Research Article

Proliferation of Cyber Situational Awareness: Today's Truly Pervasive Drive of Cybersecurity

Hafiz Muhammad Jamsheed Nazir and Weihong Han

Cyberspace Institute of Advanced Technology, Guangzhou University, Guangzhou 510006, China

Correspondence should be addressed to Weihong Han; hanweihong@gzhu.edu.cn

Received 7 August 2021; Accepted 20 December 2021; Published 25 January 2022

Academic Editor: Kuo-Hui Yeh

Copyright © 2022 Hafiz Muhammad Jamsheed Nazir and Weihong Han. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Situation awareness (SA) issues necessitate a comprehension of present activities, the ability to forecast, what will happen next, and strategies to assess the threat or impact of current internet activities and projections. These SA procedures are universal, domain-independent and can be used to detect cyber intrusions. This study introduces cyber situation awareness (CSA), its origin, conception, aim, and characteristics based on an analysis of function shortages and development requirements. Furthermore, we discussed the CSA research framework and examined the research history, which is the essential aspect, and assessed the present issues of the research as well. The assessment approaches were divided into three methods: mathematics model, knowledge reasoning, and pattern recognition. The study then goes into detail regarding the core idea, assessment procedure, strengths, and weaknesses of novel approaches, and then, it addresses CSA from three perspectives: model, knowledge representation, and assessment methods. Many common approaches are contrasted, and current CSA application research in the realms of security, transmission, survivability, and system evaluation is discussed. Finally, this study summarized the findings of the present from technical and application systems, outlined CSA's future development directions, and provided adversary activities and information that can be used to improve an organization's SA operations.

1. Introduction

The rapid development of the internet has brought convenience and fun to life; however, with it, many threats and risks have appeared in the network environment, and cyberspace security problems exist in every corner of people's lives [1]. As a complex giant system, the internet has various networking methods [2]. The proliferation of new networks, such as sensor networks, AdHoc, and space-based networks, makes the topological structure [3]. It is complicated to fix. Network equipment is inconsistent, heavy, and extremely common. Information flows more frequently, network traffic increases, and network load increases. New applications [4]: with the constant appearance of VoIP, P2P, grid, and other applications, an overlay network is formed that overrides the transmission network; the network is constantly subject to failures, attacks, the threat of disasters and emergencies; availability, security, and survivability are

facing severe challenges; network operating conditions are changing rapidly [5]. With the rapid growth of the internet with swift expansion, complexity and uncertainty also increase, and the ability to make meaningful descriptions of its characteristics decreases accordingly. Traditional network management functional units, being in an independent working state and lacking an effective information extraction and information fusion mechanism, are unable to establish connections between the network resources, and global information performance capabilities are poor. The massive network management information not only fails to strengthen management but increases the burden on network administrators [6]. Modern network management must be able to change rapidly and dynamically. In a complex environment, uncertain network management information is efficiently organized, analyzed, and evaluated. The detailed information of the managed objects is provided, which improves the network administrator's understanding

of the entire cognition and understanding of network operation status, providing diversified and personalized management services, assisting commanders to make decisions quickly and accurately, and making up for the current situation [7]. Figure 1 depicts the model of situational awareness.

Bass was the first to propose the concept of CSA in 1999 and point out that “based on converged network SA” will surely become the development direction of network management [1, 8]. The network situation refers to the operation status of various network equipment and network behavior. The current state and changing trends of the entire network are constituted by factors, such as user behavior [9]. The key situation that emphasizes the relationship between the environment, dynamics, and institutions is a state, trend, aggregate, and macroconcept, a single situation or state called situational awareness [10]. It refers to the acquisition, understanding, evaluation, display, and future development of elements that can cause changes in the network situation in a large-scale network environment trend prediction. The possible concept of action and military requirements as an integral part of data fusion-level 2 fusion is an important part of the decision-making process section [11]. The goal of CSA is to integrate the early theories of attacks and network management in a complex environment with real-time dynamic changes to efficiently organize various information, integrate the existing indicators that indicate network characteristics, display the macro and status of the network, become an administrator to strengthen the network with the ability to understand, and provide decision-making support for high-level commanders. Compared with traditional network management systems, CSA has the following characteristics [12]:

- (1) Uses data fusion technology, comprehensively considers the diversified factors affecting the network, provides a comprehensive and a macroscopic view of the network status, strengthens the understanding and control of the network, and gradually, the response of network administrators
- (2) Becomes a platform for integrated unit network management, changes the current relationship between the independent work of each unit network management, and realizes information sharing
- (3) Provides support for potential risks and plans

This study systematically and comprehensively introduces the research progress of network SA, and the main contributions are as follows: the paper discussed the CSA research framework, and the research content summarized the characteristics and existing problems. Also, it outlined the classification method for situation assessment and pointed out the lack of unified and standardized evaluation standards that are the sources of complexity and diversity of current evaluation methods and their governance. Furthermore, the three aspects of the CSA model, knowledge representation (KR), and evaluation methods are systematically discussed. The study also focuses on the current status of applied research, the common problems, and the

shortcomings of existing research. In-depth analysis and comparison of multiple evaluation methods pointed out the future research trend, summarized the essence of CSA from the multiple aspects of technology, function, and system.

The remainder of this paper is organized as follows: Section 2 provides an overview of network SA, and Section 3 presents cyber SA models. Section 4 presents the network situation assessment method, and Section 5 discusses cyber SA application approaches in more detail. Finally, Section 6 offers an interesting summary and future trends in this useful field.

2. Overview of Network Situation Awareness

The majority of businesses companies, at least one cloud computing service in 2019, and cloud data centers are predicted to process 94% of workloads by the end of 2021. Moving information technology (IT) infrastructure to specialist cloud providers has clear financial and operational benefits. Security concerns have arisen as a result of the large amounts of private and sensitive data held in cloud computing infrastructures. In this section, we clarify the interesting research content of CSA by proposing a network SA research framework. Also, the paper proposes a method to classify existing evaluation methods and summarize the historical process of technological development. Finally, it analyzes the characteristics of existing research and points out the problems.

2.1. Research Framework. As a part of data fusion, CSA does not exist in isolation [1]. It acquires various network management data from level 1 integration and provides situation information for level 3 integration for threat analysis and decision support, and it is closely related to other integration levels. The data communication between the layers is frequent and the methods are connected. There is no clear boundary, and it exists as a whole [13]. Therefore, CSA research includes many aspects, and its overall research framework is shown in Figure 2. The CSA research framework summarizes the CSA research content, embodies the closing-the-loop concept, highlights the essence of dynamic looping and continuous refinement, and emphasizes the important role of feedback. As shown in Figure 2, CSA research content is extensive, including its own functional refinement, theoretical methods of key technologies, communication, and interaction with other integration levels. The current research mainly focuses on three aspects: the CSA model, KR, and evaluation methods [14]. Among them, the CSA model is the focus of research. It is relatively mature and unified. The evaluation method is the core of CSA research. It mainly studies the application of existing theories in situation assessment, while there is relatively little research on KR. According to different application fields, network situations can be divided into security situations, topology situations, and transmission situations. The survivability situation, etc., starting from Bass, and most researches are carried out around security situation [15].

FIGURE 1: Model of situational awareness.

FIGURE 2: Cyberspace situational awareness framework.

