
White Rose Research Online

eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk
Universities of Leeds, Sheffield and York

Deposited via The University of Leeds.

White Rose Research Online URL for this paper:
https://eprints.whiterose.ac.uk/id/eprint/239082/

Version: Accepted Version

Proceedings Paper:
Wang, Z., Fei, M., Xiong, Y. et al. (2024) Dynamic Attack Path Prediction and Visualization for Industrial Cyber-Physical Systems Under Cyber Attacks. In: 2024 43rd Chinese Control Conference (CCC). 2024 43rd Chinese Control Conference (CCC), 28-31 Jul 2024, Kunming, China. IEEE, pp. 9005-9010. ISBN: 979-8-3503-6690-7. ISSN: 1934-1768. EISSN: 1934-1768.

https://doi.org/10.23919/ccc63176.2024.10662270

© 2024 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

Reuse
Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

Takedown
If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.



Dynamic Attack Path Prediction and Visualization for Industrial Cyber-Physical Systems Under Cyber Attacks

Zijin Wang[1], Minrui Fei[*], Yao Xiong[1], Aimin Wang[1]

1. Shanghai Key Laboratory of Power Station Automation Technology, School of Mechatronic Engineering and Automation, Shanghai University, Shanghai, 200444, China
E-mail: zijinwang@shu.edu.cn; mrfei@staff.shu.edu.cn; xiongyao@shu.edu.cn; amwang@shu.edu.cn

Abstract: The accurate and effective prediction of network attack paths has become a crucial concern in the realm of network security, given the inherent uncertainty and subjectivity associated with network attack methods. To solve this problem, this paper proposes a visualized dynamic attack path prediction scheme for industrial cyber-physical systems (ICPSs). The method combines the Bayesian attack graph with the knowledge graph and considers the topology of the digital twin layer to make it closer to the actual situation. In addition, node dynamic reachability probabilities are considered to provide support for the interpretation of the prediction results. The simulation results demonstrate that the proposed scheme is more flexible and scalable than the static attack graph. These improvements enable more accurate prediction of the network attack path and enhance the network's security protection ability.

Key Words: Industrial cyber-physical systems (ICPSs), Cyber-attack paths, Knowledge graph, Attack dynamic prediction, Digital twin

1 Introduction

The use of advanced information and communication technologies (ICTs) in modern industrial control systems (ICSs) has led to the development of industrial cyber-physical systems (ICPSs). These systems are crucial to the operation of various industries, including power generation, gas pipelines, and urban railway transportation systems. However, the increasing size of networks and the growing popularity of network technologies have resulted in an increase in cybersecurity risks. As a result, network attack path prediction has become an essential component in ensuring the security of ICPSs as an active defense method [1-2].

As network attackers' intrusion methods evolve from simple one-step attacks to multi-step attacks, one effective method for cybersecurity researchers to predict the path of network attacks is to identify the correlation relationship between network nodes, simulate the attack process, and visualize the attack process. Therefore, C. Phillips and P. Swiler proposed the concept of the attack graph in 1998 [3]. K. Liu *et al.* [4] proposed the Network Attack Profit Chart (NAPC) model for predicting network attacks. The model addresses problems such as difficult-to-predict network attack behavior, inaccurate prediction of network redundant path attacks, and a single prediction index for attack path assessment. The NAPC model has faster convergence. However, when using current algorithms to identify the attack paths of network nodes, the detected node attack paths are chaotic, and there is a large error in the detection of node attack paths. On this basis, V. Bartos *et al.* [5] proposed a system, the Network Entity Reputation Database System (NERDS), to characterize network entities and their likelihood of future malicious behavior. The system takes into account all available information regarding the network entity, such as its IP address, and employs machine learning to determine the likelihood of malicious behaviour. The results have been consistently positive. However, this method relies on static analysis and cannot adjust the association probability generation based on attack behavior and defense measures. X. Liu [6] proposed an attack intent prediction method based on an attack graph, mapping the maximum likelihood of a host or vulnerability attack, further obtaining the probability of intrusion under each attack path of the network, and constructing network attack node paths based on the intrusion likelihood detection results detection model, from which the shortest path to realize the attacking intent is found. X. Chen *et al.* [7] proposed a probabilistic graph model for inferring and predicting internal attack intent. The model considers the probability of attack occurrence, success, and the confidence level of observed events to measure the uncertainty of the attacker's behavior. It enables the identification of the maximum probability attack path of the attack target, thereby guiding network protection. While these methods allow for dynamic analysis, they do not take into account the dynamic updating of the attack graph. In general, current research relies on static analysis, which fails to consider the impact of attackers' capabilities on subsequent attack paths. This compromises the accuracy of attack path prediction. Additionally, existing attack graph technologies focus mainly on building attack prediction models for static networks, making it difficult to adapt to dynamic changes in network states. This limitation hinders the active updating of attack graphs and also inhibits real-time prediction of attacker behaviour.

