Hindawi
Scientific Programming
Volume 2022, Article ID 8587906, 11 pages
https://doi.org/10.1155/2022/8587906

Research Article

Design and Analysis of Network Virus Defense System Based on Multimodal Data Mining Technology

Jinhuan Hou[1] **and Fuqiang Ding**[2]

[1]*Department of Information Engineering, Shandong University of Engineering and Vocational Technology, Shandong 250200, China*
[2]*Assets and Equipment Management Office, Shandong University of Engineering and Vocational Technology, Shandong 250200, China*

Correspondence should be addressed to Jinhuan Hou; hjh0519@suet.edu.cn

Received 6 December 2021; Accepted 31 December 2021; Published 25 January 2022

Academic Editor: Sheng Bin

This study analyzes and designs a network security management system based on the K-means algorithm. Through the investigation of network security managers, this article uses structured modeling technology to derive the system's logical business functions, which are data acquisition function, data analysis function, and post-processing function. This study uses a three-tier architecture to design the system, describes the system data business processing flow, and gives the key flow of the K-means algorithm application. In addition, the system database is designed to provide support for system data processing. This study proposes a dual-network text matching model based on local interactive components. This model composes two modalities of single-modal short text data: a positional structural modal for describing local interactions and a global semantic understanding modal for extracting global semantic information. Two heterogeneous networks are used to extract two modalities. The principle of complementarity of modalities is realized by constructing the differences of each modal. At the same time, through the attention mechanism, the position information of the position structure modal is transferred to the global semantic understanding modal to obtain consistent comprehensive information so as to realize the principle of modal consistency. On this basis, by designing low-order interaction functions and high-order interaction functions, and using long- and short-term memory networks, we have, respectively, improved the position structure modal and global semantic understanding modal. The theoretical analysis and simulation results of the model show that the propagation characteristics of the network worm are completely determined by the threshold $R0$ required for its existence. When $R0 > 1$, the network worm will become popular even with a small initial infection number; conversely, even if the initial infection machine is large, the network worm will eventually become extinct. The research results of this article also show that the introduction of mobile devices has an important impact on the defense technology of network worms. To curb network worms that use both the network and mobile devices to spread, and to increase the proportion of machines that have never been infected with the virus, the most effective method is to control the number of mobile devices and reduce the possibility of cross-infection between mobile devices and machines. The simulation results show that, regardless of whether dynamic isolation and antivirus software are considered, the local area network-based choke method given in this study can effectively curb intelligent network worms that have slow scanning rates and clear scanning targets. In addition, the choke method presented in this article can determine the choke threshold without affecting the normal network scanning behavior of users in the local area network.

1. Introduction

With the continuous update and popularization of network information, people have more and more opportunities to use the network in life and work, which also makes network security a major issue that is urgently needed at the moment [1, 2]. Particularly in recent years, large-scale computer virus infections with great destructiveness have emerged one after another [3]. When the network virus is successfully activated, it will cause system data to be damaged and system resources

to be taken, interfere with system operation, damage computer hardware, and steal user information and data and other vicious behaviors [4]. In addition, computer viruses can also cause heavy workloads on optical drives, hard drives, printers, etc., thus shortening the working life. These malicious behaviors not only caused huge economic losses to the society, but also caused a certain amount of social panic, which seriously affected people's work and life [5]. The form of network security is the result of the interaction between the attacker (network virus) and the defender (network system administrator) to attack and defend. Due to the dynamic spread of network viruses, network security defense strategies are also dynamically changing. At present, the government and enterprises adopt traditional defensive measures, such as firewall technology, antivirus software, and intrusion detection system, in response to network security issues [6]. These are passive network security defense strategies that are implemented after an attack has occurred.

More and more frequent network viruses swallow up network resources, causing widespread network paralysis in a short period of time [7]. More and more hackers use the network to attack the target system, and by implanting viruses and other means to steal the target system's data and destroy the normal operation of the target system, it seriously threatens the data security of the computer system and severely reduces it [8]. As a network virus defender, you must proactively formulate defense strategies, adjust strategies in time according to the spread of network viruses, detect possible problems and loopholes in the entire system at any time, and predict possible attacks and the degree of harm. The purpose of the operating system security model is mainly to use the mandatory access control mechanism to improve the confidentiality and integrity of the operating system, without considering the virus defense function as the primary indicator, resulting in insufficient virus defense capabilities of the operating system developed therefrom [9]. The operating system is the cornerstone of computer system security. The lack of the ability of the operating system to defend against viruses directly affects the ability of the entire computer system to defend against viruses [10]. Therefore, it is of great significance to study operating system virus defense strategies and improve the ability of the operating system to defend against viruses.

Aiming at the current situation and existing problems of network security management, this article analyzes and designs the network security management system. We exported the main functions of the network security management system, which are data collection function, data analysis function, and security defense decision-making function. At the same time, it is designed for the network security management system database. We make full use of location information and semantic information and propose a double network text matching model (DNTMM) based on local interactive components. In the DNTMM, we construct the position structure modality and the global semantic understanding modality to obtain the complementarity of short text information and introduce the attention mechanism to achieve the interaction between the two modalities and obtain the consistency between the modalities' information. In the position structure mode, we enhance the position information through low-order interaction functions and high-order interaction functions. As far as we know, this is the first time that multimodal learning methods have been applied to short text matching tasks. Aiming at the characteristics of Stuxnet and Flame viruses, using mobile devices as a medium that can be cross-infected with machines, a network worm propagation model based on two means of media and network propagation is proposed. The dynamic method is used to derive the threshold conditions for the unpopularity of network worms in the model. On this basis, a strategy to control this type of network worm is given. Aiming at the problem that the threshold value is difficult to determine based on the fixed period choke method, this study uses the characteristics of the aggregation of access to network sites by LAN users and proposes a choke method based on the local area network, which solves the problem of the threshold value that is difficult to determine in the existing choke method.

2. Related Work

Regarding the interaction between network viruses, related scholars have studied the multi-worm propagation model, established an interaction model between two types of network worms that include a competitive relationship, and studied the influence of the competition-antagonism relationship on the spread of the two worms [11]. Relevant scholars analyzed the principle of the spread of social network worms based on the ideas of social engineering [12]. By studying the parameters that affect user action choices, they proposed a game model based on microscopic elements on the basis of user security thinking [13]. Based on the user's personal operating habits and behavior characteristics, a social network model that can analyze the spread of social network worms is established [14]. Researchers conducted research on worm countermeasures based on benign network worms [15]. The research results show that the introduction time of benign worms is a key factor affecting the effectiveness of benign worms on malicious worms.