2.2. Related Study. As a part of data fusion, cyber-SA first introduced the data fusion model, modified and refined the original model, established the CSA model, and clarified the function of the CSA. There are not many studies on the model [16]. Endsley's viewpoint is essentially a top-down driven mental paradigm divided into three major components: perception, comprehension, and projection. SA, according to Endsley, is the perception of environmental factors, the interpretation of their meaning, and the projection of their status "to enable decision superiority," to Endsley's definition. The JDL data fusion paradigm, on the other hand, takes a more data-centric, bottom-up approach. Level-0, subobject identification, level-1, object identification, level-2, situation assessment, level-3, threat assessment,

level-4, and process refinement make the fusion model. Situation assessment is defined by the JDL data fusion model as the evaluation and prediction of relationships among things. After 2006, the CSA model is relatively complete and mature, and the related research has not made significant progress. Evaluation is the core of CSA function, and its research is vast. The MM method was first used for situation assessment. It focuses on different factors that affect the network situation and integrates multiple factors [17]. The situation factor reflects the state of the situation from multiple perspectives; however, MM only solves the multiattribute fusion and does not involve the fusion of multisource data. The model used by MM is fixed, and only definite evaluation results can be obtained, ignoring the

uncertainty factor. To solve the two major problems of MM, the KR method appeared. On the one hand, KR uses fuzzy sets, probability theory, evidence theory, etc., to process uncertain information [18]. On the other hand, multisource and multiattribute information is gathered by reasoning. The KR method represented by Bayesian networks became a hot research topic [16]. At this point, a large number of documents have emerged, and repeated research is serious. How to obtain inference rules and prior probabilities, especially for a new research direction, such as CSA, has become the biggest challenge in front of KR, after the PR method came into being. PR has gradually attracted attention since 2005 with its strong learning ability. It has been trained on practice samples or historical data to mine the knowledge of the situation mode division—scientific and objective [19]. Most of the existing evaluation methods will more or less introduce data mining. It also reveals the trend of the comprehensive use of multiple methods. The development of evaluation methods embodies the exploration process of “problems appear to $\rightarrow$ solve problems.” Research on the representation of knowledge is rare and started relatively late. Damage assessment is essential for securing company networks and systems. If security officers have a thorough grasp of the consequences and the impact of cyberattack actions, they will be better positioned to make the right cyber-defense decisions and execute the right cyber-defense operations. As a result of this methodology, a novel production environment damage assessment architecture arises. Even though this method does not cover all abstraction levels, it indicates that damage assessment across layers may be done in complex software systems.

2.3. Classification of Assessment Methods. At the core of CSA research, there are various evaluation methods [15, 20]. These methods can be roughly divided into three categories: (1) the evaluation function based on the mathematical model to establish the mapping relationship between the situation factor set R and the situation space $\theta = f(r_1, r_2, \dots, r_n)$ is presented, where $r_i \in R (1 \leq i \leq n)$ is the situation factor [15]. The situation contains many conflicts and incommensurable and uncertain complex situation factors. These factors have a hierarchical structure and can be divided and refined layer-by-layer. The authors of [21] discussed the composition and structure of network situation factors. Traditional and general multiobjective decision-making theory and utility-related theoretical methods, such as maximum membership principle, distance deviation method, scoring method, and multiattribute utility function, can be used for situation assessment. The most commonly used methods are weight analysis, formula method, and set pair analysis. The MM method establishes clear mathematical expressions. The model is easy to understand and can establish a continuous situation space, giving a favorable or unfavorable judgment. The discontinuous results are convenient for the comparison of the advantages and disadvantages of SA [22]. However, there is no unified scientific method for the construction of the evaluation function and the selection of parameters and generally rely on domain knowledge and expert experience

which inevitably contains subjective opinions and lacks scientific and objective basis. In addition, situation assessment uses natural language to express knowledge in most cases, and this kind of knowledge cannot be easily transformed into mathematical expressions that are easy to be processed by machines. Therefore, the establishment of a mathematical model for natural language conditional statements has also become the difficulty of the method [23].

2.4. Knowledge Reasoning-Based Method. The basic idea of knowledge reasoning (KR) is to receive level 1 fusion output under the premise of known empirical knowledge and prior probability, step by step through a certain relationship based on the real-time monitoring of data information [24]. Reasoning to obtain judgments about the current situation, the situation space can be divided and the classification or classification results can be given [25]. The KR method can be divided into logical reasoning based on production rules, reasoning based on graph models, and probabilistic reasoning based on evidence theory. Different methods include fuzzy reasoning, Bayesian networks, Markov processes, and the DS evidence theory [15, 16]. The KR method can simulate the way of human thinking compared with the MM method. The application of knowledge is integrated into the reasoning process. It has a certain degree of intelligence, similar to the process of problem-solving by experts. The results of the assessment establish a discrete situation space that can determine the pros and cons of the situation or specify the type of situation. It is clear at a glance, which is convenient for understanding and grasping the situation [26, 27]. The difficulty of this method is how to obtain the knowledge needed to build the model. If it is based on experience, it will be strongly subjective. If it is through machine learning, the related research is still relatively small, and how to learn the knowledge of “simulating human thinking” is even more difficult. It can be seen that the advantages of this method have become the biggest obstacle. In addition, this method maintains a large number of inference rules. The space overhead and inference costs are high. How to deal with large-scale problems is another need. Another question that needs to be considered is how to deal with large-scale problems.

The method based on pattern recognition (PR) is divided into two stages: template establishment and pattern matching. In the first stage, the situation model is established [28]. Version, based on the division of the situation space, identifies all possible situation states. There is no unified standard for the division, and the situation can be divided into a different situation that is graded for the same type [29]. Figure 3 shows the cyber SA model.

At the second stage of pattern matching, the correlation between the measured and the template data is calculated. If the correlation coefficient reaches a predetermined threshold, it is considered that the matching is successful, and the situation state is determined. The establishment of a template is the focus of the PR method, and the key is to choose a classification method. In addition to relying on expert experience and domain knowledge, machine learning is also

FIGURE 3: Cyber situational awareness model.

the main means of classification, which is obtained from training samples or cases for the knowledge of classification. The representative methods include case-based reasoning, neural networks, gray relational analysis, rough set theory, and cluster analysis. These methods are also generally used for pattern matching [29]. The PR method introduces a machine learning mechanism that is scientific and objective and can easily obtain the knowledge of situation division from historical data or cases. The method has a large amount of calculation and has good results in a non-real-time environment. However, it may not meet the requirements in a real-time environment. Some studies use heuristic algorithms to improve high efficiency. In addition, as the classification knowledge is obtained from historical data through machine learning, it is difficult for the machine to give an intuitive explanation, which is not conducive to understanding [16].

2.5. Research Characteristics and Existing Problems. Network management requirements and broad application prospects have jointly established the important position of CSA, and related research has continued to deepen. From the preliminary exploration of existing CSA, it can be seen that related research has the following characteristics [30, 31]: (1) Network security SA is researched in other areas, such as traffic, faults, topology, and survivability, and these are rarely involved. (2) System architecture is the focus of research and is relatively mature. Although there are differences, it accepts the joint directors of laboratories (JDL) data, the design ideas of the fusion model, and the Endsley SA model. (3) The representation of the network system is based on a hierarchical structure. The representation of uncertain information is mostly simple grading and is converted into discrete data. (4) The evaluation of the method is mainly based on weight analysis [15]. Some

studies have tried to introduce the mathematical method of data fusion into CSA.

The existing research has paid attention to the existing network management technologies and failed to integrate the security situation of each unit, and it cannot achieve a comprehensive assessment and presentation of the overall situation [15, 32]. Due to the lack of comprehensive and systematic research on the entire CSA system, the concept of “situation” is too narrow and does not reflect the overall and macroscopic characteristics of the situation. Secondly, KR is not sound enough. Although the use of a hierarchical structure to represent a network system is simple, intuitive, and easy to analyze, it cannot show the intricacies between network elements. Relationships are not conducive to mining potential situational information within multisource and multiattribute data. In addition, in terms of information representation, how to select and expand the feature measurement used for situation assessment and establish a reasonable and complete indicator system remains to be studied [25, 32]. Thirdly, the situation assessment lacks a unified standard. It is partly because the concept of a situation is relatively abstract. As for what kind of situation is good and how good it is, the degree is often just a feeling. Scoring and grading lack a scientific basis and are not persuasive. However, the deeper reasons are neither clear nor convincing. The formal definition of the situation and the lack of indicators and methods to measure the pros and cons of the evaluation results cannot form a consensus on the situation and the situation assessment [33]. The theoretical methods in the field of data fusion have been applied to the situation evaluation stage, and constantly changing new methods have been introduced. It includes a lot of repeated research and the use of certain mathematical methods to increase the depth of the theory. The evaluation criteria are unified, and the evaluation methods have clear goals and directions [34].

