of network participants. P2P networks belong to the category of distributed networks, which are application layers built on top of IP networks. The participants in the network are called peer nodes, which can serve as both providers and recipients of resources. Some hardware resources, such as processing power, network connectivity, and storage capacity, can be shared

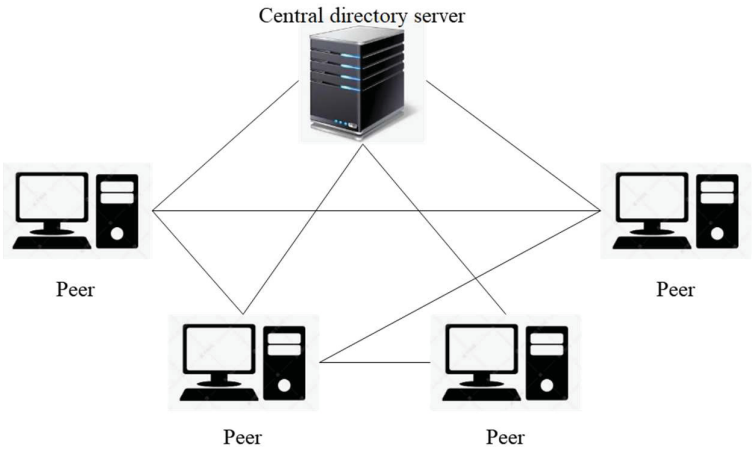


Figure 1 Centralized directory P2P network model.

as resources. In P2P networks, the services and content provided by shared resources between peer nodes can be accessed by nodes in the P2P network without passing through other intermediate entities.

The centralized directory P2P network model uses a central directory server (as shown in Figure 1) to manage peer nodes in the network, which retains the characteristics of centralization, which is also called a non-pure P2P structure. The topological structure of this model and the process of user registration are similar to the traditional C/S model but have their characteristics. The peer nodes store all the information about the services they provide. The central directory server only retains the index information of registered resources, unlike the previous server, which needs to store the corresponding resource content. Moreover, the servers in the P2P network have the ability to interact with peer nodes and peer nodes, which changes the traditional C/S model and adopts the "monopoly" method.

To control the transmission of search messages, TTL (Time To Live) reduction can be used. Figure 2 shows the flooding algorithm for unstructured P2P.

In practical applications, the centralized mode of the centralized directory network model is vulnerable to direct attacks. The pure P2P network model solves the problem of anti-attack, but it has the problem of poor search and scalability. In view of the above reasons, the layered P2P network model absorbs the advantages of the above two and divides the nodes into ordinary nodes and super nodes according to different capabilities. The status of these

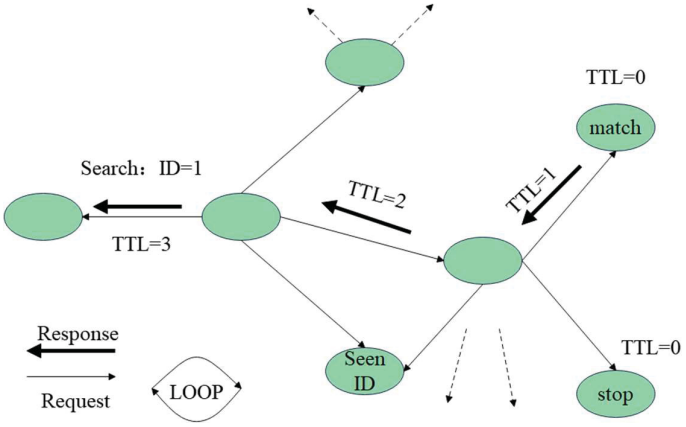


Figure 2 Unstructured P2P flooding algorithm.

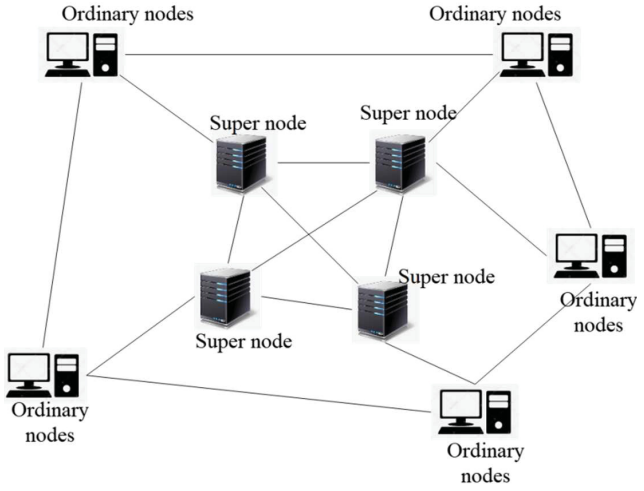


Figure 3 Hierarchical P2P structure.

two types of nodes in resource sharing is the same; that is, they can provide services to other resource-requesting nodes. The difference is that the super nodes store the information of other nodes in the network, and the resource discovery algorithm is only forwarded between super nodes, and then the super nodes forward the query request to the ordinary nodes. Therefore, a high-speed forwarding layer is formed between the super nodes, and the super nodes and the ordinary nodes in charge form a hierarchical relationship (so called a hierarchical network model), and its principle is shown in Figure 3.

