Application Analysis of Network Security Situational Awareness Model for Asset Information Protection

Yuemei Ren*, Xianju Feng
Henan Polytechnic Institute, Nanyang, China

***Abstract*—The popularity of the Internet makes the network develop rapidly. However, the network security threat is more complex and hidden. The traditional network security alarm system has the problems of low accuracy and low efficiency when dealing with huge redundant data. Therefore, the research comprehensively considers the network security problems, proposes a network security situational awareness model for asset information protection combined with knowledge graph, establishes an asset-based network security knowledge graph, utilizes attribute graphs to complete the network attack scenario discovery and network situational understanding, and verifies the effectiveness and superiority of the model. The experimental results show that the research-proposed model detects an average of 9706 attacks out of 10000 attacks. For 100 high-risk level attacks, the number of detections is higher than 98. The average correctness, recall, and false alarm rates of the research proposed model are 99.48%, 99.04%, and 0.86%, respectively. In addition, when the model is running, its maximum memory usage is only 22.67%, and the time to complete the attack detection at the same time is 258.4s, both of which are much lower than the comparison algorithms. Finally, the research-proposed model is able to effectively reflect the impact of attack events on the posture of asset nodes. The proposed cybersecurity situational awareness model is of great theoretical and practical significance for improving organizational cybersecurity, innovating cybersecurity solutions, and maintaining the security of asset information in the digital era.**

***Keywords*—Asset information protection; cyber security; situational awareness; knowledge graph; attack scenarios**

I. Introduction

Research background: With the advent of the digital age, the highly networked society makes information exchange more convenient. However, the field of network security also suffers from increasingly severe challenges, and various cyber criminals and cyber spies emerge in an endless stream [1]. Network intrusion and attacks tend to be distributed, large-scale and indirect. With the increasing network scale, traditional network security products have become increasingly difficult to meet people's needs for network security [2]. In this context, Network Security Situation Awareness (NSSA) has become a key element to protect information assets and ensure network security [3]. As a mechanism to comprehensively observe, understand and predict cyberspace entities and events, NSSA can provide timely and accurate threat intelligence, enabling it to deal with potential cyber threats more effectively [4]. Among them, the protection of asset information is one of the core tasks of cybersecurity, and the proper management and protection of assets is crucial for maintaining the normal operation of the organization and information security [5]. However, traditional network security alarm systems often face the problems of alarm aggregation and alarm correlation analysis when dealing with large-scale network data, and are susceptible to a large number of redundancies and false alarms, which reduces the accurate identification of real network threats [6].

Research method: Therefore, the research is oriented to asset information protection and proposes a Knowledge Graph-based Network Security Situational Awareness (Knowledge Graph-NSSA, KG-NSSA) model, which constructs a network security knowledge graph and introduces the techniques of attribute graph mining and similarity computation in order to accomplish the process of attack discovery and attack correlation with more accuracy, thus further solving the scenario in network attack discovery and posture understanding. The research aims to provide a more accurate and comprehensive NSSA, improve the perception level of cyber threats, and provide innovative and more effective solutions for the field of cyber security.

Research contribution: The research contribution is to use network security knowledge graph to integrate basic network events, general network security knowledge and attack characteristic events, use attribute graph mining technology to reveal potential threats and abnormal behaviors in the network, introduce similarity calculation to quantify the similarity between network events, and help identify attack events with common characteristics. KG-NSSA model can effectively reflect the impact of attack events on asset node situation, provide real-time updated network security situation, and provide support for network security management and decision-making. It has important theoretical and practical significance for improving organization network security level, innovating network security solutions, and maintaining asset information security in the digital era.

Content partitioning: The research is divided into six sections, Section II introduces the current worldwide research on network security situational awareness and other contents. Section III mainly provides a detailed description of the knowledge graph construction process and other contents in the KG-NSSA model. Section IV gives detail about the network security situational awareness. Results and discussion is given in Section V. Finally, Section V concludes the paper.

II. Related Work

With the increasing digitization of society, individuals, enterprises and government agencies rely on networks for

their daily activities and business operations. As a result, cyber security has become a key factor in maintaining social stability and personal privacy, and organizations and individuals at all levels are striving to strengthen cyber security measures to defend themselves against increasingly sophisticated cyber attacks. Chen addressed the problem of the expanding network scale and the continuous evolution of attack techniques, while the traditional perceptual prediction accuracy is limited, and proposes a cyber attack prediction model based on radial basis function neural network, and optimizes the model through simulated annealing and hybrid hierarchical genetic algorithm so as to improve the accuracy of attack prediction [7]. Tan et al. proposed an innovative honeypot network technology based on threat detection and situational awareness for the future application of AI Internet of Things (IoT) in Industry 4.0 as well as the security threats and attacks faced by the AI IoT, which improves the level of security threat perception and enhances the AI IoT's overall security and resilience against attacks [8]. Liu et al. proposed an innovative approach based on big data and artificial intelligence for the reinforcement needs of information security situational awareness systems, utilizing long and short-term memory recurrent neural networks in deep learning techniques for information security situational prediction, thereby improving the accurate prediction capability of information security situational prediction [9]. Hamdaoui et al. proposed an innovative approach based on threat detection and honeypot networking for the cybersecurity issues present on IoT devices by proposing a blockchain-based distributed protocol that allows IoT devices to communicate with each other in a distributed manner and uses self-recovery/self-healing mechanisms to ensure robustness against device failures and malicious behaviors, thereby ensuring the security, resilience, and reliability of the network [10]. Junejo et al. proposed a lightweight trust model to address the growing security threats in the Internet of vehicles due to their dependence on infrastructure, computing, dynamic nature and control technologies. The model enhances the trust of the network by identifying dishonest nodes and revoking their credentials in a man-in-the-middle attack scenario. Thus, higher authenticity, privacy, accuracy, security and trusted information sharing can be achieved [11]. Ahmed et al. made a comprehensive analysis of location privacy attacks and their solutions to the problem of location privacy protection when two or more vehicles are wirelessly connected to realize data exchange in the Internet of Things environment, so as to improve the location privacy protection of the Internet of Things and ensure the security of data exchange [12]. Memon et al. proposed a novel dynamic path privacy protection scheme to meet the needs of user path privacy protection in location-based services, designed for continuous query service in the road network environment, while hiding the identity of users in dynamic path privacy and providing untraceable attributes of the initiator. Thus, the anonymity of user identity, location information and service content in LBSs can be effectively protected [13].

