Frontiers of Information Technology & Electronic Engineering
www.jzus.zju.edu.cn; engineering.cae.cn; www.springerlink.com
ISSN 2095-9184 (print); ISSN 2095-9230 (online)
E-mail: jzus@zju.edu.cn

OntoCSD: an ontology-based security model for an integrated solution of cyberspace defense#

Dandan WU[†§‡1], Jie CHEN[§2,3], Ruiyun XIE[3], Ke CHEN[1]

[1]*School of Computer Science, Chengdu College of University of Electronic Science and Technology of China, Chengdu 610731, China*

[2]*School of Cybersecurity, Northwestern Polytechnical University, Xi'an 710000, China*

[3]*China Electronics Technology Cyber Security Co., Ltd., Chengdu 610000, China*

[†]E-mail: wudd_2023@163.com

Received Sept. 27, 2023; Revision accepted Feb. 7, 2024; Crosschecked Aug. 26, 2024

Abstract: The construction of an integrated solution for cyberspace defense with dynamic, flexible, and intelligent features is a new idea. To solve the problem whereby traditional static protection methods cannot respond to various network attacks or security demands in an adversarial network environment in time, and to form a complete integrated solution from "threat discovery" to "decision-making generation," we propose an ontology-based security model, OntoCSD, for an integrated solution of cyberspace defense that uses Web ontology language (OWL) to represent the ontology classes and relationships of threat monitoring, decision-making, response, and defense in cyberspace, and uses semantic Web rule language (SWRL) to design the defensive reasoning rules. OntoCSD can discover potential relationships among network attacks, vulnerabilities, the security state, and defense strategies. Further, an artificial intelligence (AI) expert system based on case-based reasoning (CBR) is used to quickly generate a detailed and comprehensive decision-making scheme. Finally, through Kendall's coefficient of concordance (W) and four experimental cases in a typical computer network defense (CND) system, which reasons on represented facts and the ontology, OntoCSD's consistency and its feasibility to solve the issues in the field of cyberspace defense are validated. OntoCSD supports automatic association and reasoning, and provides an integrated solution framework of cyberspace defense.

Key words: Cyberspace defense; Integrated solution; Ontology; Case-based reasoning (CBR); Computer network defense (CND)
https://doi.org/10.1631/FITEE.2300662 **CLC number:** TP393; TP18

1 Introduction

Cyberspace defense plays an important role in cyberspace security. The concept of traditional static network protection methods lies in how to improve the robustness of the system and how to maintain the safety of the system by constantly updating anti-virus software, the firewall, intrusion detection, and vulnerability repair. Network attacks have become more sophisticated and persistent, so traditional static security protection methods, such as boundary defense, vulnerability detection, and rules matching, have been overwhelmed trying to deal with them. Moving target defense (MTD) was put forward to change the asymmetric situation of the defenders in terms of time and resources. MTD technology adopts a new idea, in which constantly moving and changing security mechanisms and defensive strategies increases the difficulty and cost of attacks, limiting the exposure of vulnerabilities and the opportunities to be attacked, to achieve effective protection of targets rather than pursuing the establishment of a perfect system.

§ These two authors contributed equally to this work
‡ Corresponding author
# Electronic supplementary materials: The online version of this article (https://doi.org/10.1631/FITEE.2300662) contains supplementary materials, which are available to authorized users
ORCID: Dandan WU, https://orcid.org/0000-0001-5214-387X

Researchers have carried out in-depth research and exploration in MTD technology. The dynamic defense mechanism proposed by some researchers can improve the security of the whole system by increasing the attack difficulties and limiting the exposure of vulnerabilities to effective attacks. It has the advantage that it can dynamically implement and deploy security defense solutions to change from static to dynamic and passive to active, and fully take advantage of the complexity of the time, space, and physical environment of the attack threats to protect the system and respond to complex network attacks and cyberspace defense requirements in time. It can protect software and devices from network attacks and reduce the risk of data breaches.

Network mimic defense (NMD) is a hot topic at present. According to Wu Jiangxing (Ji et al., 2022; Ma et al., 2022), "The existing system must acquire knowledge of attack sources, characteristics, behaviors, mechanisms, and other factors to implement effective defense. However, the existing information systems and defense architecture are essentially static, similar, and deterministic, with transparent and fragile architecture, which is becoming the largest security black hole in cyberspace." By integrating active defense and passive defense technologies, mimic defense technology can help improve situations in which it is easy to attack but difficult to defend.

Iannacone et al. (2015) constructed an overall ontology model, which is a knowledge view method including 15 entities and 115 attributes, to promote the integration of various structured and unstructured data sources in the cyberspace defense field. Zhu et al. (2017) presented a security knowledge base ontology model, which includes classes, relationships, attributes, and reasoning rules. Based on the unified cybersecurity ontology (UCO), a vulnerability knowledge graph was constructed, and the reasoning behind the vulnerability hiding relationship was realized (Qin SZ and Chow, 2019).

Researchers combine artificial intelligence (AI) algorithms (such as machine learning, reinforcement learning, deep learning, and optimization algorithms) and ontology methods to generate multi-attribute decision-making schemes. The AI expert system method based on case-based reasoning (CBR) deals with the highly complex and knowledge-intensive defense decision-making problem by dimension reduction, and simplifies the difficult problem as much as possible, which can make sure that the decision-making scheme is reliable and maneuverable.

To solve the issue whereby traditional static protection methods for network information systems cannot respond to various network attacks and demands in a timely manner, and a complete integrated solution from "threat discovery" to "decision-making generation" will be formed, this work presents an ontology-based security model, OntoCSD, for an integrated solution of cyberspace defense; this is used for supporting automatic association and reasoning and providing an integrated solution framework for cyberspace defense.

Our contributions are as follows:

1. OntoCSD integrates ontology classes, relationships, and reasoning rules, which can provide a dynamic, flexible, and intelligent defense mechanism. We have also designed a double-layer knowledge reasoning model. OntoCSD is different from other models that can reason only some simple point-to-point risks and defensive measures. This standardized design expression is conducive to realizing integrated solutions quickly, accurately, and intelligently.

2. We are committed to relying on the Protégé platform and its compatibility with multiple reasoning engines (such as HermiT, Pellet, and myCBR) to form integrated solutions, from ontology modeling and threat reasoning to decision-making scheme generation. This type of collaborative simulation and verification can optimize design, evaluate the network security situation, and improve defense efficiency.

3. The consistency and feasibility of OntoCSD are validated by Kendall's coefficient of concordance (W) and four experimental cases in a typical computer network defense (CND) system. It is proved that OntoCSD can support automatic association and reasoning, and provide an integrated solution framework for cyberspace defense.

2 A security defense mechanism for cyberspace

We designed an intelligent security defense mechanism (see supplementary materials, Section 1) formed with a "monitoring-decision-response-defense" framework, which is based on a software-defined security

platform and coordination with a management center; the security service is enabled on demand (Fig. 1). It realizes an efficient and intelligent integrated solution within a changing environmental situation, improves the confidentiality of network communication, and ensures the availability and integrity of network resources.

A typical CND system is taken as an example to illustrate the security defense mechanism for cyberspace proposed in this work. The system includes a management center, a backbone network node, a user subnet, a network monitoring system, and a security platform (such as the backbone network security platform, boundary security platform, and terminal security platform). In this system, the network monitoring system completes threat warning, the backbone network security platform implements security protection for backbone network communication links, and the boundary/terminal security platform implements security protection for terminals and servers.

3 Related works

3.1 Application of ontology in cyberspace defense

The application of ontology theory to the construction of a system model is a method that researchers are keen to adopt at present. The ontology model based on threat analysis is the first step in cyberspace defense research. Researchers have carried out in-depth explorations of threat-based ontology model construction.

