Risk Analysis for Railway Signaling Safety Data Network Based on Extend Bayesian Attack Graph

Yuanyuan Liao¹, Jian Wang¹,²,³, *, Kaiyuan Tian¹, Xuyu Wang⁴, Baigen Cai¹,³

¹School of Electronics and Information Engineering, Beijing Jiaotong University, Beijing 100044, China
²State Key Laboratory of Rail Traffic Control and Safety, Beijing 100044, China
³Beijing Engineering Research Center of EMC and GNSS Technology for Rail Transportation, Beijing Jiaotong University, Beijing 100044, China
⁴Beijing Huatie Information Technology Co., Ltd, Beijing 100081, China

*Corresponding author e-mail: wangj@bjtu.edu.cn

Abstract. The railway signaling safety data network is a key network to ensure the safety of train operation and improve transportation efficiency. Its information security is closely related to the safe and efficient operation of trains. This paper proposes an extend Bayesian attack graph model for risk analysis of railway signaling safety data network. Based on the traditional Bayesian attack graph, the model introduces the protection node, and describes the causality among network attack, protection and network state through the Bayesian network. First of all, based on the analysis of the vulnerabilities and possible malicious attack paths in the railway signaling safety data network, the extend Bayesian attack graph model is established with the system functional safety accidents as the target nodes. Then calculate the probability of functional safety accidents by Bayesian network, and combine the impact of different accidents to evaluate the risk of system information security. In addition, according to different attacks and the implementation of different protective measures, evidence node state can be set up, and the extend Bayesian attack graph model is able to use Bayesian inference to analyze the system risk.

1. Introduction

As the core of ensuring train operations safety and improving transportation efficiency, railway signaling system is closely related to railway traffic system. Railway signaling system consists of signaling safety data network, centralized traffic control network and centralized signal monitoring network in China. Centralized Traffic Control system (CTC) and Centralized Signal Monitoring system (CSM) use internal network to communicate. The equipment in Chinese Train Control System (CTCS) and equipment in Computer Interlocking System (CBI) communicate with each other through the signaling safety data network, and transmits data with onboard equipment through GSM-R (Global System for Mobile Communications for Railway). Data and commands transmitted in the signaling safety data network are related to the traffic safety directly.

With the prompt level of information and intelligence, the convergence of industrial control systems and information technology is getting closer. It brings many security risks at the same time. In 2010, the
Stuxnet virus attacked Iran’s nuclear centrifuges [1]. In 2015, Ukrainian power companies suffered network intrusion resulting in more than 30 power plant shutdowns [2]. Israel’s national grid suffered a large network attack in 2016[3]. In 2017, ransom ware WannaCry swept the world. In 2018, semiconductor company TSMC suffered from viral infections, causing its three major production bases to be affected [4]. As a kind of industrial control system, the information security problems for railway signaling safety system are becoming more and more serious. It requires to establish a suitable risk analysis model with attacks and defenses to accurately describe behavior and network status.

In recent years, attack tree, attack graph, Petri net, analytic hierarchy process, Bayesian network have been studied in risk analysis. Schneier et al analysis security risk by attack tree [5]. An attack tree-based assessment approach is used for communication-based train control systems (CBTC) [6]. Arpan Roy et al combined countermeasure nodes and basic attack tree to form an attack countermeasure tree (ACT) for comprehensive analysis of attack and mitigation events [7]. And ACT is used to analyze active defense of train control system [8]. Besides, in the field of rail transit information security analysis, Fu chunchuan[9] evaluated the information security risk of train control center based on the component model based on security attributes; Kuang Xiangqi[10] and Ma Yangyang[11] evaluated the information security risk of CBTC system using Bayesian attack graph model; Wang Hongwei[12] described the robustness and recovery ability of CBTC system under the information security attack through elasticity.

