

Postprint of Intrusion Path Prediction Based on Incomplete Information Multi-stage Game

Authors: Yang Junnan, Zhang Hongqi, Zhang Chuanfu, Yang Chao

Date: 2018-05-20T00:00:00+00:00

Abstract

As the intrusion progresses, the information available to the intruder gradually increases, and based on this new information, the intruder can identify superior intrusion paths and make corresponding adjustments. To enable the defender to accurately predict intrusion paths, we first establish a dynamic defense graph based on hypergraph theory and propose a dynamic defense graph update method to predict the intruder's information updates; then, we construct an incomplete information multi-stage game model to predict the intruder's path adjustments at different stages; finally, we design a game-based dynamic defense graph path prediction algorithm to forecast the complete intrusion path. Experiments present typical examples of intrusion path prediction, and analysis of the results demonstrates the rationality and accuracy of the model.

Full Text

Preamble

Title: Intrusion Path Prediction Based on Incomplete Information Multi-Stage Game

Authors: Yang Junnan^{1,2}, Zhang Hongqi^{1,2}, Zhang Chuanfu^{1,2}, Yang Chao^{1,2}

¹ Information Engineering University, Zhengzhou 450001, China;

² Henan Province Key Laboratory of Information Security, Zhengzhou 450001, China

Abstract: As an intrusion progresses, the information available to the intruder gradually increases, enabling the discovery of better intrusion paths and subsequent adjustments. To enable defenders to accurately predict intrusion paths, this paper first establishes a dynamic defense graph based on hypergraph theory and proposes an update method for the dynamic defense graph to predict the intruder's information updates. Then, it establishes an incomplete information multi-stage game model to predict the intruder's path adjustments at different

stages. Finally, it designs a game-based dynamic defense graph path prediction algorithm to forecast complete intrusion paths. The experiment presents a typical example of intrusion path prediction, and analysis of the results demonstrates the rationality and accuracy of the model.

Keywords: game theory; defense graph; incomplete information; multi-stage; path prediction

0 Introduction

Intruders and defenders exhibit information asymmetry, as neither side can fully understand the other, which significantly limits the applicability of complete information game models. To address this information limitation, some scholars have employed incomplete information game models. Game theory studies strategic confrontation and competitive decision-making among players, which aligns with the opposing objectives, strategic interdependence, and non-cooperative nature of the relationship between network intruders and defenders. Applying game theory to intrusion behavior prediction has become a research hotspot.

When using game theory for network security research, game information is a critical issue. Some scholars adopt complete information game models. For instance, Lye and Wing define a complete information static game model using network recovery time after intrusion as payoff to analyze network security. Lin et al. employ complete information dynamic game theory to convert attack graphs into network game trees for studying proactive defense technologies. However, in actual networks, intruders and defenders cannot fully understand each other's information, making complete information game models difficult to apply in practice.

To solve the information limitation problem, some scholars use incomplete information game models. For example, Zhang et al. establish an incomplete information static game model for vulnerability risk analysis. Although this approach uses an incomplete information game model, it only conducts a single static game when predicting intrusion behavior, assuming that intruders will not change their intrusion strategies during the intrusion process. In reality, intruders have limited information collection capabilities and cannot fully understand the target network before intrusion. They can only formulate intrusion strategies with relatively high payoffs based on available information. As the intrusion progresses, intruders gain deeper understanding of the target network, continuously discover higher-payoff intrusion paths, and adjust their strategies accordingly. Therefore, actual intrusions consist of multiple stages, with intruders possessing different information about the target network at different stages and adjusting strategies at each stage to obtain greater payoffs.

Incomplete information dynamic games consider factors of information updates

and strategy adjustments. For example, Zhang et al. establish an attack-defense signaling game model and design an optimal defense strategy selection algorithm, where intruders adjust their intrusion strategies after receiving defense signals released by defenders. However, this only updates information about defense signals that intruders possess, without updating vulnerability information about the target network. Since vulnerability information is a key factor for intruders to formulate intrusion strategies, the lack of consideration for vulnerability information can cause significant errors in intrusion path prediction.