Liu JX et al. (2020) analyzed the situational elements of cyberspace according to situational awareness and built an ontology model. The sub-fields of entities, relationships, and events are included. In addition, other ontology models have often been applied to network security situations. Si et al. (2015) constructed the domain ontology of network security situation element knowledge, including some key classes such as network environment, network vulnerabilities, network attacks, network security events, and the relationships between them. Hua and Chen (2014) proposed a domain

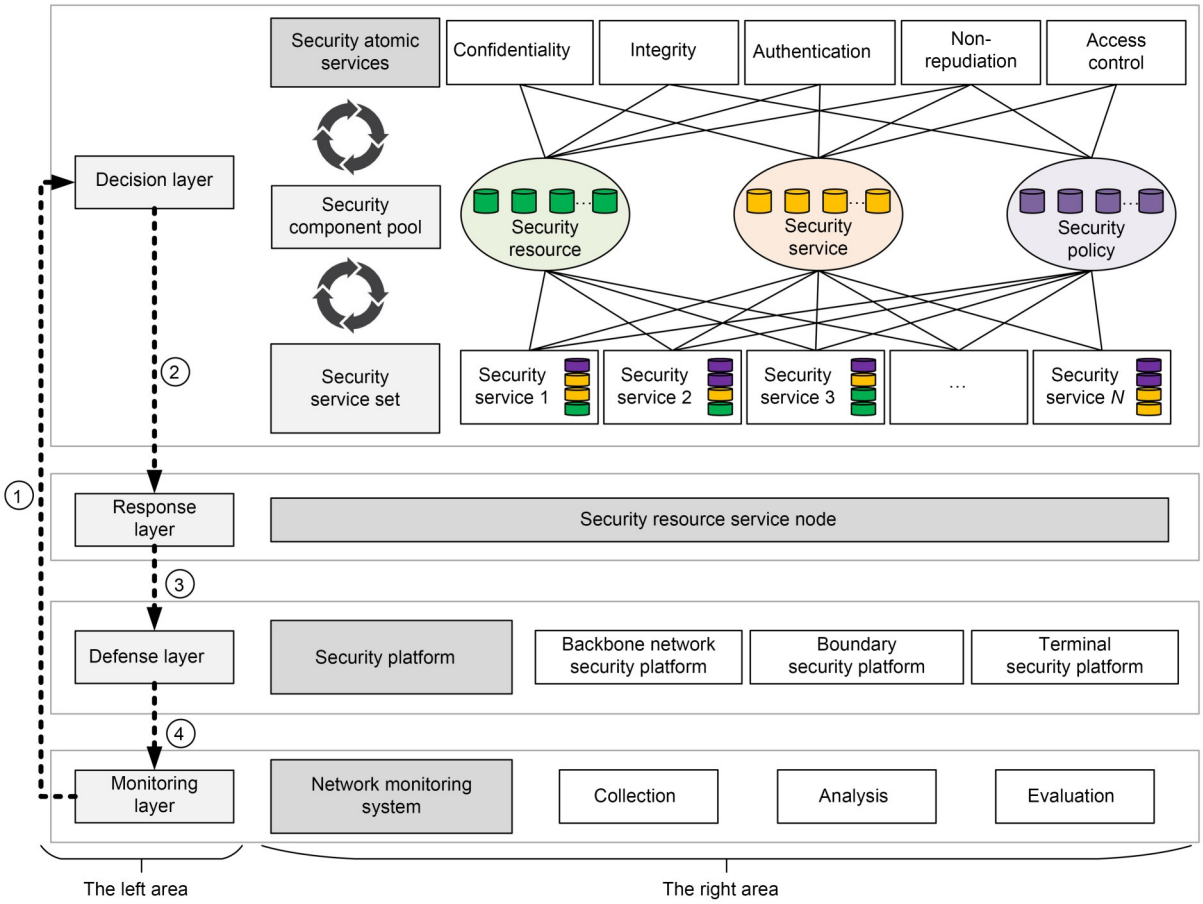

Fig. 1 The security defense mechanism

ontology including four sub-field classes, i.e., environmental security, hardware security, software security, and data security. In the field of cyberspace confrontation, the combat entities and different types of interaction behaviors involved in cyberspace warfare were abstracted into the nodes and edges of the network model. Then, the ontology-based entity description model and the behavior description model were established (Deng et al., 2014).

To carry out effective network security analysis and continuous monitoring, Merah and Kenaza (2021) regarded cyber threat intelligence (CTI) as a useful, updated, and structured source by introducing the concept of structured threat information expression (STIX). The ontology-based semantic knowledge model can provide valuable threat information according to the security alerts reported by an analyzer. Another CTI ontology reasoning model based on a unified knowledge concept was proposed to aggregate, represent, share, reuse, and analyze network attack information, which provides useful threat information and overcomes the past unsolved problems of ontology reasoning (Liu ZJ et al., 2020).

To detect changeable network attacks, Jia et al. (2018) presented a network security knowledge base based on five groups of ontology models and some reasoning rules. Machine learning was adopted to extract entities to construct the models, and the Stanford named entity recognizer (NER) was used to train it. It has good performance in the overall identification of software and vulnerabilities. However, the accuracy of recognition should be improved further.

The difference of network topology results in the difference of network threat analysis elements, such as mobile cloud environments, network shooting range, the Internet of Things (IoT), and satellite networks. For security threats to mobile cloud computing in the cloud and mobile environment, an intelligent ontology model was proposed to detect malware applications for application authority (Nisha and Bhanu, 2021), which reduces the learning load and computational complexity of the existing data, and provides a knowledge base. It solves the problem that the traditional malware detection methods cannot detect new types of malware or rapidly changing malware behavior.

In view of the fact that the existing threat modeling methods are suitable for only specific application targets and scenarios, the business requirements of automatic threat identification and adaptive security protection cannot be completely covered in practical applications. An ontology model was constructed that includes a threat meta-model, threat meta-model association, and threat ontology. It addresses issues such as asset uniqueness, recovery prediction of threat validity points and attack paths, and dynamic updating of the knowledge base (Gong et al., 2020).

To address the difficulties caused by the heterogeneity of security and open knowledge bases for threat assessment in IoT, Zhang SQ et al. (2022) proposed a risk analysis ontology model named risk analysis of IoT supply chain ontology (RIoTSCO), which includes six top-level classes—platform, vulnerability, weakness, attack pattern, campaign, and event. With secure multi-source knowledge reasoning rules, the model can sense the high vulnerability components in the IoT system, complement the context semantic information between the supply chain of the IoT and threat intelligence, and respond to threat intrusion automatically. Similar to this model, from the point of integrating vulnerability knowledge, the cyberspace security knowledge graph (SEPSES) was constructed based on ontology theory, which integrates public knowledge such as common weakness enumeration (CWE), common vulnerabilities and exposures (CVE), and common attack pattern enumeration and classification (CAPEC) (Kiesling et al., 2019).

To improve the security defense of satellite cyberspace, the satellite cyberspace situational awareness ontology model, named ontology of cyber situational awareness for satellite (OntoCSA4Sat), was proposed (Liu B et al., 2023). Reasoning rules, threat analysis, and mitigation of threat action were constructed, which can automatically infer and discover vulnerabilities to support space-based asset protection. However, the OntoCSA4Sat knowledge base is far from complete enough to answer all the questions. In addition, the cyberspace mimic defense (CMD) mechanism combined with biological characteristics is challenging. It uses heterogeneity to change the similarity and singularity of the system. It is expected that the hidden vulnerabilities will not be exploited by using a dynamic heterogeneous architecture. To deal with this, a security ontology (including attackers, threats, and vulnerabilities) based modeling method for CMD systems

(SOCMD) was proposed to build dynamic heterogeneous redundancy (DHR) items and the inner relationships (Zhang BW et al., 2020). However, the model is still in its early stages and has certain limitations in terms of safety judgment.

Table 1 shows a comparison of the research carried out by different researchers on cyberspace defense.

3.2 Application of AI algorithms to decision-making scheme generation

AI algorithms are widely used to solve the problem of decision-making scheme generation in cyberspace defense. Machine learning, reinforcement learning, deep learning, optimization algorithms, and other AI algorithms for decision-making scheme generation are a significant hot research direction, especially the integration of ontology theory into decision-making models to achieve knowledge retrieval and case retrieval, and generate accurate decision-making solutions.