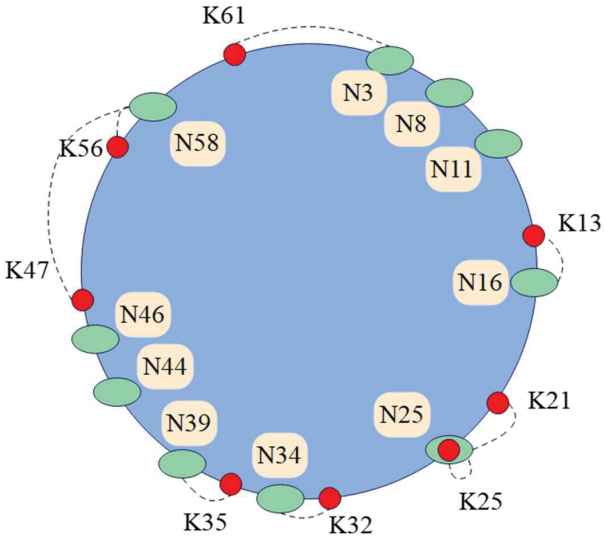


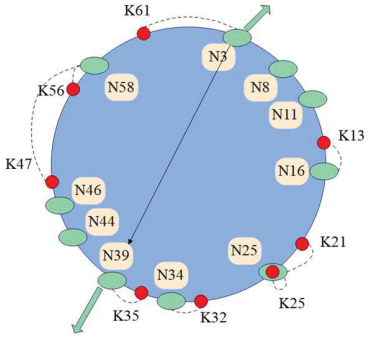
Figure 4 The identifier space of m = 6.

In a hierarchical network, a super node and several ordinary nodes form an autonomous cluster, the nodes in the cluster query messages autonomously, and each cluster is queried through the query mode of the pure P2P network model.

A Chord identifier ring is illustrated in Figure 4. If the length of the ring identifier is m = 6, the space size of the identifier is $2^6 = 64$, and the value range of the node and resource identifiers is an integer in [0, 63]. In this paper, 8 resources are given in advance, and the keyword identifiers are obtained by hashing these resources first. It is represented by K, which are K13, K21, K25, K32, K35, K47, K56, K61 respectively. There are 10 nodes on the Chord ring, and N3, N8, N11, N16, N25, N34, N39, N44, N46, N58 are respectively obtained as identifiers of the 10 nodes after hash operation.

The identifiers of the above nodes are organized in a clockwise order from small to large. According to the principle of the successor node of the keyword, the successor node of the keyword identifier K13 is N16, and the information of K13 is stored by N16. Similarly, the successor node of K21 is N25, and N25 stores the information of K21. In particular, the successor node of K25 is N25; at this time, the keyword Key and the node identifier are equal in value. On the Chord ring, there is no node identifier larger than 61, so in the ring structure, N3 should act as the successor node of K61 and save the relevant information of K61.

Starting identifier	Identifier interval	Node
$3+2^0$	[4,5)	N8
$3+2^1$	[5,7)	N8
$3+2^2$	[7,11)	N8
$3+2^3$	[11,19)	N11
$3+2^4$	[19,35)	N25
$3+2^5$	**[35,4)**	**N39**



Starting identifier	Identifier interval	Node
$39+2^0$	[40,41)	N44
$39+2^1$	[41,43)	N44
$39+2^2$	[43,47)	N44
$39+2^3$	[47,55)	N58
$39+2^4$	**[55,7)**	**N58**
$39+2^5$	[4,40)	N8

Figure 5 N3 finds resources with Key = 50.

In Figure 5, it is assumed that node N3 receives a request to locate a resource with a Key value of 50. Step 1: N3 first determines whether it is the successor node of 50. Because 50 is not within range of N3 and its predecessor node N58, N3 is not a successor node of 50. The second step: the algorithm finds whether the successor node of 50 is recorded in the pointer table of N3, and after querying, 50 is not between the start flag of any entry in the pointer table of N3 and its successor identifier. Step 3: The algorithm searches for the predecessor node closest to 50 in the pointer table of N3, which is N39 after the search. N3 passes the lookup message to N39 and notifies it to lookup 50's successor nodes. N39 starts from the first one and then continues to find the forward node of 50, and the process is also the same as the above three steps. In the fifth entry of the pointer table of N39, it is found that 50 is between the start identifier 47 and its successor node N58, so N58 is the node that stores Key50, and the positioning is successful.

2.2 Bloom Filter Algorithm

Bloom Filter is a data structure based on random numbers (or hashes) that stores members in less space and queries members with higher efficiency. The Bloom Filter algorithm is widely used in applications that express large datasets and improve query efficiency, such as email filtering, web filtering, and word matching. The Bloom Filter algorithm is a spatially efficient random data structure that succinctly represents a dataset as a single digit group and can quickly determine whether an element belongs to that set.

The Bloom Filter algorithm, as both a network transmission information and a data structure, has greater advantages in comprehensiveness compared to traditional network algorithms. This article introduces it to reduce network congestion and maintain network stability when updating server information replicas.

For example, to represent a set $S = \{e_1, e_2, e_3, \ldots e_{n-1}, e_n\}$ of n elements, we will use an m-bit bit array V to represent it. In the process of element representation and judgment, k independent hash functions $h_1, h_2, h_3, \ldots, h_k$ will be used, and the value range of these hash functions is $[1, \ldots, m]$. Specific steps are as follows:

(1) Initialization of Bloom Filter structure: the algorithm sets each bit of the bit array V of m = 12 bit (each bit is identified by B1, B2, B12, respectively) to 0 (Figure 6(a)).

(2) Insert elements into Bloom Filter: The algorithm uses k hash functions to hash the element e_x, and sets the position of $h_j(e_x)$ in V to 1, where $1 < x < n$, $1 < j < k$. A bit in V may be set to 1 multiple times, and only the first change will be valid. For example, the algorithm takes k=3, inserts two elements of e_1, e_2 into V, sets the position corresponding to the arrow in the figure below to 1, and the ninth bit is set to 1 twice. However, it is only valid when e_1 is inserted, that is, the bit is set to 1. When e_2 is inserted, this bit is already 1, so there is no change (Figure 6(b)).