In recent years, with the rapid development of emerging technologies such as “cloud-big-thing-mobile-intelligence-chain”, the realization of digital twins has become possible, and many scholars at home and abroad have carried out a large number of research and practice on digital twins and achieved corresponding results [8-10]. With the wide application of digital twins, the attack surface has also been expanded. Most of the existing methods for predicting the attack paths are constructing graph models in the cyber layer

[*]This work was supported in part by the 111 Project under Grant D18003.

and physical layer [11], which do not take the digital twin layer into account and cannot better adapt to the changes of the actual situation.

In the context of the rapid development of artificial intelligence, the knowledge graph is widely recognized as an important component of AI technologies and systems [12]. Knowledge graph has powerful functions and good query performance and is highly scalable and formally similar to attack graph models.

Motivated by the above descriptions, the visual dynamic attack path prediction scheme for ICPSs under cyber attack is proposed in this paper. The main contributions of which are as follows:

1) A method for dynamically predicting visual attack paths in real-time cyber security has been designed. This method can be applied to three different scenarios: alteration of network structure, change of host node information, and emergence of new vulnerability nodes.

2) Combining the digital twin layer, the cyber layer, and the physical layer in ICPSs is more realistic, which in turn improves the efficiency and safety of ICPSs.

2 Real-time Attack Path Prediction Method Construction

2.1 Data collection, extraction, and analysis

(1) The Bayesian attack map

Bayesian attack graphs are an extension of attribute attack graphs that incorporate Bayesian probabilistic inference [13].

Definition 1: the Attribute attack map and the Bayesian attack map.

Both are directed acyclic graphs, the Attribute attack graph can be denoted as $AG(S,A,E,R)$ and the Bayesian attack graph can be denoted as $BAG(S,A,E,R,P)$, where S is the set of states of all attribute nodes S_i. A is the set of all atomic attacks A_i. E is the set of directed edges connecting all attribute nodes in the attack graph, which is used to denote the attack relationship between two and two nodes. R is the set of relationships between the parent node and the child node, $R =< S_i, R_i >$, where $R_i \in \{AND, OR\}$, which denotes that there is an AND and OR relationship between the parent and the child node. And AND means that S_i can be attacked only if all the parent conditions of S_i are satisfied, and OR means that S_i may be attacked only if any one of all the parent conditions of S_i is satisfied. P denotes the weight on the directed edge, i.e., the probability of launching an attack on the child node if the parent node has already been compromised.

Definition 2: Dynamic Reachability Probability P_c.

The dynamic reachability probability of a node indicates the probability of an attacker successfully attacking the node from the starting attack when the intrusion detection system has not detected the attack. The computation of the dynamic reachability probability depends on several factors, including the reachability probability of the parent node, the likelihood of attacking the node after compromising its parent, and the probability of the node being compromised. When an attack is detected on node S_i the reachability probabilities of both its parent node S_j and its child node S_k will be affected. The updated reachability probability at this point is referred to as the dynamic reachability probability of the node. Specifically, the dynamic reachability probability of node S_i is set to 1, indicating that it has already been compromised.

The dynamic reachability probability of the parent node S_j is calculated as formula (1):

$$P_c(S_j|S_i) = \frac{P(S_i|S_j) \cdot P_b(S_j)}{P_b(S_i)} \quad (1)$$

As the relationship between a node and its parent is categorized as AND and OR:

For the parent node with AND relationship, the dynamic reachability probability of the child node S_k is calculated as formula (2):

$$P_c(S_k \mid S_i) = P_a(S_k) \cdot \prod_{j=1}^{n} P_c(S_j) \cdot P(S_k \mid S_j) \quad (2)$$

Where $P(S_k|S_j)$ denotes that after the parent node S_j is compromised, attacking the node S_k probability, i.e., the weights of the edges between the nodes, it is mainly negatively correlated to the cost of vulnerability exploitation of the child node S_k. $P_a(S_k)$ indicates the probability that the node S_k vulnerability is successfully exploited to obtain permission, which can quantify the vulnerability's utilization rate based on vulnerability scoring $Score$, $P_a(S_k) = \frac{Score}{10}$. $P_b(S_j)$ denotes the static reachability probability of the node S_j.