Knowledge graph is a graphical structure for organizing and representing knowledge, including entities, relations, and attributes, which is an innovative way of representing and organizing knowledge for better organizing, linking, and utilizing a large amount of information, and thus triggered the attention of many scholars. Li et al. proposed a novel heterogeneous graph neural network framework based on attention mechanism to address the embedding of knowledge graphs with heterogeneity. This framework preserves the inherent structure of knowledge graphs while effectively handling the heterogeneity of entities and relationships in knowledge graphs, making the representation learning of knowledge graphs more accurate and targeted [14]. Goel et al. proposed a novel temporal knowledge graph complementation model for the problem of containing temporal facts in the knowledge graphs as well as for the challenges of knowledge graph complementation, by introducing a non-synchronous entity embedding function, which is a new model of knowledge representation and organization equips the static model with the ability to provide entity features at any point in time by introducing an asynchronous entity embedding function, which results in superior performance of the knowledge graph [15]. Mohamed et al. addressed the problem of high false positive prediction rate in predicting drug-target interactions by proposing a novel computational method based on the knowledge graph that utilizes a biomedical knowledge base to create the knowledge graph of entities that are associated with a drug and its potential targets, thus enabling a more comprehensive understanding and prediction of the drug's mechanism of action [16]. Aiming at the limitations of approximate reasoning and reasoning mechanism, Long et al. proposed a new fuzzy knowledge graph pair model, including new representation methods and approximation algorithms, to improve the performance of finding new record labels, and thus provide a new way to solve the decision and classification problems in fuzzy systems [17].

To summarize, researchers provide new solutions for securing networks from various aspects, in addition to the fact that knowledge graphs have been applied in various domains. However, in the face of complex NSSA, few studies have combined it with knowledge graph, resulting in incomplete knowledge system of NSSA and hindering the improvement of its ability to deal with security events. Therefore, the research proposes the KG-NSSA model. The research overcomes the shortcomings of the traditional alarm aggregation process and alarm correlation analysis process which are susceptible to a large number of redundancies and false alarms, and completes the attack discovery and attack correlation through attribute graph mining and similarity computation, which can effectively reflect the specific cyber-attack behaviors and mine the attack scenarios, and thus is innovative.

III. Network Security Situational Awareness Modeling for Asset Information Protection

The study first constructs a cybersecurity knowledge graph containing three parts, and then details a feasible scheme for attack scenario discovery and situational understanding based on the cybersecurity knowledge graph, thus completing the establishment of the KG-NSSA model.

A. Knowledge Graph-based Network Security Situational Awareness Modeling

Since assets are the core of network security situational awareness, the KG-NSSA model constructed in the study starts from the aspect of asset information protection. The inputs of the KG-NSSA model include external data and internal data from the monitored network, and the construction of the network security knowledge graph is also accomplished on the basis of preprocessing of external data and internal data. The study divides the network security knowledge graph into three parts, which are Basic Network Event Graph (BNEG) that combines the actual network traffic information, General Network Security Knowledge Graph (GNSKG) that combines the asset information, and General Network Security Knowledge Graph (GNSKG) that covers the actual network traffic information. GNSKG) that combines asset information, and Attack Characteristic Event Graph (ACEG) that covers multi-step attack characteristic events, which the study collectively refers to as Asset-based Network Security Knowledge Graph (ANSKG). ANSKG), which is schematically shown in Fig. 1.

As can be seen from Fig. 1, in the ANSKG constructed in the study, BNEG utilizes techniques such as crawlers to obtain external knowledge and automate its construction, and a portion of the information of the ACEG comes from multi-step attack characterization knowledge, which can also be supplemented by the GNSKG [18]. The data of the BNEG, on the other hand, comes from the traffic sensors deployed in the monitored network, which turn the network traffic information into the basic network event information [19]. The generic GNSKG and ACEG store relatively stable security knowledge information, while the BNEG stores real-time traffic information from relatively more active monitored networks. The KG-NSSA model performs graph mining via ACEG to match attack events, and performs situational assessment of asset information in the BNEG via the GNSKG. The construction of the ANSKG is divided into three steps, the first of which is to perform the layered structure design, i.e., the schema layer-data layer (Schema-Data) layered structure. The construction of the Schema layer considers three issues, which are the construction of the domain, the construction of the type, and the determination of the attributes, where the type is contained in the domain. In ANSKG, domains correspond to schemas, i.e., the study splits three domains, corresponding to the three schemas in Fig. 1, and the domains show independent relationships with each other. And the construction of types and the determination of attributes need to be decided according to the actual needs, in which the types contain correlations between them.

Based on the abstract Schema layer, the study can obtain the concrete Data layer, in fact, the construction of ANSKG is the process of using the Schema layer to populate the Data layer. In the Data layer of BNEG, the "edges" mainly play the role of simple association, while most of the attribute information is contained in the entities, so the Data layer of BNEG uses the two-dimensional relational table representation of the traditional database. For the Data layer of ACEG and GNSKG, the study firstly gives an example, as shown in Fig. 2.

In Fig. 2(a), nodes 1, 2, and 3 represent the network nodes, where "ICMP_PING" and "ICMP_REPLY" denote the communication behaviors between the nodes. There is a characteristic of the ACEG in the ANSKG that a single attack signature event is a weakly connected branch that constitutes the entire ACEG. The study denotes the ACEG by E and a single attack characterization event by G_i , where i satisfies the condition shown in Eq. (1).

