Network Security Situation Prediction Technology Based on Fusion of Knowledge Graph

Wei Luo
School of Artificial Intelligence, Chongqing Three Gorges Vocational College, Chongqing, 400155, China

***Abstract*—It is difficult to accurately reflect different network attack events in real time, which leads to poor performance in predicting network security situations. A knowledge graph-based entity recognition model and entity relationship extraction model was developed for enhancing the reliability and processing efficiency of secure data. Then a knowledge graph-based situational assessment method was introduced, and a network security situational prediction model based on self-attention mechanism and gated recurrent unit was constructed. The study's results showed that the constructed prediction model achieved stable mean square error values of approximately 0.0127 and 0.0136 after being trained on the NSL-KDD and CICIDS2017 datasets for 678 and 589 iterations, respectively. The mean square error value was lower due to fewer training iterations compared to other prediction models. The model was embedded into the information security system of an actual Internet company, and the detection accuracy of the number of network attacks was more than 95%. The results of our study indicate that the method used in the study can accurately predict the network security situation and provide technical support for predicting network information security of the same type.**

***Keywords*—Knowledge graph; network security situation; gated recurrent unit; Bayesian attack graph; relationship extraction; relationship recognition**

I. INTRODUCTION

As the evolution of information technology, network security issues are becoming increasingly prominent. The constantly emerging new technologies have made the situation of network security threats more complex and ever-changing. For traditional network security defense systems, relying solely on security devices is no longer sufficient to cope with constantly evolving network attack methods [1-2]. Therefore, a new security concept-network security situational awareness has emerged. Network security situational awareness is a new security technology that can estimate and forecast the security situation of the network environment by integrating network monitoring devices to collect data, applying data mining and other technologies. Compared to traditional network security defense technologies, network security situational awareness solutions possess more proactive situational capture, evaluation, and prediction functions [3]. Chen et al. proposed a network security situation prediction model based on radial basis function (RBF) neural network to address the problem of traditional network security situation awareness prediction methods being relatively single. They optimized the RBF using simulated annealing algorithm and hybrid hierarchical genetic algorithm. The results showed that the optimized RBF neural network performed well in predicting 15 samples [4]. Ruan Z. et al. established a particle swarm optimization model for predicting network security by optimizing the parameters of the support vector regression (SVR) model through particle swarm optimization. The SVR model was then used to predict the network security situation. The experimental results showed that this method effectively predicted network operation security to a certain extent [5]. However, the current network security situational awareness solutions face some difficult problems. Firstly, secure data are often multi-source and heterogeneous, making them difficult to process and analyze effectively and quickly. Secondly, traditional situational analysis methods cannot capture the key information connections between past and current moments well, resulting in uncertainty in prediction results [6]. To address these issues, a knowledge graph (KG)-based network security situation prediction technology is proposed. This method utilizes KG technology to construct a network security data graph, improving data processing efficiency and reliability through the association relationship between entities. Meanwhile, it introduces a situation assessment method architecture based on KG to construct a situation prediction model that integrates self-attention mechanism and gate recurrent unit (GRU).

One of the innovative points of the research is the introduction of a situation assessment method based on KG, which can better understand and evaluate the network security situation by utilizing the structured information of KG. The second innovation is that the self-attention mechanism and GRU are used to construct a network security situation prediction model, which improves the accuracy of security situation prediction.

The article is divided into four sections. The first mainly discusses the current research status of domestic and foreign experts and scholars on KG technology and network security situation. The second section mainly discusses the integration of situational awareness data and the construction of a network security situational assessment and prediction model that integrates KG. The third section mainly discusses the setting of experimental environment and model parameters and designs corresponding experiments for verifying the effectiveness of the research and construction model. The fourth section mainly analyzes the experimental results and clarifies the shortcomings of the research method.

II. RELATED WORKS

With the continuous updates of network technology, traditional network security technologies can no longer meet people's needs. Researching new network security technologies to maintain network security becomes an urgent issue that needs to be solved. Therefore, experts and scholars around the world have conducted research on general network security situational awareness technology. Sun. J et al. proposed a TCAN BiGRU prediction model for enhancing the accuracy of network security situation prediction. This model could learn and extract effective features related to network security from historical network data, and these features were used to predict network security situations. The outcomes showcased that the determination coefficients constructed in the study reached 0.999 on both datasets [7]. Liu. Q et al. proposed a network security situation detection method based on fuzzy neural networks to address the complexity and uncertainty of the Internet of Things in smart cities. This method used fuzzy neural networks to process network data and judges the current network security situation based on the characteristics and behavior patterns of the network data. The outcomes indicated that this method possesses good accuracy and robustness in network security situation detection [8]. Lin. P et al. developed a network security situation assessment method based on text SimHash technology. This method collected text data related to network security, such as articles and blogs, and used SimHash to calculate text similarity, establishing a clustering model based on the K-Means algorithm to classify and summarize network security events. The results indicated that this method was efficient in maintaining network security [9]. Jian. Li et al. presented a security situation assessment model based on evidence theory for addressing security threats and attacks in the Internet of Things environment. By utilizing data fusion and information fusion technologies, different types of security information were fused to obtain more comprehensive and accurate security situation information. The results indicated that this method improved the perception ability of the security situation of the Internet of Things [10].

Network security situation prediction requires the integration of data from various sources. KG technology can effectively structure the representation of information from different data sources, making the correlation and connection between data more clear. Sun. C et al. presented a KG-based method to predict attacks on day 0. This method utilized knowledge in the network security to construct a KG that includes information such as vulnerabilities, attacks, and threat intelligence. Then it analyzed the relationship between known vulnerabilities and known attacks and predicted the path of the 0-day attack. The results showed that this method predicted possible 0-day attack paths [11]. Chen. Y Y and others artificially evaluated potential attack paths in wireless sensor networks, used Bayesian attack graphs to establish attack paths, and calculated the probability of successful attack on each path. The results showed that the methods used in the study could identify and evaluate the security risks present in wireless sensor networks [12]. When constructing a prediction model, GRU had a more concise model structure and could better capture long-term dependencies, providing more accurate predictions. Song. T et al. constructed a deep learning model based on BiGRU and attention mechanisms for predicting tropical cyclone paths in the Northwest Pacific region. The results showed that the model could effectively extract features from historical meteorological data and make accurate predictions for future meteorological changes [13].

On the grounds of the above research, in-depth research on network security situation assessment is meaningful in information security. The prediction of network security situation needs a large amount of data to support, and the construction of network security KG can effectively solve this problem. A prediction model that combines BiGRU and attention mechanisms can better capture the long-term dependencies of network information and provide accurate predictions. On the grounds of this background, a situation prediction model based on GRU and self-attention mechanism was constructed on the basis of network security KG.