For a parent node with an OR relationship, the dynamic reachability probability of the child node S_k is calculated as formula (3):

$$P_c(S_k \mid S_i) = P_a(S_k) \cdot \left[1 - \prod_{j=1}^{n} \left(1 - P_c(S_j) \cdot P(S_k \mid S_j)\right)\right] \quad (3)$$

(2) Steps in Bayesian attack graph construction

The Bayesian attack graph serves as a powerful tool for visualizing the exhaustive range of attack methods that can be employed to achieve a specific attack goal. This visualization capability empowers network security officers to identify and comprehend all potential attack processes that may manifest within the network. Consequently, the construction of a Bayesian attack graph model assumes paramount importance as it forms the foundation for realizing network security prediction. The construction process of the Bayesian network attack graph model entails two distinct steps, which are delineated as follows:

a) The acquisition of network environment information: This entails employing the Nmap topology scanning tool to perform a comprehensive topological scan of all hosts within the network. The objective is to obtain a clear understanding of the reachability relationships among the hosts, subsequently generating a comprehensive list of host information and a detailed table depicting host reachability within the network. Subsequently, the vulnerability scanning tool Nessus is employed to conduct a thorough scan of each host node within the target network, thereby extracting crucial vulnerability information associated with each host. In addition, to construct a comprehensive and high-quality attack graph, this paper collects vulnerability information from several structured and semi-structured data platforms, including NVD [14], CNVD [15], and research reports. The collected information serves as the data basis for

constructing the attack graph. The network environment information is formally represented to build a Bayesian network attack graph.

b) Building Bayesian Attack Graphs: The logical editing language Datalog is used to describe the system information, which mainly includes network topology information, network host information, and vulnerability information. By outputting the collected network environment information to the MulVAL inference engine, the network with thousands of machines can be analyzed in seconds to generate an attribute attack graph, which is then combined with the Bayesian network to generate a Bayesian attack graph.

2.2 Dynamic prediction method for network attack paths

To overcome the limitations of static attack graphs in capturing the dynamic nature of network structures, evolving host information, and emerging vulnerabilities, we propose a novel approach that integrates attack graphs with knowledge graphs. This integration allows for the effective representation and analysis of the network's evolving attack landscape.

Our approach utilizes attack graphs as input, which are then processed to generate comprehensive knowledge graphs encapsulating all potential attack paths. This integration enables us to incorporate the evolving network structure, updated host information, and newly discovered vulnerabilities into the attack path generation process. To achieve this, we use the Cypher query language within the graph database Neo4j. By using Cypher statements, we can traverse the information provided by the attack graphs and identify and update attack paths based on changes in host information and the presence of new exploitable vulnerabilities that meet the predefined conditions. This algorithmic implementation ensures the efficient and accurate generation of updated attack paths.

Fig. 1 presents a brief summary of the steps involved in the attack path update process. It demonstrates the integration of attack graphs and knowledge graphs to capture the dynamic nature of network structures and vulnerabilities. This approach enables network security professionals to stay informed about evolving threats and adjust their defense strategies accordingly.

The relationship between the Bayesian Attack Graph and the Knowledge Graph in this paper is as follows: the Bayesian Attack Graph serves as the input for the attack path prediction method, providing the necessary information for the prediction. The Knowledge Graph represents the attack path prediction results and is updated dynamically to enable dynamic prediction of the attack path.

The whole modeling process is shown in Fig. 2.