The flow-based virus detection method is a very effective virus detection method. It makes full use of the difference in flow when there is no virus attack and when there is a virus attack to realize worm detection. Several features in the network can be composed of vectors, and different amounts of changes targeted by different virus attacks can be used to diagnose whether they have been attacked by viruses [16]. The difficulty of the flow-based detection method lies in the determination of the thresholds of various parameters, which affects the accuracy of the detection results. Setting the threshold too low will easily cause too many false alarms, and if it is set too high, it will easily cause leakage. The report fails to have the proper detection function. Compared with other virus detection technologies, the flow-based virus detection technology is not only effective for known viruses, but also can detect unknown viruses in a timely manner, changing the passive situation of virus prevention [17].

Relevant scholars have proposed the use of virtual honeypots to detect and block malicious code attacks [18].

The main implementation is to place multiple virtual honeypots at border gateways or places vulnerable to network worm attacks, and the honeypots can be shared with each other. The rule generator of network intrusion detection system (NIDS) is used to generate matching rules for network viruses. When the network worm scans the address space of vulnerable hosts according to a certain scanning strategy, the honeypot can capture the data of the network worm scanning attack and then adopt the characteristics to determine whether there is a network worm attack [19].

Through years of observation and research on viruses, people have found that viruses have some common special behaviors, and in normal programs, these behaviors are relatively rare, and this method of using the unique behavioral characteristics of viruses to detect viruses is called heuristics [20]. The advantage of this technology is that it can find most unknown viruses quite accurately, but the disadvantage is that it cannot identify the type of virus, it is difficult to implement, and it may alarm by mistake [21]. Behavior judgment can detect and kill unknown viruses. With this technology, most new viruses and variants can be discovered without the need for a virus database. The core idea of behavior judgment technology is to terminate the operation of the virus and deal with it before it causes substantial damage to the system [22–24]. Unknown virus detection is a comprehensive judgment of suspicious files based on virus behavior characteristics. After scanning files with virus characteristics, the system will give a suspicious probability [25, 26]. The higher the probability, the greater the possibility that the suspicious file is a virus.

3. Method

3.1. System Overall Architecture Design. The network security management system based on data mining can assign an account to the user. This account can enter the system through the login terminal, connect to the network security data acquisition module, realize data analysis and mining, export the results of system mining, and further configure and initiate defense measures. After loading the network security management system based on the K-means algorithm, you can save the relevant business data in the data, use the advanced Web logic business processing service for processing, transmit the message content in real time, and then use the advanced enterprise service bus for processing. The network security management system based on the K-means algorithm can integrate various types of network log data, UDP data, TCP data, equipment operation data, etc., and can divide the results of data analysis into different levels to provide users. Human-computer interaction allows users to directly communicate with the system and realize powerful information processing and processing services, and output-related operational service capabilities, further improving the friendliness and simplicity of interactive services of the network security management system, as shown in Figure 1.

The application layer of the network security management system mainly provides users with an integrated operation service function, which can integrate application software such as data acquisition functions, data analysis functions, and post-processing functions. It uses the ESB to monitor each application's data access request and then sends it to the big data center. Every logical function of the application layer can be implemented using Web services, which can encapsulate data requests and feed them back to users.

The data layer can provide a location for information storage, processing, and management for the network security management system, forming distributed and scalable application software, and flexibly using database server resources. The big data center has introduced technologies such as index, index type, ESB, and data type. These technologies can expand and enrich database functions.

3.2. System Module Function Design. With the development and improvement of Internet technology, more and more fields have begun to introduce network security management systems. Therefore, to improve network security defense capabilities, it is necessary to effectively solve the problem of large amounts of data and excessive data. The network security management system uses the advanced K-means algorithm to build a data analysis engine that can mine potential viruses or Trojan horse genes in the data to build a highly automated network security management system. The system can handle unstructured data resources, realize application data analysis, and result report generation. The network security management system based on the K-means algorithm is shown in Figure 2.

The back-end logical business processing of the network security management based on the K-means algorithm mainly includes two components: an offline module and a database. Among them, the offline module can complete the current functions of network raw data collection, preprocessing, and data mining, so it is also called the network data acquisition stage. The front-end real-time interaction part is mainly composed of online modules, which can realize real-time interaction between users and network security management rules. The offline module can preprocess, integrate, and mine data resources such as network transmission data, equipment operating data, UDP data, and TCP data collected by the online module and find potential viruses or Trojan horses in the data set and store them in the Trojan horse. These rules and knowledge can effectively support the recommendation service work of the automatic recommendation engine of the online module. The offline module of the network security management system is the basic support module of the online module, which can realize the Trojan horse or virus security management function.

In the process of implementing the network security management system, this study adopts the principle of modular design. The back-end logic business processing and the front-end real-time interaction are logically separated, and the data mining and analysis functions in the back-end logic business processing are implemented to ensure the adaptation of the network security management system. Modern large-scale network data real-time transmission needs to ensure the real time and accuracy of network security management system services.

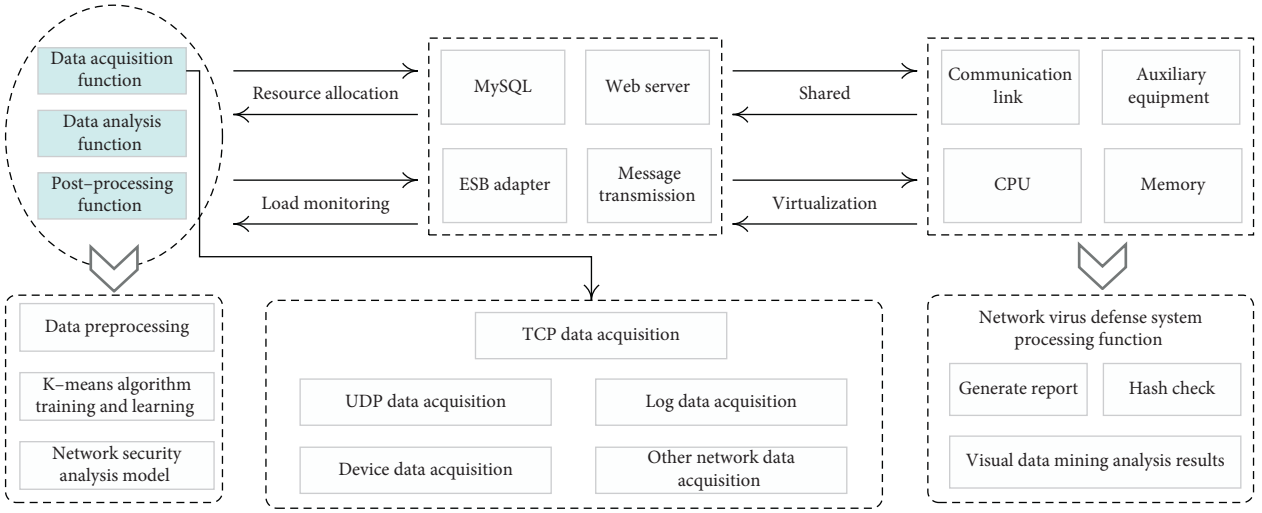

Figure 1: Overall architecture of the network security management system.

