Research on Network Security Threat Analysis Technology Based on Ontology

Weifa Zheng[1], Zitao Cai[2], Peiyu Cheng[2], Jingwen Yan[3], Yanjun Xiao[4]

[1]*Network and Information Office, Guangdong University of Finance and Economics, Guangzhou 510320, Guangdong, China*
[2]*School of Information, Guangdong University of Finance and Economics, Guangzhou 510320, Guangdong, China*
[3]*Engineering institute, Shantou University, Shantou 515063, Guangdong, China*
[4]*NSFOCUS, Haidian District 100089, Beijing, China*

Abstract

In order to solve the problem that most of the existing network security knowledge exists in isolation and the entity relationship is not comprehensive, this paper collects and analyzes the network security domain knowledge based on Ontology, constructs an ontology model named CDO for network security domain, including the definition of class, entity and relationship, and then establishes knowledge mapping through the open structured and semi-structured data, such as STIX, CVE, CWE, CPE and CAPEC. Finally, taking a large network attacked by SQL injection as an example, the knowledge map constructed in this paper is used to analyze network security events. Through examples, we can see that the knowledge map of this paper can effectively improve the efficiency of analysis techniques for network attacks.

Keywords

Network security, Network threat, Ontology, Knowledge map

1. Introduction

In recent years, with the popularization and rapid development of network applications, various network security incidents are common, and network attacks have become more complex. Analyzing network attacks requires the integration of various knowledge from multiple perspectives. As an efficient form of knowledge organization such as entities and concepts, knowledge map can give full play to its advantages of knowledge integration [1]. Many scholars carry out automatic construction of network security knowledge map based on NLP technology, for example, Li Tao proposed the extraction technology of threat intelligence entity by multi feature fusion [2]. In recent years, some scholars have begun to introduce ontology into the construction of knowledge map in the field of network security. Lian Longying [3] proposed an ontology-based method for the construction of knowledge map in cyberspace security, mainly explaining the construction process of knowledge map from ontology layer modeling, data layer mapping and storage layer visualization, but did not elaborate on the content of knowledge map.

In order to solve the problem that most of the existing network security knowledge exists in isolation, and the entity relationship is not comprehensive, this paper collects and analyzes the network security domain knowledge based on ontology, constructs the ontology model of the network security domain, and then establishes the network security domain knowledge map through open structured and semi-structured data, and carries out network security event analysis. Practice shows that the research work of this paper has good significance for network security event analysis and reasoning.

2. Network Security Intelligence Data

To build a knowledge map in the field of cybernetwork, we need to rely on a variety of structured, semi-structured and unstructured data and other knowledge sources. Structured data, such as Structured

ICCEIC2022@3rd International Conference on Computer Engineering and Intelligent Control
EMAIL: * Corresponding author: jwyan@stu.edu.cn (Jingwen Yan)

Threat Information Expression (STIX) [4]. Semi structured data, such as Common Vulnerabilities & Exposures (CVE) [5], Common Weakness Enumeration (CWE) [6], Common Platform Enumeration (CPE) [7], Common Attack Pattern Enumeration and Classification (CAPEC) [8] and ATT&CK [9]. Unstructured data, such as cybersecurity blogs on the internet, cybersecurity reports, etc. The following mainly introduces STIX and CAPEC.

STIX is a language used to exchange network threat information. It is a structured language used to describe network threat information, which can be shared, stored and analyzed in a consistent way [10]. STIX consists of some key structures such as observation representation, characteristic indicators, safety events, methods and processes, and utilization goals. Application scenarios include: collaborative threat analysis, automated threat intelligence exchange, automated threat detection and response, etc. STIX provides an important reference for ontology design in the field of network security.

CAPEC is a classification data set of commonly used attack types. Currently, more than 500 attack types are listed. CAPEC proposes corresponding attack modes for various vulnerabilities in software design. Each attack mode includes the name and category of the mode, attack premise, and related vulnerabilities [11]. CAPEC provides a basis for the design of ontology attack classes in the field of network security.

3. Construction of Ontology Model in Network Security Domain

Based on the knowledge source of network security domain, this paper defines the model of the cybersecurity ontology as a quad model $CDO = (C, I, R, A).C = \{c_1, c_2, \ldots, c_m\}$ represents the collection of network security domain classes, m represents the number of network security domain classes, and c_x refers to a certain type of network security field, such as attack, vulnerability, vulnerability, etc. $I = \{i_1, i_2, \ldots, i_n\}$ represents the entity set, n represents the number of entities, and i_x represents an entity of a certain network security domain class, such as a specific vulnerability in the vulnerability class. $R = \{r_1, r_2, \ldots, r_l\}$ represents the set of relationships between classes, l represents the number of relationships between classes in the network security field, r_x refers to the relationship between a certain type and class, such as the utilization relationship between attack class and vulnerability class. $A = \{a_1, a_2, \ldots, a_k\}$ represents the constraint set of classes and relationships, k represents the number of constraints, and a_x represents a specific constraint. For example, it is a fact that the network bandwidth belongs to the attack target. The following describes the quad model.