Based on the above analysis, defenders need to solve two key problems to accurately predict intrusion paths: First, the information update problem. Defenders need to predict the vulnerability information that intruders master at different intrusion stages. Most current game models analyze from a third-party perspective, assuming that the third party knows all information about both attack and defense sides. However, defenders have limited information and need scientific methods to predict updates to the vulnerability information mastered by intruders. Second, the strategy adjustment problem. Whenever intruders obtain new information about the target network, they will adjust strategies to gain more payoffs. Defenders need to predict each adjustment by intruders to accurately forecast the final intrusion path.

To address these problems, this paper first establishes a dynamic defense graph model based on hypergraph theory and designs vulnerability update templates to build a vulnerability update library. The vulnerability update library, combined with the target environment, updates the dynamic defense graph, solving the problem of predicting intruders' information updates. On this basis, it establishes an incomplete information multi-stage game model to predict intruders' strategy adjustments at different stages, and designs a dynamic defense graph path prediction algorithm based on incomplete information multi-stage games.

1 Dynamic Defense Graph

1.1 Concept of Dynamic Defense Graph

Using graphs for network security analysis is an effective approach. Current graph models have two problems: First, most graph models only contain intrusion information, lacking relevant defense information, which is unfavorable for comprehensive network security analysis. To fully evaluate network security, Jiang et al. introduce defense strategy nodes into graphs to define defense graphs. However, defense graphs can only simply reflect the relationship between intrusion paths and defense strategies, not the specific relationship between vulnerabilities and corresponding defense measures. Jiang et al. use bipartite graphs to describe the relationship between system states and defense strategies, but this traditional graph theory method, while capable of representing multivariate relationships between nodes, causes inconvenience in research and processing such as data connectivity and data clustering due to heterogeneity between nodes.

Second, current graph models are static. During the intrusion process, intruders continuously deepen their understanding of the target system and adjust their intrusion strategies when discovering higher-payoff intrusion paths. This is a dynamic process that static graph models cannot effectively represent.

Based on the above analysis, this paper first introduces hypergraph theory to improve the defense graph. The hypergraph method proposed by Berge can well represent multivariate relationships between nodes while avoiding difficulties in data processing caused by node heterogeneity. The defense graph is defined as a two-layer model: the upper layer represents vulnerabilities that intruders can exploit, indicating possible intrusion paths; the lower layer represents defense measures corresponding to each vulnerability, described using hypergraphs. Then, vulnerability update relationships are proposed to define the dynamic defense graph, predicting updates to target system vulnerabilities mastered by intruders during the intrusion process.

Definition 1 (Dynamic Defense Graph). A dynamic defense graph is defined as $DDG = \{N, D, L, E\}$, where: - $N = \{n_1, n_2, \dots, n_n\}$ is the set of vulnerabilities - $D = \{d_1, d_2, \dots, d_m\}$ is the set of defense measures - L is the set of directed edges representing the exploitation relationships between vulnerabilities - E is the set of hyperedges representing the defense measures available for each vulnerability

$N, D, L,$ and E are dynamically changing sets, where $N = N + f(n)$ represents the updated relationship between vulnerabilities, and $f(n)$ represents new vulnerabilities added to the dynamic defense graph when intruders penetrate vulnerability n . When new vulnerabilities are added to N , sets $D, L,$ and E need to be updated accordingly. [Figure 1: see original paper] shows an example of a dynamic defense graph in a certain state.

1.2 Dynamic Defense Graph Update

The key to updating the dynamic defense graph is determining the update relationships between vulnerabilities—that is, when an intruder penetrates a vulnerability node, determining which new vulnerabilities will be added to the dynamic defense graph. Currently, there is no research on vulnerability update relationships. This paper proposes a method to predict vulnerability updates.