In view of the heterogeneity and dynamic change of networks, researchers began to focus on using ontology and CBR to deal with the decision-making problem in the cyberspace defense field, which has achieved good results. The productivity and quality of a decision-making scheme cannot be demonstrated comprehensively due to the lack of a knowledge base or the involvement of all security issues, especially human influence on cyberspace defense decision-making schemes. To solve this problem, Solic et al. (2015) proposed an ontology model based on an improved evidence reasoning algorithm and a simple reflective

Table 1 Performance comparison of multiple types of defense strategies' optimization models

Reference	Model	Application field	Technical theory	Tool	Advantages & disadvantages
Zhang SQ et al., 2022	RIoTSCO	IoT situational awareness	Ontology; OWL; SWRL	Protégé; Pellet	It can be used to sense the high-vulnerability components, complement the context semantics, and respond automatically.
Zhang BW et al., 2020	SOCMD	Network mimic defense	Ontology; OWL; SWRL	Protégé	It can be used to describe the security architecture of CMD clearly and track the security state changes by using different DHRE configurations. However, there are limitations in safety judgment.
Jia et al., 2018	A knowledge graph for cybersecurity	Dynamic network attack recognition	Ontology; knowledge graph; linear chain CRF; path sorting algorithm	Stanford NER	It has good performance in software and vulnerability identification. However, the recognition accuracy needs to be further improved.
Liu B et al., 2023	OntoCSA4Sat	Satellite cyberspace situational awareness	Ontology; OWL; SWRL	Protégé; Pellet	It can be applied for the automatic association and reasoning of multi-source information. However, the knowledge base is not complete.
Nisha and Bhanu, 2021	–	Network attack detection	Ontology; OWL; SWRL	Protégé; Pellet	It can be used to detect new types of malware or rapidly changing malware behavior and provide a knowledge base.
Gong et al., 2020	Threat model	Cyberspace situational awareness	Ontology	–	It can realize threat analysis, find the starting point of threat, and restore and predict the attack path.
Merah and Kenaza, 2021	Semantic knowledge model	Network threat intelligence	Ontology	–	It provides valuable threat information according to the security alerts reported by an analyzer.

"–" indicates that this index is not mentioned in this model. IoT: Internet of Things; OWL: Web ontology language; SWRL: semantic Web rule language; CRF: conditional random field; CMD: cyberspace mimic defense; DHRE: dynamic heterogeneous redundancy entity

intelligent agent algorithm. Further, knowledge reasoning was transformed into the hard decision-making problem of agent path selection in knowledge graphs based on reinforcement learning methods (Qin PD et al., 2018; Zeng et al., 2018), and the reasoning of chain rules was realized using a walking strategy.

The CBR method is one of the techniques based on successful historical cases to predict schemes for a new case (Hameed et al., 2023), which creates a growing pattern of learning. To improve knowledge query and case retrieval in the emergency field efficiently, a case retrieval method with semantic similarity was established, which constructs a unified and standardized metro emergency plan ontology knowledge base, and uses semantic Web rule language (SWRL) to design the expression of emergency plan knowledge and reasoning rules (Zhang ZH et al., 2022).

Combined with the constructed subway operation accident case, CBR significantly enhances the use efficiency of emergency plan knowledge. To improve the retrieval ability of problems, Penadés et al. (2011) put forward a structured knowledge module to finish the compilation and generation of a scheme. In terms of semantic reasoning, Zhang L (2012) provided a strong theoretical basis for decision-making in subway emergencies and argued that emergency response programs can be generated on the basis of rules and CBR.

To address the difficulties posed by the large amount of information input to air traffic controllers (ATCs) in providing air traffic management (ATM) service decisions, Insaurralde and Blasch (2022) proposed a situation awareness (SAW) decision support system ATM reasoner (SAWDAR) approach to support the ATM decision-making process. The results showed that SAWDAR can support stakeholders (ATCs and pilots) in deciding whether or not to allow an aircraft to take off independently and dynamically.

For the fuzzy decision-making problem of road network selection, Guo X et al. (2021) used a case ontology model to formally express knowledge and eliminate case noise and conflict. This method is loyal to expert experience, which can reduce the difficulties of decision-making and improve the selection accuracy. However, due to the lack of perfection of the ontology database, the accuracy of results still needs to be strengthened. Li and Zhang (2022) built a case retrieval model combining ontology and CBR with the diversity of high-speed railroad equipment maintenance cases and the diversity of case knowledge combinations and structures. The results showed that the method can improve the accuracy of case retrieval.

4 Construction of an ontology-based security model, OntoCSD, for an integrated solution

Ontology can describe knowledge at the semantic level and can be seen as a universal conceptual model that describes knowledge in a certain discipline or domain, including basic terms (concepts) within a certain discipline and their relationships. It has been widely used in knowledge engineering, knowledge management, multi-intelligent system, Web semantics, network security, and other scenarios (Gao et al., 2012).

4.1 An ontology-based security model: OntoCSD

Web ontology language (OWL) and SWRL are adopted to construct a security model for cyberspace defense.

The definition of OntoCSD based on ontology is

$$O=(C, R, I, D, A), \tag{1}$$

where $C=C_c \cup C_I$ is the set of entities of the ontology. $C_c=\{C_1, C_2, C_3, C_4, C_5\}$ consists of classes representing entities that describe a set of objects, including the information system ontology class, vulnerability ontology class, network attack ontology class, defensive measures ontology class, and security state ontology class, while the set C_I is constituted by instances.

$R=\{\text{relationship about}(C_1, C_2)\}$ is the set of the relationships between two classes, which define a concept hierarchy such as "SubClassOf(C_1, C_2)."

I denotes the collection of relationships between an ontology class and its instances. For example, routers, security platforms, and terminals are all instances in the information system ontology class. For the vulnerability ontology class, vulnerabilities and security risks are its instances.

$D=\{(c_i, \text{datatype})|c_i \in C_n, C_n \in C_c, n=1, 2, 3, 4, 5\}$ is the set of properties of ontology entities and the basic datatype.

A={condition(C_x) →conclusion(C_y)|C_x, C_y ∈C, x≠y} is a set of axioms and rules that can infer new knowledge through some reasoning rules. It is used to make a security risk judgment.

The whole ontology classes in Fig. 2 and the relationships based on the concept of a security defense mechanism are established. Five major classes and 11 relationships are defined. The box with a yellow dot represents a class and the grey box on the line connecting two classes represents their relationship (Fig. 3). The labels Indicates, hasVulnerability, Mitigates, EquippedWith, ExploitedWith, CompromisedBy, SubClassOf, AttributedTo, Uses, CommunicateTo, and LackOf are included for the description of the relationships between ontology classes (see supplementary materials, Section 2).

4.2 Threat discovery reasoning

Knowledge reasoning can discover the implicit relationships among system assets, vulnerabilities, security mechanisms, and threats based on network topology. To simulate the attack scene, the corresponding reasoning rules are designed using SWRL. The form of the SWRL rule is as follows:

$$A_1, A_2, \cdots, A_m \rightarrow B_1, B_2, \cdots, B_n. \quad (2)$$

$R_{\text{cyber-security}}$ is a set of reasoning rules and $\text{KB}_{\text{cyber-security}}$ is a knowledge base of threat reasoning; the reasoning rule set is defined as follows:

$$R_{\text{cyber-security}}=\left\{R_1, R_2, \cdots, R_n \middle| R_n \in \text{KB}_{\text{cyber-security}}\right\}, n\geq 0, \quad (3)$$

where R_n is a reasoning rule.

So, different kinds of reasoning rules such as finding a potential risk in entities based on actual needs can be designed. For example, reasoning rules for entity security vulnerabilities in systems can be designed:

{

ISComponents(?ISC)^hasVulnerability(?ISC, ?vul)^Vulnerability(?vul)^Attacker(?attacker)^Uses(?vul, ?Network)->CompromisedBy(?ISC, ?attacker)

}.