The Bayesian attack graph has extensive research and application in the field of Internet and industrial control system risk analysis. It combines the attack graph and the uncertainty value of the state transition by Bayesian theory. Poolsappasit et al use Bayesian attack graph to evaluate the risk by calculating the expected profit and loss value of each attribute in the Bayesian attack graph [13]. Chen Xiaojun et al considered the uncertainty of monitoring results, and used the probabilistic attack graph for complex internal attacks to infer the attack intention and calculate the possible highest probability attack path [14]. Gao Ni et al. proposed a dynamic evaluation model based on Bayesian attack graph, and the methods of static and dynamic risk assessment based on Bayesian attack graph posterior probability [15]. Zhang Qi et al. analyzed industrial control systems by multimodel-based incident protection and risk assessment through the modeling and analysis of the attack multiple levels [16].

The Bayesian attack graph of Internet considers the host or network status in the information security domain. The target state node is the different levels of authority of the host or network communication interruption. However, malicious attacks against railway signaling systems are mostly aimed at impairing functional safety, such as malfunctions of signaling equipment, emergency braking of trains, and even accidents. At the same time, the railway signaling system equipment adopts the fault-safe design principles, and the key equipment are double-redundantly configured. With the particularity in modeling analysis, traditional Bayesian attack network is not entirely applicable to railway signaling safety data network. It is necessary to conduct targeted research on the railway signaling safety data network risk analysis method. Based on the fail-safe characteristics of the functional safety domain, this paper establishes an extend Bayesian attack graph model to realize risk analysis of railway signaling safety data network. This method quantifies the system risk value by the probability of the occurrence of the accident and the influence after the accident. At the same time, the method adds protection nodes on the basis of the Bayesian attack graph, so as to evaluate the impact of different protection nodes on the system risk, which can be used to help the formulation of the protection strategy.

The rest of the work is organized as follows. Section 2 presents a brief introduction of railway signaling safety data network and the security threats. In section 3, we proposes the extend Bayesian attack graph model. A simulation experiment is conducted as an example to demonstrate the effectiveness of the proposed method in section 4. Finally, we make the conclusion remarks in section 5.

2. Security Status of Railway Signaling Safety Data Network
The railway signaling safety data network is applied in the CTCS-2 and CTCS-3 train control systems. The overall structure of the network is shown in Figure 1.
Figure 1. Structure of Railway Signaling Safety Data Network

The railway signaling safety data network consists of network equipment, application equipment and network management system. Network equipment includes industrial Ethernet switch equipment and repeaters; application equipment includes Train Control System (TCC), Computer based Interlocking System (CBI), Temporary Speed Restriction Server (TSRS), Radio Block Center (RBC), Network Management Server, etc. The network equipment completes the communication between application devices; the network management system includes a network management server and many remote terminals.

To complete the train operation control and safety protection functions, the signaling safety data network communicates with the external system to obtain the data required for control and output the control signal. The communication boundary between the railway signaling safety data network and the external system includes the connection boundary between each maintenance terminal and the CSM network, the connection boundary between the CBI or TCC and the CTC system, the connection boundary between the TSRS or RBC and the CTC system, and the wireless connection of the GSM-R and balise. Among them, the GSM-R wireless network is working with the wireless common channel, which increases the information security vulnerability of the railway signaling safety data network [17]. At the same time, the railway signaling safety data network is a safety critical network while the connected CSM is not a safety critical system. The connection between them should be isolated or encrypted by security protection equipment.

The signaling safety data network is divided into two VLANs: “management” and “service”, which respectively correspond to network management data transmission and service data transmission of signaling safety data network control devices. Devices between two VLANs cannot directly communicate. Therefore, the network is divided into two areas. The area corresponding to the management VLAN includes network management server remote terminals, network management servers, and firewall devices. The service VLAN corresponding area includes equipment of TCC, CBI, TSRS, RBC, and respective maintenance terminals and operation computers.

The railway signal security data network belongs to the physical isolated network, but the access of mobile equipment, the access of maintenance personnel's computer and the attack of internal person may pose a threat to it. At the same time, some devices have common vulnerabilities with the use of windows, Linux and other operating systems, or general hardware and software devices. In this case, it is particularly important to evaluate the system security situation and provide reference for maintenance personnel and managers.

3. Extended Bayesian attack graph model
A large number of hosts and network components in the railway signaling safety data network are running complex and diverse application software and services. The network attack is the process of changing the
host or network to the target state by exploiting the vulnerability of the assets, so we can use the uncertainty of attack to describe cyber risks.