Based on Li et al.'s attack template model, this paper proposes a vulnerability update template model to construct a vulnerability update library. The AGML (attack graphs modeling language) is used to formally describe vulnerability update templates. The vulnerability update template is formally represented as a triple $VulnerabilityUpdate = \langle Vul, Pre, Eff \rangle$, where: - Vul is the vulnerability entity, described as $\langle VulID, OS, App \rangle$. $VulID$ is the unique identifier of the vulnerability, represented using standard CVE numbers. OS is operating system information, described as $\langle OS_name, OS_version \rangle$. App is application information, described as $\langle App_name, App_version \rangle$. - Pre is the prerequisite set for vulnerability discovery - Eff is the consequence set of vulnerability exploitation

Vulnerability detection has two forms: host-based, where the vulnerability detector directly logs into the target host to detect vulnerabilities after obtaining certain control rights; and network-based, where the vulnerability detector remotely detects vulnerabilities on the target host using the network. Therefore, both the prerequisite set and consequence set are divided into host and network parts: $Pre = Pre_host, Pre_net$, $Eff = Eff_host, Eff_net$, where Pre_host and Pre_net have an “OR” relationship—only one condition needs to be satisfied to detect the vulnerability—but the predicates within Pre_host and Pre_net have an “AND” relationship. The specific prerequisite and consequence sets are shown in .

Target environment construction formally describes information about hosts, networks, and vulnerabilities in the target network. This paper uses Sheyner’s target environment construction method to formally describe the target network environment. The target environment, combined with the vulnerability update library, can determine the update relationships between vulnerabilities $f(n)$. Using $f(n)$ to update the vulnerability set N , and then updating sets D , L , and E accordingly, completes the dynamic defense graph update. Since automatic construction of dynamic defense graphs is not the focus of this paper and is limited by space, its construction technology will not be detailed here but will be discussed in future research.

2 Incomplete Information Multi-Stage Game Model

Intruders are often sequentially rational, pursuing maximum payoff. However, due to information limitations, they can only find the best intrusion path under current conditions at the beginning of the intrusion, rather than the objectively best path. As understanding of the target system deepens during the intrusion process, intruders continuously adjust toward intrusion paths with higher payoffs. At different stages of intrusion, intruders master different information, and to maximize payoff at each stage, they adjust strategies at different stages.

With the development of network security technology, most defenders can discover all vulnerabilities in their networks (excluding 0-day vulnerabilities), and each vulnerability has one or more optional defense measures. However, due to cost considerations of capital and time, defenders do not deploy all optional defense measures in reality, and only a few vulnerabilities pose major threats to the system. Therefore, defenders conduct targeted defenses under limited costs.

The model makes the following assumptions: - **Assumption 1:** Intruders are greedy, always pursuing maximum return with minimum cost. - **Assumption 2:** Intruders are rational and will not launch attacks for already obtained privileges; their intrusions only develop in the direction of privilege escalation. - **Assumption 3:** Intruder type is private information, but defenders have prior judgment about the probability distribution of intruder types. This prior judgment, defender type, and defender strategy set are common knowledge to both

intrusion and defense sides. - **Assumption 4:** Defenders are intelligent and can discover all vulnerabilities in their network (excluding 0-day). - **Assumption 5:** Defenders pursue “moderate security,” seeking maximum security payoff under limited cost.

2.1 Model Establishment

Definition 2 (Incomplete Information Multi-Stage Game Model). The incomplete information multi-stage game model is defined as $IIMG = \{N; S; P; U\}$:
 - $N = \{n, n_d\}$ is the set of players, which can be individuals or groups with common interests. Most confrontations can be viewed as two-player games between intruder n and defender n_d .
 - $S = \{S, D\}$ is the strategy set, where:
 - $S = \{a_1, a_2, \dots, a_s\}$ represents the intruder’s strategy set after exploiting vulnerability n .
 - $D = \{d_1, d_2, \dots, d_D\}$ represents the defender’s strategy set after the intruder exploits vulnerability n .
 - $P = \{P_1, P_2, \dots, P_A\}$ represents the defender’s prior belief about intruder types, where $P = (p_1, p_2, \dots, p_A)$ represents the defender’s prior belief after the intruder exploits vulnerability n , with $p_1 + p_2 + \dots + p_A = 1$.
 - $U = \{U, U_d\}$ is the payoff function set, where:
 - $U = \{u_1(n, s), u_2(n, s), \dots, u_s(n, s)\}$ represents the intruder’s payoff after exploiting vulnerability n .
 - $U_d = \{u_{d1}(n, s), u_{d2}(n, s), \dots, u_{dD}(n, s)\}$ represents the defender’s payoff.