A graphical example of a reasoning rule is shown in Fig. 4. A solid line represents a known relationship, and a dotted line represents an unknown relationship. The semantic rule means that if there are vulnerabilities in a system that can be exploited by certain network attack tools equipped by attackers, there is a potential risk of being compromised by the attacker.

Fig. 2 The OntoCSD model

4.3 Decision-making scheme generation based on CBR

The process of scheme generation based on external threats has the characteristics of high complexity

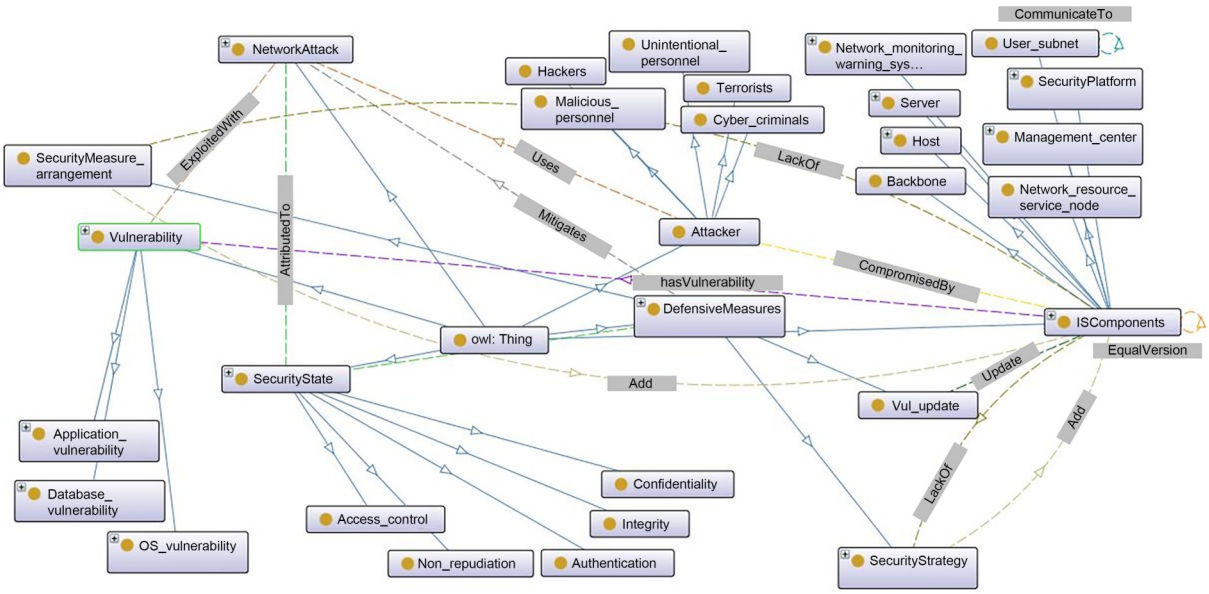

Fig. 3 The classes and their relationships in OntoCSD (References to color refer to the online version of this figure)

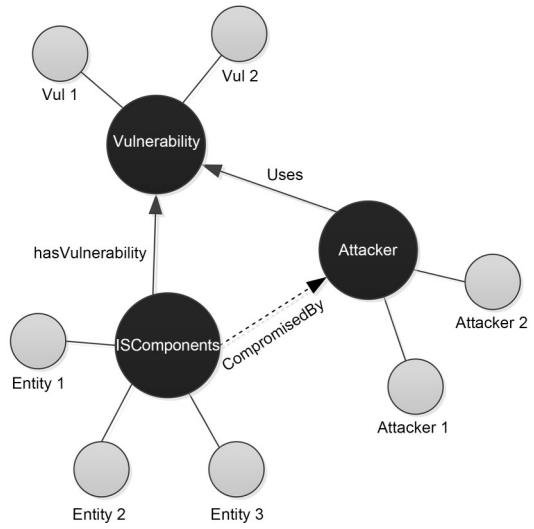

Fig. 4 A graphical example of a reasoning rule

and being knowledge-intensive. CBR is an example of the combination of problem solving and learning; it can take advantage of the specific knowledge of specific problem situations (cases) that have been experienced before. The CBR method can reduce the dimension of difficult high-dimensional problems, simplify knowledge acquisition, and improve the quality of solutions. We propose a decision-making mechanism based on CBR. Through the construction of the case base, the knowledge mining of expert cases is carried out by using an analogical reasoning idea, and the optimal decision-making scheme is generated by case matching. If the similarity value cannot meet the requirements, the rule reasoning is transferred to modify the design (Guo M et al., 2014; He et al., 2020) (see supplementary materials, Section 3).

4.4 Framework for an integrated solution

OntoCSD focuses on the double-layer reasoning process for an integrated solution in Fig. 5 based on ontology, including network threat reasoning (rule reasoning) and decision-making reasoning (case reasoning), which can achieve an integrated solution for cyberspace defense.

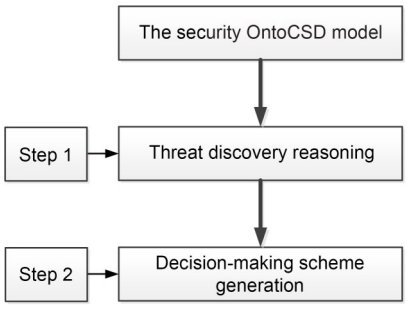

Fig. 5 The double-layer reasoning framework for an integrated solution

Step 1: According to the input of network threat information, the implicit security risk data will be obtained by threat reasoning rules.

Step 2: The output of threat reasoning is applied to the generation of a decision-making scheme as the input data; the best scheme will be obtained by calculating the similarity value. Moreover, the administrator can make local adjustments to this scheme.

5 Experimental results and discussion

We take the existing network attack and vulnerability data as the main threat database to design different experiments to test the performance of OntoCSD. The consistency of the ontology-based security model is validated by Kendall's coefficient of concordance (W). Furthermore, to evaluate the consistency and feasibility of reasoning about the problem of cyberspace defense, four security threat cases in a typical CND system are selected for representation and reasoning experiments.

5.1 Data sources

In this work, we take the existing network attack technology and vulnerability data as the main body, and the risks caused by the unreasonable design of network security policies and security measures are taken into account as input data to construct the knowledge form as the threat knowledge base for modeling and analysis. The threat type, main attack mode, CVE number, common vulnerability scoring system (CVSS), threat level, description, potential risk device, and so on, are included.

5.2 Experiments and validation

5.2.1 Kendall's coefficient of concordance (W)

Kendall's W is used to validate the consistency of the ontology-based security model.

Kendall's W can measure the degree of concordance or association that individuals have in relation to k variables and has values between 0 and 1. Kendall's W values that are close to 0 indicate strong disagreement between individuals, while values close to 1 indicate strong agreement.

We invited 12 experts to assist in the consistency validation of the proposed OntoCSD. They are all authoritative experts in the field of cyberspace defense and have mastered the professional knowledge of ontology development. These experts validated the consistency of OntoCSD by evaluating the terms required, i.e., identified concepts, relations used, taxonomy, properties, instances, constants, concepts dictionary, ad hoc binary relations, instance attributes, class attributes, axioms, and rules. We used 14 questions aiming at complying with the aspects that methodology determines in the ontology for data integration in OntoCSD (Silva and Rafael, 2023). The proposed hypotheses are as follows:

H0: There is no agreement among experts;

H1: There is agreement among experts.

The significance of Kendall's W determines whether the null hypothesis is rejected or not. If the significance of the P value is greater than the preset alpha level of 0.05, then the null hypothesis is accepted; it will be rejected if the significance is less than or equal to the alpha level. Table 2 shows that the significance of the P value of the overall data is 0.000^{***}, which presents significance at the horizontal level and rejects the null hypothesis, so the data present consistency. Meanwhile, Kendall's W is 0.695, so the degree of correlation is highly consistent among the experts.

5.2.2 Typical experimental cases

The consistency and feasibility of the network threat reasoning and decision-making scheme generation of OntoCSD are verified in detail through four different experimental cases designed in a typical CND system.