3.1. The Collection of the Classes in the CDO Model

This paper defines the important classes of network security ontology from five aspects: network assets, vulnerability, attacks, observation indicators and intelligence.

Network assets are the basic environment of various applications in the managed network, including hosts, network devices and security devices. Host is a general term for equipment and software, including servers, personal computers and portable terminals. It is usually composed of hardware, operating systems, software and running services, and communicates with other hosts on the Internet through the network. Therefore, the important class set of network assets *Class (Assets)={Host, Network Equipment, Safety Equipment, Hardware, OS, Software, Network, IP, Port}*.

Vulnerability is the weakness of network assets, which is easy to be attacked and utilized. It exists in the host's hardware, operating system, software and application services, and is usually disclosed in the form of vulnerabilities and defects. The important class set of vulnerability is *Class (Fragility)={Vulnerability, Weakness}*.

Attacks mainly include attackers, attack methods, attack tools, attack events and attack consequences. Attackers include individuals, groups or hacker organizations. Attackers take advantage of the vulnerability of network assets to launch attacks. The attack methods include the techniques and means used in the attack, such as XSS cross site script attack, SQL injection, etc. Attack tools include normal software and malware. Attackers use normal tools to conduct network attacks and hide attacks. An attack event indicates an event generated by an attacker using an attack method and tool. The attack consequence is used to describe the effect of the attack, such as unauthorized access, authorization, denial of service, etc. The important class set of attacks is represented as *Class (Attack)={Attacker, Mean, Tool, Malware, Consequence}*.

Observation indicators are the characteristics of network assets after being attacked, including host characteristics and network characteristics. For example, after the host is attacked, a new main table entry is added. The important class set of observation indicators is represented as *Class (Indicator)={Host Feature, Network Feature}*.

Intelligence mainly refers to network security incident intelligence. The intelligence in this paper refers to comprehensive network security information obtained from network security alarm information, network security forums and websites, especially some unpublished vulnerability information. Through security event intelligence, operators can get the overview of external threats and details of security events, and then conduct targeted protection against current security events. The important class set of information dimension is represented as *Class (Intelligence)={Event, Threat intelligence}*.

The design of ontology model in network security domain mainly considers three aspects: modeling basis, characteristics and functions. In the CDO model, network assets are the basis of attacks, vulnerability is the basis of attacks, attack dimension is the attack behavior, and observation indicators are the features displayed after the attack results. Network assets are vulnerable and easy to be used by attacks. Attack elements act on network assets, and observation indicators reflect the impact of attacks. The four dimensions form an attack loop that can be used to fully describe attacks. As a supplement to vulnerability, intelligence is more conducive to identifying attack threats. Therefore, these five aspects are used as the basis for modeling. the CDO model is a knowledge system that builds a network security knowledge map as a knowledge map. On the one hand, it provides support for building a network security knowledge map using network security knowledge, and on the other hand, it provides knowledge for network security knowledge map completion and link prediction.

3.2. The Set of Relationships between Classes in the CDO Model

Network attacks usually consist of a variety of behaviors. The relationship set of the CDO model is expressed as *Relation={subClassOf, hasHardware, hasOS, hasSoftware, hasIP, hasPort, hasVulnerability, hasWeakness, explore, use, include, involved, lunch, attach, associate, generate, indicate, cause}*.The network security ontology relationship established in this paper is shown in Figure.1.

3.3. The Attributes of the Important Class in the CDO Model

After determining the important class set and relationship set in the network security domain, we will introduce several important classes and their attributes.

- The attributes of hardware, OS and software class. The class of host, network device, and security device include hardware, operating system, and software, which are usually developed by a manufacturer or organization, and updated and iterated according to the product life cycle, so they have different versions. This paper refers to the general properties of hardware, operating system and software described in the generalplatform enumeration library CPE. The properties of these three classes are summarized as *Attributes (Hardware | OS | Software)={id, CPE_names, vendor, product, version, release date}*.