Definition 3 (Pure Strategy Nash Equilibrium). In game $IIMG = \{N; S; P; U\}$, the strategy pair (a_j, d_q) is a Nash equilibrium for intruder type j exploiting vulnerability n , then a_j^* and d_q^* are the optimal strategies for the intruder and defender respectively:
 $u(a_j, d_q) \geq u(a, d_q), a \in A, p = 1, 2, 3, \dots$
 $u_{d_q}(a_j, d_q) \geq u_{d_q}(a, d_q), d_q \in D, q = 1, 2, 3, \dots$

Definition 4 (Mixed Strategy Nash Equilibrium). In game $IIMG = \{N; S; P; U\}$, the intruder’s strategy probability distribution is $\sigma_a = (p_1, p_2, \dots, p_s)$, where $0 \leq p_i \leq 1$ and $\sum p_i = 1$. The defender’s strategy probability distribution is $\sigma_d = (q_1, q_2, \dots, q_D)$, where $0 \leq q_i \leq 1$ and $\sum q_i = 1$. The intruder’s expected payoff is:

$$U_a(\sigma_a, \sigma_d) = \sum_{i=1}^n \sum_{j=1}^s p_i q_j u_a(d_i, a_j)$$

The defender’s expected payoff is:

$$U_d(\sigma_a, \sigma_d) = \sum_{i=1}^n \sum_{j=1}^s p_i q_j u_d(d_i, a_j)$$

A strategy pair (σ_a, σ_d) is a Nash equilibrium if:
 $U_a(\sigma_a, \sigma_d) \geq U_a(\sigma_a, \sigma'_d), \sigma'_d \in \Sigma_d$
 $U_d(\sigma_a, \sigma_d) \geq U_d(\sigma'_a, \sigma_d), \sigma'_a \in \Sigma_a$

2.2.1 Nash Equilibrium Solution

Nash Equilibrium Existence Theorem (Nash, 1950): Every finite game has at least one Nash equilibrium (pure or mixed strategy). Each static game with incomplete information in this model is obviously a finite game, so each game must have a Nash equilibrium.

Based on the payoff functions, we obtain the payoff matrices between high-ability, medium-ability, and low-ability intruders and defenders. Inputting the payoff matrices into the game tool Gambit for equilibrium solving yields mixed strategy Nash equilibria.

Pure strategy Nash equilibrium is a special case of mixed strategy Nash equilibrium where one strategy has probability 1 and others have probability 0. Therefore, the solution results are expressed in mixed strategy form.

2.2.2 Related Probability Solution

According to the definition of Nash equilibrium, when the defender adopts the Nash equilibrium strategy, the intruder can only obtain maximum payoff by also adopting the Nash equilibrium strategy. Defenders can use this to calculate relevant probabilities and predict intrusion paths.

Transfer Probability t : The intruder selects intrusion strategies according to Nash equilibrium. The first vulnerability of each intrusion strategy is the node the intruder will choose to penetrate next. If different intrusion strategies have the same first vulnerability, the transfer probabilities need to be superimposed. After the intruder penetrates a node, the vulnerability set N is updated. If new vulnerabilities are added to N , both intrusion and defense strategies change, the Nash equilibrium is broken, and a new Nash equilibrium must be solved. The transfer probability calculation is shown in Equation (1). If no new vulnerabilities are added to N , the previous Nash equilibrium remains valid, and the previous equilibrium results can be used to continue calculating transfer probabilities, as shown in Equation (2).

Definition 5 (Transfer Probability). t represents the transfer probability from vulnerability n to vulnerability n .

Definition 6 (Exploitation Probability). p represents the probability that vulnerability n is exploited by the intruder.

An intrusion path consists of different vulnerabilities in a certain order. The intrusion path prediction problem can be transformed into solving the exploitation probability and transfer probability. When an intruder successfully penetrates a vulnerability, the dynamic defense graph updates, possibly adding new vulnerabilities and defense measures, causing intrusion and defense strategies to update and the Nash equilibrium to be broken. Both sides enter a new stage and must readjust their strategies through game play. The specific process of the IIMG model is shown in [Figure 2: see original paper].