With four different cases designed in the given network topology, the potential security hidden danger, potential attack path, and security risk coverage are analyzed by setting different threat reasoning rules. Taking the results of threat reasoning as input data, we have verified the consistency, effectiveness, and maneuverability of the decision-making scheme obtained by the CBR algorithm under different network attacks and different network topology scales. To demonstrate the main performance advantages of OntoCSD, the experiments are carried out based on a real CND system. The topological graph in Fig. 6 includes four backbone network nodes and three user subnets; each of the user subnets contains corresponding computer terminals or data center terminals.

Protégé (version 5.5.0), developed by Stanford University, is used as an ontology platform that can provide a graphical and interactive knowledge ontology

Table 2 Test statistics

Questionnaire item	Rank average	Median	Kendall's W	χ^2	P
1. All items necessary for the construction of the ontology are present.	10.833	5			
2. The concepts used in the ontology model are adequate.	6.333	4			
3. The relations used in the ontology represent a type of association between concepts in the cyberspace field.	5.167	4			
4. The taxonomies used in the ontology construction process establish the concepts that define their hierarchy.	11.500	5			
5. The taxonomic system used in the ontology is clear, consistent, flexible, comprehensive, and practical.	5.167	4			
6. The properties that describe each concept of the taxonomy are specified.	11.417	5			
7. All instances are identified in the ontology.	5.750	4	0.695	108.38	0.000***
8. Do you agree with the constants (numerical values that do not change over a prolonged period of time) used in the ontology?	5.333	4			
9. All concepts of the domain, its relations, instances, and class and instance attributes are included.	4.417	4			
10. All ad hoc binary relations are described in detail in the binary relation diagram.	4.792	4			
11. All instance attributes are described in detail.	6.333	4			
12. All class attributes are described in detail.	11.417	5			
13. The formal axioms used in the ontology are Boolean expressions that are always true to define constraints in the ontology.	10.875	5			
14. The reasoning rules used in the ontology are used to infer knowledge.	5.667	4			

Kendall's W: Kendall's coefficient of concordance; χ^2: chi-square distribution. *** represents the significance level of ≤0.001

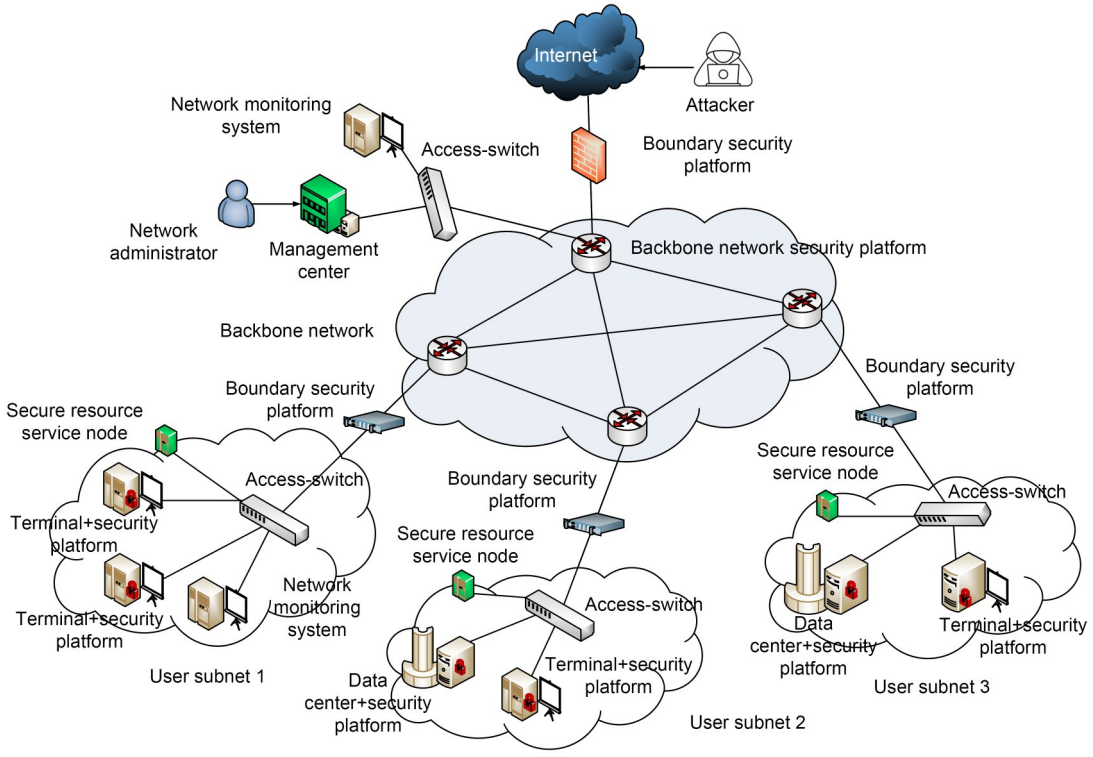

Fig. 6 A real computer network defense (CND) system

development environment. The OWL class is used to establish the ontology model. Protégé can be compatible with different inference tools for joint simulation operations, such as HermiT, Pellet, and myCBR. In our experiment, the HermiT (version 1.4.3.456) engine is selected as a threat reasoning tool to verify the consistency of the knowledge. The implicit knowledge and correlation can be deduced correctly with SWRL rules. The myCBR workbench is ideal for solving problems ranging from threats to decision-making generation for its simulation process, which is consistent with the CBR theoretical approach. So, myCBR is selected as the decision-making reasoning engine to provide task-oriented configuration, knowledge modeling, case base processing, and case similarity comparison.

1. Potential risk analysis

We design the reasoning statement of SWRL for four different experimental cases to verify the potential risk information.

Case 1: Vulnerabilities in the entity of a CND system have the potential risk of being breached by an attacker.

Case 2: Vulnerabilities in the configuration of an entity in a CND system are attacked by a certain means of network attack, such as operating system vulnerability; the entity device configured in the system is also at risk of being attacked.

Case 3: The security policies/measures configured on the system entity are at risk of being broken by an attacker because they do not meet the security protection requirements. If the system has an operating system with vulnerabilities, and the security policy against this vulnerability does not exist, it will be attacked. If the backbone switch has a redundant interface and no interface control strategy, the adjacent backbone node is at risk of being attacked by illegal access. The local area network terminal has a universal serial bus (USB) interface and has not carried on the authentication measure to the interface.

Case 4: The risk coverage of the CND system determines that if there is a potential risk of attack in one subnet of the system, other subnets connected to it also have potential risk.

According to the four different experimental cases above, we design the reasoning rules (Table 3), which can further derive new knowledge about potential threats through relevant input information, for example, by obtaining the current version of the operating system running on computer terminal 1 in subnet 1 and some potential vulnerabilities (including threat type, main attack mode, CVE number, CVSS, threat level, and description) in the operating system, and then infer the impact of threat actions.

Table 3 Reasoning rules

Rule No.	Reasoning rule
1	ISComponents(?ISC)^hasVulnerability(?ISC, ?vul) ^Vulnerability(?vul)^ExploitedWith(?Vul, ?NA) ^NetworkAttack(?NA)^Uses(?attacker, ?NA) ->CompromisedBy(?ISC, ?attacker)
2	ISComponents(?ISC_1) ^CompromisedBy(?ISC_1, ?attacker) ^EqualVersion(?ISC_1, ?ISC_2) ->CompromisedBy(?ISC_2, ?attacker)
3	ISComponents(?ISC)^LackOf(?ISC, ?str) ^DefensiveMeasures(?str)^SecurityState(?sta) ->AttributedTo(?ISC, ?sta)
4	ISComponents(?ISC_1)^hasVulnerability(?ISC_1, ?vul)^CommunicateTo(?ISC_1, ?ISC_2) ->hasVulnerability(?ISC_2, ?vul)

Fig. 7 shows the newly deduced information about the threat situation and the knowledge of vulnerabilities, security policies, and security measures. The items with a light yellow background represent implicit knowledge discovered through reasoning rules. We conduct reasoning analysis on all risks in three stages.