III. Construction of a Network Security Situation Prediction Model on the Grounds of the Fusion of KG

This section mainly elaborates on the recognition model and extraction model construction method based on network security KG, and then introduces the situation assessment indicators and quantitative standards based on network security KG. Finally, based on the situation assessment, a situation prediction model based on GRU-Self-Attention was proposed.

A. *Integration Analysis of Situational Awareness Data on the Grounds of KG Fusion*

A brief introduction is provided to the technical models involved in building this model, as shown in Table I.

TABLE I. Brief Description of Key Technology Models

Model	Introduction
BERT	A Transformer-based pre-training model that can transform network security information text into word embedding vectors
GRU	A recurrent neural network model used for processing temporal data, divided into reset gates and update gates. The impact of the previous state of door control on the current state is reset, and the impact of the previous state of door control on the current state is updated.
BiGRU	A recurrent neural network model composed of the forward GRU and backward GRU models. Compared to the GRU model, the BiGRU model can better handle long-term dependencies in temporal data.
Entity relationship extraction model based on self-attention mechanism	In entity relationship extraction, the attention mechanism calculates the attention weights between each position and entity in the input sequence, and models the relationships between entities based on this. The self-attention mechanism, on the other hand, is a low-level model that extracts semantic relationships between entities from text through shared entity recognition.

Ensuring the accuracy and completeness of data is crucial in network security situation prediction. Faced with large-scale and complex network data, data integration operations are required, including data association, cleaning, and normalization, for ensuring the quality and availability of network data. The research divides network situation prediction into two steps: data integration and data analysis, as shown in Fig. 1.

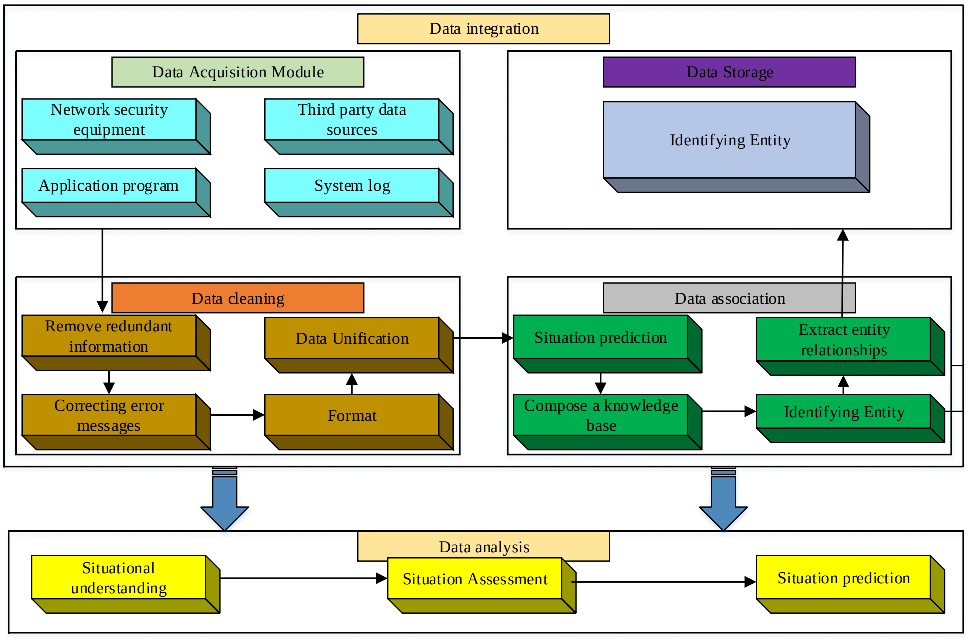
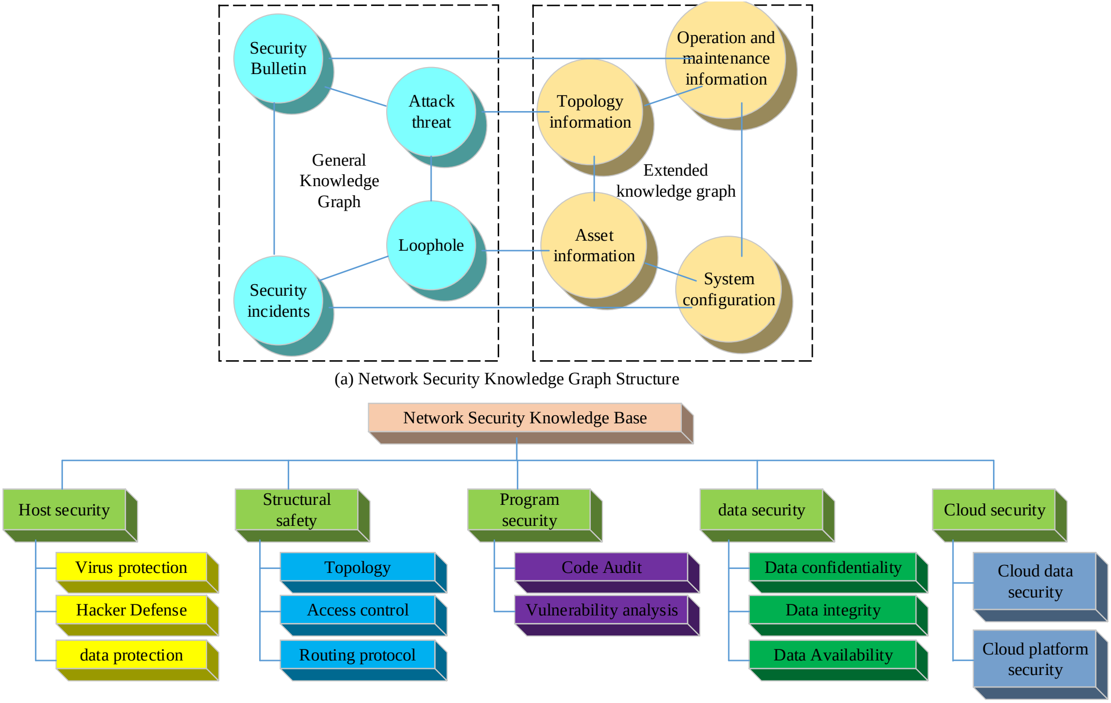


Fig. 1. Specific process of network security situation prediction.

In Fig. 1, data integration consists of four modules: collection, storage, cleaning, and association. The data collection of network security can be carried out through network device logs, security device logs, malware samples, network traffic data, security vulnerability databases, and other channels. Data cleaning includes removing redundant information, correcting error information, standardizing formats, and normalizing data. The data association module needs to be applied to KG technology, which can integrate various security data, enhance the accuracy of network security detection, and achieve network security threat prediction. At present, common network security issues include identity forgery, unauthorized access, and denial of access, all of which involve the association between network entities. Therefore, attackers need to establish network connections before conducting network security intrusions. Therefore, the study focuses on the connections between network entities and constructs a network security data KG, with a specific structure shown in Fig. 2(a).