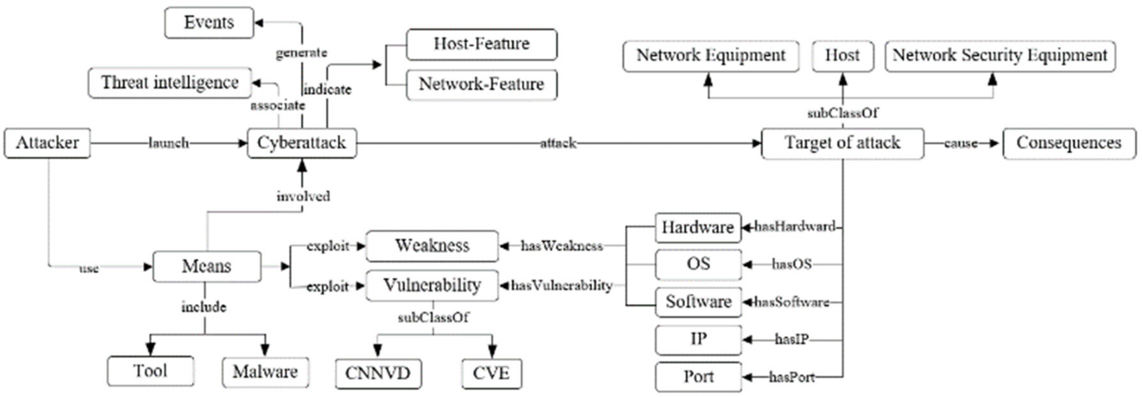

Figure 1. Network Security Ontology Relationship Diagram

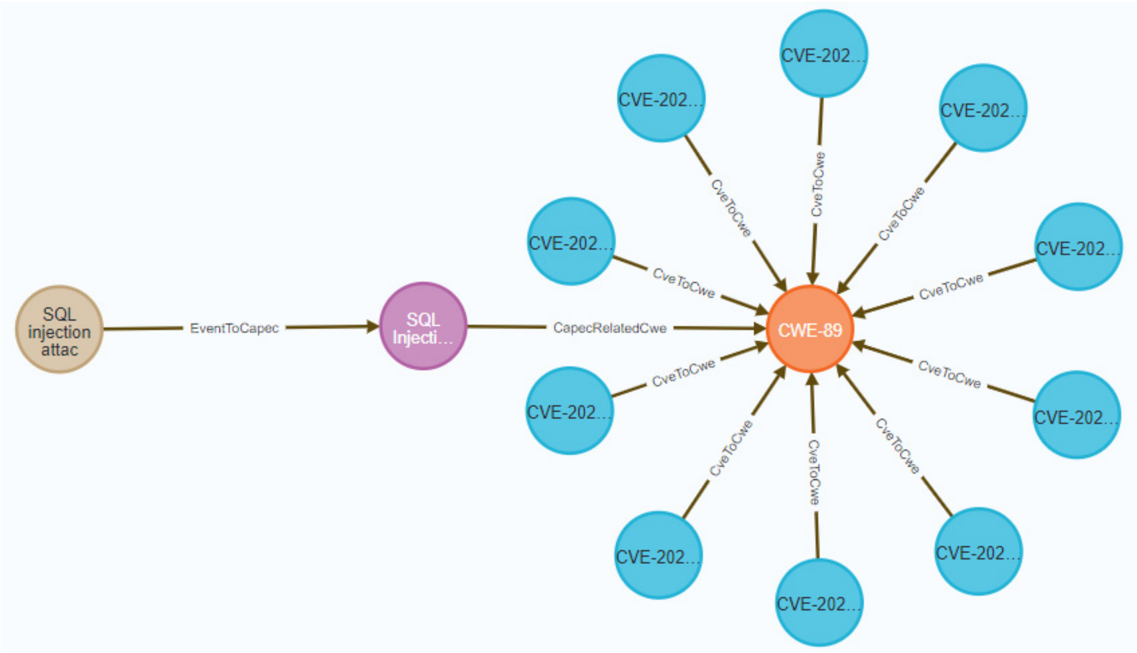

Figure 2. SQL Injection Attack Knowledge Map.

- The attributes of vulnerabilities class. The attributes of vulnerabilities mainly refers to the National Vulnerability Database of Information Security (CNNVD). They are summarized as *Attributes (Vulnerability)={cnnvdi, cveid, name, vul_type, threat_type, description, announcement, affected_entity, publish_date, update_date, reference}*.
- The attributes of CVE class. The CVE attributes aim to establish the association relationship between CNNVD, CWE, CPE and other vulnerability databases, so as to facilitate the identification of attacks. They are summarized as *Attributes (Vulnerability)={cveid, description, cvss_level, cweid, cpes, create_date}*.
- The attributes of weakness class. They are based on CWE, and summarized as *Attributes (Weakness)={cweid, name, platform, desc, mode, sequence, likelifed_of_exploit, detection_method, potential_missification, affected_resource, related_attack_pattern}*.

```
MATCH p= (a: Cve) - [r1: CveToCwe] -(b:Cwe) -
[r2:CapecRelatedCwe] - (c:Capec)-[r3:EventToCapec]
- (d:Event) where d.eventname="SQL injection attack"
RETURN p limit 10
```

- The attributes of attack class. They are based on CAPEC and ATT&CK, and summarized as *Attributes (Mean)={id, name, desc, likelihood_of_mean, mean_occurrence, execution_flow, prerequisite, required_skill, mean_mission, cwe_id, mean_action, platform, required_permission, capec_id, att_ck_id, common_target, common_group, common_tool, affected_service}*.
- The attributes of security event. They mainly refer to the alarm collected by various security devices such as firewall, and summarized as *Attributes(Event)={event, enent_name, sip, sport, dip, dport, desc}*.