$$\begin{cases} i \in N^+ \\ 1 \le i \le M \end{cases} \tag{1}$$

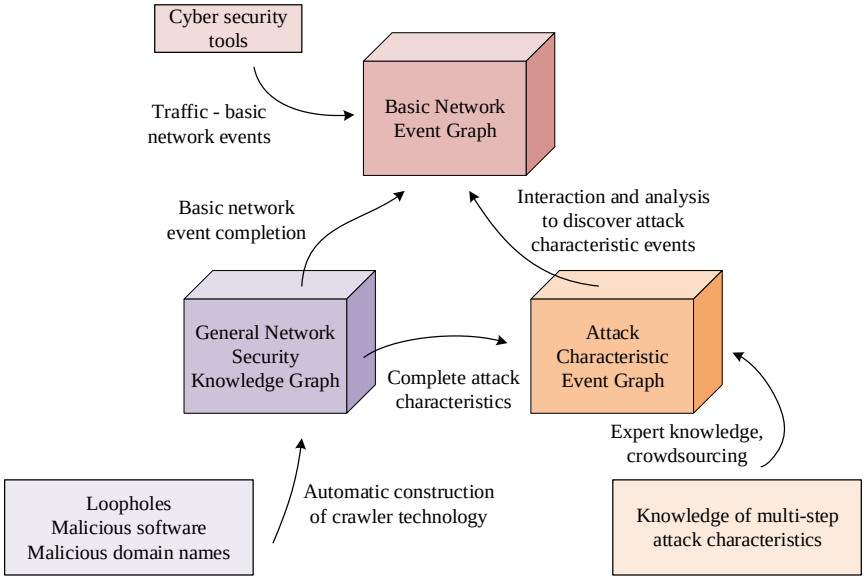
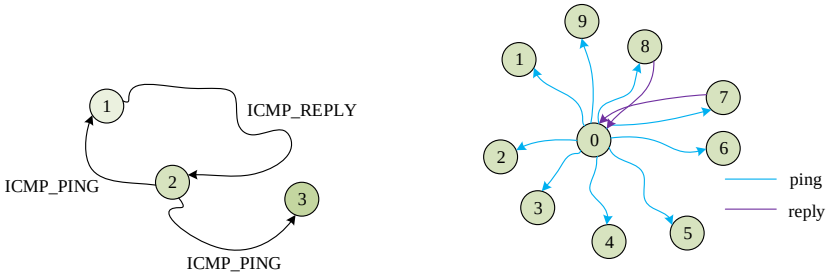


Fig. 1. Asset-based network security knowledge graph diagram.



(a) Example of Characteristic Event Graph Data

(b) Example of General Network Security Knowledge Graph Data

Fig. 2. Data examples for both graphs.

In Eq. (1), N^+ is the set of positive integers and M denotes the total number of weakly connected branches. Then the relationship shown in Eq. (2) exists in ACEG.

$$\begin{cases} G = \cup_{i=1}^{M} G_i \\ G_i \cap G_j = \varnothing \left(1 \le i \le j \le M\right) \\ E = G \end{cases} \tag{2}$$

In Eq. (2), G denotes the concatenation set of weakly connected branches, and $G_i \cap G_j$ denotes the intersection set of weakly connected branches G_i and G_j. The purpose of studying such design of ACEG is to facilitate the KG-NSSA model to traverse each weakly connected branch of the attack feature event mapping when performing attack behavior discovery at a later stage. In Fig. 2(b), the Data layer of GNSKG focuses more on the amount of data, and the amount of data is larger compared to ACEG, reflecting the real network communication situation. In Fig. 2(b), the example given in the study shows that node 0 launches a "ping" request to other nodes, and only nodes 7 and 8 respond with "reply", then nodes 7 and 8 are alive.

The second step of ANSKG construction is data acquisition. Since the data of ANSKG is divided into internal and external data, in which the external data is dominant and more complex, the study proposes an automated approach to construct the external data. The third step of ANSKG construction is the data preprocessing, which is aimed at integrating the data collected during the data acquisition phase with the characteristics of multi-source and heterogeneity to make it meet the requirements of ANSKG. The process of data acquisition and data preprocessing is shown in Fig. 3.

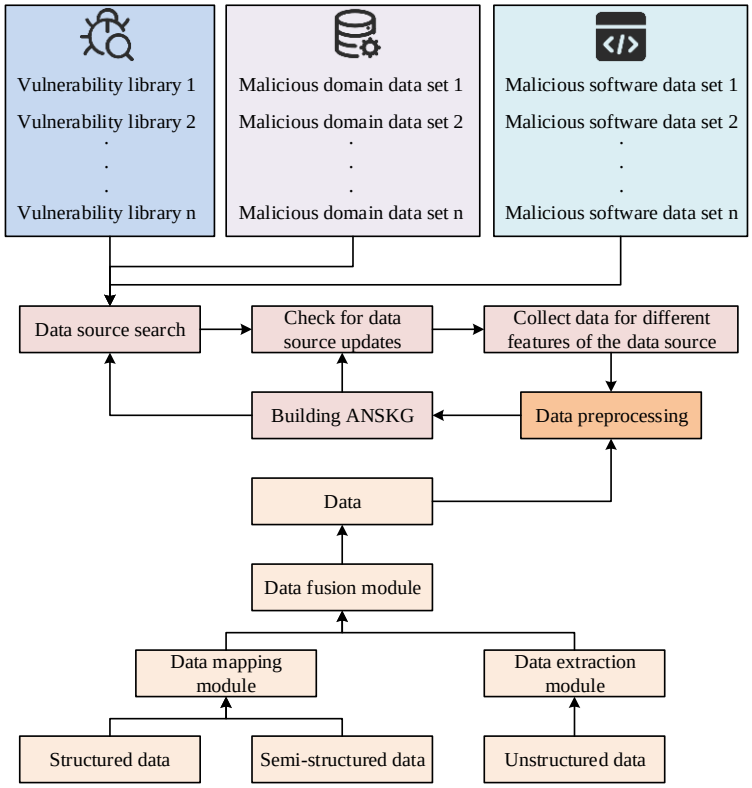


Fig. 3. Process of data acquisition and data preprocessing.

As can be seen from Fig. 3, ANSKG's data collection focuses on automated construction, searching new data sources and checking updates through iterative loops. In this step, the research adopts download link, crawler module and other methods. The collected original data is multi-source and heterogeneous, so it is necessary to integrate and transform these data through data preprocessing to meet the requirements of the ontology model of network security knowledge graph. The pre-processing process includes three key modules: data mapping, data extraction and data fusion, which are mapping, extraction and fusion module respectively. The data mapping module processes structured and semi-structured data and converts it into a format that matches the ontology model. The data extraction module applies natural language processing technology to extract entities and relationships from unstructured texts. The data fusion module is responsible for integrating data from different sources, performing tasks such as entity disambiguation, attribute alignment, and attribute value fusion. Data preprocessing is dedicated to integrating multi-source heterogeneous data, processing structured, semi-structured and unstructured data through data mapping module, and iteratively adjusting to ensure that the data layer meets the requirements of ANSKG. ANSKG provides the functions of query and retrieval, data management and graph calculation on the technical level.

B. *ANSKG-based Scenario Discovery and Situational Understanding*

The establishment of ANSKG lays the foundation for the KG-NSSA model, and in order to further improve the KG-NSSA model, the research addresses the two issues of attack scenario discovery and cybersecurity posture understanding for detailed discussion. Firstly, to understand what is scene discovery and posture understanding, the former refers to the analysis of various events, data and behaviors on the network in order to identify and understand potential threats or attack processes, and the latter refers to the perception, understanding and assessment of various factors and events in the network environment, as well as the real-time monitoring and response to threats and vulnerabilities [20-21]. On this basis, the basic flow of the KG-NSSA model is established, as shown in Fig. 4.