(3) Query elements in Bloom Filter: The algorithm finds whether an element e belongs to the set S, and only needs to check whether all the positions of $\mathrm{h_j(e)}$ in V corresponding to the k hash functions are all 1. If all are 1, then e may belong to the set S. If one bit is 0, then e must not belong to the set S, where $1 < j < k$ (Figure 6(c)).

The Bloom Filter structure is built by mapping each element in the set $S = \{e_1, e_2, e_3, \ldots e_{n-1}e_n\}$ to a bit array with k hash functions. At this time,

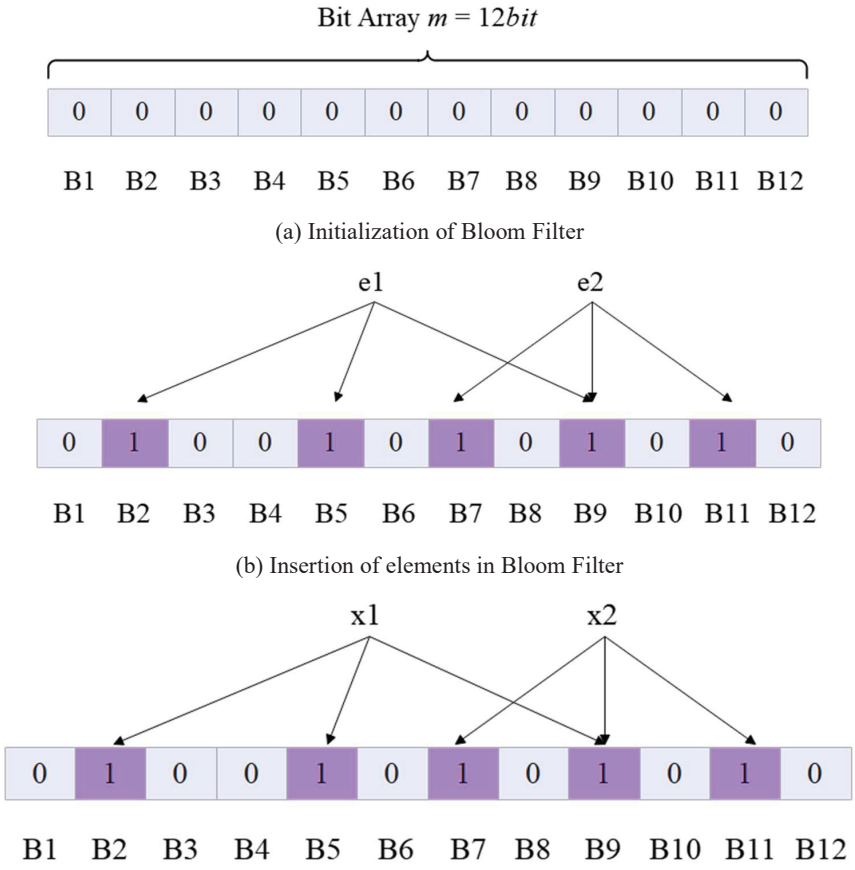


Figure 6 Working principle diagram of bloom filter.

the probability that a bit in the bit array is still 0 is:

$$p = \left(1 - \frac{1}{m}\right)^{kn} \approx e^{-kn/m} \tag{1}$$

Under the premise that the hash function is completely random, 1/m represents the probability that a certain hash function is mapped to this bit, then the probability that k hash functions are not mapped to this bit is (1 − 1/m). Because when all n elements in S are mapped, a total of kn mappings are done. If a bit is still 0, it means that kn hash operations have not selected the bit, so the probability that a bit is still 0 is $(1 - 1/m)^{\mathrm{k}}$.

To simplify the operation, we will use the approximate formula commonly used to calculate e.

$$\lim_{x\to\infty}\left(1-\frac{1}{x}\right)^{-x}=e \tag{2}$$

The false positive rate f of Bloom Filter in the query is estimated as:

$$f=(1-e^{-kn/m})^k \tag{3}$$

If $p = e^{-kn/m}$, then the false positive rate of Bloom Filter can be expressed as:

$$f=(1-p)^k=(1-p)^{(-ln\ p)(m/n)}=(e^{-ln\ p^* ln(1-p)})^{(m/n)} \tag{4}$$

In order to minimize the error rate f, the method of derivation of f can be used, it can be seen that when

$$k=\frac{1\ nm}{n}\approx\frac{om}{n} \tag{5}$$

f gets the minimum value. That is to say, when the relationship of kn = 0.7 m is established, the false positive rate of Bloom Filter is the lowest.

Therefore, the false positive rate of the Bloom Filter algorithm is related to the length m of the bit array used, the number k of the hash function used (in the actual process, an integer), and the size of the set represented by n. It is k, m, and n that jointly determine the misjudgment rate.