(a) Network Security Knowledge Graph Structure

(b) Architecture of Network Security Knowledge Base

Fig. 2. Network security KG and knowledge base structure.

In Fig. 2(a), the constructed KG consists of a general KG and an extended KG. The general KG includes obtained security and vulnerability information, etc. It can supplement new vulnerability and attack knowledge in real-time based on changes in network security information. The extended KG, on the other hand, contains network structure information such as network nodes and network operations, which is built specifically for specific networks and has strong targeting. The construction of the network security knowledge base in the KG mainly consists of three parts: entities, relationships, and attributes. The specific structural design is shown in Fig. 2(b). The KG network security repository is mainly divided into five main entities: host security, data security, etc. These five entities each protect the security of their respective networks within their respective scope, while closely interconnected to form a secure network system.

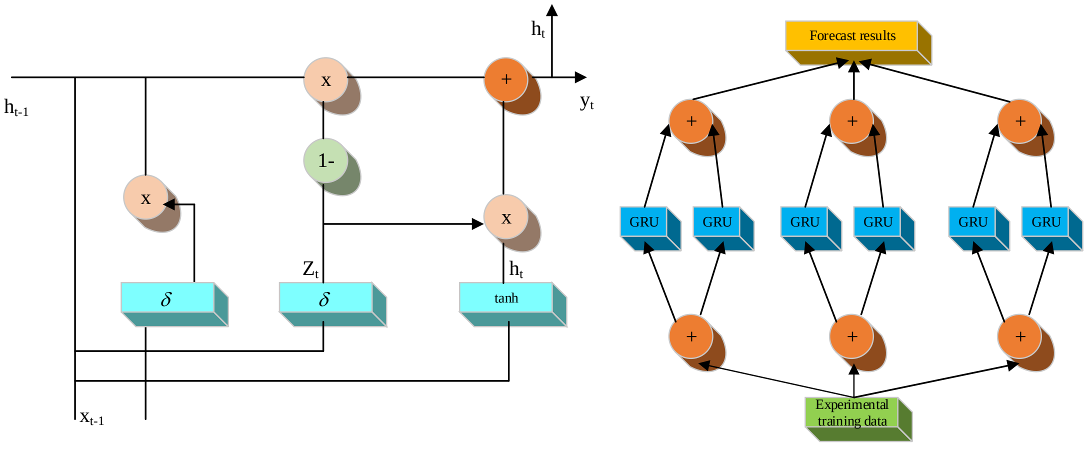
For the recognition of named entities in the network security KG, a recognition model combining feature templates and bidirectional recurrent neural networks is studied. This recognition model uses network security ontology relationships for filtering and feature template generation, and then the input network security information text is transformed into a word embedding vector through the BERT model. Then the word embedding vector is combined with local context features to form the input of a bidirectional recurrent neural network. Finally, by training a bidirectional recurrent neural network, the corresponding semantic features can be obtained. In the entire recognition process, there are two key technologies, one of which is the extraction of feature templates. Research will set the required template as Eq. (1).

$$x[-3,0],x[-2,0],x[-1,0],x[0,0],x[1,0],x[2,0],x[3,0] \quad (1)$$

In equation (1), $x[row,col]$ represents the semantic character in the row-th row and col-th column of the monitoring window. A monitoring window is a window used to extract contextual information, centered around the current character to be recognized. Then it sets the feature function $f_j(y_{i-1},y_i,x,i)$, which is the current position marker, the next stage marker, the current semantic character, and the current position marker. The characteristic function can be summed at different positions to obtain Eq. (2).

$$f_k(y,x)=\sum_{i=1}^{n}\lambda_i f_k(y_{i-1},y_i,i) \quad (2)$$

In Eq. (2), λ_i serves as the weight value of the feature function. The higher the feature score, the higher the corresponding label score, and the more accurate the final prediction result. The second is a bidirectional recurrent neural network, which is composed of two GRUs. The bidirectional recurrent neural network model is showcased in Fig. 3(a), and the GRU model is shown in Fig. 3(b).



(a) Gate control loop unit

(b) Bidirectional recurrent neural network

Fig. 3. Model structure of bidirectional recurrent neural network and gated recurrent unit.

GRU consists of a reset gate and an update gate. The reset gate is responsible for the impact of the previous state on the existing state, while the update gate is responsible for the impact of the previous state on the current state [14]. The operation method for resetting gate r_t and updating gate z_t is shown in equation (3).

$$\begin{cases} r_t=\delta\left(W_r\bullet\left[h_{h-1},x_t\right]\right) \\ z_t=\delta\left(W_z\bullet\left[h_{h-1},x_t\right]\right) \end{cases} \quad (3)$$

In equation (3), x_t serves as the input at time t, and h_{t-1} serves as the hidden state value at time. W_r and W_z are the weight matrices of two gates. δ is the activation function. After passing through GRU, the network security data information can calculate the candidate state value $\hat{h}_t$ and the hidden state value h_t, as shown in Eq. (4).

$$\begin{cases} \hat{h}_t=\tanh\left(W_{\hat{h}}\bullet\left[r_t\bullet h_{t-1}\bullet x_t\right]\right) \\ h_t=\left(1-z_t\right)\bullet h_{t-1}+z_t\bullet\hat{h}_t \end{cases} \quad (4)$$

In equation (4), $W_{\hat{h}}$ serves as the weight matrix of $\hat{h}_t$. The input data are filtered through GRU to obtain the target data as shown in Eq. (5).

$$y_t=\delta\left(W_0 h_t\right) \quad (5)$$

In Eq. (5), W_0 is the weight matrix of the output value. δ is the activation function, and the Sigmoid function is selected as the activation function for this model.

For the extraction of entity relationships in the network security KG, this study constructs an entity relationship extraction model based on the self-attention mechanism. This model first transforms the input network security information text into a word embedding vector, then sequentially uses a bidirectional recurrent neural network and a self-attention model, and finally extracts the entity relationships of network security information [15]. The self-attention model adopts the query key value (QKV) mode. It sets the input sequence as X, and then embeds it into words to obtain A, as shown in Eq. (6).

$$X = \left[x_1, x_2, \dots x_N\right] \in R^{D_x \times N}, \quad A = \left[a_1, a_2, \dots a_N\right] \in R^{D_a \times N} \tag{6}$$

Then it projects A onto three different spaces, namely the query matrix Q, key matrix K, and value matrix V, as showcased in Eq. (7) [16].

$$\begin{cases} Q = \left[q_1, q_2, \dots, q_N\right] \in R^{D_q \times N} \\ K = \left[k_1, k_2, \dots, k_N\right] \in R^{D_k \times N} \\ V = \left[v_1, v_2, \dots, v_N\right] \in R^{D_v \times N} \end{cases} \tag{7}$$