3.1. Extend Bayesian Attack Model

This paper considers the impact of security on functional safety while modeling the network attack behavior, and combines the impact of attack behavior and mitigation measures on network risk. Figure 2 shows an example of extended Bayesian attack graph. The extended Bayesian attack graph is represented by the five-tuple $<S, A, M, E, P>$:

1. $S = S_v \cup S_i$ is a collection of state nodes. $S_v$ is a set of state nodes indicating the cybersecurity attack phase, including the vulnerability information of each host and server resources, the rights obtained by the attacker, whether it is in a normal working state, and the network connection state. $S_i$ is a collection of state nodes that represent the functional safety phase, including the normal/failure status of each host, server, and network device in the functional safety domain and the functional safety consequences that may ultimately result.

2. $A = \{a = s_{pre} \rightarrow s_{post}|s_{pre}, s_{post} \in S\}$ is a collection of atomic attack behaviors. The atomic attack node has a value of $T$ indicating that the atomic attack occurred. $a$ indicates the transition of the state, $s_{pre}$ is the initial state node, and $s_{post}$ is the target state node.

3. $M$ is a set of protection measures against attack events in the network. The value $T$ of the protection node indicating that the protection is enabled and by default it is not enabled.

4. $E$ is a set of directed edges of connected nodes, indicating the logical sequence relationship between state nodes and attack nodes. The probability value of a directed edge describes the uncertainty of atomic attack.

5. $P$ is a set of conditional probability distribution tables for each node, indicating the logical relationship and probability with its parent node set.

3.2. Extended Bayesian Attack Graph Probability Calculation

Probability of successful atomic attack $P(s_j|a_i)$ describes the probability that an attack node $a_i$ successfully reaches the state node $s_j$, which is related to the difficulty of performing an atomic attack. The probability of successful atomic attack is defined according to the access vector (AV), the access complexity (AC), and the authentication (Au) in the CVSS [18]

$$P(s_j|a_i) = 2 \times P_{AV} \times P_{AC} \times P_{Au}$$ (1)
where $P_{AV}$, $P_{AC}$ and $P_{Au}$ are the scores of AV, AC, and Au in the CVSS 2.

There are a large number of attacks without vulnerabilities on the railway signaling safety data network. The probability of attack success is related to the attack type, attack scenario, attack tool, password complexity, protocol type, etc. So determine the value according to expert knowledge. Attack probability indicates the probability of launching an attack under a state node, depends on whether there is open vulnerability information, attack methods, and attack tools.

When the parent node of the state node has a corresponding protection node, the probability of successful attack is reduced. Use $P_m$ indicates the probability of the attack $a_i$ failure if the protection node $m_i$ is true, the probability of successful protection $P_m = 1 - P(a_i | m_i)$ . When an attack node has one corresponding protection node, the probability of attack success of the attack node $P(s_j | a_i, M_i) = P(s_j | a_i)(1 - P_m)$ . When a protection node corresponds to multiple attack nodes, the default multiple attack nodes do not affect each other. The logical attack is similar to the corresponding one, $P(s | a_i, m) = P(s | a_i) \times (1 - P_m)$ . When an attack node has multiple corresponding protection nodes, the attack success probability calculation formula is

$$P(s_j | a_i, M_i) = P(s_j | a_i) P(a_i | m_{i1}, m_{i2}, \ldots, m_{im})$$

$$= P(s_j | a_i)(1 - P_{m1})(1 - P_{m2}) \cdots (1 - P_{mn})$$

(2)

Considering the influence of the protection node, the conditional probability calculation corresponding to the different logical dependencies between the state node and the attack node is divided into the following two cases. When the logical dependency between the parent and child nodes is AND, the state node conditional probability table takes the value as shown in formula (3):

$$P(s_j | A_i, M_i) = \begin{cases}
0, & \exists a_i \in A_i, a_i = F; \\
\prod_{a_i \neq 0} P(s_j | a_i, M_i), & \text{Others}
\end{cases}$$