We assume that the Bloom Filter before compression is m bits, z is the optimal number of bits after Bloom Filter compression, and p is the probability that a bit in z is still 0. H(p) is the entropy function and

$$H(p)=-p\ log_2 p-(1-p)log_2(1-p) \tag{6}$$

Then, the Bloom Filter of m bits becomes mH(p) bits after compression. For a given n and z, it is necessary to choose the appropriate k and m, and in the case of satisfying $mH(p)\le z$, the false positive rate f of Compressed Bloom Filter can be minimized. Here is the entire push to process:

When m = z, at this time p = 1/2, this is the standard Bloom Filter structure, then

$$f\le 0.6185^{z/n} \tag{7}$$

The parameter k used at this time to design the Compressed Bloom Filter has the worst effect. Therefore, the parameters need to be adjusted to make

$$k=\alpha m/n \tag{8}$$

Therefore, $p = e^{-\alpha}$ can be obtained, and an appropriate one can be chosen that satisfies $m = z/H(e^{-\alpha})$ according to m and n. In order to ensure the integrity of the structure, the larger m is selected, the better, then the misjudgment rate of compressed Bloom Filter can be expressed as:

$$f = (1 - e^{-\alpha})^{\alpha z/(nH(e^{-\alpha}))} \tag{9}$$

Since the relationship between n and z n < z is fixed, the following formula can be maximized:

$$\beta = f^{(n/z)} = (1 - e^{-\alpha})^{\alpha/H(e^{-\alpha})} \tag{10}$$

If $\alpha = -ln\ x$, then there is:

$$\beta = (1 - x)^{-ln\ x/H(x)} = \exp\left(\frac{-ln\ x \cdot ln(1 - x)}{(-\log_2 e)(x ln\ x + (1 - x) ln(1 - x))}\right) \tag{11}$$

In formula (11), the larger the logarithmic value is, the larger the value of β is, so the smaller Y is at this time.

$$Y = \frac{x}{ln(1 - x)} + \frac{1 - x}{ln\ x} \tag{12}$$

The algorithm takes the derivative of γ with x as a variable:

$$\frac{d}{dx} = \frac{1}{ln(1 - x)} - \frac{1}{ln\ x} + \frac{x}{(1 - x) ln^2\ (1 - x)} - \frac{1 - x}{x ln^2\ x} \tag{13}$$

When x = 1/2, $\beta = 0$. When x < 12, $\beta < 0$. When $x > 1/2, \beta > 0$. Therefore, when x = 1/2, the probability of misjudgment rate of Compressed Bloom Filter is the largest, and the corresponding is $\alpha = ln\ 2$, then $z = mln\ 2$, and the space saving rate is about 30%.

2.3 Resource Information Registration

In order to register all resource information on the information server, the false positive rate f of Bloom Filter is as small as possible. If this method is adopted, the length m of the bit array must be determined by the number of registered resources stored in the GIS (Geographic Information System) that stores the largest amount of information. If it is assumed that the maximum number of resources is $n_{\max}$ f, according to the relationship among m, n, and

k when Bloom Filter obtains the minimum error rate f, it can be known that the length m of the bit array using this method is:

$$m = k \cdot n_{\max}/ln\ 2 \tag{14}$$

If it is assumed that the number of storage resources on the information server is n_p, then the length m of the bit array of Bloom Filter can be expressed as:

$$m = k.n_p/1n2 \tag{15}$$

The biggest advantage of designing the bit array length m of the Bloom Filter in this way is to maximize the space utilization of each information server without causing waste. In this paper, considering the computational complexity, program implementation, and other factors, the first method is chosen to construct the Bloom Filter structure of the information server. That is, all grid information servers use a uniform length of the bit array length m as the Bloom Filter.

The hash function is a function that establishes a correspondence between the element's key k and the element's storage location p, so that:

$$p = f(k) \tag{16}$$

The array combined by the storage location p is called a hash table. The element of k is directly stored in the unit whose address is f(k). If the element whose key is k is searched, the storage location f(k) can be calculated by using the hash function f. The algorithm can then fetch the element directly from that address.

The number of hash functions used to construct the Bloom Filter structure is determined according to the length m of the bit array of the Bloom Filter and the maximum number of resource registrations n that the information server node can bear. According to the formula

$$k = m1n2/n \tag{17}$$

When m/n = 10, the false positive rate of Bloom Filter is the lowest. At this time, the required number of hash functions is 7ln2, which is rounded up to 7. Therefore, this paper also follows this rule, and all Bloom Filter structures use 7 hash functions to hash resources.

In this paper, an information server node needs a copy of the Bloom Filter structure of other information server nodes cached in its routing table. Taking Figure 7 as an example, in this Chord ring, the identifier bit length of the node

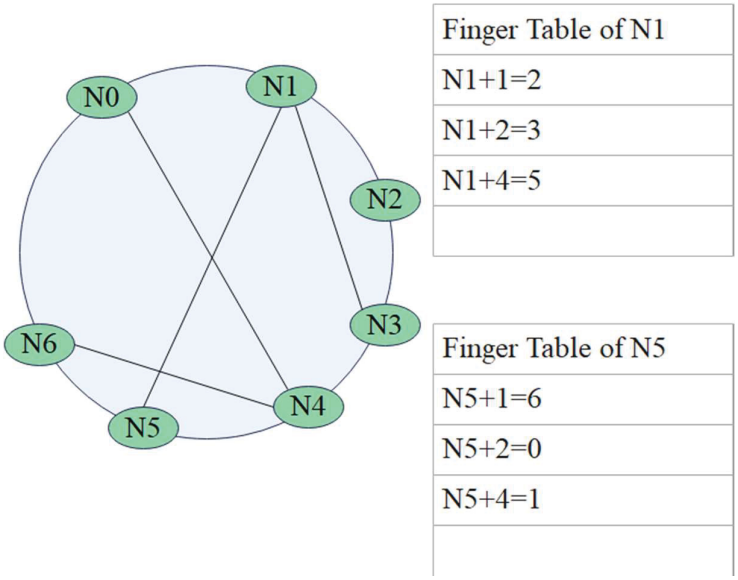


Figure 7 Chord ring and routing table with m = 3.

is m'3, and the total number of nodes is 7. If there are nodes N2, N3, and N5 in the routing table of N1, then N1 should cache the Bloom Filter structures of nodes N2, N3, and N5. N1 needs to send a request to cache its Bloom Filter structure to its neighbor nodes N2, N3, and N5. After the neighbor node receives the request, it must pass its own Bloom Filter structure to N1. Similarly, N6, N0, and N1 will pass their own Bloom Filter structure to the N5 node cache.