The fourth point is that the existing research focuses on the situation alone and is relatively independent. The vertical (level 1/3) and horizontal (level 4/5/6) levels lack integration with other levels. It is not integrated into the data fusion system. It can be seen from the research framework in Figure 1 that the current three key research areas are level-2 integration, the key technology of integration, without involving communication, interaction, system management, etc., which makes the research of each layer separate from each other, and it is difficult to directly use level-1. The result of the fusion is used as the original data for the measurement of SA features, and it does not cater to the needs of level-3 fusion upwards, which is an improvement for threat analysis and decision-making [35]. The current network SA is still in the initial stage of research. Although it has received general attention from the academic community, it has become a hot topic in the field of data fusion. However, most researches only involve a certain aspect of CSA, and they are still at the stage of theoretical exploration. Comprehensive and in-depth theoretical research and practical application deployment is the focus of current and future research work, and the most urgent need is to establish a unified evaluation system [36].

3. Cyber Situational Awareness Models

This section introduces several of the most influential CSA models. It shows the development and evolution process of the model, and finally, it summarizes the common problems of model research [15]. SA is centered on data fusion, and the establishment of its model is also based on the data fusion model. Currently, dozens of data fusion models have been proposed [32, 37], namely the Intelligent loop, JDL model, Boyd control loop, Endsley model, waterfall model, Dasarathy model, Omnibus model, and extended OODA (Observe, Orient, Decide, and Act) model. In addition, there are perceptual reasoning models that rely on associative storage and associated databases to imitate human thinking. Among them, the most influential one is the JDL data fusion model [25]. The JDL model divides the fusion into 4 levels: target refinement, situation refinement, risk refinement, and process refinement, among which SA is a higher level of level-2 fusion and the monitoring data of network elements is received from level-1 fusion downwards, as information sources for SA upwardly provide situational information for level-3 fusion for threat analysis and decision support. Blasch developed the JDL model [34, 38].

Based on four-layer integration, the fifth layer-user refinement is proposed. Level-5 integration emphasizes the role of users and requires users' knowledge and reasoning. At the same time, provide feedback for the lower layers to optimize the fusion process. At present, the JDL model has been further developed into a data fusion information group (DFIG) model [35]. The sixth layer-task management is proposed, which distinguishes information fusion and management functions. Boyd Control Cycle Model (BCCM) describes the search process with the purpose of activity. The process of observation, judgment, and decision-making and the process of finding a way consist of a total of four stages.

[20]. Among them, observation from the physical domain to the physical domain, judgment, and decision-making belongs to the perception domain, and the action from the information domain to the physical domain completes the sequence ring. The first three stages end the JDL model, and the actions cycle through the effects of the stage decisions in the real world. A mechanism to deal with multiple concurrencies and potential interactive data fusion is proposed. The Endsley model further refines the SA stage, which is divided into three levels: SA, situational understanding, and situational prediction [39]. At present, the Endsley model has received increasing attention. Salerno proposed a general SA framework on its basis and combined it with the actual characteristics of the network environment. The implementation of CSA is analyzed [39]. The authors of [40] proposed the CSA model based on the general SA reference model and combined it with the specific problems of network applications. The CSA model contains network elements, such as logs, configuration, tasks, attacks, and intrusion attempts, reflecting the characteristics of the network. Although there are a large number of SA models, there are several commonalities: (1) The model focuses on the function of SA. Except for the Dasarathy model, which focuses on tasks, the other models are focused, firstly, on data fusion, and then, they are focused on the situation. The functions of perception are divided. Although the components of different models have different names, many functions are the same. (2) Loop is the essence of SA. Each component in the model does not have a clear order; however, it is repeated iteratively. (3) Emphasize the role of feedback. Most models eventually form a closed-loop system.

3.1. Network Situation Knowledge Representation. This section discusses the KR of network situation from the two aspects of information representation and system representation and points out the deficiencies of existing research. KR [15] solves two problems: one is the representation of uncertain information. The expression of information uncertainty is by no means unique to SA. There are many related studies, and as basic research, a complete and universal solution from expression to integration is proposed by the common uncertainty theory. Probability theory, fuzzy sets, possibility theory, evidence theory, etc., have been applied in the field of SA. Moreover, uncertainty itself has many forms, and its classification and definition are different in different fields. Semantic discourse is basically like classification [41], meaning uncertainty is divided into fuzzy and ambiguity, and ambiguity can be divided into nonspecificity and conflict. The types of uncertainties that different uncertainty theories can handle are different. Fuzzy sets are the theories that deal with ambiguities, and the probability theory involves conflicts alone between the events [42]. The possibility theory expresses the nonspecificity of events, while the evidence theory describes nonspecificity and conflict. Among them, the most used fuzzy set proposed by [43] is a mathematical method to deal with the uncertainty and imprecision of human cognition. To process uncertain

semantic information, it is a powerful tool to describe the processing of fuzzy information in the human brain. Not only can it be used independently to construct representation and push the management system but also can be combined with other technologies as a general information representation technology. Fuzzy sets are recognized and widely used in the field of SA [43]. In addition, the gray theory is a theoretical method that uses gray information. It uses a whitening function to whiten various gray information to obtain the amount of gray information estimated.

Another problem is the representation of complex systems, which is also one of the challenges of CSA. To change the current situation of independent work of each unit network management, establishing CSA in the true sense relies on a unified system representation method to a large extent. It is to be able to make an accurate, comprehensive, and detailed description of the network and the premise of SA. On the other hand, as a complex giant system, the network can describe it in the face of rich content and intricate relationships, and there are few related studies. Ontology is one of the important methods; it is derived from philosophical concepts and after being introduced into the field of computer science [44]. It specifically refers to the clear definition of shared conceptual models of the formal specification of norms [45]. The authors of [44] proposed formal ontology, which emphasizes the essential concepts in the field, and time also emphasizes the relationship between these essential concepts and can express various concepts in the field and the relationship between concepts explicitly and formally. To express the semantics contained in the concept and enhance the ability to express complete systems, based on formal ontology, the study in [44] compares space-related content SNAP (spatial items of interest) and time-related content SPAN (temporal items of interest). They are both distinct and two inherent components of the connection that are considered separately. The authors of [46] used ontology to establish a situational view and analyzed both time and space to give a formal structure. However, it is too specific, i.e., it only plays the role of hierarchical structure. It is widely used in CSA and can be combined with evaluation methods to truly play a role in system description. The system representation method of the system also counts the traditional tree-like hierarchical structure. For example, the study in [47] borrowed the SNMP MIB tree-like structure model to represent the security threat classification and established a TCP/IP threat classification framework. The authors of [14] constructed a hierarchical network system security threat situation, which is divided into a network system, host, service, and attack from top to bottom. There are four levels of attack/vulnerability. In addition, the extension sets proposed in [14] focused on conflicting events in the system. Seek the internal mechanism of things, and solve incompatible problems by establishing matter-element models. Unfortunately, there are still many shortcomings in the current research on the KR of network situation [14, 46]: (1) the research lacks

continuity. However, simply pay attention to the system representation instead of applying it to subsequent evaluations. The methods of representing ontology and theoretical systems suggest a general approach to the integration of diagnostic work and hierarchies that show good results with ontology. It is because ontology and extensions are both systematic theoretical methods that can be applied everywhere. However, their theoretical system is huge, and hence, the use is complicated. Since it is not proposed for the SA problem, the presentation elements involved cannot fully reflect the characteristics of the network situation. (2) The tree-like hierarchical structure, although simple and clear, reflects the inherently hierarchical nature of the complex giant system, and it can simplify the analysis. However, it cannot express the complex relationship and content, and it cannot satisfy the need for SA. (3) As a key issue of SA, the research on information representation is insufficient, although the representation of uncertain information has been implicit in the evaluation method.

4. Network Situation Assessment Method

Network situation assessment refers to in a large-scale network environment based on the level-1 fusion of various network monitoring data and simple processing and based on domain knowledge and historical data, with the help of certain mathematical tools or mathematical models, after analysis and reasoning [48]. To make a reasonable explanation for the current state of the entire network composed of various network resources, network operations, and user behaviors, situation assessment emphasizes the relationship information between the entities and determines the convergence method of situation factors. The favorable or unfavorable judgment result, in short, is the mapping from the situational factor set to the situational space [27]. The situational factor set refers to the set of factors that can cause changes in the network situation, and it is a subset of the set of monitoring indicators. The evaluation method is the focus of SA and even data fusion. Hence, it has attracted much attention, and the theoretical research is relatively mature [48]. Some researchers innovatively and theoretically introduce the field of situation assessment. Others expand traditional methods and comprehensively apply a variety of theories to analyze uncertain information. To improve the accuracy of situation assessment, it is also necessary to compromise between factors, such as time expenditure and assessment cost. Among many assessment methods, the traditional method includes Bayesian technology, knowledge-based methods, artificial neural networks, and fuzzy logic technology. The new theories introduced include set pair analysis, DS evidence theory, rough set theory, gray relational analysis, cluster analysis, etc. [15, 27]. These methods can be roughly divided into three categories: methods based on mathematical models, methods based on knowledge inference, and methods based on pattern recognition. Figure 4 represents the network situation assessment methods.