In Eq. (7), $R^{D \times N}$ represents the range of different matrices. The matrix operation method is shown in Eq. (8).

$$\begin{cases} Q = W^q A & W^q \in R^{D_q \times D_a} \\ K = W^k A & W^k \in R^{D_k \times D_a} \\ V = W^v A & W^v \in R^{D_v \times D_a} \end{cases} \tag{8}$$

It sets each query vector as q_i and uses a key value pair attention mechanism for q_i to obtain the attention distribution as shown in Eq. (9).

$$\hat{b}_{1.1}, \hat{b}_{1.2}, \dots, \hat{b}_{1.N} \tag{9}$$

It uses the "scaled dot product" method to score attention. To avoid inputting values that are too large or too small, $\sqrt{D_k}$ is used to scale them, as shown in Eq. (10).

$$B = \frac{K^T Q}{\sqrt{D_k}} \tag{10}$$

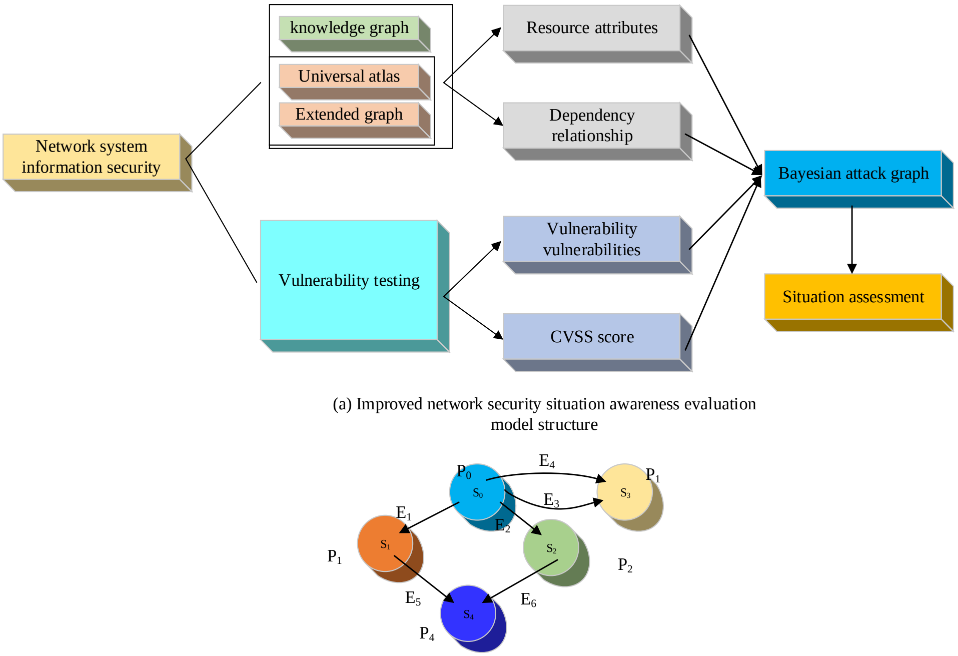
In Eq. (10), B is the proof of attention distribution. After obtaining the attention distribution matrix, it uses the Softmax function to perform column wise operations to obtain the attention distribution $\hat{B}$, as shown in Eq. (11).

$$\hat{B} = soft\max\left(B\right) \tag{11}$$

After obtaining the attention distribution $\hat{B}$, the final output can be obtained by using a weighted sum method.

B. Construction of Evaluation and Prediction Models for Network Security Situation

For enhancing the shortcomings of existing methods for evaluating network security situations, this study combines network security KG, universal vulnerability score, and Bayesian attack graph to design an improved network security situation awareness evaluation model, as showcased in Fig. 4.



(a) Improved network security situation awareness evaluation model structure

(b) Example of Bayesian attack graph

Fig. 4. Improved network security situation awareness evaluation model structure and Bayesian attack graph.

In Fig. 4(a), the evaluation model combines KG, vulnerability score, and Bayesian attack graph. It perceives and evaluates the security situation of the network by comprehensively considering factors such as vulnerability rating of vulnerabilities, accessibility of Bayesian attack graphs, and probability of attacks, and identifies potential security risks and threats. A Bayesian attack graph is a directed acyclic graph, which can be defined as Eq. (12) [17-18].

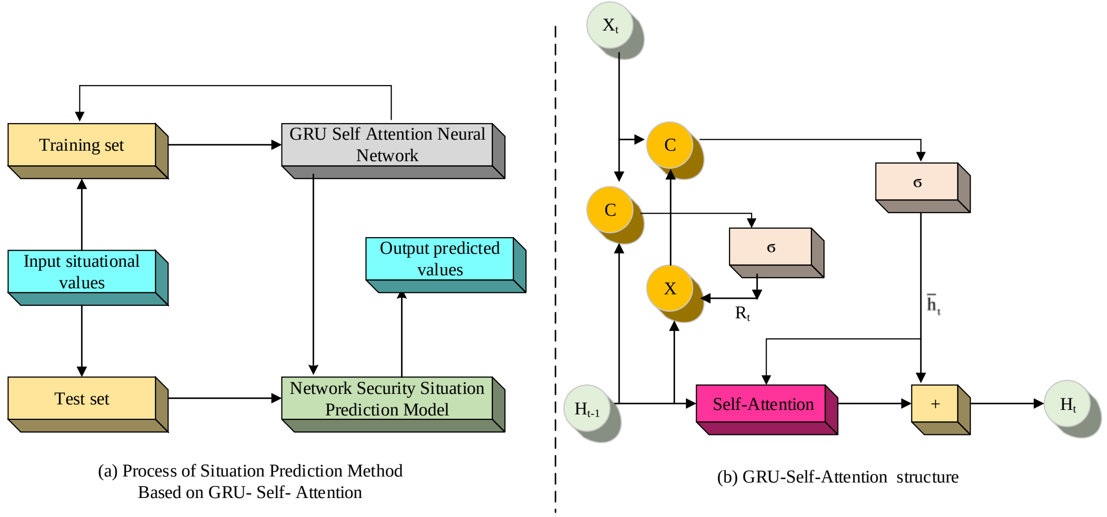
$$BAG = \left(S, E, R, P\right) \qquad (12)$$