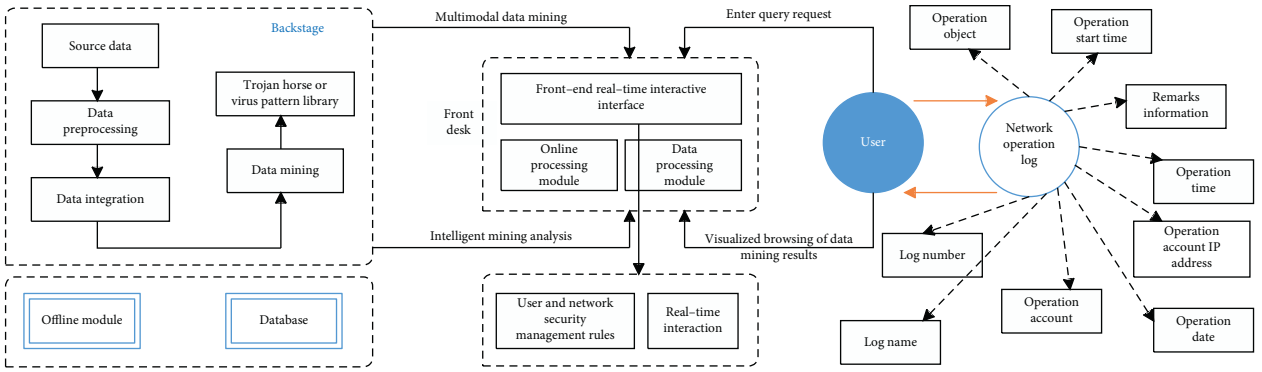

Figure 2: Network security management system structure based on the K-means algorithm.

3.3. Database Table Design. The fields of the security administrator table are the administrator code, name, gender, mailing address, contact phone number, and certificate number, as shown in Table 1.

The fields of the security policy table are security policy number, security policy name, administrator code, usage authority, whether to disable, security policy description, security policy establishment date, security policy running time, version number, etc., as shown in Table 2.

The fields of the network operation log table are log number, log name, operation account, operation date, operation object, operation duration, remarks, operation start time, and operation account IP address, as shown in Table 3.

3.4. Algorithm Design. We define the two short texts to be matched as T_1 and T_2, respectively, where $t_{i\ j} \in R_{1\times d}$ is the word vector of the jth word of T_i. Then, we can express the short text as $T_i = [tTi_1, \ldots, tTi_n]^T$, where n represents the number of words in the text T_i. We denote the DNTMM with the symbol $s(\cdot)$, the positional structure modal network with the symbol $sl(\cdot)$, and the global semantic understanding modal network with $sg(\cdot)$. In the position structure modal network, G represents the low-order interaction matrix, and $G(S)$ represents the high-order interaction matrix we proposed based on G.

Our goal is to propose a comprehensive short text matching model based on multimodal learning, which is basically divided into three purposes:

(1) This study constructs the monomodal problem of short text matching into a multimodal problem to ensure the consistency and complementarity between the modalities.

(2) This article solves the problems of insufficient information obtained by location structure, ignoring context in the extraction of semantic structure information, and the integration of location structure information and semantic structure information is too simple.

(3) The overall result needs to be better than the result of the duet model and better than the result of the current best algorithm

We define low-order interactive functions to obtain basic text position information. The low-order interaction function $G(\cdot)$ of text $T_1 \in R_{n1\times d}$ and text $T_2 \in R_{n2\times d}$ is specifically defined as follows:

Table 1: Security administrator table.

Field name	Type of data	Is it empty	Primary key	Meaning description
Coding	Integer	No	Yes	Administrator code
Gender	Character type	No	No	Gender of the administrator
ID number	Integer	Yes	No	Administrator's ID number
Name	Integer	No	No	Administrator's name
Contact number	Character type	No	No	Administrator's contact number
Mailing address	Integer	Yes	No	Administrator's mailing address

Table 2: Security policy table.

Column name	Type of data	Is it empty	Primary key
Security policy runtime	Long integer	No	Yes
Security policy establishment date	Long integer	No	Yes
Usage rights	Character type	No	No
Security policy name	Long integer	No	Yes
Version number	Integer	No	No
Whether to disable	Character type	No	No
Security policy number	Integer	No	No
Security policy description	Character type	No	Yes
Administrator code	Long integer	No	No

Table 3: Network operation log table.

Column name	Type of data	Is it empty	Primary key
Operation start time	Long integer	No	Yes
Operation account IP address	Integer	No	No
Operation date	Long integer	No	No
Log name	Character type	Yes	No
Remark information	Date type	No	No
Operation object	Integer	No	Yes
Operation time	Character type	No	No
Log number	Long integer	Yes	Yes
Operation account	Character type	No	No

$$G(T_1, T_2 : S) = h\left(T_1^T S T_2\right) - 0.5\alpha S. \tag{1}$$

Among them, $S \in R_{d\times d}$ is a symmetric matrix that reweights different dimensions in the text word vector, h is the scaling function, α is the truncation parameter, and $[\cdot]+$ is the hinge operator.

We use low-order interaction functions to calculate the initial position information between short texts. Compared with the traditional exact matching matrix, the low-order interaction function makes the interaction more general and pays more attention to the positional relationship between words in the short text. The hinge function in the low-order interaction function can be used to discard small position similarity. Since the similarity is small, it can be considered that there is no obvious positional relationship between the words. Then, we can consider that a smaller position similarity has less impact and can be ignored. In this way, the time for matrix operations can be greatly reduced. In the experiment, we initialize S as the identity matrix.