As the schema layer of network security knowledge map, the ontology is used to integrate network security data to construct network security knowledge map. The network security domain ontology established in this paper uses the general network security terminology as the knowledge system of the network security knowledge map, which can be used to link external network security data and integrate the network security knowledge data into the knowledge system to build the corresponding network security knowledge map.

4. An Example of Network Security Knowledge Map in Network Attack Analysis

This paper takes a large network as an example. On October 11, 2022, the IPS received a remote SQL injection vulnerability alarm message from the WEB service. It can be seen from the map that the source host with an IP address of 113.66.36.xxx launched a SQL Injection (CAPCE No. 66) attack against the target host. According to the map, SQL Injection mainly attacks the weakness of CWE-89, which involves CVE-2022-0153, CVE-2022-0169, CVE-2022-0190 and other vulnerabilities, there has been 161 CWE-89 vulnerabilities in 2022 alone.

After the construction of the CDO model is completed, this paper uses Neo4j as the graph database, imports CVE, CWE, CAPEC and other data into the CDO model through the Cypher language, and forms a knowledge map in the network security field.

In order to verify the analysis efficiency of knowledge maps, this paper takes SQL Injection attacks as an example to introduce how to use knowledge maps to analyze network attacks. On October 11, 2022, the intrusion detection system of a large network received an alarm information of the WEB service remote SQL injection vulnerability, and the system displayed that the source host with IP address 113.66.36.31 launched a SQL injection attack against the target host. The Cypher query statement is as follows:

The query result is shown in Figure 2. According to the graph, the attack mode of SQL injection attack event is the attack mode numbered 66 in CAPEC. This attack takes advantage of CWE-89 vulnerability in the Web application system, which involves CVE-2022-0153, CVE-2022-0169, CVE-2022-0190 and other vulnerabilities. In 2022 alone, there will be 161 CWE-89 vulnerabilities.

5. Conclusion

This paper constructs an ontology model in the network security domain, establishes ontology classes, attributes, and ontology relationships, extracts knowledge through structured and semi-structured data, and establishes a knowledge map. Finally, the model is verified through attack events. The knowledge map established in this paper can intuitively analyze the relationship between attack sources, attack targets, attack behaviors, vulnerabilities, threat indicators, etc., so as to quickly make emergency response in network security incidents.

6.Acknowledgment

This work is funded by CCF-NSFOCUS Kunpeng Scientific Research Fund.

7.References

[1] Z.Y. Ding, K. Liu, B. Liu, X. X. Zhu, “Survey of cyber security knowledge graph,” J. Huazhong Univ. of Sci. & Tech. (Natural Science Edition), Vol.49 No.7, pp.79-81, Jul. 2021.

[2] T. Li. Research on Key Technologies for Construction and Application of Threat Intelligence Knowledge Graph. Ph.D. thesis, Information Engineering University, Henan, China.2020.

[3] L.Y. Lian, “Ontology-based construction method for cyberspace security knowledge graph,” Journal of Heilongjiang University of Science & Technology, Vol.31 No.2, pp.254-258, Mar.2021.

[4] Structured Threat Information Expression, http://stixproject.github.io, last accessed 2022/11/12.

[5] Common Vulnerabilities and Exposures, http://cve.mitre.org, last accessed 2022/11/12.

[6] The Common Weakness Enumeration Specification, http://cwe.mitre.org, last accessed 2022/11/12.

[7] Common Platform Enumeration, https://cpe.mitre.org, last accessed 2022/11/12.

[8] Common Attack Pattern Enumeration and Classification, http://capec.mitre.org, last accessed 2022/11/12.

[9] MITRE ATT&CK, https://attack.mitre.org, last accessed 2022/11/12.

[10] Y. Liu, H.F. Zhang, L. Zhang, “Research on a collaborative method of penetration testing based on STIX information interaction,” Network and Information Security,Vol.37 No.12, pp.1-5,Dec.2018.

[11] H.B. Xiao. Software Security Knowledge Graph Completion Based on Relation Reasoning. M.D. thesis, Tianjin University, Tianjin, China.2019.