FIGURE 4: Network situation assessment methods.

4.1. Method Based on the Mathematical Model. The mathematical model (MM) method comprehensively considers various factors for situation assessment, and its goal is to evaluate the network situation from different perspectives. It introduces three commonly used MM assessment methods [49].

4.1.1. Formulation-Based Method. The formulation-based method originated from the traditional situation assessment. To illustrate the behavior of war, mathematical tools are used to establish a war model so that the enemy and the forces of our opponents can be presented [16]. The classic expression of state loss and the calculation of the coefficients in the equation reveals the basic law of the effectiveness of the two parties' declining forces. As the most representative combat model in the field of CSA, mathematical formulas are also used to build models to explain network behavior and status. The authors of [25] discussed the relationship between throughput and latency and evaluated the effectiveness of resource allocation strategies, where $\text{power} = \text{throughput}^a / \text{delay}$ ($0 > a > > 1$).

4.1.2. Weight Analysis Method. The weight analysis method is the most commonly used evaluation method, and its evaluation function is usually an exponential expression, which is determined by the situation factor and its importance weight [50]. The authors of [15] adopted a bottom-up, partial first, and overall strategy to establish a hierarchical network security threat situation quantitative assessment model. Based on the statistics of the severity of the alarm and its network bandwidth consumption rate, the importance of using a layer-by-layer convergence method for attacks, services, hosts, and the entire network is provided. The weights are weighted to calculate the threat index, and the security threat situation is evaluated. The weight analysis method is the most typical MM method, and it has advantages and disadvantages. The key is to obtain the importance weight of the situation factor. It is the most prominent advantage, i.e., the result of level-1 fusion is directly used as the parameter of the situation assessment function, which narrows the distance between the data fusion levels [50].

4.1.3. Set Pair Analysis. Set pair analysis is a kind of quantitative analysis on the determination of the sameness, difference, and opposition of uncertain systems proposed [51]. The so-called set pair H refers to the pair composed of two sets with a certain connection. The connection number is an important concept of set pair analysis. The general form is $U = A + B_i + C_j$. A , B , and C are, respectively, the measures of the identity, difference, and opposition of the research object, and the connection number connects the three measures to form a same-different-contrary system (a certain-uncertain system) [51]. In addition to the connection number, the set pair analysis also defines the potential of the set pair $shi(H) = a/c$. Analysis can be used for situation assessment to construct a quantitative analysis model because each contact component in the contact number contains the situation information of the system.

The basic idea of using set pair analysis for situation assessment is as follows.

Firstly, determine the expression of the connection number U of the system. Give the calculation method of the similarity, difference, and opposite connection degrees A , B , and C , and establish the set pair analysis model for situation assessment. Analyze the value of $shi(H)$, which can be observed whether the system has a unified, opposite, or evenly matched trend in the same-different-contrary relationship, and further analyze the relationship between A , B , and C . According to the principle of permutation and combination, the establishment is based on $U = A + B_i + C_j$ system situation state table. A , B , and C three-dimensional system states have 12 states. Finally, the number of connections of the system in a specific environment is calculated according to the set pair analysis model, and the state of the system state is determined by looking up the state table. The connection number form is flexible and can be easily expanded according to needs [48]. In the system situation assessment, if the universe of discourse is divided into same, different, and reversed, it still feels rough. Then, one can expand the connection number into the form of $U = A + B_{i_1} + C_{i_2} + D_j$. Perform a four-dimensional situation analysis, and provide more possible details. Expand B and C for applications, where the boundaries and interior of the universe of discourse can be subdivided finitely and infinitely [32]. When the accuracy of the problem is high, the value analysis of i can be introduced to examine whether the uncertainty and ambiguity of the subdivision boundaries will affect the stability and reliability of the conclusion. The advantage of set pair analysis lies in the use of connection numbers to uniformly deal with multiple uncertainties caused by vagueness, randomness, intermediary, and incomplete information from the perspective of identity. It conducts situation assessment from multiple perspectives, such as differences and antagonistic measures, avoiding the limitations of using a single standard and system situation analysis based on the number of connections [38]. It is a kind of full ordering and unique, and it divides the state level of the situation, replacing the fuzzy evaluation method of "take the big and take the smallest" to avoid losing a lot of valuable information, leading to wrong conclusions [16]. Although the set pair analysis has the aforementioned inherent

advantages of the MM method, it is still inevitable. Its inherent shortcomings, how to construct the degree of similarity, difference, and opposition, have always lacked scientific basis and recognized methods, and therefore, it has become a difficult point for set pair analysis.

4.2. Method Based on Knowledge Reasoning. The knowledge reasoning (KR) method makes the full use of empirical knowledge to establish a situation assessment model. It draws on fuzzy sets, probability theory, evidence theory, etc., to deal with uncertain information, and it adopts logical reasoning and judges the network situation to complete the assessment [52]. Its goal is to process multisource and multiattribute information. KR can be further subdivided. This section introduces several classic KR methods.

4.2.1. Logical Reasoning Based on Production Rules. Logical reasoning based on production rules is an important method for artificial intelligence and expert systems. However, the classic reasoning method cannot solve the problem of situation assessment [52]. With the uncertainty of the domain, therefore, is necessary to introduce a method of representing uncertain information in logical reasoning. The authors of [26] proposed fuzzy sets, which is precisely to deal with human recognition. The mathematical method to understand uncertainty and imprecision is generally divided into three steps: fuzzification, fuzzy logic inference, and defuzzification [42]. Logical expressions are used to describe the reasoning method of the membership relationship in fuzzy sets. The research on fuzzy logic reasoning continues to deepen together with its different extensions, such as intuitionistic fuzzy sets, L-fuzzy sets, interval-valued fuzzy sets, vague sets, which have also been gradually introduced into the field of situation assessment.

Machine learning (ML) is also an artificial intelligence (AI) field that focuses on train computers to evaluate complex data and improve their performance over time. The most comprehensive set of machine learning services and infrastructure allow any developer, data scientist, or skilled practitioner to use machine learning. The SA paradigm appears to be a promising cybersecurity technique because it requires the collection, fusion, and assessment of a wide range of data from the operational environment to make predictions about future threats, such as cyber-attacks.

4.2.2. Reasoning Based on the Graph Model. Based on the reasoning of the graph model, the state transition is displayed by the establishment of a graph. The points represent preconditions or conclusions. The edges represent logical relationships between nodes. The nodes are used later to represent possible speeches, and the edges represent conditions for initiating conversion [2, 19]. The graph model implies knowledge about statements, relationships, and reasoning methods. Bayesian is probably the most used one in the evaluation of doctors' diagnostic methods, and it does not need to be done anymore. Relevant research continues to develop. The combined assessment still features the

Dynamic Basin Network's assessment accuracy and introduction of Dynamic Basin Network (DBN) to critique and move forward over time [23]. The dynamic Bayesian network represents a false model that can be regarded as a Markov process [19]. The difference is that DBN takes the state of the complex system divided into several small parts. DBN can simultaneously utilize the capabilities, space, and intelligence in the robot model and can achieve comprehensive problems and leverage. The thinking of other methods has led to strong Bayesian reasoning and deep Bayesian reasoning [25]. The reasoning based on the graph model highlights the advantages of the reasoning method. The state transition diagram is used to represent the reasoning process, which is clear and easy to understand. However, it also expands because of the difficulty of the reasoning method. The storage cost of the graph is high. How to build a graph model using machine learning has become an urgent concern for this method. Compared with the method, the graph model is based on the classical probability theory. It can use confidence to express the uncertainty, and implicitly, it solves the uncertainty representation problem.