We further propose a high-order interaction function, so that the algorithm can further mine the position structure information between short texts on the basis of the low-order interaction function. Here, we need to introduce a tensor algebra operation-Kronecker product, which is defined as follows.

The Kronecker product of $A \in R_{n1\times m1}$ and $B \in R_{n2\times m2}$ is defined as $A \otimes B$. This operation is a special form of tensor product. The specific form is as follows:

$$A \otimes B = \begin{pmatrix} a_{00}B & \cdots & a_{0(m1-2)}B & a_{0(m1-1)}B \\ \vdots & \vdots & \vdots & a_{1(m1-1)}B \\ a_{(n1-2)0}B & \cdots & \cdots & a_{(n1-2)(m1-1)}B \\ a_{(n1-1)0}B & \cdots & a_{(n1-1)(m1-2)}B & a_{(n1-1)(m1-1)}B \end{pmatrix}. \tag{2}$$

It can be seen from the definition that the matrix size of $A \otimes B$ is $n1n2 \otimes m1m2$. On the basis of the Kronecker product, we give the definition of high-order interaction function.

The high-order interaction function $G'(\cdot)$ based on the low-order interaction function $G \in R_{n1\times n2}$ is defined as follows:

$$G'(G, D) = (G \otimes G \odot G) \bullet h(E \odot D + D \otimes G). \tag{3}$$

Among them, $\odot$ is bitwise division, $E \in R_{n1\times n2}$ is a matrix of all ones, and $D \in R_{n1\times n2}$ is a position matrix satisfying $D_{i\ j} = i - j$, and $\hbar(\cdot)$ is a smoothing function.

The numerator of the high-order interaction function represents the comprehensive consideration of the low-order positional correlation between the word pairs, and the denominator represents the positional relationship between the word pairs. In this way, the location structure information is further reflected.

For two pairs of words ($t11$, $t22$) and ($t14$, $t23$), their low-order interaction function values are G_{12} and G_{43}, respectively. Then, the similarity of these two word pairs can be represented by $G_{12}{\cdot}G_{43}$, which is just a comprehensive interaction between two pairs of words based on a low-order interaction function. Next, we need to design the relative positional distance between the two word pairs as $d\ (t_{11}, t_{22})$—$d\ (t_{14}, t_{23})$. Using the commutative and associative laws of subtraction, we find that from (t11, t22) to (t14, t23), the change in location similarity is regular.

We calculate the position similarity of the feature matrix obtained by the high-order interaction function, and the specific expression is as follows:

$$\begin{aligned} s_l(T_1, T_2) &= g_{L-1}W_{L-1} - b_L, \\ s.t.g_0 &= Line(G'), \\ g_i &= relu(g_iW_{i-1} - b_i \bullet b_{i-1}) i \in [1, L-1]. \end{aligned} \tag{4}$$

Among them, W_i is the parameter matrix of MLP, b_i is the bias term of MLP, L is the number of MLP network layers, relu(·) is the activation function used by MLP, and Line(·) means stacking the high-order interactive feature matrix into 1-dimensional vector.

In the field of short text matching, we usually think that the importance of words in the text is different, so the attention mechanism can be widely used in this field. It is worth noting that we hope to pass the information of the positional structure modalities into the global semantic understanding modalities through weights to achieve interaction, so as to obtain the consistency information between the modalities.

$$a_{ij} = s_j \bullet W_i \bullet \prod_{k=0}^{n_i-1} \exp(s_k \bullet s_{k+1} \bullet W_i). \tag{5}$$

where $a_{i\ j}$ represents the attention mechanism weight of the jth word of the ith short text, W_i represents the learnable parameter matrix of the ith short text, s_j represents the characteristics of the jth word, and n_i represents the word of the ith short text number.

We need to normalize the weights of the learned attention mechanism and generate the respective integrated features of each short text:

$$\begin{aligned} \gamma_{1i} &= e^{a_{0i}} \bullet \prod_{k=0}^{n_1-1} e^{a_{0k}}, \\ \gamma_{2j} &= e^{a_{1j}} \bullet \prod_{k=0}^{n_2-1} e^{a_{1k}}, \\ h_0' &= \prod_{k=0}^{n_1-1} (h_{0k}\gamma_{0k}), \\ h_1' &= \prod_{k=0}^{n_2-1} (h_{1k}\gamma_{1k}). \end{aligned} \tag{6}$$

Among them, γ_{1i} and $\gamma_{2\ j}$ are the normalized attention mechanism weight values of the first paragraph of short text and the second paragraph of short text, respectively.

After integrating the features through the attention mechanism method, further learning is carried out in the global semantic understanding modality. We analyze the importance of the position of each word in the short text in the position structure mode, ignore those positions that are not particularly important, and transfer this important information to the global semantic understanding mode through the attention mechanism. The readers get a certain amount of interaction in information to achieve the purpose of modal consistency, and then, we can better match short texts without much deviation in global semantic understanding. The DNTMM takes into account the weighting efficiency of the attention mechanism, so the low-order interaction matrix G is selected to generate the attention mechanism weights. Compared with G, the high-order interaction matrix is more sparse, and G not only has low computational cost, but also contains location information to a certain extent, making it easier to weight words.

The DNTMM realizes the complementarity of the modals by constructing the position structure modal and the global semantic understanding modal and obtains the consistency of the modal through the attention mechanism. In the end, we use additive aggregation to make the final prediction. The formula is as follows:

$$s(T_1, T_2) = (w_1, w_2) \begin{pmatrix} s_g(T_1) \\ s_l(T_2) \end{pmatrix}. \tag{7}$$

Among them, w_1 and w_2 are the parameter matrices of the additive aggregation method.

Since the short text matching problem has different tasks, we need to use different loss functions. For classification and recognition tasks, we use the cross-entropy loss function for training, and the loss function is as follows:

$$L_{Class}(q,d) = (y-1)\log[\sigma(s(q,d))+1] + y\log[\sigma(s(q,d))]. \quad (8)$$

Among them, q and d are a pair of short texts to be matched, y is the true mark of the matching correlation between q and d, and $\sigma(\cdot)$ is the sigmoid activation function. For the question and answer (QA) recommendation task, we use the triple loss function for training, and the loss function is as follows:

$$L_{QA}(q,d-,d+) = Max\{(y-1)\log[\sigma(s(q,d+))+1] - y\log[\sigma(s(q,d-))], -1\}. \quad (9)$$

Among them, q represents the short text question, and $d+$ and $d-$, respectively, represent the positive and negative answers of the short text.