2.4 Resource Information Query

Resource query is a basic function of grid information service, and it is also the most commonly used service. If all users who join the grid system need a certain resource, they need to send a query request for the resource to the grid information server. When an information server node receives a query request, it can process the request according to the Chord protocol within the grid system.

The query process of this model is shown in Figure 8.

According to the above query steps, for the query request, the time spent T is the matching time of the Bloom Filter structure on the GIS (t1), the matching time in the resource database (t2) and the transmission time in the

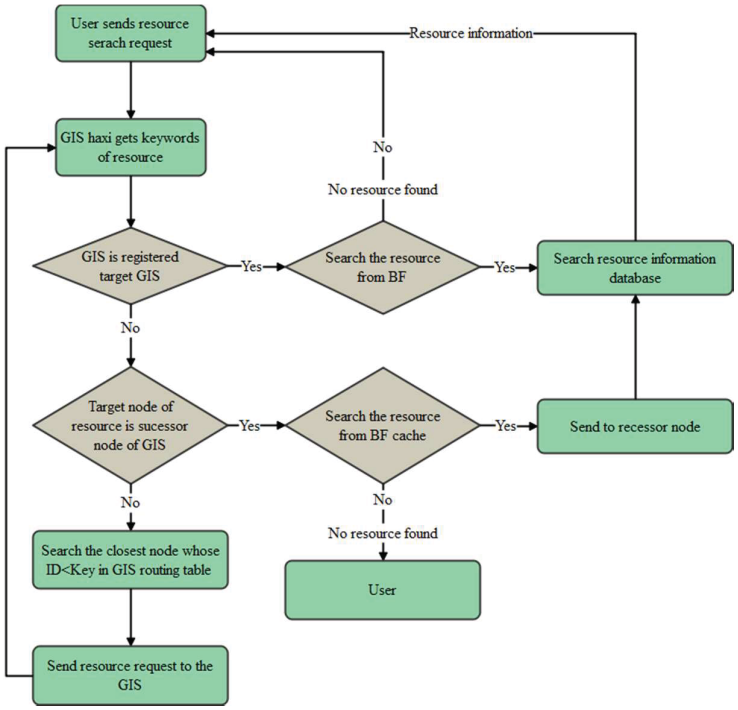


Figure 8 Resource query flow chart.

network (t3). The relationship between the three is

$$t1 < t2 < t3 \tag{18}$$

If it is assumed that there is a total of count query requests to be processed, there are x resources that can be successfully matched on the Chord ring based on Bloom Filter, y resources that fail to match, and z resources that are misjudged successfully, that is,

$$Count = x + y + z \tag{19}$$

Then, the total time T1 spent is:

$$T1 = x(t1 + t2) + yt1 + z(t1 + t2) + t3\ Count \tag{20}$$

If you search for count resources on a common Chord ring, the time T2 spent is:

$$T2 = t3\ Count + t2\ Count \tag{21}$$

At this point, the comparison between T1 and T2 is transformed into the comparison between $t1\,Count$ and yt2, that is, the time difference is related to the total number of queries, the number of failures, and t1 and t2. Of course, in the actual query process, the time spent and the specific value of the resources involved in the test also have a great relationship.

3 Campus Network Security Intrusion Detection System Based on Feature Segmentation and Deep Learning

The network intrusion detection model based on feature segmentation and cascaded random forest is shown in Figure 9. FS-CRF includes three modules: data set preprocessing module, feature segmentation module and cascaded random forest module.

The feature segmentation module will perform fine-grained segmentation on the data features in the standard dataset. After being processed by a single-layer random forest, the vectors will be re concatenated to generate transformed feature vectors as a new enhanced representation of the original data features. The cascaded random forest module realizes the classification

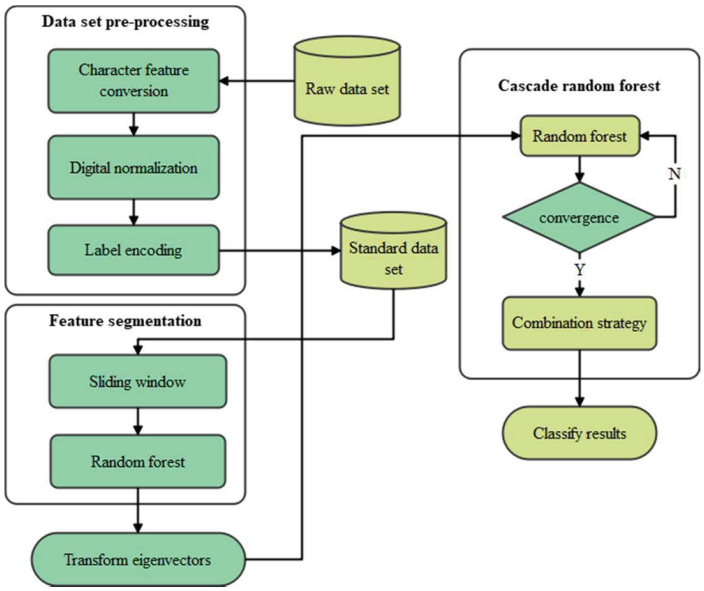


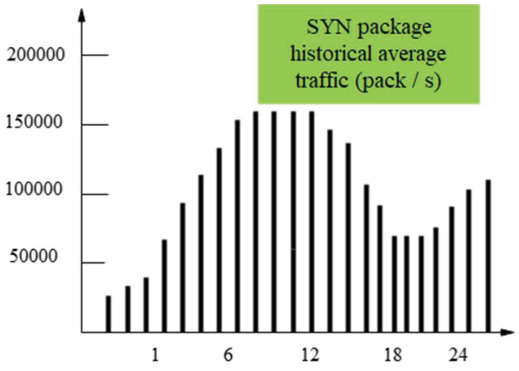
Figure 9 Campus network security intrusion detection system based on feature segmentation and deep learning.