4.2.3. Probabilistic Reasoning Based on Evidence Theory. The evidence theory established by [35], referred to as DS theory for short, is an important method for uncertain reasoning. The DS theory processes the difference between uncertainty and ignorance. It does not calculate the probability of a proposition; however, it calculates the probability that the evidence may support the proposition and gives the letter of information. Any measure and likelihood measure has a certain degree of scepticism. The DS theory gives evidence of the basic probability allocation (BPA) and focal element (FE) [49]. The introduction of the trust function $Bel(A)$ represents the full degree of trust given to proposition A by the evidence, and the likelihood function $Pl(A)$ represents the process that the evidence does not doubt the proposition A degree. Its core is the DS evidence synthesis rule, which is a method to combine the evidence from two independent information sources into a new piece of evidence, and it can be extended to the case of multiple pieces of evidence. The authors of [52] introduced the DS theory into the computer field for network situation assessment. The analysis process is as follows: firstly, establish the logic between the evidence and the proposition. The relationship between the entity and the situational factors is the convergence method of the situational state, and the basic probability distribution is determined. Then, based on the evidence that comes, i.e., each event occurs the reported information, use the evidence synthesis rules for evidence synthesis, obtain a new basic probability distribution, and send the synthesized results to the decision logic for judgment. The proposition with the greatest confidence is regarded as the candidate proposition. When there are continuous events, this process can continue until the confidence of the candidate proposition.

The degree exceeds a certain threshold, and the evidence meets the requirements, i.e., the proposition is considered to be valid, and the situation is in a certain state. The use of the

DS theory for situation assessment overcomes the shortcomings of using probability to describe the uncertainty. It requires neither an accurate understanding of the probability distribution nor does it need to explicitly express uncertainty [31]. By establishing the corresponding relationship between the proposition and the set, the uncertainty problem of the proposition is transformed into the uncertainty problem of the set, and the trust measure and likelihood measure of the information is given. When the prior probability is very high and is difficult to obtain, DS theory is more effective than probability theory. Another major advantage of using the DS theory is that the form is flexible and changeable. Related research combines the DS theory with fuzzy logic, neural network, and expert system, and it further improves the accuracy of reasoning [42]. The disadvantage of this method is that the computational complexity is high, and the contradiction and conflict are ignored because of standardization in practical applications. The conflict information is lost, and it is not suitable for the case of high conflict evidence. The latter has also become the research of the DS theory. Hot spots and various improvement measures have been proposed to fill this gap [52].

4.3. Method Based on Pattern Recognition. The pattern recognition (PR) method establishes a situation template by machine learning and completes the division of the situation after pattern matching [53]. Its goal is not to rely too much on experts and experience. Automatically, acquire knowledge and establish a scientific and objective evaluation template. This section introduces different PR methods.

4.3.1. Gray Relational Analysis. Gray system theory [54] was established in the 1980s and is a theoretical method for dealing with uncertain information. The basic idea of Gray relational analysis is to judge whether the relationship is close according to the similarity of the geometric shape of the sequence curve [55]. The closer the curves are, the greater the correlation between the respective configurations is and the lower the contrast is [53]. Apply gray correlation analysis to the field of situation assessment and perform pattern correlation and pattern matching on the situation according to the degree of correlation. For hybrid analysis, firstly, construct the situational element data sequence and select the template sequence and the comparison sequence. Then, calculate the absolute difference of each factor and then calculate the two configurations of the maximum difference and minimum difference of the potential sequence. The gray correlation coefficient of the situation factor is established according to the gray correlation analysis model, and the relationship between the situation factors of the two sets of sequences is compared with connection value. The gray correlation coefficient of the situation factor is concentrated on one value. The gray correlation degree reflects the degree of correlation between the situation, and the situation is calculated according to the value of the value [53]. Gray classification for the calculation of gray correlation chooses the algebraic average of the correlation coefficients of each situation factor. Otherwise, one can choose the weighted average according to the situation. The importance of factors assigns different weights to

each factor. The idea of gray correlation analysis of the situation is simple, and a pattern matching scheme is proposed. Compare the historical data of each situation to complete the situation classification evaluation [54]. However, in the pattern matching stage, it needs to be compared with each template, and each comparison needs to calculate the gray correlation coefficient and gray correlation. The calculation complexity is high. In addition, there is no question of how to create a template sequence.

4.3.2. Rough Set Theory. Rough set theory (rough set) was proposed by a Polish mathematician [56]. It stimulates the function of human abstract logical thinking and is a new type of processing. Mathematics tool can be used to deal with fuzzy, uncertain, and incomplete information [57]. Unlike traditional processing methods, the rough set theory is based on the division of the whole domain by equivalence released on the research, and the importance of expressing the attributes of the knowledge system, the dependency between the attributes, and the descriptive characteristics of the optimal system are studied, by introducing the upper approximation sets and lower approximation sets used to describe a set [54]. The main idea is to derive the concept of the concept by the reduction of knowledge under the premise of keeping the classification ability of unchanged classification rules. The rough set theory is used in the situation assessment with the help of the decision information system. The connection between the situation factor and the situation division is established by the decision table. In the situational template stage, historical data is collected as training samples. After feature selection, information table discretization, attribute reduction, attribute value reduction, and formalization rule extraction are the 5 steps to construct a situation assessment decision table. In the pattern matching stage, when new information arrives, the current situation can be determined by the classification decision table and the state of potential [57]. Situation assessment based on rough set theory, with both expressions, learning and classification capabilities, is given. The outstanding feature is that the rough set learning ability is strong, and it has a large amount of history. The advantages of discovering hidden knowledge in data or cases are that it reveals potential laws and transforms them into logical rules. Secondly, with the help of the formal model of information systems, bring the expression, learning, and analysis of knowledge into a unified framework. There is no need to provide any prior information beyond the required processing data set, scientific or objective, avoiding the influence of subjective factors. The difficulty lies in the determination of the decision table core and the attribute reduction algorithm (seeking core and reduction). The attribute reduction is under the condition of keeping the classification ability unchanged, deleting irrelevant or unimportant knowledge [53]. It has been proved that finding all reductions and minimum reductions in the decision table is one typical NP problem with a large amount of calculation. It has a good effect in a non-real-time environment; however, it may not meet the requirements in a real-time environment, which makes a rough set theory as the focus of research [57]. With the deepening of research and theoretical breakthroughs, the situation assessment method

based on the rough set theory will also continue to be improved. It is a combination of expression and learning with the method of classification ability. It will have brighter development prospects.

4.3.3. Cluster Analysis. Clustering refers to the automatic division of data into different clusters according to different characteristics of the data. The goal is to have objects that belong to the same cluster. There is a high degree of similarity, and the objects in different clusters are quite different (dissimilarities) [58]. The difference is that the classification is a predefined classification category, and the category of the cluster depends on the data itself. Clustering techniques can be roughly divided into five categories: partitioning method, hierarchical method, and density-based. Clustering technique is divided into five types: partitioning method, hierarchical method, density-based method, grid-based method, and model-based method [23]. The type of clustering process mainly includes steps, such as data preparation, feature extraction, feature selection, similarity calculation, clustering, and the effectiveness evaluation of clustering results. At present, it has been widely used in the field of network data analysis, such as anomaly detection, key traffic matrix discovery, usage pattern mining, text search, and situation assessment. [58]. Clustering is an unsupervised machine learning method. It provides a method of pattern division for unknown data. It does not require any prior knowledge and can be self-explanatory. It is scientific and objective to identify the internal structure of the data and discover the laws hidden in the data. Although there are many related studies, there is no clustering algorithm. The method can be universally applied to reveal the various structures presented by various multidimensional data sets [26] and to find a clustering algorithm suitable for network data flow.

4.4. Comparison of Different Methods. The several novel situation assessment methods were combined with other traditional methods and compared from different angles. The width is limited, the table uses m to represent medium, and 10 indicators are selected for comparison, among which, modeling time, model space overhead, and evaluation time reflect the evaluation of the efficiency of the method and the cost of dynamically updating the model. The number of features, versatility, and scalability reflect the flexibility of the evaluation method [53]. Whether it is conducive to actual application deployment, according to the resulting form, the diagnostic method can be divided into three types, scoring classification. The source of knowledge reflects whether the evaluation method is objective and is able to evaluate the dependence on empirical knowledge and domain experts. Understandability reflects whether the evaluation method is friendly and points out the inherent shortcomings of the theory [49]. Some interesting comparison results of the situation assessment are presented in Table 1.