As can be seen from Fig. 4, after establishing the ANSKG, which contains the asset information of the monitored network and its basic traffic information, it can therefore be used for attack scenario discovery, and the study sets the attack discovery parameters and executes the cyber-attack scenario discovery method to extract cyber-attack information in the monitored network from the ANSKG, which will become the key basis for cyber-security understanding. Eventually, taking the asset information in ANSKG as a starting point and combining it with the attack event information, the network posture understanding method is executed to generate detailed posture information about the asset nodes, which will provide comprehensive support for network security decision-making. While attack scenario discovery firstly requires traversing the attack feature event graph to consider all possible attack feature events, secondly, it performs the operations of event feature extraction and feature matching, which actually constitutes the GNSKG, and finally, it mines the attack events and restores the attack scenarios. When the features are extracted, the study introduces attribute graph mining for attack discovery since ANSKG provides functions such as query and retrieval, data management and graph computation. The study stores the extracted features in the form of a five-tuple data set, and sets two other parameters, namely, the time window size parameter TIME_WINDOW and the K-value parameter, the former of which restricts the time interval of the attack discovery analyzed object and thus reduces the amount of data currently analyzed, and the latter of which is used to designate the suspected malicious nodes in the basic event graph of the network. The study needs to find and identify the malicious nodes and record the attack events. The specific process is shown in Fig. 5.

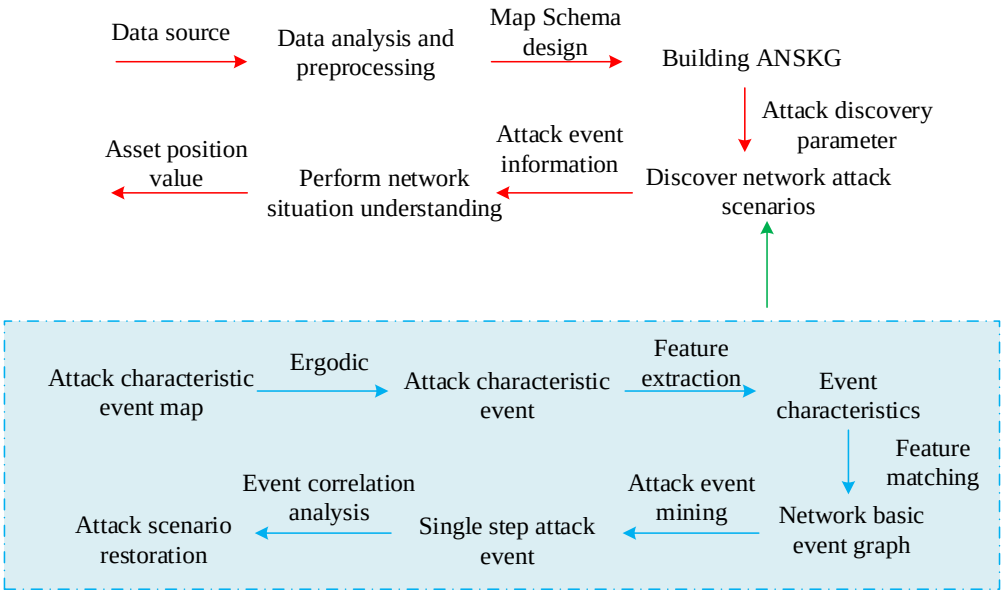


Fig. 4. Basic flow of KG-NSSA model.

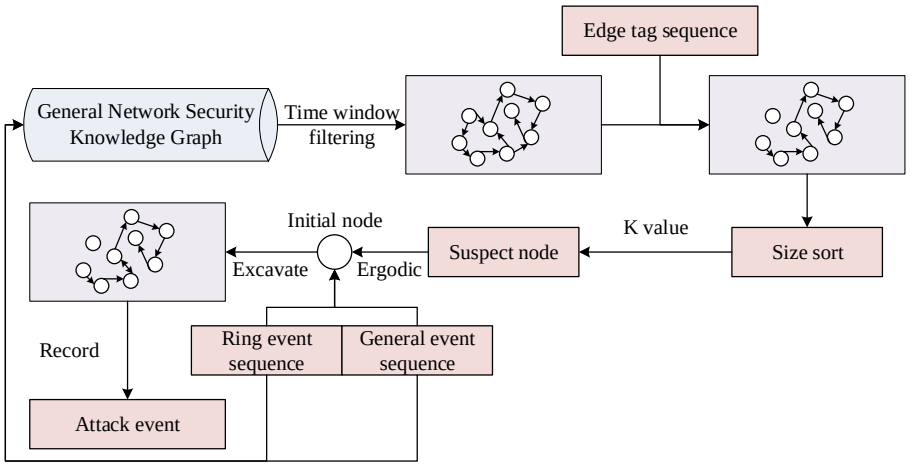


Fig. 5. Schematic diagram of identifying malicious nodes and recording attack events.

As can be seen from Fig. 5, the first step in the KG-NSSA model to find and identify malicious nodes and record the attack events is to extract the ringed event sequences and general event sequences from the quintuple data to construct the edge event sequence matching conditions. Subsequently, subgraphs within the current time window are separated from the basic event graph of the network by time window filtering to extract side event labels. The subgraphs are exported using the edge event labels and sorted by node degree, and the highly ranked nodes are selected as suspicious nodes, which are further traversed to determine the malicious nodes according to the matching conditions and record the information of the relevant nodes affected by their attacks. After obtaining the cyber attack events, it is necessary to find the possible correlation relationship between the cyber attack events and discover the attack scenarios, the study adopts the attack correlation method based on the similarity calculation of attribute graph. Let the correlation between any two attack events E_i and E_j be shown in Eq. (3).

$$Corr\left(E_i, E_j\right) = \sum_{k=1}^{4} \omega_k \cdot C_k, 0 \le i \le j \le n \tag{3}$$

In Eq. (3), C_1 , C_2 , C_3 , C_4 represent the temporal correlation metric between attack events, the spatial correlation metric between attack events, the service correlation metric between attack events, and the type correlation metric between attack events, respectively, and ω_k is the corresponding weight of the four correlation metrics. C_1 As shown in Eq. (4).