Stage 1: The potential vulnerabilities are obtained based on rule 1. (1) The backbone network platform is at risk of being attacked due to the Cisco Small Business RV Series router. (2) Terminal computing platform 1 in user subnet 1 may be vulnerable to permissions and access control attacks initiated by attackers due to the installation of the Microsoft Windows HTTP protocol stack. The risk level is critical, with a CVSS of 9.8. (3) The data center in user subnet 2 may be vulnerable to permissions and access control attacks initiated by attackers due to the installation of the Oracle WebLogic Server. The risk level is medium, with a CVSS of 6.3. (4) The data center in user subnet 3 may be vulnerable to permissions and access control attacks initiated by attackers due to the installation of the Apache CouchDB database system. The risk level is critical, with a CVSS of 9.8.

Fig. 7 Security risk inference results ("border" in the software is equivalent to "boundary"). References to color refer to the online version of this figure

Stage 2: The security risks caused by a lack of security measures or insufficient security policies are obtained based on rule 2. Due to the lack of boundary security platforms (such as firewalls), user subnet 1 poses significant security risks to the system's security status, such as confidentiality, integrity, identity authentication, and access control.

Stage 3: The potential security hazards and risks found because of the interrelation or the same configuration are obtained based on rule 3 and rule 4. First, there are vulnerabilities in user subnet 1, and it is inferred that user subnet 1 has exploitable vulnerabilities and risks of being attacked by attackers. Second, terminal computer 2 in user subnet 1 faces the same permissions and access control attacks as computer 1 due to the installation of the same Microsoft Windows operating system version. The risk level is critical, with a CVSS of 9.8. At the same time, due to lacking the security policy "USB interface protocol authentication measures," the security status has identity authentication risks. Third, terminal computer 1 in user subnet 3 faces the same permissions and access control attacks as computer 1 due to the installation of the same Microsoft Windows operating system version. The risk level is critical, with a CVSS of 9.8. At the same time, due to lacking the security policy "video encryption service," the security status is at risk of confidentiality. User subnet 1 and user subnet 2 can be interconnected under normal conditions. Due to the potential risk of permissions and access control attacks initiated by attackers, user subnet 2 faces the same risk.

We present all the potential risk information around the vulnerabilities and the lack of security policies/measures in Table 4. It shows all the new knowledge of potential security risks obtained by the reasoning rules in four different experimental cases. Furthermore, we can realize the whole chain of threat analysis through the reasoning of OntoCSD.

From the above analysis, it can be concluded that OntoCSD can obtain the results of threat analysis correctly; thus, the consistency of OntoCSD is verified.

2. Decision-making scheme generation

The results of the potential threat analysis in this subsection are used as the known threat input data here.

By comparing the equipment configuration and security threat information of each entity and weight values that are closely related to the system security threat, 26 characteristic attribute elements are established first (Fig. 8). We define the data type and value range for each of the characteristic attribute elements.

The myCBR workbench provides graphical modeling support, makes a predefined distance function and predefined similar behaviors (constant, single step,

Table 4 List of main security risks of system entities

Entity	Configuration	Security risk	Main attack mode	CVE number	CVSS	Threat grade
1. Backbone network security platform 4	Cisco Small Business RV Series	Vulnerability	Permissions and access control attacks	CVE-2022-20705	9.8	Critical
2. Boundary security platform in subnet 1	Lack of boundary security platforms	Security measures	Multiple attack routes	–	–	Critical
3. Computer terminal 1 in subnet 1	Microsoft Windows HTTP protocol stack	Vulnerability	Permissions and access control attacks	CVE-2022-21907	9.8	Critical
4. Computer terminal 2 in subnet 1	Microsoft Windows HTTP protocol stack	Vulnerability	Permissions and access control attacks	–	–	Critical
	Lack of "USB interface protocol authentication measures"	Security policy	Identity authentication risks	–	–	Critical
5. Computer terminal in subnet 2	Microsoft Windows HTTP protocol stack	Vulnerability	Permissions and access control attacks	CVE-2022-21907	9.8	Critical
6. Data center in subnet 2	Oracle WebLogic Server	Vulnerability	Permissions and access control attacks	CVE-2022-21280	6.3	Medium
7. Computer terminal in subnet 3	Microsoft Windows HTTP protocol stack	Vulnerability	Permissions and access control attacks	CVE-2022-21907	9.8	Critical
	Lack of "video encryption service"	Security policy	Information deciphering	–	–	Confidentiality risk
8. Data center in subnet 3	Apache CouchDB database system	Vulnerability	Permissions and access control attacks	CVE-2022-24706	9.8	Critical

"–" indicates that this index is not mentioned in this model. CVE: common vulnerabilities and exposures; CVSS: common vulnerability scoring system; USB: universal serial bus

and polynomial similarity reduction) for numerical data, and offers table functions and taxonomical functions for symbolic values. Table functions allow similar values to be defined for each value pair, while classification rules contain similar values for a subset of values.

It should be noted that this part focuses mainly on the application of the CBR method to the Protégé platform to build the complete process from potential risk discovery to decision-making scheme generation. So, the process of constructing characteristic attribute elements and weights is omitted here due to the 26 characteristic attribute elements selected in this experiment not fully representing the elements of other systems.

As for the weight values, any security weakness will lead to the collapse of the entire system from the perspective of the barrel effect; they are considered equal in the local similarity calculation process due to the difficulty in quantifying the weight values between different security feature elements based on existing security proofs.

Here, we assume that the global similarity threshold is set to 0.9 and the weight values are set to be equal. According to the experimental results, the case with the highest similarity value will be selected, and the scheme corresponding to this case will be the ideal one.

The similarity calculation is conducted between the new case and the other 11 cases in the current case library (Fig. 9). The results indicate that the similarity between dmt #0 and the new case is 0.96, and that the similarity between dmt #6 and the new case is 0.9. The similarity between dmt #1 and the new case is 0.83, and the similarity between dmt #7 and the new

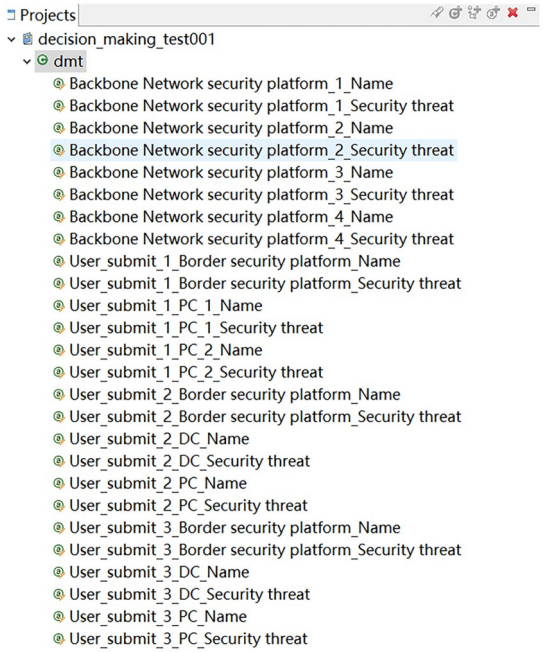

Fig. 8 The 26 characteristic attribute elements ("border" in the software is equivalent to "boundary")

case is 0.78. Following the principle of prioritization, we will output the defense decision-making scheme of dmt #0 as the ideal scheme. The scheme can even be adjusted to achieve better defense goals as wanted.

In conclusion, Section 5.2 has verified the capability of the CND system to achieve the integrated solution. So, its consistency and feasibility are validated.

5.3 Advantage analysis

The knowledge area coverage and problem-solving capacities of OntoCSD and other ontology models in the domain of cyberspace defense are compared in Table 5.