(3)

When the logical dependency between the parent and child nodes is OR, the state node conditional probability table takes the value as shown in formula (4):

$$P(s_j | A_i, M_i) = \begin{cases}
0, & \forall a_i \in A_i, a_i = F; \\
1 - \prod_{a_i \neq 0} [1 - P(s_j | a_i, M_i)], & \text{Others}
\end{cases}$$

(4)

The conditional probability of the attack node describes the possibility of launching an attack under the set of state nodes. According to the logical dependency between the parent and child nodes, the conditional probability of the attacking node is calculated in two cases. The corresponding conditional probability transfer table is shown in Table 3 and Table 4.
| $s_1$ | $s_2$ | ... | $s_n$ | $P(a_j = T)$ |
|------|------|-----|------|-------------|
| F    | F    | ... | F    | $1 - \prod_{a_j \neq 0} [1 - P(a_j | s_j)]$ |
| Others | | | | 1 |

### 4. Case studies

#### 4.1. Simulation environment

This simulation experiments mainly analyzes the network environment composed of network management system, signaling safety data network switch and equipment of TSRS. CBI, TCC and RBC are analyzed as parallel subsystems, and the extended Bayesian attack graph and analysis results of them are not included in this paper.

![Simulation network diagram](image)

*Figure 3. Simulation network diagram*

Equipment in Railway signaling safety data network adopts fail-safe design principles. It improves the safety of the system with the control mode changes and degradation mechanism, and has a favorable impact on security at the same time.

#### 4.2. Extend Bayesian Attack Graph

Firstly, we perform asset identification on the railway signaling safety data network, perform statistics on the types, quantities, and values of hardware and software devices in the network. Then we use vulnerability scanning and penetration test to conduct vulnerability analysis, and consider functional safety accidents caused by information security attacks. At the same time, we analyze implementable protection measures based on network topology and device characteristics.

Combining the network topology, network configuration and the relationship between vulnerability and network, the state, attack node set and edge set of Bayesian attack graph are established. According to the transition probability of each edge and the logical dependence of the parent-child nodes, the conditional probability table of each node is determined, and the establishment of the extended Bayesian attack graph is completed.

Combined with the simulation experiment environment in Figure 3, we analyzed the possible attack path and corresponding network state after different attacks with the TSRS failure as the attack goal. At the same time, we considered the possible protection nodes. The extend Bayesian attack graph is established and demonstrated in Fig. 4.
Figure 4. The Extend Bayesian Attack Graph for TSRS
4.3. Risk Analysis

According to the seriousness of the number of casualties caused by accidents, delay time, property loss, social impact, etc., the impact of a successful attack is analyzed. The accident is divided into train collision or derailment, emergency braking, degraded operation, equipment non-safety function failure, and confidential information leakage. Five categories of accident impact measures [19].

\[
I = \sum_{\alpha \in \mathcal{E}_i} \sum_{k=1}^{N_{\alpha}} i_k \times n_k + \sum_{\beta \in \mathcal{E}_q} \sum_{m=1}^{N_{\beta}} i_m \times n_m
\]  

(5)

where \( \mathcal{E}_i \) indicates train collision or derailment, emergency braking, deceleration operation event, \( n_k \) indicating the impact of single train when different events occur, \( i_k \) indicating the total number of vehicles in which the event occurred, such as the total number of deceleration vehicles including vehicles decelerating due to degraded operation and collision with the front train. The number of vehicles that are decelerated based on the fault-safety principle in the case of derailment and emergency braking. \( \mathcal{E}_q \) indicates the device data leakage and non-safety function failure events. The value of \( i_m \) indicates the impact of different events on a single device and \( n_m \) indicates the total number of devices that have this event.

The sum of the product of the prior probability of each accident node \( P(s) \) and the impact \( I(s) \) is used as the system risk.

\[
Risk = \sum_{s \in \mathcal{S}_f} P(s) \times I(s)
\]  

(6)

Set the conditional probability table according to the method in section 3. Since the vulnerability information in the extended Bayesian attack graph exists in the real network, the prior probability of setting each vulnerability state node is \( P(s_v) = 1 \). Combined with the probability calculation method in section 3, the prior probability of each node in the extended Bayesian attack graph is obtained. In this experiment, the TSRS fault is the most important functional safety event node and the functional safety impact of TSRS fault is far greater than other accident nodes. So the risk in this experiment is directly proportional to the TSRS fault event probability. In the following analysis, we mainly consider the probability of accidents in different states.