In Eq. (12), S is the set of conditions, $S_i = \{0,1\}$, where S_0 indicates that the attacker has not occupied the node, and S_0 indicates that the attacker has already occupied the node. E is the set of directed edges in the attack graph, which represents both the causal relationship between network nodes and the attacker's exploitation of vulnerabilities. R is the relationship between the conditional node and its incoming edge, represented by $\left(S_j, d_j\right), d_j \in \{AND, OR\}$. AND represents that all incoming edges of S_j have been successfully attacked before the attacker can occupy the S_j node. OR represents a successful edge attack in S_j, allowing the attacker to occupy the S_j node. P is the set of probabilities that conditional nodes can reach, and P_i is the probability that the attacker occupies S_i. The constructed Bayesian attack diagram is shown in Fig. 4 (b), where S_0, S_1, S_2, S_3, S_4 are conditional nodes. $E_1, E_2, E_3, E_4, E_5, E_6$ are directed edges. P_0, P_1, P_2, P_3, P_4 are the probability that the attacker will occupy S_0, S_1, S_2, S_3, S_4. When evaluating the network security situation, some indicators (such as CPU utilization, memory usage, etc.) can be directly obtained by collecting data through corresponding devices. However certain evaluation indicators need to be quantified to be visualized, and the quantified indicators can be found in Eq. (13) [19].

$$\begin{cases} Degree = \dfrac{\sum_{i=1}^{N} Status_i * Scorei}{N} \\ Threa_t = \dfrac{\sum_{i=1}^{N} 2^{level_i}}{N} \\ R_i = \dfrac{Form_i}{\sum_{i=1}^{n} Form_j} \\ Rate = \dfrac{Event_t}{Event_{t-1}} \end{cases} \qquad (13)$$

In Eq. (13), $Degree$ represents the security level of the operating system kernel. $Score_i$ is the operating system kernel security score of the i -th host. $Status_i$ is the operating status of the i -th host. N is the number of hosts. $Threat$ is the average threat level of security incidents. $level_i$ is the safety level. N is the number of events. R_i is is the distribution of security event types. $Form_i$ serves as the number of i -th security incidents. $Rate$ serves as the rate of change of safety events. $Event_t$ is the number of event triggers in the t -th time period. $Event_{t-1}$ is the number of event triggers in the $t-1$ -th time period.

After accurately evaluating the network security situation, situation prediction can predict potential security events that may occur in the future of the information network based on the evaluation information. A GRU-Self-Attention network security situation prediction model is constructed for this study, and the prediction process is shown in Fig. 5(a).



(a) Process of Situation Prediction Method Based on GRU- Self- Attention

(b) GRU-Self-Attention structure

Fig. 5. Operation process and composition structure of GRU-Self-Attention network security situation prediction model.

In Fig. 5(a), the GRU-Self-Attention prediction model divides the obtained dataset in the potential evaluation into a test set and a training set in a 3:7 ratio. It reuses the training set for training the prediction model, saves the trained parameters, and finally verifies the prediction model using the test set. Fig. 5(b) shows the structural diagram of constructing GRU-Self-Attention, where R_t is the reset gate of GRU. $\bar{h}_t$

is the candidate set for GRU. H_t and H_{t-1} represent the hidden information of the current and past time steps, respectively. X_t serves as the input at the current time. The evidence flow for the operation of this prediction model is shown in Eq. (14).

$$\begin{cases} R_t = \delta\left(W_R \bullet \left[H_{t-1}, H_t\right]\right) \quad (14.1) \\ \bar{h}_t = \tanh\left(W_{\bar{h}} \bullet \left[R_t \times \left(H_{t-1}, H_t\right)\right]\right) \quad (14.2) \\ h_{cat} = cat\left(\bar{h}_t, H_{t-1}\right) \quad (14.3) \\ H_{t-1} = seft - Attention(h_{cat}) \quad (14.4) \\ y_t = W_o \bullet H_t \quad (14.5) \end{cases} \qquad (14)$$

The GRU-Self-Attention prediction model first initializes the reset gate of the GRU, as shown in Eq. 14.1. Then it constructs a candidate set, as shown in Eq. 14.2, and constructs input data, as shown in Eq. 14.3. Next, it uses the self-attention mechanism for learning the correlation between H_{t-1} and $\bar{h}_t$, and constructs a new matrix, as shown in Eq. 14.4. Finally, it updates the hidden state of the current time step and outputs the prediction result.

IV. Experimental Analysis

This section first elaborates on the setting of experimental environment, model parameters, etc., and then designs experiments to verify the effectiveness of entity recognition models and entity relationship extraction models. Afterwards, the average absolute error and root mean square error (RMSE) of the GRU-Self-Attention prediction model were tested for the predicted trend values. Finally, an experiment was designed for testing the practical application effect of the GRU-Self-Attention prediction model.

A. *Performance Analysis of Entity Recognition Models and Entity Relationship Extraction Models*

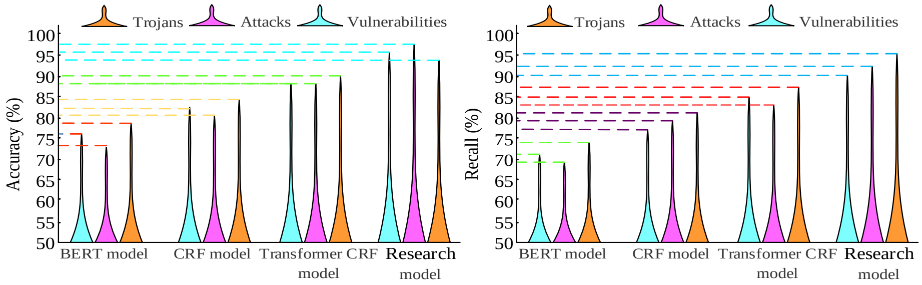
For ensuring the accuracy and reliability of constructing a network security KG, it is necessary to verify the effectiveness of the entity recognition model and entity relationship extraction model studied and constructed. To this end, the research used crawler software to crawl malicious code, security vulnerability information, network intrusion information and other network security text data from major security vendors, hacker communities and other websites on the Internet, a total of 34582 pieces. At the same time, it set the network environment and model parameters required for the experiment, as showcased in Table II.

TABLE II. Basic Hardware Environment and Model Parameters for the Experiment

Project	Parameter
Operating system	Windows10
System PC side memory	16G
CUP	Intel Core i9
Storage	256GB SSD
Graphics card	NVIDIA GGTX 1060
Development tool	Pycharm3.6, Anaconda3
Model Optimizer	Stochastic Gradient Descent
Hidden layer size	0.0001
Batch size	100
Epoch size	40
Dropout	0.5

The selected dataset was divided into three types of named entities: vulnerabilities, attacks, and Trojans, and the recognition model constructed in the study was used to detect these four types of named entities. Additionally, to demonstrate the effectiveness and superiority of the model, the currently popular entity recognition model was selected and compared with the research model. The selected models include the BERT model in reference [20], the CRF model in reference [21], and the Transformer-CRF model in reference [22]. The accuracy and recall results of the four models are showcased in Fig. 6.