From the perspective of the system entity, OntoCSD includes not only the platform, vulnerability, network attack, security state, and other entities of the system, but also the security service components and defense strategies (security measures/policies) related to the cyberspace defense field. The reasoning rules can reason and discover the hidden risk knowledge of CND

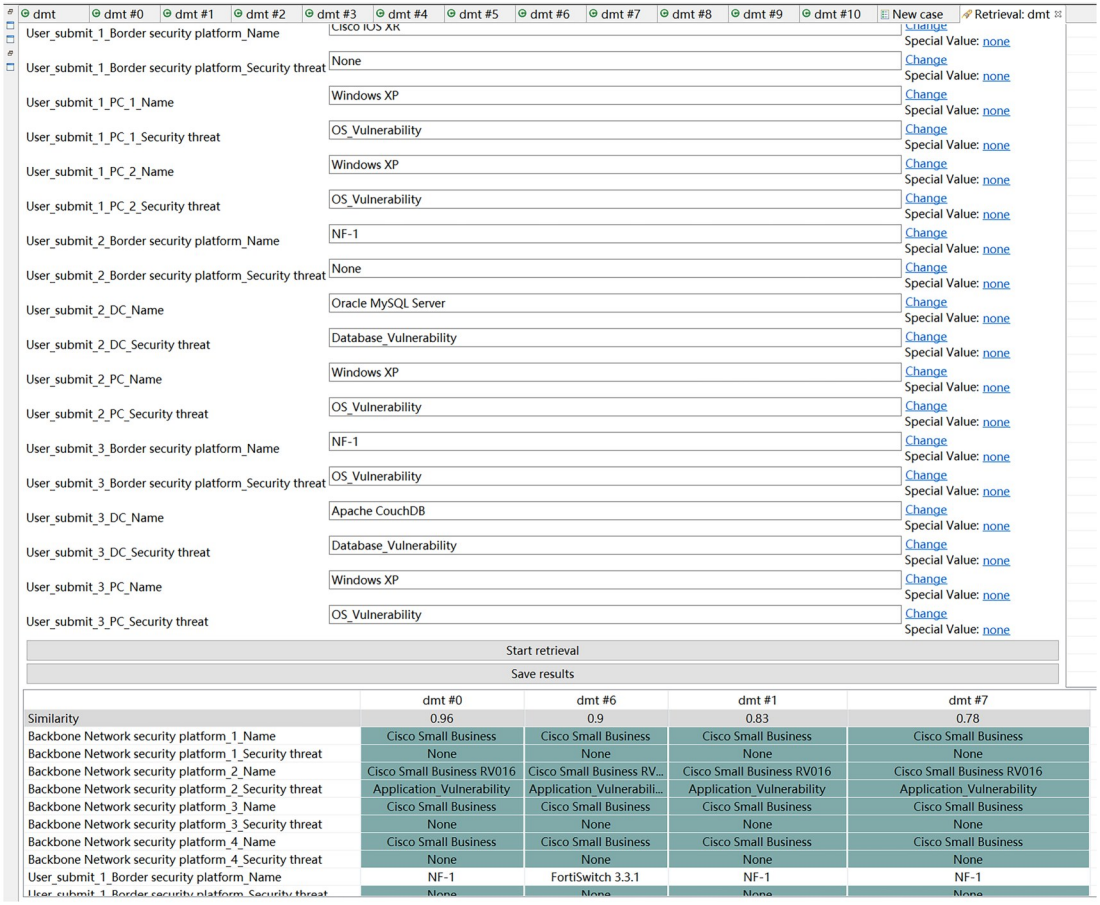

Fig. 9 Calculation results of graph similarity data ("border" in the software is equivalent to "boundary")

Table 5 Comparison of capacities between the proposed model and others

Ontology model	System entity	Reasoning rules	Consistency verification	Defense decision-making scheme
Zhang SQ et al. (2022)'s	System entities; vulnerability; weakness; network attack; campaign; event	Rules for security state and defensive measures	–	The reasoning rules are designed to identify potential security risks and automate mitigation measures to deal with threat events.
Zhang BW et al. (2020)'s	System entities; CMD components; DHR components; vulnerability; network attack; security state	Attacker capability assumption rules; multi-mode arbitration mechanism and combination rules	Experiments that are carried out based on the actual mimic Web server and mimic router validate the effectiveness of SOCMD in the security modeling and security evaluation of CMD systems.	–
Jia et al. (2018)'s	System entities; asset; vulnerability; network attack	Determining the relationship between two entities based on the path sorting algorithm	Accuracy, recall, and F1 score are used to verify the effectiveness of the model in experiments. The accuracy needs to be further improved.	–
Liu B et al. (2023)'s	System entities; satellite components; campaign; network attack	Rules for vulnerabilities and defensive measures	The consistency of the OntoCSA4Sat knowledge base and the feasibility of reasoning to solve the problem of satellite cyberspace threat situational awareness are proved through an experiment.	Some simple point-to-point defensive measures can be given.
Gong et al. (2020)'s	System entities; assets; vulnerability; network attack; defensive measures	Rules for threat analysis based on threat, attack, path, and attack defense view	–	–
Merah and Kenaza (2021)'s	System entities; intrusion set; alert; vulnerability attack pattern; attack tool; malware; event	Rules for the most dangerous vulnerabilities and exposed nodes, threats, and attack defense action plans	Through the experimental test, we are able to demonstrate the usefulness of such an approach.	–
OntoCSD	System entities; security server component; vulnerability; network attack; security state; defense strategy	Rules for vulnerabilities; rules for security measures/policies; rules for security state in CND systems of different scales	Kendall's coefficient of concordance (W) and four experimental cases in a typical CND system are used to verify the consistency and feasibility of OntoCSD.	The complete process of an integrated solution is obtained.

"–" indicates that this index is not mentioned in this model. CMD: cyberspace mimic defense; DHR: dynamic heterogeneous redundancy; CND: computer network defense

systems of different scales. In terms of consistency verification, we design four experimental cases in a typical CND system to verify the consistency and feasibility of OntoCSD. Even Kendall's *W* is used to verify the consistency of OntoCSD, which has not been mentioned in other models. In decision-making scheme generation, we obtain the complete process of an integrated solution. However, either other models do not propose defensive solutions, or the proposed solutions rely on only some simple point-to-point defensive measures generated by the Protégé platform, which cannot be called the integrated solution of the whole system.

To summarize, OntoCSD is more comprehensive and valuable than the other ontology models in the field of cyberspace security in four dimensions: system entity, reasoning rules, consistency verification, and whether to generate defense decision-making schemes. Furthermore, OntoCSD can adapt to the changing environment, such as power networks, task networks, service networks, supply chain networks, and other complex networks. Network managers can design reasoning rules according to the actual environment and security requirements. Different reasoning rules that meet different defense needs can significantly improve the reasoning capabilities of OntoCSD and have adaptability in virtual network adversarial environments.

6 Conclusions and future work

In this work, an ontology-based security model (OntoCSD) of cyberspace defense for an integrated solution is proposed, which links attackers, threats, vulnerabilities, threat perception, response, decision-making, defense, and resource components. It describes and represents the integrated solution clearly based on central side management control from the ontology model, threat reasoning, and decision-making scheme generation. The threat reasoning rules and the CBR method are used for threat reasoning and defensive scheme generation. Finally, the consistency and feasibility of OntoCSD are validated by experiments. So, OntoCSD can support automatic association and reasoning, and provide an integrated solution framework of cyberspace defense.

In addition, it should be made clear that the research into constructing an integrated solution with dynamic, flexible, and intelligent characteristics is a new challenge. The OntoCSD model is still in its initial stages. The mentionable thing is that the defensive scheme generated using a similarity-based CBR method is currently used for only small-scale CND systems, while network topology security schemes targeting large-scale nodes (such as hundreds of nodes or more) require further analysis and argumentation. The established defense case base needs to be updated in real time and evaluated for defense performance based on the latest developments and changes in network threats, which should be conducted from multiple dimensions such as efficiency, robustness, availability, security, and reliability of system security operation. These issues will be the main focus of our research in the future.