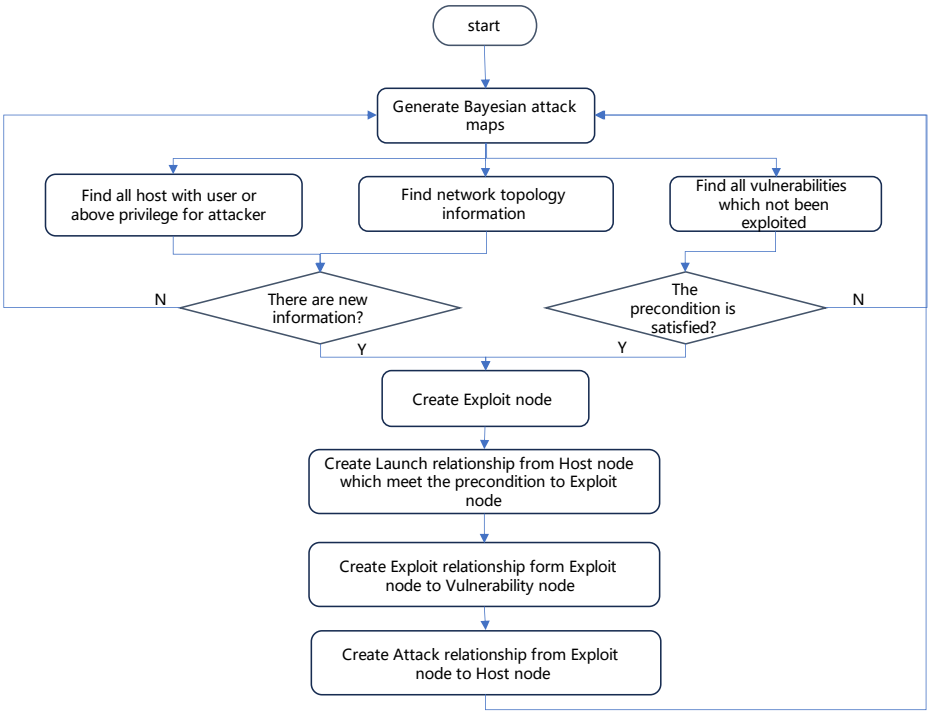

Fig. 1: Dynamic update flowchart of the attack path

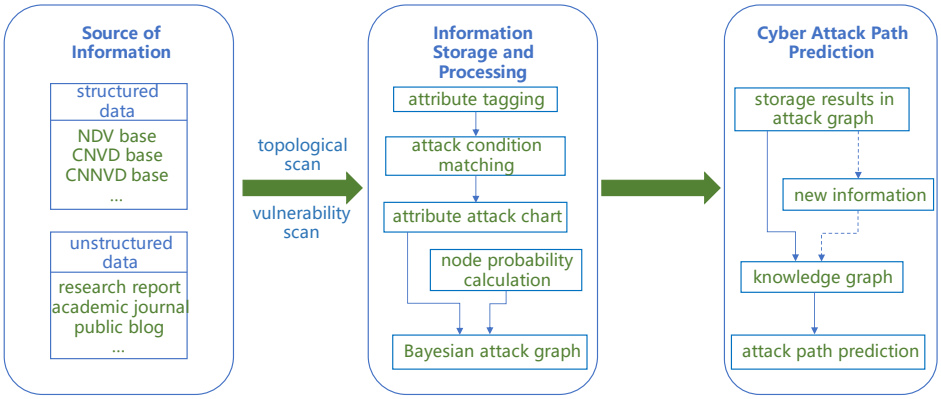

Fig. 2: Visual dynamic attack path prediction method

3 Experimental simulation analysis

To verify the effectiveness of the network attack path prediction method proposed in this paper, we designed a LAN environment for experimentation. Firstly, we constructed the experimental network using the Bayesian attack graph construction method introduced in section 2.1 to store and process the experimental network data. Finally, the Bayesian attack graph is imported into the Neo4j graph database, and the dynamic prediction of the attack path is achieved using the Cypher query language.

3.1 Experimental environment

The system comprises a database server, digital twin server, monitor, serial server, wind turbine, photovoltaic array, photovoltaic inverter, wind turbine controller, and electrical control cabinet [16]. The paper presents an ICPS, as depicted in Fig. 3. To validate the method, the experiment scene's network topology is constructed based on the actual diagram, as shown in Fig. 4.

Fig. 3: Microgrid physical platform

We use Nmap to scan the topology and device information, at the same time, there are some vulnerabilities in the devices in the network topology, we check the vulnerabilities of the hosts under the experimental network through the vulnerability scanner Nessus, at the same time, each attribute is comprehensively obtained from the National Open Vulnerability Database. The information for each node and vulnerability is shown in Table 1.