5. Cyber Situational Awareness Application Approaches

Since Bass first proposed the concept of SA, most of the relevant research has been carried out around the security situation, and there is also a small amount of literature involving traffic analysis and information advantage measurement [9]. With increasing attention to network survivability, research on network survivability SA has emerged. At present, CSA application research is carried out around the core part of SA, and also, it focuses on the three key contents of the model, KR, and evaluation methods. Research in the security field is relatively in-depth. Theoretical methods have been applied in the evaluation, and some methods have achieved situation prediction [59]. Nevertheless, the combination of hierarchical structure and weight analysis is still the mainstream of the evaluation method. Other fields still stay at the data level, such as the transmission situation focusing on detection tool development and visualization research [9]. The survivability situation is still in the qualitative evaluation stage, and the abstraction from data to information to knowledge has not been realized [60]. It also cannot reflect the advantages and necessity of introducing SA in network management. The application has brought new weather as a specific case of the CSA system evaluation standard research, although only a set of specific evaluation measures and calculation formulas are given [59]. There is no clear standard for situation assessment; however, a new subject direction is proposed, which is the mature standard of CSA. In short, the current research is a preliminary attempt at the practical application of CSA. The integration of system design, situation, and system evaluation system and other topics remain to be studied.

5.1. Research by Tim Bass. Bass's research on CSA focuses on the security situation. There is a hierarchical structure based on multisensor data fusion intrusion detection system (IDS) proposed, as well as IDS data and IDS model presented, which are based on the air traffic control (ATC) SA theory [1, 9]. The IDS must use an out-of-band model, according to the mining model. Bass makes a distinction between declarative and procedural knowledge when it comes to KR, according to his theory. It is a specific case of declarative knowledge, expressed by patterns, algorithms, and mathematical transformations [59]. Bass also borrows the SNMP MIB model to represent security threat classification and establishes TCP/IP threats classification framework. In terms of evaluation, Bass proposes an information assurance (IA) mechanism for prevention, detection, and correction [1]. A risk identification model for the three major risk elements of importance, vulnerability, and threat. Quantitative analysis of cyber risk metrics is the risk of attack at the target's attack speed. Network security will be used in unified risk management and has a total of nine phases as we implement specific prototype systems [11].

TABLE 1: Comparison of different situation assessment methods.

Approach	Result method	Modeling time	Model space	Knowledge source	Feature number	General	Scalability	Comprehensibility	Theoretical shortcoming
Maths model	Math	Shorter	Smaller	Experience	Smaller	Lower	Harder	Easier	Subjective
	Weighed	Shorter	Smaller	Experience	Smaller	Lower	Harder	Easier	Subjective
	Pair set	Shorter	Smaller	Identical different country, experience	Medium	Medium	Easier	Medium	No uniform method
Knowledge reasoning	Fuzzy	Medium	Medium	Expert+	Medium	Medium	Medium	Easier	Subjective
	Markov	Medium	Big	Expert+, probability	Medium	Medium	Medium	Easier	Probability hypothesis
	Bayesian	Medium	Big	Expert+, probability	Medium	Medium	Medium	Easier	Probability hypothesis
	Evidence	Medium	Medium	Probability+, evidence combination	Medium	Medium	Medium	Easier	Lose conflict information
Pattern recognition	Gray	Medium	Medium	Experience+, gray relation degree	Bigger	Medium	Medium	Medium	Subjective
	Neural	Longer	Medium	Machine learning	Bigger	Higher	Easier	Harder	Knowledge recognition
	Rough set	Longer	Medium	Machine learning	Bigger	Higher	Easier	Harder	N-P hard
	Clustering	Longer	Medium	Unsupervised learning	Bigger	Higher	Easier	Harder	Low accurate

5.2. Security Situation Awareness. The realization of intrusion detection, intrusion rate calculation, intruder identity, intruder behavior identification, situation assessment, threat assessment, and other functions have become a new development in the field of information security [59]. The security situation pays attention to the confidentiality, availability, and integrity of the network and applies data fusion technology to processes, including firewall logs and intrusion detection logs. Heterogeneous and multisource data include virus logs, network scanning, illegal outreach, device status, and real-time alarms. According to the data source, the network security SA can be divided into system configuration information based on the dynamic assessment and close integration of the environment and the dynamic grasp of risks and based on system operation information [53]. The authors of [37] used a hierarchical structure to establish a threat assessment model combined with the weight analysis method and calculated the threat index layer-by-layer. The authors of [52] used the improved DS evidence theory to combine information from multiple data sources that perform fusion, used weight analysis to gather the situation elements and node situation layer-by-layer to calculate the network security situation, and further, they combined the actual performance information to correct the node security situation value [22]. In addition, the network security trend prediction is realized based on time series analysis.

5.3. Transmission Situational Awareness. Traffic has always been the focus of network management [61]. From the perspective of SA, the transmission situation focuses on information advantage, visualization, and performance evaluation [62]. It refers to the ability to process information and disseminate information permanently. The ability to process and disseminate a continuous flow of complete and accurate information using diagnostic information methods is a method of formula weight analysis [63]. The existing internet-level network traffic visualization tool provides the entire network. Traffic information makes it easy to find network attack behaviors and extract attack behavior characteristics [61]. Among them, the spinning cube of potential doom system uses points to represent the network traffic information in a three-dimensional space: the x -axis represents the network, the y -axis represents all possible source IPs, and the z -axis represents the port number, which improves the ability to transmit SA [64]. However, visualization takes the raw traffic data as objects rather than high-level abstract situation information and still stay at the data level. In some domestic research on network transmission performance evaluation and transmission SA, the relationship is relatively close. Among them, the authors of [65] established a network performance index system and gave a formal description framework [27] and others used formulas where the evaluation function is used as a standard to expand the performance evaluation of network transmission control and evaluate network performance based on weight analysis.

5.4. Survival Situational Awareness. Survivability refers to the ability of the system to complete its critical tasks promptly in the event of an attack, failure, or accident. Discussing the availability of the network from the same perspective, the scope is not limited to the network facility itself. After the 9/11 incident, network survivability has received widespread attention. SA is the analysis and evaluation of the survivability of the network, focusing on the three main factors of integrity, availability, and confidentiality, and no breakthrough has been made [27].

The security analyst must synthesize several pieces of data similar to criminal forensics to analyze the basic reasons for global attack occurrences on the internet. It can be a time-consuming and lengthy, and it is an informal approach that strongly relies on the analyst's knowledge and covers a wide range of assault scenarios. As a result, we are developing a multidimensional knowledge discovery and data mining methodology that will help us gain a more systematic understanding of emerging internet dangers, leading to improved CSA.

5.5. Cyber Situational Awareness System Performance Evaluation. Different from the above CSA research that focused on different areas of network management, we studied the testing and evaluation of SA systems from a higher perspective [9]. The goal is to quantify the benefits of CSA purely from the perspective of data visualization or load reduction, measure the performance or efficiency of the system, and determine whether it meets the task requirements. The authors of [59] established a standardized fusion system performance evaluation method by introducing the concept of measurement and proposed a minimum set of measures, and the calculation formula is clarified. The authors of [41] applied the evaluation method to CSA, introduced four-dimensional measures of reliability, purity, cost, and time, and established an evaluation function by specific case analysis to evaluate the performance of the CSA system.

6. Concluding Remarks and Future Trends

Cybercrime has become more sophisticated and organized in recent years, and organizations and/or countries have been targeted. This study puts forward the CSA problem system based on fully studying the relevant work of CSA, and it clarifies the research content of CSA and the problem system, correspondingly. The technical system of CSA is analyzed, focusing on the three aspects of model, KR, and evaluation method. The application of CSA is discussed, including security, transmission, survivability, system evaluation, and other fields. Each field involves many aspects, and special technologies are used to solve specific problems. Finally, the analysis of future development trends is done. The CSA's exploration has the research of level 1–3 integration, such as network monitoring. It is still in its infancy, and its degree of difficulty is higher. It is because the network's own flexible organization and growth mode have increased complexity and uncertainty, providing the

knowledge and management of the network. Finally, research on network management is relatively mature, providing a wealth of information for situational awareness. Data fusion, which is also quite extensive, proposes a general system framework and diagnostic model, which provides a theoretical basis for situational awareness.