4. Results and Analysis

4.1. Defense against the Spread of Network Worms. This section simulates and analyzes the control strategy of network worms based on mobile devices and network propagation. The results show that the convergence of the simulation algorithm is guaranteed. Since this section focuses on the impact of mobile devices as a means of spreading network worms on the spread and control of network worms, only the parameters related to mobile devices are simulated, including the number of mobile devices, the infection rate of mobile devices, and the recovery rate of mobile devices. Given that the elimination rate of mobile devices is usually much smaller than the recovery rate, we have not simulated the impact of the elimination rate on network worms.

At the initial moment, the number of machines already immune to network worms (i.e., the value of R) is large. This is mainly because scanning network worms can only infect machines with a specific vulnerability, and machines without the vulnerability are immune to the virus. Therefore, the immunized person at the initial moment is not a machine recovered from an infected person, but a machine without loopholes.

In the study of network worms, in addition to paying attention to the number of infections and peaks at each moment, a more noteworthy quantity is the cumulative number of infections at the last moment or the number of machines that have not been infected with the virus at the last moment. Before, we only considered the number of infections at each moment. In contrast, in this section, we simulated the number of machines that were not infected with network worms at the last moment.

Figure 3 shows the relationship curve between the proportion of susceptible machines at the last moment and the proportion of mobile devices and machines. Figure 4 shows the relationship curve between the proportion of susceptible machines at the last moment and the infection rate of mobile devices. Figure 5 shows the relationship curve between the proportion of susceptible machines at the last moment and the recovery rate of mobile devices.

For network worms that can be spread with the help of mobile devices, the scan rate or infection rate on the network is relatively small. As can be seen from Figures 3–5, for a small network infection rate, there are a threshold for the proportion of mobile devices and machines, a threshold for mobile device infection rate, and a threshold for mobile device recovery rate. The number of infected machines is close to the total number of susceptible machines at the initial moment. This means that network worms are directly extinct.

Figure 3 shows that controlling the number of mobile devices (corresponding to reducing the value of M-to-N ratio) helps to curb the spread of network worms. However, controlling the number of machines will have the opposite effect. In addition, Figures 4 and 5 show that reducing the chance of cross-infection of mobile devices or disinfecting mobile devices in time can also help contain network worms.

4.2. Simulation Analysis of LAN Choke. Assume that there are no patches for network worms and the spread rate is constant. The assumption that there is no patch is easier to understand, because the choke method is a method to automatically contain the spread of the worm before analyzing the characteristics of the network worm. It is obviously impossible to develop a patch without analyzing the characteristics of the network worm. The constant spread rate is a common assumption for scanning network worms.

In the simulation, we assume that the number of all LAN network users is 400, and there is only one proxy server (corresponding to a given public network reachable IP address). Assume that the susceptibility rate of the proxy server in the entire network is 0.005, and when a proxy server is infected by a network worm, the internal machines are susceptible to the network worm with a probability of 50%. Since the spread of network worms in the LAN is much faster than the speed of external infections, we ignore the time spent by the worms in the LAN.

The threshold MS is 120000~240000, respectively. The scanning frequency of the suspicious network is set to the same scanning frequency as Bell Labs. Figure 6 shows the simulation results.

Figure 6 shows that the number of infected LANs during the choke cycle is not more than 1,400, which has a significant choking effect. This article further examines the influence of the value of Sf (the scan rate specified for suspicious networks) on the spread of network worms. Figure 7 shows the propagation situation when Sf is 150~750.

It can be seen from Figure 7 that the number of LAN infections at the peak point is less than 1150; the number of LAN infections will fluctuate with the increase in Sf. Figures 6 and 7 both show the number of infected LANs, not the

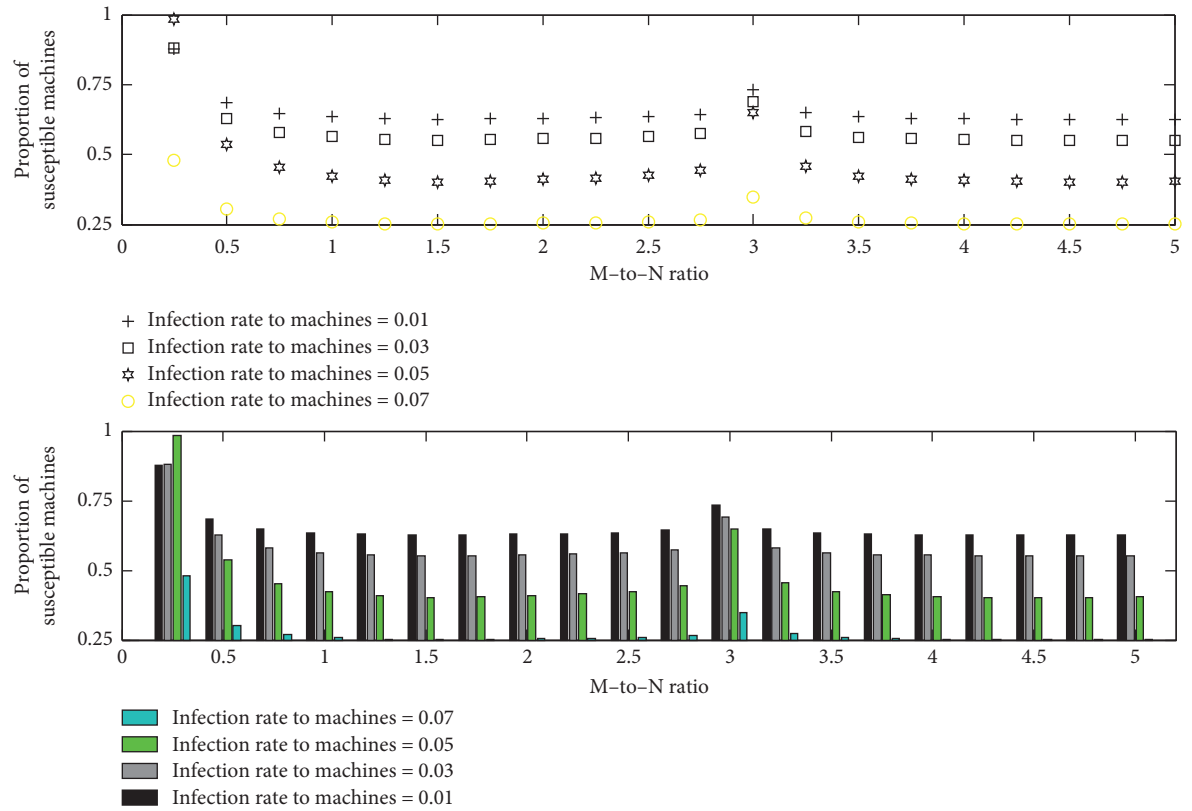

Figure 3: Relationship between the final proportion of susceptible machines and the value of the M-to-N ratio.