Fig. 6(a) showcases the accuracy test results of four entity recognition models. This indicates that the detection accuracy of the research model for vulnerability, attack, and Trojan named entities is 95.9%, 97.1%, and 93.7%, respectively, with the highest recognition accuracy among the four models. Fig. 6(b) shows the recall test results of four entity recognition models. This indicates that the recall rates of the research model for vulnerability, attack, and Trojan named entities are 90.3%, 92.7%, and 95.9%, respectively, which are also the highest among the four models. This indicates that the entity recognition model constructed in the study has significant advantages. This is because the model incorporates feature models and text embedding vector representation methods to optimize model performance, thereby enhancing the accuracy and recall.



(a) Model accuracy detection results (b) Model recall detection results

Fig. 6. Accuracy and recall test results of four recognition models.

3698 sentences with multiple network security entities from the collected 34582 network security text data were selected to test the accuracy of the entity extraction model constructed. Similarly, to verify the superiority of the extraction model constructed, SVM, CRF, and SRL models were selected and compared with the research model. The results are shown in Fig. 7.

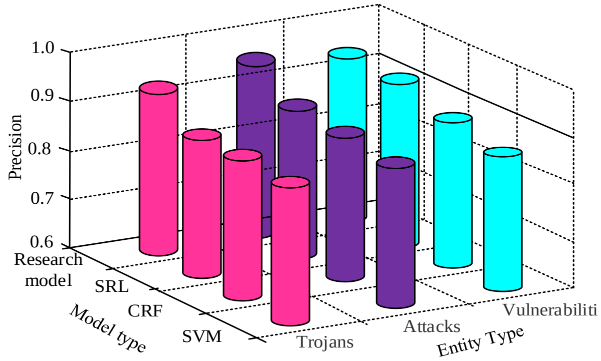
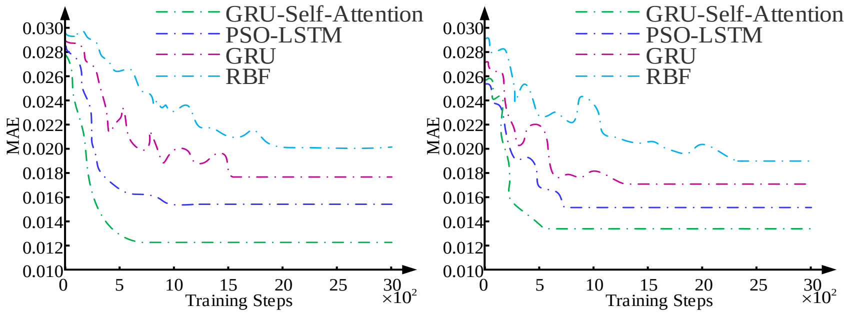


Fig. 7. Accuracy test results of four extraction models.

In Fig. 7, for the types of vulnerabilities, the extraction accuracy of the research model, SVM, CRF, and SRL are 0.92, 0.82, 0.80, and 0.76. For attack types, the extraction accuracy of the four models is 0.94, 0.85, 0.83, and 0.81, respectively. For the types of vulnerability attacks, the extraction accuracy of the four models is 0.93, 0.90, 0.88, and 0.81, respectively. This indicates that the extraction model used in the study has the highest accuracy, as it adds a self-attention layer, which can better combine long sentence information and improve model performance.

B. *Performance Analysis of GRU-Self-Attention Prediction Model*

For verifying the GRU-Self-Attention prediction model, the NSL-KDD and CICIDS2017 datasets were selected as network information sources. Then the network security situation quantification method mentioned in section 2.2 was used to quantify the selected dataset, which was used as the experimental dataset. Similarly, the dataset was allocated in a ratio of 3:7 between the test set and the training set, and the experimental environment was consistent with that in section 3.1. The model network structure adopted a "input layer-hidden layer-output layer" approach, with a learning rate of 0.001 and a training period of 3000. In addition, GRU model, PSO-LSTM model, and RBF model were selected as controls. The first step is to test the average absolute error of the predicted situation value of the model, as shown in Fig. 8.



(a) NSL-KDD dataset

(b) CICIDS2017 dataset

Fig. 8. Test results of the average absolute error of the model's predicted situational values.

Fig. 8(a) shows that in the NSL-KDD dataset, the mean square error (MSE) values predicted by the four models decrease with increasing training step size. The MSE values of RBF and GRU models are relatively large, and the curves fluctuate repeatedly during the training process until they stabilize after about 1500 iterations. The MSE values of PSO-LSTM and GRU-Self-Attention are continuously decreasing without any fluctuations. The PSO-LSTM model reaches a stable MSE value of approximately 0.0156 after training for about 1000 times. The GRU-Self-Attention model achieves a stable MSE value of approximately 0.0127 after training for approximately 678 times. Fig. 8(a) shows that in the CICIDS2017 dataset, the MSE value of the GRU-Self-Attention model is still the lowest, about 0.0136, and the training frequency is the least, about 589 times. The GRU-Self-Attention model constructed in the study can achieve good prediction results on different datasets, with lower MSE values and fewer training steps. Next, based on the NSL-KDD dataset, the neural network of the model was tested, and the results are shown in Fig. 9.

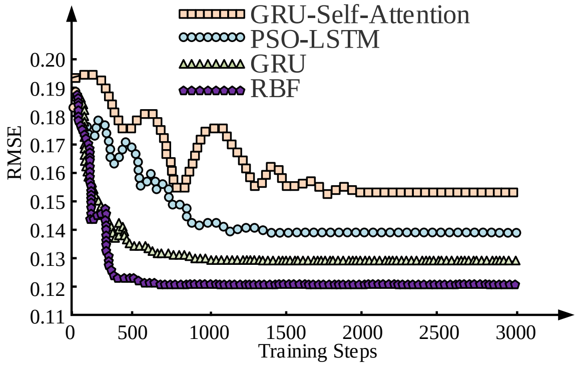


Fig. 9. Test results of neural network of model predictive situation values.

Fig. 9 shows that the RMSE values predicted by the four models decrease with the increase of training step size. The RMSE values of the RBF model fluctuate multiple times and have a large amplitude as they tend to stabilize. At around 1945 training sessions, the RMSE value tends to stabilize, with an RMSE value of approximately 0.154. In contrast, the GRU model curve fluctuates slightly, but the fluctuation amplitude is small, stabilizing after approximately 1389 training sessions, with an RMSE value of approximately 0.141. The PSO-LSTM and GRU-Self-Attention models require less training to achieve stable RMSE values. The PSO-LSTM model is trained approximately 689 times, with a stable RMSE value of 0.129. The GRU-Self-Attention model is trained approximately 524 times, with a stable RMSE value of 0.121. The GRU-Self-Attention model has a higher degree of fitting to the training data and less computational time.

C. Application Analysis of GRU-Self-Attention Prediction Model

The above experiment has proven the feasibility of the network security situation prediction method constructed in the research. The information security system of a large Internet company was selected as the experimental object, and the GRU-Self-Attention prediction model was embedded into the company's security system. It detects the number of network information attacks, as shown in Fig. 10.