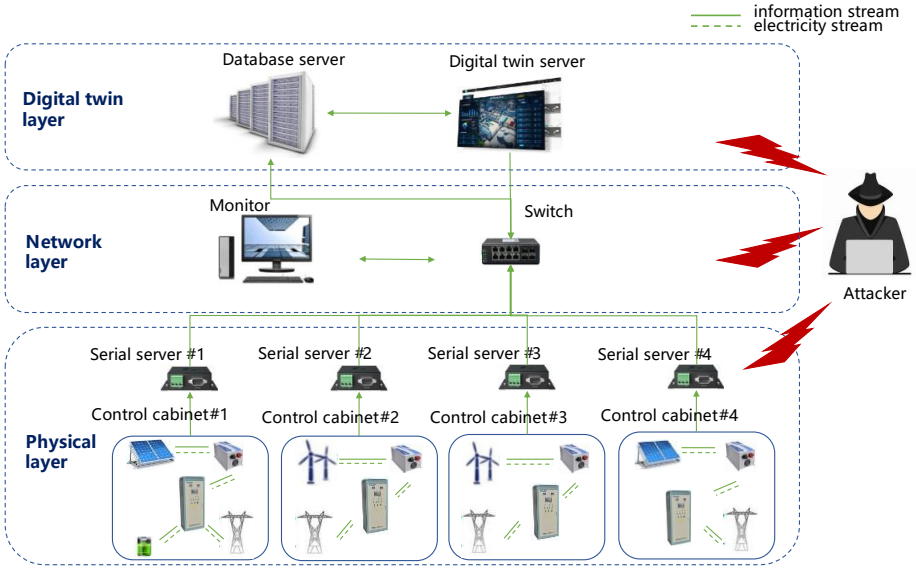

Fig. 4: The actual topology

Table 1: Network host and vulnerability information

Network equipment	ID	Vulnerability	ID	CVSS score	Attack complexity
Database server	S1	CVE-2023-21949	V1	3.7	HIGH
Digital twin server	S2	CVE-2017-5638	V2	9.8	LOW
Monitor	S3	CNVD-2018-06928	V3	5	LOW
Switch	S4	CVE-2022-40224	V4	7.5	LOW
Serial server	S5	CVE-2023-4204	V5	9.8	LOW
Breaker	S6	CVE-2022-25156	V6	8.1	HIGH
Automatic voltage regulator	S7	CVE-2022-44808	V7	6.5	HIGH
Controller	S8	CVE-2023-29107	V8	5.3	LOW
Expandable device	S9	CNVD-2023-57643	V9	3.3	LOW

3.2 Results

We use OMNET++ software to simulate the above network system model, simulate a network attack in ICPS, and visually present the predicted path of the network attack (as shown in Fig. 5). In this process, the simulated attack method is random, which can be a cyclic attack or a repeated attack.

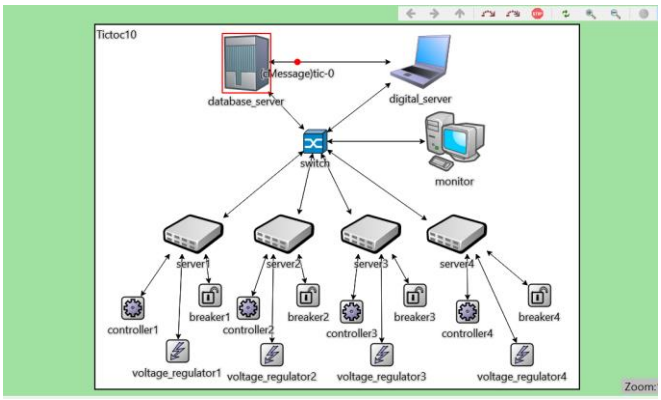

Fig. 5: Network attack path prediction visualization

The system topology diagram may appear simple, but it can become very complex when used to create a multi-path attack diagram. To ensure clarity, this article avoids listing all possible paths associated with the attack. Instead, the attack flow is depicted in Fig. 6, with the newly formed attack path represented by a red line. If an insecure mobile maintenance device, such as a USB flash drive or a laptop, is illegally accessed and vulnerability V9 exists, attackers can exploit the device's vulnerability to carry out attacks and create a new attack path.