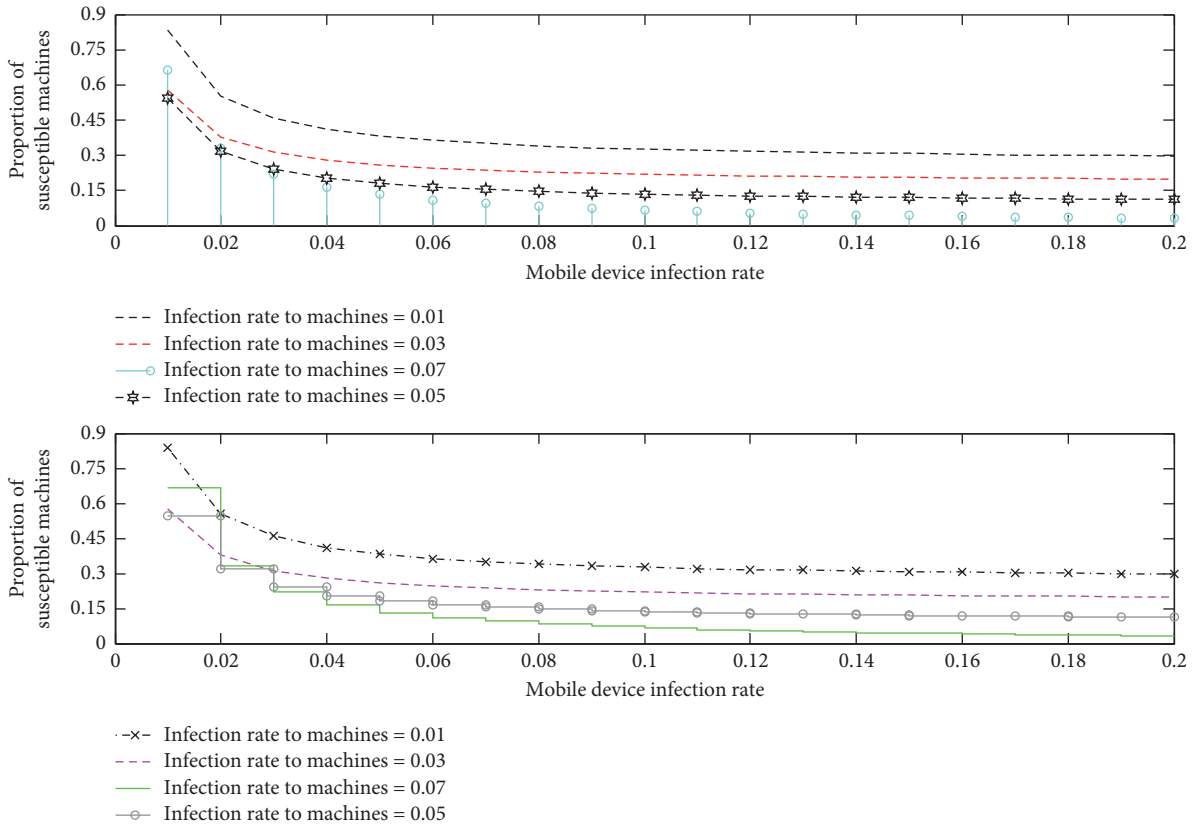

Figure 4: Relationship between the final proportion of susceptible machines and the infection rate of mobile devices.

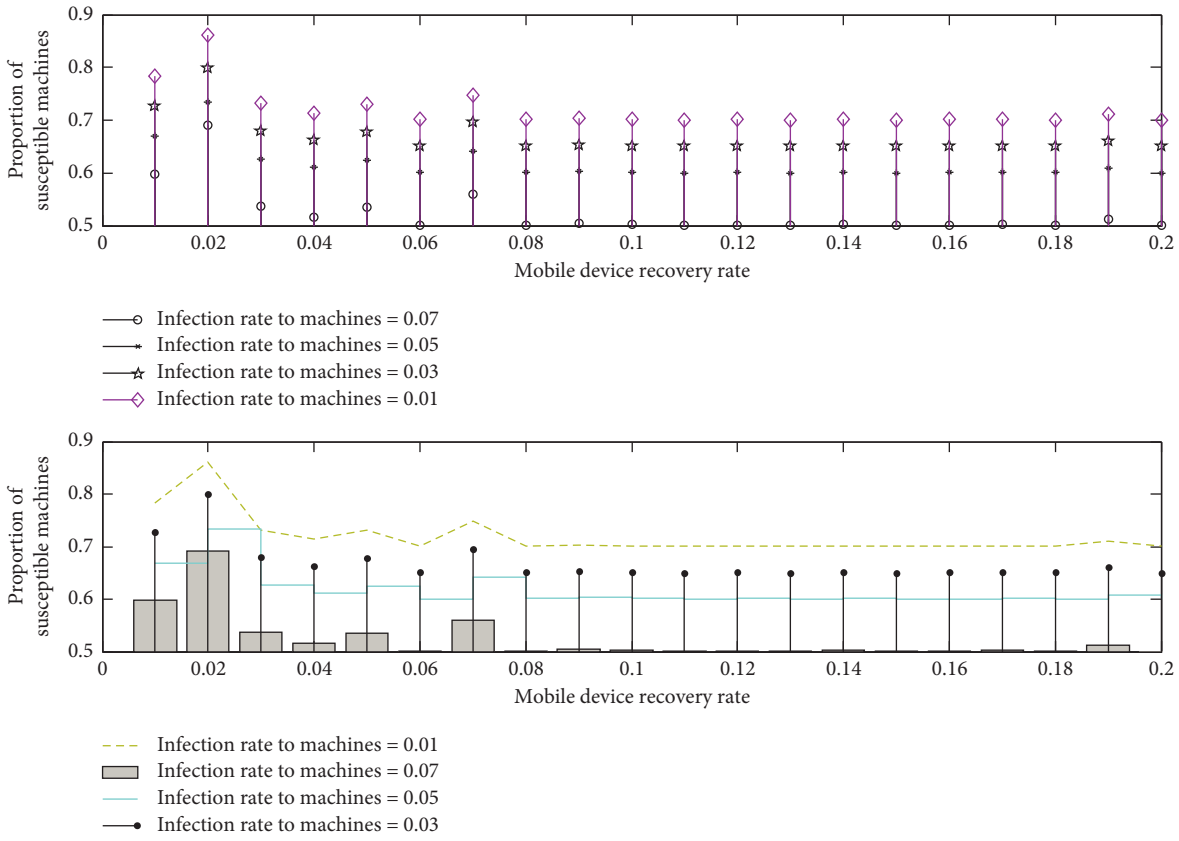

Figure 5: Relationship between the final proportion of susceptible machines and the recovery rate of mobile devices.