and prediction of various behaviors in network traffic data, and finally adopts a voting method in the combination strategy section to obtain the final intrusion detection result

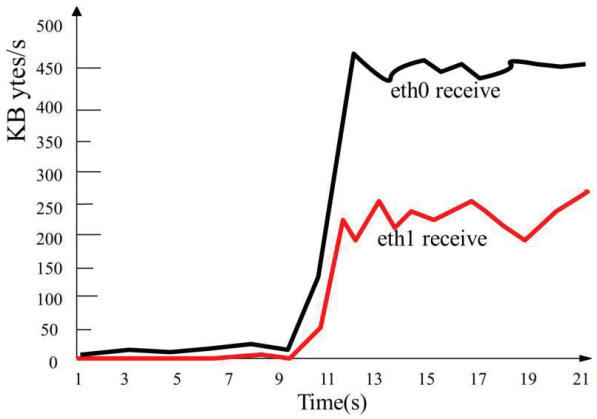
The dataset selected in this article is CICIDS2018, which was developed by the University of New Brunswick and the Canadian Institute of Cybersecurity, and is currently the most advanced network security dataset. These two datasets are close to the real network environment, not only containing the latest network attack scenarios, but also meeting all the standards of real-world network attacks. The data of CICIDS2018 includes 78 dimensional features and 13 traffic labels including normal behavior, and is divided into training and testing sets in a 3:2 ratio

On the basis of the above model, the effect of the model proposed in this paper is verified. In the test, TCP packets with SYN flag were detected, and the historical average traffic of SYN packets was 23210 packets/second. After running the program for a period of time, SYN Flooding tool was used for testing, and after the attack, it was found that the traffic of SYN packets rapidly increased to an abnormal state of 100000–130000 packets/second, The campus network contains servers of several application websites, and the server of the network training center is selected as the experimental object. In this paper, the experimental analysis is carried out in combination with the simulation research. Figure 10(a) is a graphical representation of several groups of data after the attack. The traffic growth in the figure is abnormally obvious, so the abnormal traffic phenomenon clearly indicates that there is an intrusion.

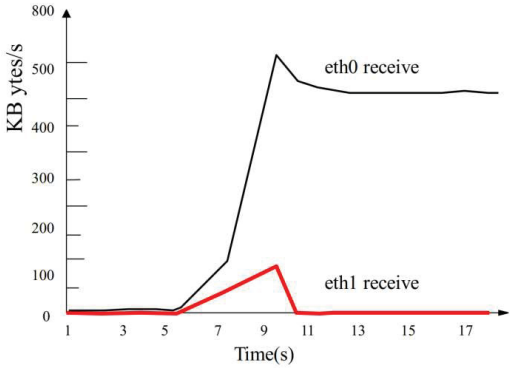
In terms of the detection of known attack intrusion, this paper tests the denial of service attack with the model proposed in this paper. The first request delay processing mechanism is used to prevent attacks. Figure 10(b) is the traffic diagram of the gateway when the intrusion detection system is not activated when the gateway is attacked. The eth0 network card is connected to the external network, and the ethl network card is connected to the internal network. Figure 10(c) activates the intrusion detection system. Comparing Figures 10(b) and 10(c), it can be found that the attack is resisted outside the gateway firewall. The response time for accessing the server through the gateway in the local area network is less than 1 second. When subjected to SYN Flooding attacks, if the intrusion detection system is activated, the response time for the first visit is 3 seconds, and the response time for subsequent visits is less than 1 second. However, accessing without starting the intrusion detection system always times out. Server response time isshown in Table 1.



(a) Abnormal traffic of attacks



(b) Gateway traffic when the campus network is attacked



(c) IDS response when the campus network is attacked

Figure 10 Simulation of the attack process.

Table 1 Server response time

Visits	1	2	3	4
When there is no attack	<1.0S	<1.0S	<1.0S	<1.0S
Start the IDS system	3.0S	<1.0S	<1.0S	<1.0S
Do not start the IDS system	Timeout	Timeout	Timeout	Timeout

Table 2 Comparison results of campus network security defense effectiveness

	Evaluate Results
The method of this article	96.5
The method of reference [21]	88.7
Method of reference [22]	91.2

After running the program for a period of time, the SYN Flooding tool was used for testing. After the attack, it was found that the traffic of SYN packets rapidly increased to an abnormal state of 100000–130000 packets/second, and related alarm messages were generated, indicating that the system has an ideal detection ability for intrusion behaviors such as denial of service attacks.

In order to further study the progressiveness of the model in this paper in campus network security intrusion detection, the model proposed in this paper is compared with the existing research. The design experiment compares the model in this paper with literature [21] and literature [22], and evaluates its defense attack effect through expert evaluation. The quantitative evaluation method is the hundred point system. The higher the score, the better the effect of resisting attacks. Finally, the evaluation results are shown in Table 2 below.