When the attacker scans and detects the host's openness and exploits its vulnerability to gain authority, thereby launching a network attack, the probability of each node in the network being attacked also changes. In the original experimental environment, assuming that the attribute node S4 has been taken by the attacker, the experimental attribute node's static/dynamic reachable probability comparison is shown in Fig. 7.

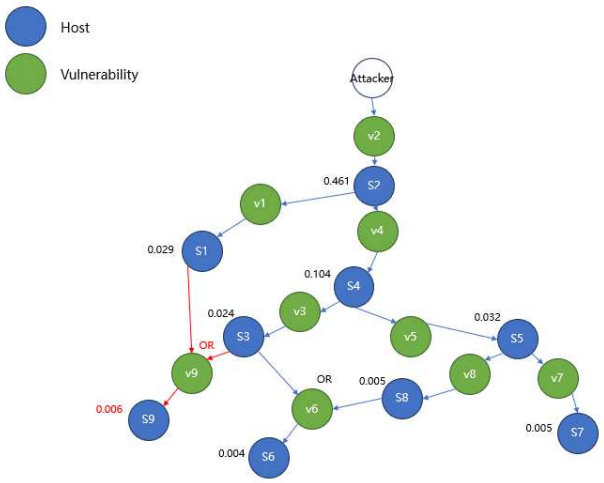

Fig. 6: Network attack prediction path graph

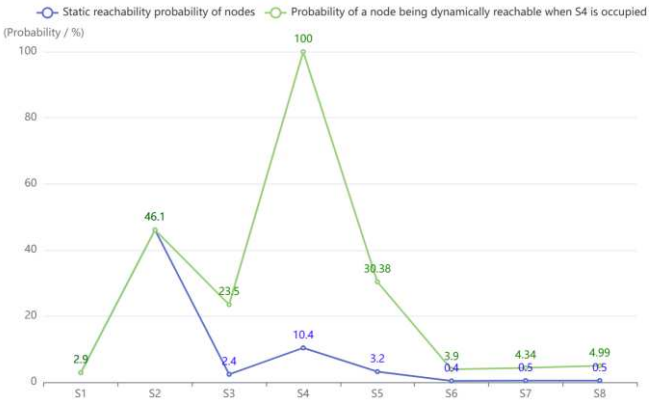

Fig. 7: Comparison of dynamic/static reachability probabilities of nodes

3.3 Analysis of simulation results

(1) Flexibility

Through the above experiments, we found that when the attacker takes a non-traditional way to attack, such as the attack packet moving from the digital twin server to the database server and then back to the digital twin server, this is a new attack path that the static attack graph fails to predict. Therefore, the static attack graph cannot continue to predict the attacker's next move. Instead, our proposed solution adapts by learning these new behaviors and updating the network attack prediction pathways, as shown in Fig. 5. Consequently, our approach offers enhanced adaptability, effectively anticipating and countering the dynamic nature of network attacks with greater resilience.

(2) Scalability

It can be seen from the above experiments that the static attack graph cannot be changed when the experimental topology is changed or new vulnerabilities appear, and the strict specification of the static attack graph makes it possible only to redefine the attributes and generate a new attack graph. In contrast, the method proposed in this paper is scalable and only needs to add new edges to the original.

(3) Predictive analytics and accuracy

The attack path diagram depicted in Fig. 7 highlights the notable increase in the dynamic reachability probability of nodes S3 and S5 when node S4 is targeted. Moreover, the dynamic reachability probability of node S5 surpasses the others, indicating a higher likelihood for the attacker to target node S5 in the subsequent step. It is worth noting that the static attack graph remains unchanged and fails to incorporate the specificities of the situation. Consequently, the proposed prediction method can effectively anticipate the attacker's next move with a higher degree of accuracy compared to the static attack graph.

In conclusion, the proposed methodology entails the integration of the attack graph with the knowledge graph, followed by pre-processing the data utilizing the attack graph to mitigate the impact of class imbalance. Subsequently, the processed data is combined with the knowledge graph to facilitate the dynamic prediction of the attack path. The approach also takes into account the dynamic reachable probability of nodes, thereby rendering the prediction analysis more rational and precise. Based on the results obtained from the aforementioned experiments, it can be inferred that the proposed approach exhibits superior levels of flexibility and scalability, and is capable of accurately predicting the next behavior of the attacker, thereby establishing its potential as a viable method for enhancing cybersecurity measures.