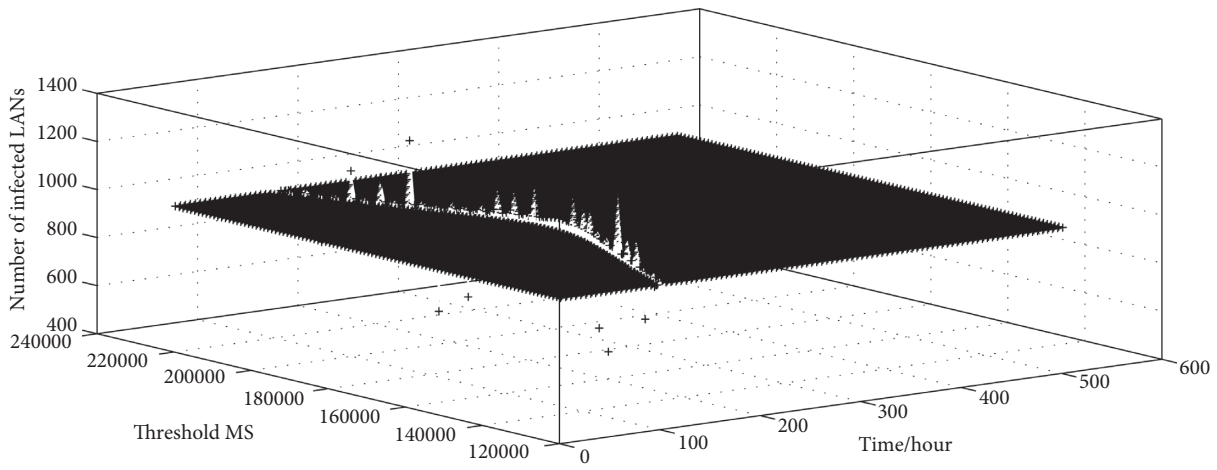

Figure 6: Number of LANs infected during the choke cycle, and the initial number of infections is 1.

number of machine nodes. However, in this article, the number of machine nodes in the local area network is fixed, so it is easy to change the number of infected machine nodes. In addition, according to the preset maximum number of LANs (or the number of machine nodes) allowed to be infected, it can provide a basis for the selection of MS and Sf. After the suspicious network is identified, it is clear that stronger isolation measures can be taken. In addition, random scans of antivirus software may also find existing system vulnerabilities and apply patches.

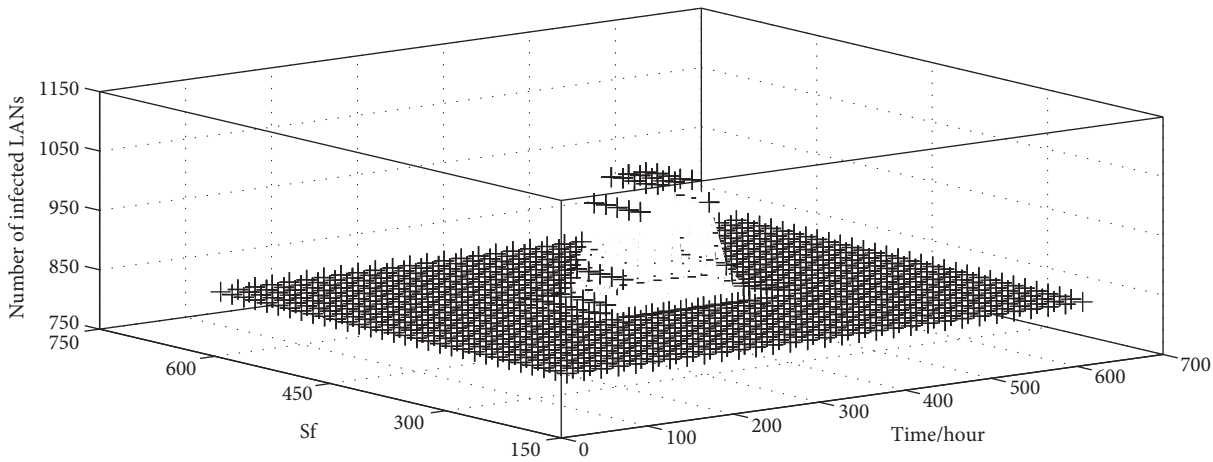

Figure 7: Number of LANs infected during the choke cycle, and the initial number of infections is 1, MS = 140000.

5. Conclusion

This study analyzes the situation of network security management and derives the functions of the network security management system, which are data acquisition function, data analysis function, and post-processing function. At the same time, the overall structure and detailed functions of the system were designed, and the database was modeled. In this study, the monomodal short text data are composed of two modalities: the position structure modal used to describe local interactions and the global semantic understanding modal used to extract global semantic information. The two modalities are extracted by two heterogeneous networks. The DNTMM realizes the principle of complementarity of modalities by constructing the differences in each modal. At the same time, through the attention mechanism, the position information of the position structure modal is transferred to the global semantic understanding modal to obtain consistent comprehensive information so as to realize the principle of modal consistency. In the position structure mode, this study uses the Kronecker product to construct a high-order interaction function to further obtain the position information of the position mode. Aiming at network worms that can spread through the network and through mobile devices, this study proposes a network worm propagation model based on media and network propagation. Based on the proposed mathematical model, the dynamic method is used to derive the threshold conditions for the unpopularity of network worms in the model, and the influence of the use of media on the threshold conditions for the unpopularity of network worms is analyzed. For control, both network transmission and mobile devices can be used. Media-borne network worms provide theoretical guidance. Aiming at intelligent network worms that have both slow scanning rate and specific scanning targets, this study proposes a local area network-based choke method. Using the aggregation of users in the same local area network to access network IP addresses and the randomness of intelligent network worms accessing network IP addresses, the cumulative effect over a period of time solves the problem of scanning based on the scanning rate and the number of scanning in a fixed period. The difficulty of determining the choke threshold in the choke method (for intelligent network worms) provides a feasible solution for deploying the choke method on the LAN border routers, which can effectively curb the spread of intelligent network worms.