$$C_1\left(E_i, E_j\right) = \frac{1}{et_j - st_i + 1} \tag{4}$$

In Eq. (4), et_j and st_i represent the end time of table E_i and the start time of table E_j , respectively. Eq. (4) calculates the time correlation, which measures how close two attack events are in time. C_2 As shown in Eq. (5).

$$C_2\left(E_i, E_j\right) = \frac{\left|V_i \cap V_j\right|}{\left|V_i \cup V_j\right|} \tag{5}$$

In Eq. (5), V_i and V_j represent the influence range of E_i and E_j , respectively. Eq. (5) calculates the spatial correlation degree, which measures the overlap of the spatial scope of the impact of two attack events. C_3 As shown in Eq. (6).

$$C_3\left(E_i, E_j\right) = \frac{\left|P_i \cap P_j\right|}{\left|P_i \cup P_j\right|} \tag{6}$$

In Eq. (6), P_i and P_j denote the set of server-side port numbers of E_i and E_j , respectively. Eq. (6) calculates the service correlation, which measures whether two attacks use the same service or port. C_4 As shown in Eq. (7).

$$C_4\left(E_i, E_j\right) = \frac{1}{2} \cdot a_{ij} \cdot \omega_4 + \frac{1}{2} \cdot b_{ij} \cdot \omega_4 \tag{7}$$

In Eq. (7), a_{ij} denotes the distance metric relationship between the attackers of E_i and E_j , and b_{ij} is a bool variable that denotes the type relationship between the tags of E_i and E_j . a_{ij} The value range is [0, 1], as shown in Eq. (8).

$$a_{ij} = \begin{cases} 1, d_{ij} = 0 \\ \dfrac{1}{d_{ij}^2}, 0 < d_{ij} \le N \\ 0, d_{ij} > N \end{cases} \tag{8}$$

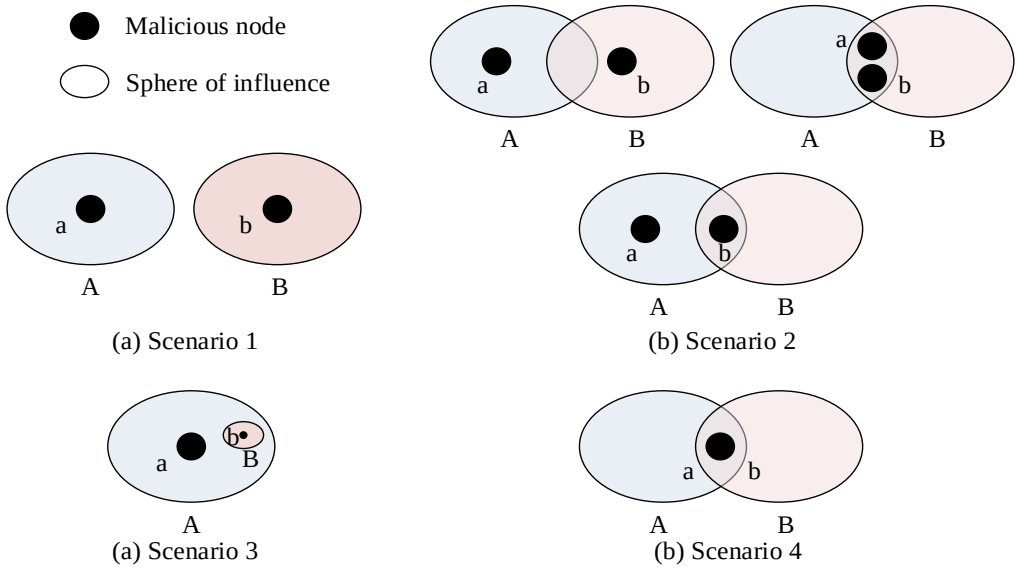


Fig. 6. Four scenarios of situation understanding.

In Eq. (8), d_{ij} denotes the distance between the attackers of E_i and E_j, and N denotes the threshold of the distance. b_{ij} It takes the value of 0 or 1. Specifically, b_{ij} is 0, which means that there is no actual relationship between the tags of E_i and E_j, and vice versa b_{ij} is 1. Eq. (3) to (8), combined, provide a method for the KG-NSSA model to quantify the similarity between different attack events, which helps to identify attack events with common characteristics, thereby improving the accuracy of network security situation awareness. Finally, the KG-NSSA model needs to be situationally understood for cybersecurity, and the study considers four scenarios, as shown in Fig. 6.

Fig. 6(a) indicates that the malicious nodes are different and the spheres of influence do not intersect. Fig. 6(b) indicates that the malicious nodes are different and the influence ranges intersect. Fig. 6(c) indicates that the malicious nodes are different and the influence ranges are included. Fig. 6(d) indicates that the malicious nodes are the same. Different malicious node scenarios create different risks. Further, the study borrows the PageRank algorithm for situational understanding, specifically, the PageRank value of asset nodes is calculated based on GNSKG, and then the converged PageRank value is used as the base value of the node weights, and based on the risk information of cyber-attack events, the study considers to take into account the impact of different risk levels, and the weights of the nodes are corrected, and in addition, for malicious nodes, the weights are additionally corrected to increase their influence. At the same time, the KG-NSSA model can be further corrected for posture based on the threat intelligence information in ANSKG. Through this process, the KG-NSSA model is able to update the risk posture of asset nodes in real time, quantify the impact of cyber-attacks into specific node weights, and provide a more comprehensive and accurate quantitative assessment of the network security posture.

IV. Network Security Situational Awareness Model Simulation Experiment and Analysis

In order to verify the validity and superiority of the KG-NSSA model proposed in the study, the study uses the DARPA2000 dataset from MIT for simulation verification. The study specifies the experimental environment, the sample LLDos 1.0 attack scenarios in the dataset, and the experimental parameters, as shown in Table I.