From the experimental results in Table 2, it can be seen that the campus network security model proposed in this paper can play an important role in the security protection of campus networks, and it also has certain advantages compared to existing research.

This article will verify the network security information service efficiency of our model, BloomFilter grid information service model (our model), and Chord grid information service model (existing model) from the following three aspects.

Figure 11 reflects the variation of request matching rates between existing models and our model with the number of information server nodes on the ring, given a fixed number of resource registrations (100000), a total of 2000 resource query requests, and a resource out of bounds rate of 20%. Among

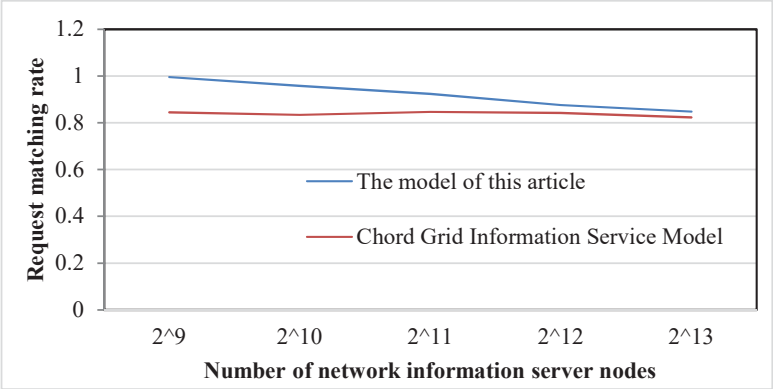


Figure 11 Changes in resource request matching rate with the number of information server nodes.

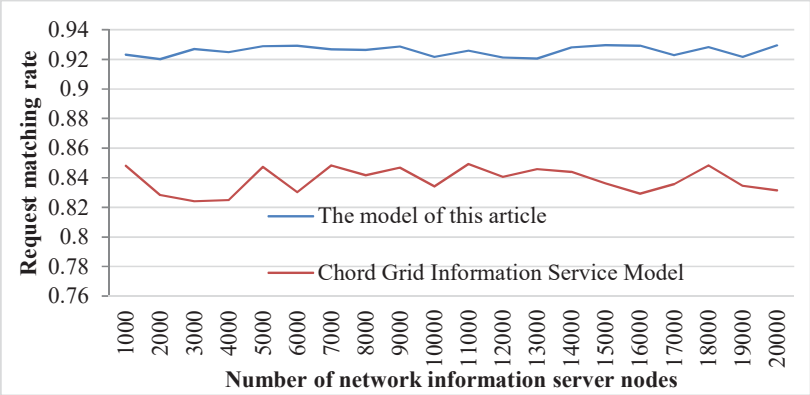


Figure 12 The change of request matching rate with the number of resource query requests.

them, the out of bounds rate refers to the percentage of resources that have not been registered on the server compared to the total number of query resources.

When the out of bound rate is constant, the relationship between the request matching rate of existing models is basically consistent with the out of bound rate. However, the request matching rate of our model decreases with the increase of the number of nodes, and ultimately is comparable to the request matching rate of existing models. This figure reflects that the information service performance of the model proposed in this article is better

Figure 12 shows the variation of resource request matching rate with the number of resource query requests, assuming a total of 211 information server

nodes and 100000 registered resources on the ring. From this graph, it can be seen that the request matching rate of the proposed model is higher than that of existing models, which also indicates that the GIS performance of the proposed model is better.

Through the above research, it can be seen that the campus network security intrusion detection model based on feature segmentation and deep learning proposed in this paper can effectively improve the effect of campus network security monitoring.

4 Conclusion

With the development of network applications, in order to have a secure network system, first of all, we should spend a lot of energy to formulate a thorough network security strategy, which is the most important step. In the implementation of network security, technology is only a means, and we should first have a clear concept of network security management, then formulate detailed security strategies, and finally select appropriate products for specific implementation and construction of security systems. The security threat of the campus network may come from outside or inside. Therefore, when designing a network security strategy, it is necessary to take security precautions at the boundary of the campus network to resist external intrusions and attacks, and to control the security access of various devices inside the campus network. In this way, legal or illegal users within the campus network can be prevented from damaging the campus network due to mistakes or intentional damage. This paper combines feature segmentation and deep learning technology to construct a campus network security intrusion detection model. Through experimental research, we can see that the campus network security intrusion detection model based on feature segmentation and deep learning can effectively improve the effect of campus network security monitoring. The method proposed in this article can not only be applied to campus network security, but also to the network security management of enterprises and other units, with certain scalability.

The feature segmentation model can effectively reduce the detection time for processing large-scale network traffic. Therefore, feature segmentation models have significant advantages over individual learning algorithms in improving detection rate, reducing false alarm rate, and reducing detection time. Currently, feature segmentation methods are receiving increasing attention in the field of network intrusion detection. Therefore, studying network

intrusion detection technology based on feature segmentation learning is of great significance.

Due to the limitations of experimental conditions, this paper only used single machine simulation for the proposed model. Due to the influence of bandwidth, delay and other conditions in the actual network, it is impossible to verify the actual efficiency of this model in the actual grid. The next step is to improve this model by using the improved Chord protocol with fewer search steps to organize information server nodes in the grid and the Bloom Filter structure that supports element deletion to store resource registration information. The focus is on strengthening the improvement of resource range search and expanding the scale of the experiment

References

[1] Li, D., Cai, Z., Deng, L., Yao, X., and Wang, H. H. (2019). Information security model of block chain based on intrusion sensing in the IoT environment. *Cluster computing*, *22*(1), 451–468.