Fig. 10(a) shows the actual detection results of network information attacks. This indicates that during the detection process, the website is subjected to 110, 100, 106, 100, 109, and 102 malicious code attacks, DOS attacks, other types of attacks, web attacks, virus attacks, and vulnerability attacks, respectively. Fig. 10(b) showcases the detection results of the original prediction system of Internet companies. This indicates that the system predicts 90, 78, 83, 82, 101, and 84 malicious code attacks, DOS attacks, other types of attacks, web attacks, virus attacks, and vulnerability attacks, respectively. Fig. 10(c) showcases the detection results of the GRU-Self-Attention prediction model. This indicates that the model predicts 104, 98, 102, 99, 104, and 98 malicious code attacks, DOS attacks, other types of attacks, web attacks, virus attacks, and vulnerability attacks, respectively. The detection results indicate that the GRU-Self-Attention prediction model greatly improves the detection accuracy of the original system against network attacks. Compared with actual results, the detection accuracy is over 95%. It retests the detection time of network information attack events, and the results are shown in Fig. 11.

Fig. 11(a) shows the variation in time taken by the original prediction system to detect 1000 network security attack events. This indicates that as the number of attack events grows, the original prediction system takes more and more time, and the increase in time is becoming larger and larger. Detecting 1000 network security attack events takes approximately 11.7 minutes. In Fig. 11(b), the time growth of the GRU-Self-Attention prediction model is far less than that of the original system, with a detection time of about 1.2 minutes for 200 network security attack events. It detects 1000 network security attack events, taking only about 3.8 minutes.

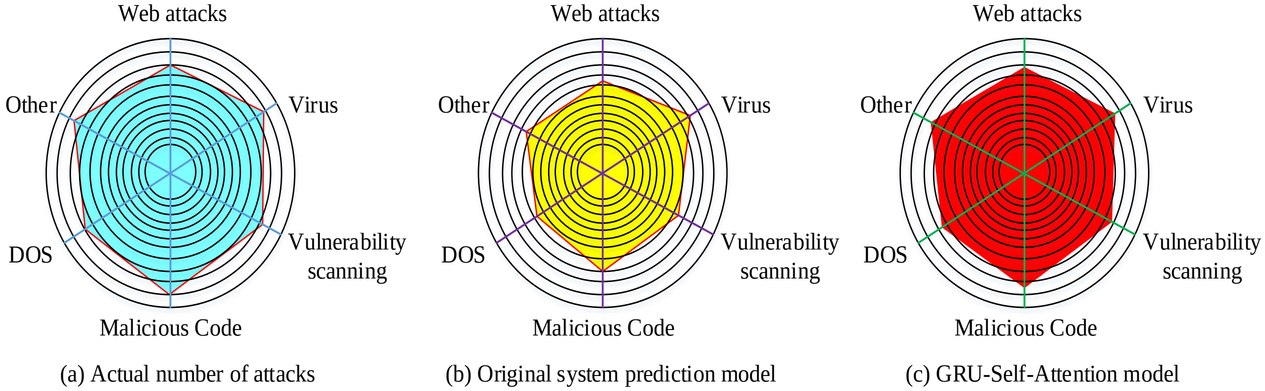


Fig. 10. Detection results of network information attack by prediction model at this time.

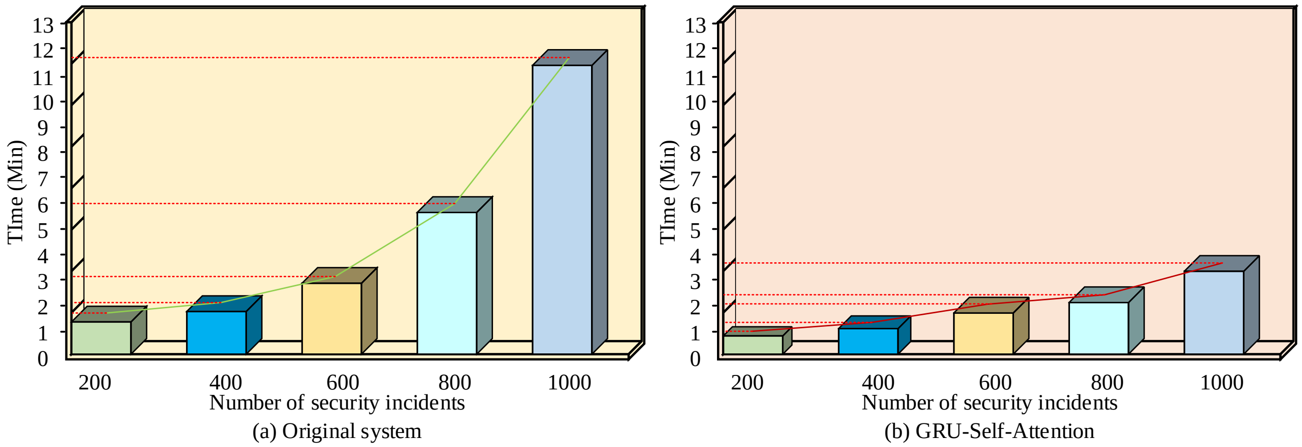


Fig. 11. Prediction time for 1000 network security attack events.

Finally, the prediction performance of the GRU-Self-Attention prediction model on network security situation values was verified, as shown in Fig. 12.

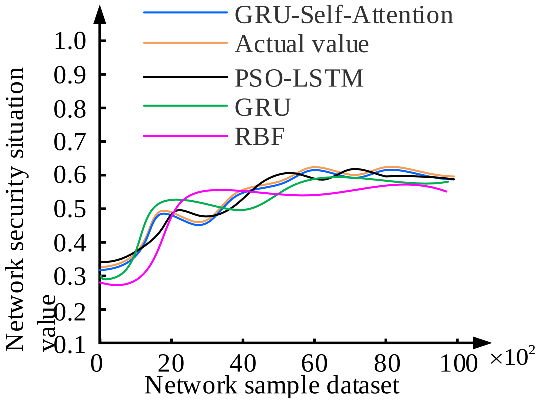


Fig. 12. Network security situation prediction value.

In Fig. 12, the actual potential value of the network ranges from 0.3 to 0.6 and increases with the increase of sample data. From the curve trend of the graph, the GRU-Self-Attention model constructed in the study has little difference between the predicted values of the network potential and the actual potential values. This indicates that the prediction model can make accurate predictions of the network potential values.