TABLE I. Experimental environment and process

Experimental environment	
Configuration item	***Configuration details***
Processor	Intel® Core™ i5-4200U CPU@1.60GHz
Internal memory	8.00G
Hard disk	500.00G
Operating system	Ubuntu 16.04 LTS (Xenial Xerus)
Attack scenario phase	
Phase 1	A remote node initiates IPsweep to detect living nodes
Phase 2	A probe is sent to the living node to obtain information about the host running the sadmind daemon
Phase 3	The target host is hacked through the vulnerability of the sadmind daemon and the root execution permission of the target host is obtained
Phase 4	Install DDoS malware on the target host
Phase 5	Launching DDoS attacks
Experimental parameter	
TIME_WINDOW	20.00min
K	3.00
ω_1	0.20
ω_2	0.20
ω_3	0.20
ω_4	0.40
α	0.85

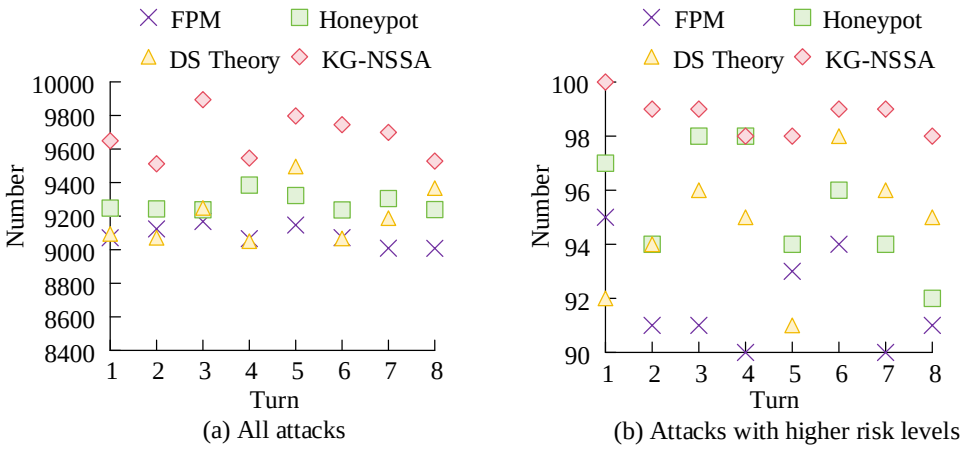


Fig. 7. Statistics on the number of detected attacks.

The specific environment configuration of the experiment, the five stages of the attack scenario, and the parameter values can be obtained from Table I, where α is the parameter in the PageRank algorithm. As a result, the ANSKG constructed by the study contains a total of 276210 edges and 34167 vertices, and in addition, the study sets up five attack features, which are named as L1-L5. The study is based on the Frequent Pattern Mining (FPM) method, Honeypot Technology. Honeypot), and DS Evidence Theory method (Dempster-Shafer Evidence Theory, DS Theory) as comparison algorithms. First of all, the study sets 10,000 attacks for 8 rounds, in which there are 100 attacks with high risk level, and the number of attacks detected by KG-NSSA model and comparison algorithm are counted, and the results are shown in Fig. 7.

As can be seen from Fig. 7(a), for 10,000 attacks, the KG-NSSA model detects an average of 9,706 attacks in eight rounds, while FPM, Honeypot & DS Theory detect an average of 9,112, 9,307, and 9,260 attacks, respectively. As can be seen in Fig. 7(b), for attacks with higher risk levels, the KG-NSSA model detects more than 98 attacks on average, while all three comparison algorithms, FPM, Honeypot & DS Theory, average less than 95 attacks. Further, the study statistically analyzes the correct rate, recall rate and false alarm rate of the detected attacks, and the results are shown in Fig. 8.

From Fig. 8(a), it can be seen that the KG-NSSA models are all detected correctly above 99%, while FPM, Honeypot & DS Theory are only detected correctly in the range of 97%-99%. From Fig. 8(b), it can be seen that the average recall of KG-NSSA model is 99.04%, while the average recall of FPM, Honeypot & DS Theory are 97.12%, 97.36% and 97.15%, respectively. As can be seen from Fig. 8(c), the false alarm rate of KG-NSSA model is as low as 0.26% and as high as 1.67%, which is much lower than the three comparison algorithms of FPM, Honeypot & DS Theory. This shows that the KG-NSSA model not only detects a large number of attacks, but also has an extremely high correct rate. Further, since a large amount of data and malicious nodes are generated when performing attacks, the study compares the memory footprint with the algorithm runtime, and the results are shown in Fig. 9.

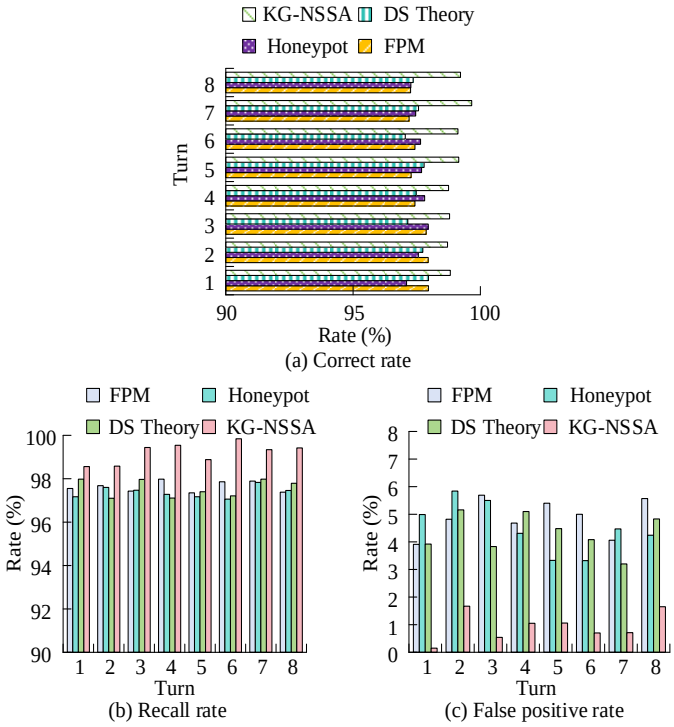


Fig. 8. Statistical results of correct rate, recall rate and false positive rate.

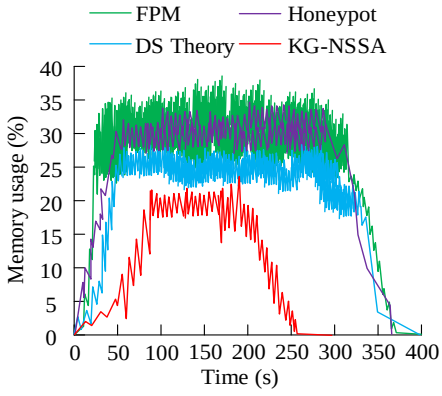


Fig. 9. Comparison between memory usage and running time.