Data Availability

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

Conflicts of Interest

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

References

[1] I. J. Cruickshank and K. M. Carley, "Analysis of malware communities using multi-modal features," *Ieee Access*, vol. 8, pp. 77435–77448, 2020.

[2] S. Peng, G. Wang, Y. Zhou et al., "An immunization framework for social networks through big data based influence modeling," *IEEE Transactions on Dependable and Secure Computing*, vol. 16, no. 6, pp. 984–995, 2017.

[3] R. Priyadarshini, R. K. Barik, H. C. Dubey, and B. K. Mishra, "A survey of fog computing-based healthcare big data analytics and its security," *International Journal of Ambient Computing and Intelligence*, vol. 12, no. 2, pp. 53–72, 2021.

[4] P. Wang, L. T. Yang, J. Li, J. Chen, and S. Hu, "Data fusion in cyber-physical-social systems: state-of-the-art and perspectives," *Information Fusion*, vol. 51, pp. 42–57, 2019.

[5] W. Zhang, J. Li, Y. Wen, and Y. Luo, "Toward a wearable crowdsource system to monitor respiratory symptoms for pandemic early warning," *IEEE Network*, vol. 35, no. 3, pp. 56–63, 2021.

[6] A. Dushimimana, T. Tao, R. Kindong, and A. Nishyirimbere, "Bi-directional recurrent neural network for intrusion detection system (IDS) in the internet of things (IoT)," *International Journal of Advanced Engineering Research and Science*, vol. 7, no. 3, pp. 524–539, 2020.

[7] T. G. Kim, B. J. Kang, M. Rho et al., "A multimodal deep learning method for android malware detection using various features," *IEEE Transactions on Information Forensics and Security*, vol. 14, no. 3, pp. 773–788, 2018.

[8] S. Shayaa, N. I. Jaafar, S. Bahri et al., "Sentiment analysis of big data: methods, applications, and open challenges," *IEEE Access*, vol. 6, pp. 37807–37827, 2018.

[9] D. Goel and A. K. Jain, "Mobile phishing attacks and defence mechanisms: state of art and open research challenges," *Computers & Security*, vol. 73, pp. 519–544, 2018.

[10] G. Donkal and G. K. Verma, "A multimodal fusion based framework to reinforce IDS for securing Big Data environment using Spark," *Journal of information security and applications*, vol. 43, pp. 1–11, 2018.

[11] D. Wang, W. Yu, and X. Zou, "Tensor-based mathematical framework and new centralities for temporal multilayer networks," *Information Sciences*, vol. 512, pp. 563–580, 2020.

[12] W. Tang, B. Hui, L. Tian, G. Luo, Z. He, and Z. Cai, "Learning disentangled user representation with multi-view information fusion on social networks," *Information Fusion*, vol. 74, pp. 77–86, 2021.

[13] J. Kusyk, M. U. Uyar, and C. S. Sahin, "Survey on evolutionary computation methods for cybersecurity of mobile ad hoc networks," *Evolutionary Intelligence*, vol. 10, no. 3, pp. 95–117, 2018.

[14] D. Bao, O. Ganbaatar, X. Cui et al., "Down-regulation of genes coding for core RNAi components and disease resistance proteins via corresponding microRNAs might be correlated with successful Soybean mosaic virus infection in soybean," *Molecular Plant Pathology*, vol. 19, no. 4, pp. 948–960, 2018.

[15] X. Wang, S. Yin, H. Li, J. Wang, and L. Teng, "A network intrusion detection method based on deep multi-scale convolutional neural network," *International Journal of Wireless Information Networks*, vol. 27, no. 4, pp. 503–517, 2020.

[16] S. Mostafavi and W. Shafik, "Fog computing architectures, privacy and security solutions," *Journal of Communications Technology, Electronics and Computer Science*, vol. 24, pp. 1–14, 2019.

[17] Y. Lu, "Artificial intelligence: a survey on evolution, models, applications and future trends," *Journal of Management Analytics*, vol. 6, no. 1, pp. 1–29, 2019.

[18] Q. Zhu, L. Li, and C. Gan, "Modeling and analysis of the impact of adaptive defense strategy on virus spreading," *IAENG International Journal of Applied Mathematics*, vol. 48, no. 2, pp. 1–6, 2018.

[19] C. Zuo, "Defense of computer network viruses based on data mining technology," *International Journal on Network Security*, vol. 20, no. 4, pp. 805–810, 2018.

[20] M. Xiao and M. Guo, "Computer network security and preventive measures in the age of big data," *Procedia Computer Science*, vol. 166, pp. 438–442, 2020.

[21] Q. A. Dang and M. T. Hoang, "Numerical dynamics of nonstandard finite difference schemes for a computer virus propagation model," *International Journal of Dynamics and Control*, vol. 8, no. 3, pp. 772–778, 2020.

[22] S. Armenia and G. Tsaples, "Individual behavior as a defense in the "war on cyberterror": a system dynamics approach," *Studies in Conflict & Terrorism*, vol. 41, no. 2, pp. 109–132, 2018.

[23] Y. Wu, H. Huang, Q. Wu, A. Liu, and T. Wang, "A risk defense method based on microscopic state prediction with partial information observations in social networks," *Journal of Parallel and Distributed Computing*, vol. 131, pp. 189–199, 2019.

[24] V. MadhuSudanan and R. Geetha, "Dynamics of epidemic computer virus spreading model with delays," *Wireless Personal Communications*, vol. 115, no. 3, pp. 2047–2061, 2020.

[25] G. Chenquan, F. Qingdong, Z. Qingyi, Z. Zufan, Z. Yushu, and X. Yong, "Analysis of computer virus propagation behaviors over complex networks: a case study of Oregon routing network," *Nonlinear Dynamics*, vol. 100, no. 2, pp. 1725–1740, 2020.

[26] L. Lin, "A detection and defense technology for information-stealing and deceitful trojan viruses based on behavioral features," *International Journal on Network Security*, vol. 20, no. 5, pp. 983–987, 2018.