Based on the above analysis, we believe that future research trends are important. The points are in the following aspects:

Research on the CSA evaluation system

To establish a CSA evaluation system, the following three phases must be completed: first, clarify the definition of the situation, reach a consensus on the division and grading of the situation, and establish a formal description. Secondly, formulate metrics to evaluate the accuracy of the situation assessment methods, select specific metrics, and establish standardized accuracy measurement methods. Thirdly, focus on accuracy and integrate other system evaluation indicators to establish CSA system evaluation standards.

Research on accurate and efficient evaluation methods

There are three types of assessment methods that are not mutually exclusive but seek solutions for different problems, for example, the evidence theory and fuzzy logic. Their combination can improve the ability to deal with ambiguity, while the neural network or rough set theory based on clustering will effectively reduce the sample space and reduce complexity.

Research on network system representation

It is necessary to explore innovative representation methods suitable for network systems to stay true to the ground.

Research on the actual deployment of the system

The important role of CSA in strengthening management and supporting decision-making in practical applications is demonstrated.

Research on the human-computer interaction mechanism

The establishment of a human-computer interaction mechanism includes the following tasks: (i) friendly human-computer interaction interface, (ii) perceive domain knowledge and expert advice at any time, and modify the model in time, (iii) realize the conversion from natural language to mathematical expressions that can be processed by the computer, and (iv) combine the data mining technology with domain knowledge to realize the important function of data fusion level-5-user refinement.

Data Availability

Data have been cited within the paper and are also available upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

The listed author(s) are highly grateful to Cyberspace Institute of Advanced Technology, Guangzhou University, for providing all the necessary facilities under the project by NSFC (No. 61972106), the Key R&D Program of Guangdong Province (No. 2019B010136003), and the National Key Research and Development Plan (Grant No. 2019QY1406).

References

- [1] T. Bass, "Multisensor data fusion for next generation distributed intrusion detection systems," in *Proceedings of the IRIS National Symposium on Sensor and Data Fusion*, April 1999.
- [2] M. Aledhari, R. Razzak, and R. M. Parizi, "Machine learning for network application security: empirical evaluation and optimization," *Computers & Electrical Engineering*, vol. 91, Article ID 107052, 2021.
- [3] C. K. Shiva, B. Vedic, R. Kumar, S. Mahapatra, and S. Raj, "Impacts of Computational Techniques for Wireless Sensor Networks," *Nature-Inspired Computing for Smart Application Design*, Springer, Singapore, pp. 87–103, 2021.
- [4] S. K. S. Durai, B. Duraisamy, and J. Thirukrishna, "A novel approach for QoS enhancement with revision scheme using SeDSR protocol in wireless sensor networks," *Wireless Personal Communications*, vol. 120, pp. 1–17, 2021.
- [5] D. Antunes and M. Esteve, "Hybrid situational awareness against cyber-attacks," in *The Russian Federation in Global Knowledge Warfare*, pp. 233–245, Springer Cham, 2021.
- [6] J. Xu, "Research on cyberspace mimic defense based on dynamic heterogeneous redundancy mechanism," *Journal of Computer and Communications*, vol. 9, no. 7, pp. 1–7, 2021.
- [7] K. J. Lippert and R. Cloutier, "Cyberspace: a digital ecosystem," *Systems*, vol. 9, no. 3, p. 48, 2021.
- [8] B. Uchendu, J. R. C. Nurse, M. Bada, and S. Furnell, "Developing a Cyber Security Culture: Current Practices and Future Needs," *Computers & Security*, vol. 109, Article ID 102387, 2021.
- [9] T. Bass, "Intrusion detection systems and multisensor data fusion," *Communications of the ACM*, vol. 43, no. 4, pp. 99–105, 2000.
- [10] B. M. Vancik, *Simulated Environment for the Demonstration of Cyber Situational Awareness*, Faculty of Informatics Masaryk University, Brno, Czechia, 2019.
- [11] M. Bokang, P. Panti, and F. Mapimele, "Cyber protect: a situational awareness platform," in *Proceedings of the ICCWS 2021 Sixteenth International Conference on Cyber Warfare and Security*, February 2021.
- [12] M. Darwish, A. Ouda, and L. F. Capretz, "A cloud-based secure authentication (CSA) protocol suite for defense against Denial of Service (DoS) attacks," *Journal of Information Security and Applications*, vol. 20, pp. 90–98, 2015.
- [13] P. Wang, L. T. Yang, J. Li, J. Chen, and S. Hu, "Data fusion in cyber-physical-social systems: state-of-the-art and perspectives," *Information Fusion*, vol. 51, pp. 42–57, 2019.
- [14] Y. Zhuo, Q. Zhang, and Z. Gong, "Cyberspace situation representation based on niche theory," in *Proceedings of the*