As can be seen from Fig. 9, for the KG-NSSA model, its memory occupancy rate reaches up to 22.67%, and its time for completing the attack detection is 258.4 s. For FPM, its memory occupancy rate floats between 25% and 38%, and it completes the attack detection at about 37.72.6 s. The average memory occupancy rate of the FPM algorithm is about 1.5 times higher than the FPM algorithm, but it is more stable. For Honeypot, its average memory usage is close to that of the FPM algorithm, but its floating interval is smaller and more stable. For DS Theory, its average memory occupancy is around 25%, but the time to complete the detection is close to 400 s. It can be seen that the KG-NSSA model has absolutely excellent performance. On this basis, the study discusses the network posture understanding results of KG-NSSA model, and its results are shown in Table II.

TABLE II. NETWORK SITUATION UNDERSTANDING RESULTS OF KG-NSSA MODEL

Stats	**Value**				
Tag	***L1***	***L2***	***L3***	***L4***	***L5***
Start time (s)	696.02	1687.34	3194.72	44202.10	46471.14
End time (s)	696.63	2110.18	43301.71	44253.96	47279.78
Malicious node	202.77.162.213	202.77.162.213	202.77.162.213	202.77.162.213	131.84.1.31
Sphere of influence	Most nodes in the 172.16.112.0/24 network segment and so on	172.16.115.20 172.16.115.87 172.16.114.10 172.16.114.20	172.16.115.20 172.16.112.10 172.16.112.50	172.16.115.20 172.16.112.10 172.16.112.50	Large number of external server addresses

As can be seen from Table II, the success rate of the KG-NSSA model in matching the L1-L5 attack feature events reaches 100%, which provides a basis for attack association. Analyzing the attack events matched by the attack scenario discovery step, it can be seen that the attack events corresponding to L1-L4 are all initiated by node 202.77.162.213, and the influence range of the attack events is gradually narrowed, which precisely reflects the process of external malicious nodes searching for injectable nodes. l5 attack events have a wide influence range, last for a long time, and affect external servers, which is not directly associated with L1-L4 attack events from the viewpoint of malicious nodes only. In terms of malicious nodes only, there is no direct correlation with the L1-L4 attacks. Finally, the study also discusses the posture change of node 172.16.115.20, and the result of its change over time is shown in Fig. 10.

In Fig. 10, the KG-NSSA model updates the posture of the incremental part of the network basic event graph in this time period every 40s, and node 172.16.115.20 has a significant posture value when the attack event occurs (time periods 200s to 240s, 280s to 320s, 400s to 440s, 520s to 560s, and 760s to 80s), and the value of the posture has a significant The change in posture can well reflect the impact of the attack event on the posture of the asset node, whereas DS Theory only reflects the impact of the attack event on the posture of the asset node. The DS Theory only reflects the first four attack phases, but not the last attack phase, and the node posture will fall back to the normal value and remain stable in the gap between the attack phases, which cannot reflect the impact of the past attack events on the node posture. The FPM method reflects the upward trend of the node posture better, but the posture value only climbs significantly in three places, which cannot reflect all the attack phases. Honeypot's posture curve also does not reflect the progressive relationship of each attack stage in the attack scenario well, and is easily affected by the changes in the normal traffic of the nodes, leading to a certain degree of misjudgment. Therefore, the KG-NSSA model proposed in the study can effectively sense the posture of network security.

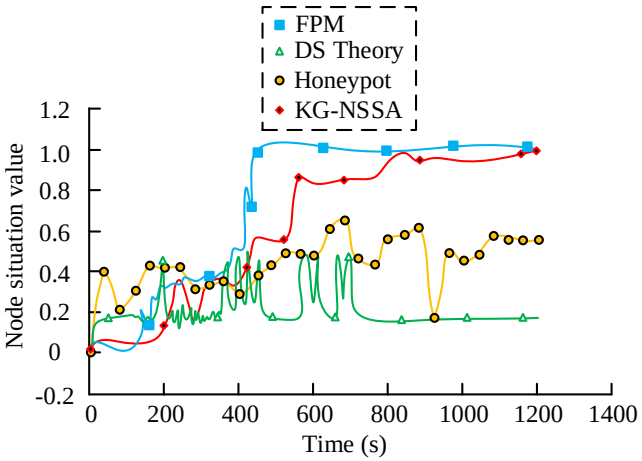


Fig. 10. Changes in situation over time.

V. RESULTS AND DISCUSSION

The proposed KG-NSSA model can effectively improve the capability of network security situation awareness by constructing ANSKG and using attribute graph mining technology and similarity calculation method. The experimental results show that the KG-NSSA model is superior to the existing comparison algorithms such as FPM, Honeypot and DSTheory in the accuracy, recall rate and false positive rate of attack detection. Specifically, the KG-NSSA model detected an average of 9,706 attacks out of 10,000 attacks, with more than 98 detections for 100 high-risk attacks. In addition, the average correct rate, recall rate and false positive rate of the model are 99.48%, 99.04% and 0.86%, respectively, showing high detection efficiency and accuracy. In terms of performance, the maximum memory usage of KG-NSSA model is only 22.67%, and the time to complete attack detection is 258.4 seconds, which is significantly lower than that of the comparison algorithm, indicating that the model has obvious advantages in resource utilization and response speed. The results of network situation understanding show that KG-NSSA model can accurately reflect the impact of attack events on asset node situation, and provide strong support for network security management and decision-making. Although the KG-NSSA model has performed well in experiments, there is still room for further improvement and expansion. First, while the current model focuses on situational awareness of the current

situation, future work could explore how to use the KG-NSSA model for predictive analysis of future cyber threats to achieve more proactive security protection. Secondly, updated machine learning algorithms can be introduced to enable the model to adaptively update the knowledge graph according to new network behaviors and attack patterns, and improve the generalization ability and adaptability of the model. Finally, the KG-NSSA model can be applied to other fields, such as industrial control systems, Internet of Things devices, etc., to explore its effectiveness and applicability in different environments.

VI. CONCLUSION

Aiming at the threats to network security brought by social development, this paper studies and applies knowledge graph technology to construct ANSKG, proposes a KG-NSSA model related to network attack scene discovery and network security situation understanding, and uses attribute graph mining method, attribute graph similarity calculation method and PageRank algorithm to improve the KG-NSSA model. Finally, the research content is verified. In the experiment, ANSKG contains 276,210 edges, 34167 vertices, and sets 5 attack features. By counting the number of attacks detected, KG-NSSA model detected 9706 attacks on average in 8 rounds, while FPM, Honeypot and DS Theory detected 9112 attacks, 9307 attacks and 9260 attacks on average, respectively. For attacks with higher risk level, the average detected attacks of KG-NSSA model was greater than 98, while the average detected attacks of FPM, Honeypot and DS Theory were all below 95. Secondly, the detection accuracy of KG-NSSA model is above 99%, while the detection accuracy of FPM, Honeypot and DS Theory is only between 97% and 99%. The average recall rate of KG-NSSA model is 99.04%, which is higher than the average recall rate of FPM, Honeypot and DS Theory. The false positive rate of KG-NSSA model was 0.26% and 1.67% respectively. In addition, the memory usage and running time of KG-NSSA model are much lower than that of the comparison algorithm, with the highest memory usage reaching 22.67% and the time to complete attack detection being 258.4s. Finally, the network situation understanding results show that the KG-NSSA model can effectively perceive the network security situation. In summary, the proposed KG-NSSA is effective and has excellent performance, and has good application potential in the field of network security situation awareness. However, the study did not analyze the situation prediction, which can be further discussed in the future.