[2] Parveen Sultana, H., Shrivastava, N., Dominic, D. D., Nalini, N., and Balajee, J. M. (2019). Comparison of machine learning algorithms to build optimized network intrusion detection system. *Journal of Computational and Theoretical Nanoscience*, *16*(5–6), 2541–2549.

[3] Pham, V., Seo, E., and Chung, T. M. (2020). Lightweight Convolutional Neural Network Based Intrusion Detection System. *J. Commun.*, *15*(11), 808–817.

[4] Subbarayalu, V., Surendiran, B., and Arun Raj Kumar, P. (2019). Hybrid network intrusion detection system for smart environments based on internet of things. *The Computer Journal*, *62*(12), 1822–1839.

[5] Molina-Coronado, B., Mori, U., Mendiburu, A., and Miguel-Alonso, J. (2020). Survey of network intrusion detection methods from the perspective of the knowledge discovery in databases process. *IEEE Transactions on Network and Service Management*, *17*(4), 2451–2479.

[6] Sharma, P., Sengupta, J., and Suri, P. K. (2019). Survey of intrusion detection techniques and architectures in cloud computing. *International Journal of High Performance Computing and Networking*, *13*(2), 184–198.

[7] Gifty, R., Bharathi, R., and Krishnakumar, P. (2019). Privacy and security of big data in cyber physical systems using Weibull distribution-based intrusion detection. *Neural Computing and Applications*, *31*(1), 23–34.

[8] Stergiopoulos, G., Chronopoulou, G., Bitsikas, E., Tsalis, N., and Gritzalis, D. (2019). Using side channel TCP features for real-time detection of malware connections. *Journal of Computer Security*, *27*(5), 507–520.

[9] Xue, Y. (2023). Machine Learning: Research on Detection of Network Security Vulnerabilities by Extracting and Matching Features. Journal of Cyber Security and Mobility, 12(05), 697–710.

[10] Safaldin, M., Otair, M., and Abualigah, L. (2021). Improved binary gray wolf optimizer and SVM for intrusion detection system in wireless sensor networks. *Journal of ambient intelligence and humanized computing*, *12*(2), 1559–1576.

[11] Bharathy, A. V., and Basha, A. M. (2017). A multi-class classification MCLP model with particle swarm optimization for network intrusion detection. *Sâdhanâ*, *42*(5), 631–640.

[12] Spathoulas, G., Theodoridis, G., and Damiris, G. P. (2021). Using homomorphic encryption for privacy-preserving clustering of intrusion detection alerts. *International Journal of Information Security*, *20*(3), 347–370.

[13] Young, C., Zambreno, J., Olufowobi, H., and Bloom, G. (2019). Survey of automotive controller area network intrusion detection systems. *IEEE Design & Test*, *36*(6), 48–55.

[14] Li, X. (2023). Construction of a Smart City Network Information Security Evaluation Model Based on GRA-BPNN. Journal of Cyber Security and Mobility, 11(06), 755–776.

[15] Rathore, M. M., Ahmad, A., Anand, P., and Rho, S. (2018). Exploiting encrypted and tunneled multimedia calls in high-speed big data environment. *Multimedia Tools and Applications*, *77*(4), 4959–4984.

[16] Park, S. T., Li, G., and Hong, J. C. (2020). A study on smart factory-based ambient intelligence context-aware intrusion detection system using machine learning. *Journal of Ambient Intelligence and Humanized Computing*, *11*(4), 1405–1412.

[17] Chen, L., Gao, S., Liu, B., Lu, Z., and Jiang, Z. (2020). THS-IDPC: A three-stage hierarchical sampling method based on improved density peaks clustering algorithm for encrypted malicious traffic detection. *The Journal of Supercomputing*, *76*(9), 7489–7518.

[18] Naveed Ahmed, N., and Nanath, K. (2021). Exploring Cybersecurity Ecosystem in the Middle East: Towards an SME Recommender System. Journal of Cyber Security and Mobility, 10(3), 511–536.

[19] Lagraa, S., Husák, M., Seba, H., Vuppala, S., State, R., and Ouedraogo, M. (2024). A review on graph-based approaches for network security monitoring and botnet detection. *International Journal of Information Security*, *23*(1), 119–140.

[20] Wei, K., Zang, H., Pan, Y., Wang, G., and Shen, Z. (2024). Strategic application of ai intelligent algorithm in network threat detection and defense. *Journal of Theory and Practice of Engineering Science*, *4*(01), 49–57.

[21] Ennaji, S., El Akkad, N., and Haddouch, K. (2023). i-2NIDS Novel Intelligent Intrusion Detection Approach for a Strong Network Security. *International Journal of Information Security and Privacy (IJISP)*, *17*(1), 1–17.

[22] He, K., Kim, D. D., and Asghar, M. R. (2023). Adversarial machine learning for network intrusion detection systems: A comprehensive survey. *IEEE Communications Surveys & Tutorials*, *25*(1), 538–566.

[23] Mohy-Eddine, M., Guezzaz, A., Benkirane, S., and Azrour, M. (2023). An efficient network intrusion detection model for IoT security using K-NN classifier and feature selection. *Multimedia Tools and Applications*, *82*(15), 23615–23633.

Biography

Zhe Chen was born in 1983 in Shandong, China. He received a bachelor's degree from Minzu University of China in 2005 and a master's degree from Xinjiang University of the Arts in 2020. His research interests include the performance and teaching of Chinese ethnic and folk dances. He is currently working at Shandong University of the Arts.