D. Discussion

With the increasing complexity and intelligence of network security threats, it has become inevitable to take more comprehensive and efficient measures to protect network security. Based on the Fusion of Knowledge Graph, this study constructs a network security data graph using KG technology and constructs a situation prediction model that combines self-attention mechanism and Gate Recursive Unit. The experimental results showed that the detection accuracy and recall rate of the entity recognition model constructed in the study were above 90%, which were higher than the BERT, CRF, and Transformer CRF models under the same experimental conditions. Compared with the research results of Chen. Z. et al. [4], there has been further improvement. This is because the constructed entity recognition model incorporates feature models and text embedding vector representation methods to optimize the performance of the model, thereby improving the accuracy and recall of the model. The accuracy of the constructed entity relationship extraction model was also above 90%. The extraction accuracy of the SVM, CRF, and SRL models under the same experimental conditions were 82%, 80%, and 76%, respectively, significantly lower than the research model. Compared to the accuracy achieved by Ruan. Z. et al. [5] is over 85%, but the method proposed in this paper is significantly higher. This is because the constructed entity relationship extraction model adds a self-attention layer, which can better combine long sentence information and improve model performance. This also indicates that the constructed network security KG has extremely high feasibility. But the prediction efficiency of the methods proposed in network security needs to be improved. In the future, it is necessary to optimize the model structure and utilize more advanced parallel computing to further improve the predictive performance of network security, providing more reliable guarantees for network security.

V. Conclusion

Situation prediction technology possesses an essential influence on mitigating network security threats. Therefore, this study optimized the entity recognition model and entity relationship extraction model of network security KG and presented a network security situation assessment method based on KG and Bayesian attack graph. Meanwhile, the situation prediction method was optimized through self-attention mechanism, and a GRU-Self-Attention prediction model was constructed. The experimental results showed that the average absolute error of the GRU-Self-Attention situational prediction model based on network security KG in predicting situational values in the NSL-KDD dataset was about 0.0127. This value was about 0.0136 in the CICIDS2017 dataset. The average absolute error in different scenarios was lower than that of the GRU, PSO-LSTM, and RBF models under the same experimental conditions. The model constructed in the study has good fit to different datasets, and the average absolute error is lower than the existing models. The GRU-Self-Attention model was embed into an information security system, which could accurately predict the number of different types of network attacks. In addition, the model detected 1000 network security attack events, which only took about 3.8 minutes. This indicates that the research plan can effectively improve the accuracy of network security situation prediction. However, building a network security KG requires a large number of resources and time in the early stages, and later research will further optimize the construction process and data processing of KG.

References

[1] Potember. R. S., Balhana .C. D., Obrst. L. J. An Introduction to Semantic Threat Analysis for Systems Security Engineering. INCOSE International Symposium, 2022, 32(1):498-513.

[2] Samuel. O. S. Cyber Situation Awareness Perception Model for Computer Network. International Journal of Advanced Computer Science and Applications, 2021, 12(1):392-397.

[3] Venkatesan. B., Chitra. S. An enhance the data security performance using an optimal cloud network security for big data cloud framework. International Journal of Communication Systems, 2021, 35(16):1-15.

[4] Chen. Z. Research on internet security situation awareness prediction technology based on improved RBF neural network algorithm. Journal of Computational and Cognitive Engineering, 2022, 1(3): 103-108.

[5] Ruan. Z. Network security prediction method based on kubernetes. Journal of Physics: Conference Series. IOP Publishing, 2021, 2010(1): 109-111.

[6] Zhang. H., Kang. K., Bai. W. Hierarchical network security situation awareness data fusion method in cloud computing environment.Journal of computational methods in sciences and engineering, 2023, 23(1):237-251.

[7] Sun. J., Li. C., Song. Y., Ni. P., Wang. J. Network Security Situation Prediction Based on TCAN-BiGRU Optimized by SSA and IQPSO. Tech Science Press, 2023, 47(10):993-1021.

[8] Liu. Q., Zeng. M. Network security situation detection of internet of things for smart city based on fuzzy neural network. International Journal of Reasoning-based Intelligent Systems, 2020, 12(3):222-227.

[9] Lin. P., Chen. Y. Network Security Situation Assessment Based on Text SimHash in Big Data Environment. International Journal of Network Security, 2019, 21(4):699-708.

[10] Jian. L. I., Dong. T., Jie. L. I. Research on IoT security situation awareness method based on evidence theory. Chinese Journal of Network and Information Security, 2022, 8(2):39-47.

[11] Sun. C., Hao. H. U., Yang. Y. Prediction method of 0day attack path based on cyber defense knowledge graph. Chinese Journal of Network and Information Security, 2022, 8(1):151-166.

[12] Chen. Y. Y., Xu .B., Long. J. Information security assessment of wireless sensor networks based on bayesian attack graphs. Journal of Intelligent and Fuzzy Systems, 2021, 41(4):1-7.

[13] Song. T., Li. Y., Meng. F., Xie. P., Xu. D. A Novel Deep Learning Model by BiGRU with Attention Mechanism for Tropical Cyclone Track Prediction in the Northwest Pacific. Journal of Applied Meteorology and Climatology, 2022, 61(1):3-12.

[14] Liu. M., Wang. X., Liang. S., Sheng .X. I., Lou. S. Single and composite disturbance event recognition based on the DBN-GRU network in 9-OTDR.Applied optics, 2023 62(1):133-141.

[15] Wang. J., Wang. X., Ma. C., Kou. L. A survey on the development status and application prospects of knowledge graph in smart grids. IET Generation, Transmission & Distribution, 2021, 15(3): 383-407.

[16] Zhou. B., Shen. X., Lu. Y., Li. X., Hua. B., Liu. T., Bao. J. Semantic-aware event link reasoning over industrial knowledge graph embedding time series data. International Journal of Production Research, 2023, 61(12): 4117-4134.

[17] Li. Z., Zhao. Y., Li. Y., Rahman. S., Wang. F., Xin. X., Zhang. J. Fault localization based on knowledge graph in software-defined optical networks. Journal of Lightwave Technology, 2021, 39(13): 4236-4246.

[18] Hebbi. C., Mamatha. H. Comprehensive Dataset Building and Recognition of Isolated Handwritten Kannada Characters Using Machine Learning Models. Artificial Intelligence and Applications, 2023, 1(3):179-190.

[19] P. Preethi and H. R. Mamatha, "Region-Based Convolutional Neural Network for Segmenting Text in Epigraphical Images," Artif. Intell. Appl., vol. 1, no. 2, pp. 119-127, Sep, 2023, DOI: 10.47852/bonviewAIA2202293.

[20] Ghourabi. A .SM-Detector: A security model based on BERT to detect SMiShing messages in mobile environments. Concurrency and Computation: Practice and Experience, 2021, 33(24)：1-15.

[21] Ghaffari. R., Golpardaz. M., Helfroush. M. S., Danyali. H. A fast, weighted CRF algorithm based on a two-step superpixel generation for SAR image segmentation. International Journal of Remote Sensing, 2020, 41(9):3535-3557.

[22] Ma. C., Zhang. C. Joint Pre-Trained Chinese Named Entity Recognition Based on Bi-Directional Language Model. International Journal of Pattern Recognition and Artificial Intelligence, 2021, 35(9):1-16.