4 Conclusion

The integration of ICSs, the Internet of Things (IoT), and the Internet has increased the need for network security. This paper proposes a real-time method for predicting the attack path of ICPSs, addressing the limitations of the static attack graph. Our approach involves constructing a Bayesian attack graph that utilizes the knowledge of Bayesian networks, a more mature research field. In addition, we combine the Bayesian attack graph with the knowledge graph to enable dynamic updates of the attack path. Moreover, we take into account the dynamic changes in nodes' reachability probability to support the interpretability of prediction results and enhance prediction accuracy.

Moving forward, in light of the ongoing optimization of the dynamic path graph of cyberattacks, we intend to undertake a more exhaustive approach to enhance the precision of attack prediction. Furthermore, we plan to implement effective mitigation measures and remedial actions for operators following the prediction of cyberattacks.

References

[1] Y. Deng, T. Zhang, G. Lou, X. Zheng, J. Jin, and Q.-L. Han, "Deep Learning-Based Autonomous Driving Systems: A Survey of Attacks and Defenses," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 12, pp. 7897-7912, Dec. 2021.

[2] M. Husák, J. Komárková, E. Bou-Harb, and P. Celeda, "Survey of attack projection, prediction, and forecasting in cyber security," *IEEE Communications Surveys & Tutorials*, vol. 21, no. 1, pp. 640-660, Mar. 2019.

[3] C. Phillips and P. Swiler, "A graph-based system for network-vulnerability analysis," *Proceedings of The 1998 Workshop on New Security Paradigms*, pp. 71-79, 1998.

[4] K. Liu, H. Wang, and Z. Shen, "Prediction of network attack profit path based on NAPG model," *The Journal of China Universities of Posts and Telecommunications*, vol. 27, no. 05, pp. 91-102, 2020.

[5] V. Bartos, M. Zadnik, S. Habib, and E. Vasilomanolakis, "Network entity characterization and attack prediction," *Future Gener Computer Systems-The International Journal of Science,* vol. 97, pp. 674–686, Aug. 2019.

[6] X. Liu, "A network attack path prediction method using attack graph," *Journal of Ambient Intelligence and Humanized Computing*, pp. 1-8, Jun. 2020.

[7] X. Chen, B. Fang, Q. Tan, and H. Zhang, "Inferring Attack Intent of Malicious Insider Based on Probabilistic Attack Graph Model," *Chinese Journal of Computers*, vol. 37, no. 1, pp. 62-72, 2014.

[8] C. Gehrmann and M. Gunnarsson, "A digital twin based industrial automation and control system security architecture," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 1, pp. 669-680, Jan. 2019.

[9] A. Murillo and S. Rueda, "Access control policies for network function virtualization environments in industrial control systems," *2020 4th Conference on Cloud and Internet of Things (CIoT)*, pp. 17-24, 2020.

[10] Y. Xiong and M. Fei, "Virtual Manufacturing + Digital Twin Development and Application Practice," *Process Automation Instrumentation*, vol. 44, no. 8, pp.1-6+14, Aug. 2023.

[11] J. Wang and Y. Guo, "Network Security Risk Assessment of Cyber Physical System Based on Attack Graph," *Science Technology and Engineering*, vol. 23, no. 28, pp. 12175-12181, 2023.

[12] B. Xue and L. Zou, "Knowledge Graph Quality Management: A Comprehensive Survey," *IEEE Transactions on Knowledge and Data Engineering*, vol. 35, no. 5, pp. 4969-4988, May. 2023.

[13] P. Zhang, "Research on Vulnerability Assessment and Cyber Attack of Information Physical System of Power Distribution Grid," *North China Electric Power University (Beijing)*, 2021.

[14] H. Booth and D. Rike, and G. White, "The national vulnerability database (nvd): overview," 2013.

[15] I. Forain, R. De Oliveira Albuquerque, and R. De Sousa Júnior, "Towards System Security: What a Comparison of National Vulnerability Databases Reveals," *2022 17th Iberian Conference on Information Systems and Technologies (CISTI). IEEE*, pp. 1-6, 2022.

[16] W. Chen, "The Research and Implementation of Intrusion Detection and Imbalanced Data Generation Method for Industrial Control Networks," *Shanghai University*, 2023.