- International Conference on Information and Automation*, pp. 1400–1405, IEEE, Changsha, China, June, 2008.
- [15] G. P. Tadda and J. S. Salerno, “Overview of cyber situation awareness,” in *Advances in Information Security*, pp. 15–35, Springer, Boston, MA, USA, 2010.
 - [16] M. J. Pappaterra and F. Flammini, “A review of intelligent cybersecurity with Bayesian Networks,” in *Proceedings of the IEEE International Conference on Systems, Man and Cybernetics (SMC)*, October 2019.
 - [17] P. H. Foo and G. W. Ng, “High-level information fusion: an overview,” *Journal of Advances in Information Fusion*, vol. 8, no. 1, pp. 33–72, 2013.
 - [18] M. Park, H. Oh, and K. Lee, “Security risk measurement for information leakage in IoT-based smart homes from a situational awareness perspective,” *Sensors*, vol. 19, no. 9, 2019.
 - [19] D. Shen, G. Chen, J. B. Cruz, E. Blasch, and K. Pham, “An Adaptive Markov Game Model for Cyber Threat Intent Inference,” *Theory and novel applications of machine learning*, Intechopen, London, UK, 2009.
 - [20] U. Franke and J. Brynielsson, “Cyber situational awareness - a systematic review of the literature,” *Computers & Security*, vol. 46, pp. 18–31, 2014.
 - [21] H. Zhang, F. Li, J. Wang, Z. Wang, C. Sanin, and E. Szczerbicki, “Experience-oriented intelligence for internet of things,” *Cybernetics and Systems*, vol. 48, no. 3, pp. 162–181, 2017.
 - [22] V. Dutt, Y. S. Ahn, and C. Gonzalez, “Cyber situation awareness,” *Human Factors: The Journal of the Human Factors and Ergonomics Society*, vol. 55, no. 3, pp. 605–618, 2013.
 - [23] S. Jajodia and M. Albanese, “An integrated framework for cyber situation awareness,” in *Theory and Models for Cyber Situation Awareness*, pp. 29–46, Springer Cham, 2017.
 - [24] C. Li and X. M. Li, “Cyber performance situation awareness on fuzzy correlation analysis,” in *Proceedings of the Third IEEE International Conference on Computer and Communications (ICCC)*, December 2017.
 - [25] C. Onwubiko, “Understanding cyber situation awareness,” *International Journal on Cyber Situational Awareness*, vol. 1, no. 1, pp. 11–30, 2016.
 - [26] L. Hao, C. G. Healey, and S. E. Hutchinson, “Ensemble visualization for cyber situation awareness of network security data,” in *Proceedings of the IEEE Symposium on Visualization for Cyber Security (VizSec)*, October 2015.
 - [27] L. Lewis, G. Jakobson, and J. Buford, “Enabling cyber situation awareness, impact assessment, and situation projection,” in *Proceedings of the MILCOM 2008-IEEE Military Communications Conference*, November 2008.
 - [28] K. Golestan, R. Soua, F. Karray, and M. S. Kamel, “Situation awareness within the context of connected cars: a comprehensive review and recent trends,” *Information Fusion*, vol. 29, pp. 68–83, 2016.
 - [29] T. Xiaobin, Q. Guihong, Z. Yong, and L. Ping, “Network security situation awareness using exponential and logarithmic analysis,” in *Proceedings of the Fifth International Conference on Information Assurance and Security*, August 2009.
 - [30] S. Tepjit, I. Horváth, and Z. Rusák, “The state of framework development for implementing reasoning mechanisms in smart cyber-physical systems: a literature review,” *Journal of computational design and engineering*, vol. 6, no. 4, pp. 527–541, 2019.
 - [31] M. Husák, T. Jirsík, and S. J. Yang, “SoK: Contemporary issues and challenges to enable cyber situational awareness for network security,” in *Proceedings of the 15th International Conference on Availability, Reliability and Security*, Dublin, Ireland, August 2020.
 - [32] G. P. Tadda, “Measuring performance of cyber situation awareness systems,” in *Proceedings of the Eleventh International Conference on Information Fusion*, IEEE, Cologne, Germany, July 2008.
 - [33] J. Brynielsson, U. Franke, and S. Varga, “Cyber situational awareness testing, Advanced Sciences and Technologies for Security Applications,” in *Combating Cybercrime and Cyberterrorism*, pp. 209–233, Springer Cham, 2016.
 - [34] M. Albanese, H. Cam, and S. Jajodia, “Automated cyber situation awareness tools and models for improving analyst performance, Advances in Information Security,” in *Cybersecurity Systems for Human Cognition Augmentation*, pp. 47–60, Springer Cham, 2014.
 - [35] M. Tyworth, N. A. Giacobe, V. Mancuso, and C. Dancy, “The distributed nature of cyber situation awareness,” in *Proceedings of the IEEE International Multi-Disciplinary Conference on Cognitive Methods in Situation Awareness and Decision Support*, March 2012.
 - [36] N. J. Cooke, M. Champion, P. Rajivan, and S. Jariwala, “Cyber situation awareness and teamwork,” *ICST Transactions on Security and Safety*, vol. 1, no. 2, p. e5, 2013.
 - [37] I. A. Cooke, A. Scott, K. Sliwinski et al., “Toward robust models of cyber situation awareness,” in *Proceedings of the International Conference on Applied Human Factors and Ergonomics*, July 2018.
 - [38] G. Tadda, J. J. Salerno, D. Boulware, M. Hinman, and S. Gorton, “Realizing situation awareness within a cyber environment,” in *Proceedings of the Multisensor, Multisource Information Fusion: Architectures, Algorithms, and Applications*, April 2006.
 - [39] I. Kotenko and E. Novikova, “Visualization of security metrics for cyber situation awareness,” in *Proceedings of the Ninth International Conference on Availability, Reliability and Security*, September 2014.
 - [40] H. Tianfield, “Cyber security situational awareness,” in *Proceedings of the IEEE International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData)*, December 2016.
 - [41] M. Canan and A. Sousa-Poza, “Cyberspace situational awareness framework,” *Strategic Cyber Defense: A Multidisciplinary Perspective*, vol. 48, p. 42, 2017.
 - [42] Z. Huang, C. C. Shen, S. Doshi, N. Thomas, and H. Duong, “Fuzzy sets based team decision-making for cyber situation awareness,” in *Proceedings of the MILCOM 2016-IEEE Military Communications Conference*, November 2016.
 - [43] C. J. Matheus, A. Boran, and D. Carr, “Semantic network management for next-generation networks,” *Computational Intelligence*, vol. 35, no. 2, pp. 285–309, 2019.
 - [44] S. Brier, “Levels of cybersemiotics: possible ontologies of signification,” *Cognitive Semiotics*, vol. 4, no. Supplement, pp. 28–62, 2009.
 - [45] A. Oltramari, L. F. Cranor, R. J. Walls, and P. McDaniel, *Building an Ontology of Cyber Security*, Citeseer, Princeton, NJ, USA, 2014.
 - [46] K. Wu and J. Brenner, “Philosophy of information: revolution in philosophy. towards an informational metaphilosophy of science,” *Philosophie*, vol. 2, no. 4, 2017.
 - [47] M. I. Yolles, “Knowledge cybernetics: a metaphor for post-normal science,” in *Organizational Learning and Knowledge:*

- Concepts, Methodologies, Tools and Applications*, pp. 2924–2948, IGI Global, Hershey, PA, USA, 2012.
- [48] P. Lif, M. Granåsen, and T. Sommestad, “Development and validation of technique to measure cyber situation awareness,” in *Proceedings of the International Conference on Cyber Situational Awareness, Data Analytics and Assessment (CyberSA)*, June 2017.
 - [49] D. Adenusi, B. K. Alese, B. M. Kuboye, and A. F. . Thompson, “Development of cyber situation awareness model,” in *Proceedings of the International Conference on Cyber Situational Awareness, Data Analytics and Assessment (CyberSA)*, June 2015.
 - [50] T. G. Dietterich, X. Bao, V. Keiser, and J. Shen, “Machine learning methods for high level cyber situation awareness, Advances in Information Security,” in *Cyber Situational Awareness*, pp. 227–247, Springer, Boston, MA, USA, 2010.
 - [51] J. Yen, M. McNeese, T. Mullen, D. Hall, X. Fan, and P. Liu, “RPD-based hypothesis reasoning for cyber situation awareness, Advances in Information Security,” in *Cyber Situational Awareness*, pp. 39–49, Springer, Boston, MA, USA, 2010.
 - [52] P. Liu, S. Jajodia, and C. Wang, *Theory and Models for Cyber Situation Awareness*, Springer, NewYork, NY, USA, 2017.
 - [53] J. M. Bradshaw, M. Carvalho, L. Bunch et al., “Sol: an agent-based framework for cyber situation awareness,” *KI-Künstliche Intelligenz*, vol. 26, no. 2, pp. 127–140, 2012.
 - [54] E. Akyildirim, O. Cepni, S. Corbet, and G. S. Uddin, “Forecasting mid-price movement of Bitcoin futures using machine learning,” *Annals of Operations Research*, vol. 22, pp. 1–32, 2021.
 - [55] J. Yao, X. Fan, and N. Cao, “Survey of network security situational awareness,” in *Proceedings of the International Symposium on Cyberspace Safety and Security*, Springer, 2019.
 - [56] T. Rutkowski, *Explainable Artificial Intelligence Based on Neuro-Fuzzy Modeling with Applications in Finance*, Springer Nature, London, UK, 2021.
 - [57] B. Cao and Y. Liu, “Rough set decision making model based on dominance grey degree,” *Journal of Grey System*, vol. 28, no. 4, 2016.
 - [58] Y. Zhang, L. Liao, C. Xu, and M. Yang, “Hierarchical clustering of group behaviors in cyber situation awareness,” in *Proceedings of the IEEE International Conference on Software Quality, Reliability and Security Companion (QRS-C)*, August 2016.
 - [59] T. Bass, R. Zuech, and R. Gutzwiller, “Virtualized cyberspace-visualizing patterns & anomalies for cognitive cyber situational awareness,” Technical Report, ResearchGate, Berlin, Germany, 2017.
 - [60] A. Jakalan, “Network security situational awareness,” *The International Journal of Computer Science and Communication Security (IJCS)*, vol. 3, pp. 61–67, 2013.
 - [61] B. Predic and D. Stojanovic, “Enhancing driver situational awareness through crowd intelligence,” *Expert Systems with Applications*, vol. 42, no. 11, pp. 4892–4909, 2015.
 - [62] K. B. Swain, S. S. Mahato, and M. Cherukuri, “Expeditious situational awareness-based transmission line fault classification and prediction using synchronized phasor measurements,” *IEEE Access*, vol. 7, Article ID 168187, 2019.
 - [63] Y. Zhuo, Q. Zhang, and Z. Gong, “Research and implementation of network transmission situation awareness,” in *WRI World Congress on Computer Science and Information Engineering* vol. 4, , pp. 210–214, IEEE, 2009.
 - [64] K. Lakkaraju, W. Yurcik, and A. J. Lee, “NVisionIP: Netflow visualizations of system state for security situational awareness,” in *Proceedings of the 2004 ACM Workshop on Visualization and Data Mining for Computer Security*, Washington DC, USA, October 2004.
 - [65] M. Debashi and P. Vickers, “Sonification of network traffic flow for monitoring and situational awareness,” *PLoS One*, vol. 13, no. 4, Article ID e0195948, 2018.