The attacks in the simulation environment are known attacks and can be detected by intrusion detection device. Therefore, in this experiment, if network management server vulnerability X1 exploit is observed, the prior probability and the posterior probability update result of state nodes in the extended Bayesian attack graph are shown in Fig. 5.

Figure 5. The prior and posterior probability of state nodes
The nodes s3, s5, s7, s11, s13, s17, s19, s20, s22 are the nodes corresponding to the vulnerability of the host and the network, and the probability value is always 1. The state node s4, s2, s1 is the precursor nodes of attack node a1, so the posterior probability value of them is 1 when a1 occurs and is detected. The state node s6, s8, s9, s10 and so on are the successor nodes of the attack event a1, and the posterior probabilities of them also increase. At the same time, since the probability of the initial state node s1 increases, the probability that the state node of the branch of s12 and s16 is true also increases accordingly.

Comparing the prior probability with the posterior probability, after detecting that the network management server vulnerability X1 is exploited, the probability of each state node in the Bayesian attack graph increases, and the overall risk of the experimental network increases. In line with the actual risk changes cognition, the extended Bayesian attack graph proposed in this paper can effectively evaluate the network risk.

4.4. Defense Analysis

The risk assessment model of the railway signaling safety data network analyzes the system risk from the perspective of the vulnerability utilization and privilege elevation of the attacker. Based on this, the defender role is introduced into the model. The railway signaling safety data network defender develops a series of protection strategies against the attacker's attack behavior by setting the corresponding protection node status in the extended Bayesian attack graph. The effect of protective measures on the probability of successful attack is calculated according to the method in section 3. Finally, through the definition of system risk, the system risk value changes under different protection scenarios are analyzed.

![Figure 6. Defense strategies analyze](image)

From the result in Fig. 6, we can see that the probability of system accident is reduced after the implementation of various defense measures. When the defense measures are implemented separately, the protective effect of M1 and M2 is better. According to the analysis of the prior probability of nodes in Figure 4, the cumulative risk of the attack path set where the defense measures M1 and M2 are located is relatively large, so when the defense resources are limited, the path set with relatively large risk should be protected first. The difference of protective effect between M1 and M2 is caused by the different protective ability of defense measures. Through the comparative analysis of the implementation results of schemes {M1, M5}, {M2, M5}, {M1, M2}, it can be seen that the combination of defense measures of different attack branches can obtain better protection results. In practice, attack paths should be covered as much as possible. At the same time, when the scheme {M1, M5} is implemented, the cost of M2 is the same as that of M2 alone, but the benefit is reduced. Therefore, the protection effect of the defense scheme is related to the current network state. When designing the protection scheme, it is necessary to integrate the interaction of various protection measures and the current network state to achieve appropriate protection.
5. Conclusion
In this paper, an extended Bayesian attack graph model are proposed for the railway signaling safety data network to analyze the network security risk. Based on the model, the attack path and the success probability of the railway signaling safety data network are analyzed. With the analysis result of probability and the impact on functional safety domain, the risk assessment under the different network attack and defense status is studied. Finally the feasibility and correctness are verified by simulation. The simulation results show that, based on the extended Bayesian attack graph, the security risk assessment method combined with defense nodes information are consistent with the actual situation. And it can provide reference for the formulation of defense strategy.

Currently, the extended Bayesian attack graph model of this paper is mostly constructed based on expert experience. It is acceptable for the railway signaling safety data network, which network structure will not change easily. However, the network topology and equipment composition of the other systems often change and this method is less scalable for this kind of system. In the future, a more general risk analysis model will be studied and more defense strategies assessment factors will be considered.