Contributors

Dandan WU designed the research and made algorithm data analysis. Dandan WU, Jie CHEN, and Ke CHEN determined the whole research framework and searched the literature. Ruiyun XIE revised the entire research framework and technology. Dandan WU drafted, revised, and finalized the paper.

Conflict of interest

All the authors declare that they have no conflict of interest.

Data availability

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

References

Deng ZH, Lao SY, Bai L, et al., 2014. An extensible description model of cyber war system. *J Natl Univ Def Technol*, 36(1): 184-190 (in Chinese). https://doi.org/10.11887/j.cn.201401032

Gao JB, Zhang BW, Chen XH, 2012. Research progress in security ontology. *Comput Sci*, 39(8):14-19, 41 (in Chinese). https://doi.org/10.3969/j.issn.1002-137X.2012.08.003

Gong L, Si RB, Tian Y, 2020. Research on key technologies of ontology based threat modeling for cyber range. *J CAEIT*, 15(12):1139-1144, 1162 (in Chinese). https://doi.org/10.3969/j.issn.1673-5692.2020.12.001

Guo M, Qian HZ, Huang ZS, et al., 2014. Intelligent road-network selection using cases based reasoning. *Acta Geod Cartograph Sin*, 43(7):761-770 (in Chinese). https://doi.org/10.13485/j.cnki.11-2089.2014.0120

Guo X, Qian HZ, Wang X, et al., 2021. A method of road network selection based on case and ontology reasoning. *Acta Geod Cartograph Sin*, 50(12):1717-1727 (in Chinese). https://doi.org/10.11947/J.AGCS.2021.20200360

Hameed S, Elsheikh Y, Azzeh M, 2023. An optimized case-based software project effort estimation using genetic algorithm.

Inform Softw Technol, 153:107088. https://doi.org/10.1016/j.infsof.2022.107088

He HW, Qian HZ, Duan PX, et al., 2020. Automatic line simplification algorithm selecting and parameter setting based on case-based reasoning. *Geomat Inform Sci Wuhan Univ*, 45(3): 344-352 (in Chinese). https://doi.org/10.13203/j.whugis20180250

Hua HY, Chen QM, 2014. Network security situation knowledge base model based on ontology. *J Comput Appl*, 34(S2): 95-98, 107 (in Chinese).

Iannacone M, Bohn S, Nakamura G, et al., 2015. Developing an ontology for cyber security knowledge graphs. Proc 10th Annual Cyber and Information Security Research Conf, Article 12. https://doi.org/10.1145/2746266.2746278

Insaurralde CC, Blasch E, 2022. Situation awareness decision support system for air traffic management using ontological reasoning. *J Aerosp Inform Syst*, 19(3):224-245. https://doi.org/10.2514/1.I010989

Ji XS, Wu JX, Jin L, et al., 2022. Discussion on a new paradigm of endogenous security towards 6 G networks. *Front Inform Technol Electron Eng*, 23(10):1421-1450. https://doi.org/10.1631/FITEE.2200060

Jia Y, Qi YL, Shang HJ, et al., 2018. A practical approach to constructing a knowledge graph for cybersecurity. *Engineering*, 4(1):53-60. https://doi.org/10.1016/J.ENG.2018.01.004

Kiesling E, Ekelhart A, Kurniawan K, et al., 2019. The SEPSES knowledge graph: an integrated resource for cybersecurity. Proc 18th Int Semantic Web Conf, p.198-214. https://doi.org/10.1007/978-3-030-30796-7_13

Li HL, Zhang ZH, 2022. Ontology-based knowledge management model for high-speed railway onboard equipment maintenance cases. *Railw Stand Des*, 66(2):149-155 (in Chinese). https://doi.org/10.13238/j.issn.1004-2954.202011230003

Liu B, Yi JC, Yao L, et al., 2023. Situational awareness ontology modeling for threat from space cyber operations. *Syst Eng Electron*, 45(3):745-754. https://doi.org/10.12305/j.issn.1001-506X.2023.03.15

Liu JX, Guo JX, Song LY, 2020. Study on cyberspace situation ontology for situation awareness. *Fire Contr Command Contr*, 45(3):90-94 (in Chinese). https://doi.org/10.3969/j.issn.1002-0640.2020.03.016

Liu ZJ, Sun Z, Chen JF, et al., 2020. STIX-based network security knowledge graph ontology modeling method. Proc 3rd Int Conf on Geoinformatics and Data Analysis, p.152-157. https://doi.org/10.1145/3397056.3397083

Ma HL, Wang L, Hu T, et al., 2022. Survey on the development of mimic defense in cyberspace: from mimic concept to "mimic+" ecology. *Chin J Netw Inform Secur*, 8(2):15-38 (in Chinese). https://doi.org/10.11959/j.issn.2096-109x.2022018

Merah Y, Kenaza T, 2021. Proactive ontology-based cyber threat intelligence analytic. Int Conf on Recent Advances in Mathematics and Informatics, p.1-7. https://doi.org/10.1109/ICRAMI52622.2021.9585984

Nisha OSJ, Bhanu SMS, 2021. Detection of malicious Android applications using ontology-based intelligent model in mobile cloud environment. *J Inform Secur Appl*, 58:102751. https://doi.org/10.1016/j.jisa.2021.102751

Penadés MC, Borges MRS, Canós-Cerdá JH, et al., 2011. A product line approach to the development of advanced emergency plans. Proc 8th Int Conf on Information Systems for Crisis Response and Management.

Qin PD, Xu WR, Wang WY, 2018. Robust distant supervision relation extraction via deep reinforcement learning. Proc 56th Annual Meeting of the Association for Computational Linguistics, p.2137-2147. https://doi.org/10.18653/v1/P18-1199

Qin SZ, Chow KP, 2019. Automatic analysis and reasoning based on vulnerability knowledge graph. Proc Int Conf on Cyberspace Data and Intelligence, p.3-19. https://doi.org/10.1007/978-981-15-1922-2_1

Si C, Zhang HQ, Wang YW, et al., 2015. Research on network security situational elements knowledge base model based on ontology. *Comput Sci*, 42(5):173-177 (in Chinese). https://doi.org/10.11896/j.issn.1002-137X.2015.5.035

Silva DV, Rafael GR, 2023. Ontology for data integration in honeynet. *Res Milit*, 13(2):4959-4972.

Solic K, Ocevcic H, Golub M, 2015. The information systems' security level assessment model based on an ontology and evidential reasoning approach. *Comput Secur*, 55:100-112. https://doi.org/10.1016/j.cose.2015.08.004

Zeng XR, He SZ, Liu K, et al., 2018. Large scaled relation extraction with reinforcement learning. Proc 32nd AAAI Conf on Artificial Intelligence, p.5658-5665. https://doi.org/10.1609/aaai.v32i1.11950

Zhang BW, Chang X, Li JH, 2020. A generalized information security model SOCMD for CMD systems. *Chin J Electron*, 29(3):417-426. https://doi.org/10.1049/cje.2020.02.017

Zhang L, 2012. Ontology-Based Digital Method and Application of Urban Rail Transit Emergency Plan. MS Thesis, Beijing Jiaotong University, Beijing, China (in Chinese).

Zhang SQ, Bai GY, Li H, et al., 2022. IoT security knowledge reasoning method of multi-source data fusion. *J Comput Res Dev*, 59(12):2735-2749 (in Chinese). https://doi.org/10.7544/issn1000-1239.20210954

Zhang ZH, Li HL, Wang QW, et al., 2022. Ontology-based knowledge modeling of metro emergency response plan and construction of case database. *Urban Mass Transit*, 25(8):17-22 (in Chinese). https://doi.org/10.16037/j.1007-869x.2022.08.004

Zhu X, Huang JM, Zhou B, et al., 2017. Real-time personalized twitter search based on semantic expansion and quality model. *Neurocomputing*, 254:13-21. https://doi.org/10.1016/j.neucom.2016.10.082

List of supplementary materials

1 A security defense mechanism for cyberspace

2 The detailed design process for classes and relationships for the OntoCSD security model

3 Decision-making scheme generation based on CBR