ACKNOWLEDGMENT

This research was supported by the Science and Technology Key Project of Henan Province (No.222102210128,232102321072), Science and Technology project of Nanyang (No. KJGG036).

REFERENCES

[1] L. Hong, H. Guo, J. Liu and Y. Zhang, "Toward Swarm Coordination: Topology-Aware Inter-UAV Routing Optimization," IEEE Trans. Veh. Technol., vol. 69, no. 9, pp. 10177-10187, September 2020.

[2] A. Karami, V. Shah, R. Vaezi and A. Bansal, "Twitter Speaks: A Case of National Disaster Situational Awareness," J. Inf. Sci., vol. 46, no. 3, pp. 313-324, March 2020.

[3] M. R. Endsley, "A Systematic Review and Meta-Analysis of Direct Objective Measures of Situation Awareness: A Comparison of SAGAT and SPAM," Hum. Factors, vol. 63, no. 1, pp. 124-150, February 2021.

[4] P. Wang and M. Govindarasu, "Multi-Agent Based Attack-Resilient System Integrity Protection for Smart Grid," IEEE T. Smart. Grid., vol. 11, no. 4, pp. 3447-3456, April 2020.

[5] G. C. Kessler, "Protected AIS: A Demonstration of Capability Scheme to Provide Authentication and Message Integrity," TransNav: Int. J. Mar. Navigation Safety Sea Transport., vol. 14, no. 2, pp. 279-286, April 2020.

[6] L. Jaeger and A. Eckhardt, "Eyes Wide Open: The Role of Situational Information Security Awareness for Security - Related Behavior," Inform. Syst. J., vol. 31, no. 3, pp. 429-472, June 2021.

[7] Z. Chen, "Research on Internet Security Situation Awareness Prediction Technology Based on Improved RBF Neural Network Algorithm," J. Comput. Cogn. Eng., vol. 1, no. 3, pp. 103-108, March 2022.

[8] L. Tan, K. Yu, F. Ming, X. Cheng, and G. Srivastava, "Secure and Resilient Artificial Intelligence of Things: A HoneyNet Approach for Threat Detection and Situational Awareness," IEEE Consum. Electr. M., vol. 11, no. 3, pp. 69-78, October 2021.

[9] X. Liu, Z. Li, Z. Tang, X. Zhang, and H. Wang, "Application of Artificial Intelligence Technology in Electromechanical Information Security Situation Awareness System," Scal. Comput. Pract. Exp., vol. 25, no. 1, pp. 127-136, March 2024.

[10] B. Hamdaoui, M. Alkalbani, A. Rayes, and N. Zorba, "IoTShare: A Blockchain-Enabled IoT Resource Sharing On-Demand Protocol for Smart City Situation-Awareness Applications," IEEE IoTJ, vol. 7, no. 10, pp. 10548-10561, October 2020

[11] M. H. Junejo, A. A. H. Ab Rahman, R. A. Shaikh, K. Mohamad Yusof, I. Memon, H. Fazal, et al, "A privacy-preserving attack-resistant trust model for internet of vehicles ad hoc networks," Sci. Programming-neth, pp. 1-21, 2020.

[12] N. Ahmed, Z. Deng, I. Memon, F. Hassan, K. H. Mohammadani, et al, "A survey on location privacy attacks and prevention deployed with IoT in vehicular networks," Wirel. Commun. Mob. Com., 2022.

[13] Memon I, Arain Q. "A Dynamic Path Privacy Protection Framework for Continuous Query Service Over Road Networks," World Wide Web, vol. 20, no. 4, pp. 639-672, Aug. 2017.

[14] Z. Li, H. Liu, Z. Zhang, T. Liu, and N. N. Xiong, "Learning Knowledge Graph Embedding with Heterogeneous Relation Attention Networks," IEEE Trans. Neural Netw. Learn. Syst., vol. 33, no. 8, pp. 3961-3973, August 2021.

[15] R. Goel, S. M. Kazemi, M. Brubaker, and P. Poupart, "Diachronic Embedding for Temporal Knowledge Graph Completion," Proc. AAAI Conf. Artif. Intell., vol. 34, no. 04, pp. 3988-3995, April 2020.

[16] S. K. Mohamed, V. Nováček, and A. Nounu, "Discovering Protein Drug Targets Using Knowledge Graph Embeddings," Bioinformatics, vol. 36, no. 2, pp. 603-610, January 2020.

[17] C. K. Long, P. Van Hai, T. M. Tuan, L. T. H. Lan, P. M. Chuan, and L. H. Son, "A Novel Fuzzy Knowledge Graph Pairs Approach in Decision Making," Multimed. Tools Appl., vol. 81, no. 18, pp. 26505-26534, July 2022.

[18] P. P. Groumpos, "A Critical Historic Overview of Artificial Intelligence: Issues, Challenges, Opportunities, and Threats," Artif. Intell. Appl., vol. 1, no. 4, pp. 197-213, January 2023.

[19] S. Zeebaree, S. Ameen, and M. Sadeeq, "Social Media Networks Security Threats, Risks and Recommendation: A Case Study in the Kurdistan Region," Int. J. Innov. Creat. Chang., vol. 13, no. 7, pp. 349-365, July 2020.

[20] A. K. Jain, S. R. Sahoo, and J. Kaubiyal, "Online Social Networks Security and Privacy: Comprehensive Review and Analysis," Complex Intell. Syst., vol. 7, no. 5, pp. 2157-2177, October 2021.

[21] Q. Zhou, M. Shahidehpour, A. Alabdulwahab, and A. Abusorrah, "A Cyber-Attack Resilient Distributed Control Strategy in Islanded Microgrids," IEEE Trans. Smart Grid, vol. 11, no. 5, pp. 3690-3701, September 2020.