Acknowledgment
The authors would like to thank the experimental environment and the actual equipment of railway signaling system provided by Signal & Communication Research Institute, China Academy of Railway Sciences. This research was supported by Ministry of Science and Technology Key R & D Program (2018YFB1201500).

References
[1] R. Langner, “Stuxnet: Dissecting a cyberwarfare weapon,” IEEE Secur.Privacy, vol. 9, no. 3, pp. 49–51, May/Jun. 2011.
[2] Sun C C , Hahn A , Liu C C . Cyber security of a power grid: State-of-the-art[J]. International Journal of Electrical Power & Energy Systems, 2018, 99:45-56.
[3] LI Zhongwei, TONG Weiming, JIN Xianji. Construction of Cyber Security Defense Hierarchy and Cyber Security Testing System of Smart Grid: Thinking and Enlightenment for Network Attack Events to National Power Grid of Ukraine and Israel [J]. Automation of Electric Power Systems, 2016,40(08):147-151.
[4] Lü Yunfeng, Rising Network Security Report and Trend Prospect in 2018[J]. Journal of Information Security Research, 2019(3):186-191.
[5] B. Schnerier, "Attack trees-modeling security threats", Dr Dobb S Journal, vol. 12, no. 1999, pp. 21-29.
[6] Dong H, Wang H, Tang T. An attack tree-based approach for vulnerability assessment of communication-based train control systems[C]//2017 Chinese Automation Congress (CAC). IEEE, 2017: 6407-6412.
[7] Roy A , Kim D S , Trivedi K S . Attack countermeasure trees (ACT): towards unifying the constructs of attack and defense trees[J]. Security and Communication Networks, 2012, 5(8):0-0.
[8] Bao Zhengtang. Active Defense of Security Risk in Train Control System[D]. Beijing:Beijing Jiaotong University.2017.
[9] FU Chunchuan,WANG Xiaomin,ZHANG Wenfang, et al. A Component Security Attribute Model Driven Information Security Risk Assessment Approach for Train Control Center [J].Journal of the China Railway Society, 2017,39(8):77-84.
[10] Kuang Xiang-Qi. Research on Risk Assessment Method of Information Security in Communication-Based Train Control Systems[D], Beijing Jiaotong University, China, 2018.
[11] MA Yangyang, WANG Xuan, KUANG Xiangqi, et al. The Security Risk Assessment of the CBTC System Based on Bayesian Attack Graph [J]. Railway Standard Design, 2019, 63(8):155-161.
[12] Hongwei Wang, Minming Ni, Shigen Gao. et al. A Resilience-based Security Assessment
Approach for Railway Signalling Systems[C]// 2018 37th Chinese Control Conference (CCC).

[13] Poolsappasit N, Dewri R, Ray I. Dynamic security risk management using bayesian attack graphs[J]. IEEE Transactions on Dependable and Secure Computing, 2011, 9(1): 61-74.

[14] CHEN Xiaojun, FANG BinXing, TAN Qingfeng, et al. In-ferring attack intent of malicious insider based on proba-bilistic attack graph model[J]. Chinese Journal of Computers, 2014, 37(1): 62-72.

[15] GAO Ni, GAO Ling, HE Yiyue, et al. Dynamic Security Risk Assessment Model Based on Bayesian Attack Graph[J]. Journal of Sichuan University(Engineering Science Edition), 2016, 48(1): 111-118.

[16] Zhang Q, Zhou C, Xiong N, et al. Multimodel-based incident prediction and risk assessment in dynamic cybersecurity protection for industrial control systems[J]. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 2015, 46(10): 1429-1444.

[17] Xu W. High-speed railway signal system information security analysis and countermeasures[C]//2017 Chinese Automation Congress (CAC). IEEE, 2017: 4862-4866.

[18] SCHIFFMAN M. Common vulnerability scoring system (CVSS)[EB/OL]. [2019-06-03]. https://www.first.org/cvss/user-guide.

[19] Bloomfield, R., Bendele, M., Bishop, P., et al. The risk assessment of ERTMS-based railway systems from a cyber security perspective: Methodology and lessons learned[C]// International Conference on Reliability, Safety, and Security of Railway Systems. Springer. 2016: 3-19.