3 Dynamic Defense Graph Path Prediction Algorithm Based on Incomplete Information Multi-Stage Game

The algorithm predicts intrusion paths on the dynamic defense graph through incomplete information multi-stage games. For narrative convenience, the intruder is added to the dynamic defense graph as the starting point, and the intrusion target is added as the target node. The target node does not participate in game analysis. If a vulnerability node is directly connected to the target node, when the intruder penetrates this node, the intrusion target is considered achieved, and the transfer probability from this vulnerability node to the target node is 1.

Algorithm 1 (Dynamic Defense Graph Path Prediction Based on Incomplete Information Multi-Stage Game)

Input: Dynamic defense graph $DDG = \{N, D, L, E\}$, initial prior probability $P = (p_1, p_2, p_3)$, vulnerability update library, target environment

Output: Transfer probability set T ; Exploitation probability set T

1. Initialize $k = 0$
2. Do:
 3. Construct intrusion strategy set A
 4. Construct defense strategy set D
 5. Use A and D to calculate payoffs for different types of intruders and defenders when selecting different strategies, then construct corresponding payoff matrices M_1, M_2, M_3
 6. Use game tool Gambit to solve for equilibrium, obtaining three mixed strategy Nash equilibria $\sigma_1^*, \sigma_2^*, \sigma_3^*$
 7. If vulnerability set N changes:
 8. Use Equation (1) to calculate transfer probability t_k from vulnerability node k to its child node r
 8. Else:
 10. Use Equation (2) to calculate transfer probability t_k from vulnerability node k to its child node r
 9. End if
 10. Use Breadth-First Search (BFS) to select the next vulnerability node and update k
 11. Update N
 12. While (intrusion target reached && remaining nodes exist)
3. If (intrusion target reached && no remaining nodes):
4. Algorithm ends
5. End if
6. While (true)

Algorithm 1 employs two measures to reduce computation: (a) Statistics from

numerous real intrusion events show that the length of effective intrusion paths implemented by real intruders is mostly within 10 steps; (b) Not every intrusion adds new vulnerability nodes to the dynamic defense graph. When the dynamic defense graph remains unchanged, the game is not replayed.

Let n be the number of vulnerability nodes in the defense graph when the algorithm ends, and m be the number of strategies for both sides at each node. The time complexity for generating or updating the defense graph is $O(n)$, the equilibrium solving time complexity is $O(n^2 \times m)$, and the path prediction algorithm time complexity is $O(n \times m)$.

4 Application Example and Analysis

4.1 Experimental Environment

This paper uses a typical scenario shown in [Figure 3: see original paper] for experiments. This scenario facilitates in-depth analysis of how information affects intruder path selection and is suitable for verifying the rationality and accuracy of the algorithm for path prediction. The intruder is located in the external network, separated from the target network by Firewall 1. The target network contains four hosts: user host, administrator host, shared file server, and database server. Firewall 2 isolates the two servers into a subnet. Firewall rule information is shown in and , host vulnerability information in , and defender strategies in .

Assume the intruder has root privilege on the local machine and aims to obtain root privilege on the database server. Since the administrator host has administrative privileges over the database server, obtaining root privilege on the administrator host is also considered achieving the intrusion target.

4.2 Path Prediction

Using Algorithm 1 for path prediction. The intruder is added as the initial node numbered s_1 , and the intrusion target as the target node numbered s_4 . The initial vulnerability set is $N = \{s_1, s_2, s_3, s_4, s_5, s_6, s_7, s_8\}$. The generated dynamic defense graph is shown in [Figure 4: see original paper], where the exploitation probability of s_1 is $t_1 = 1$. Using expert knowledge, the defender's initial prior probability for the intruder is determined as $P = (0.3, 0.4, 0.3)$.

Since there is no vulnerability change at node s_1 , there is no need to replay the game. Use Equation (2) to calculate transfer probability. According to the dynamic defense graph, the intruder's intrusion strategies are: - a : $s_1 \rightarrow s_2 \rightarrow s_3 \rightarrow s_4$ - a : $s_1 \rightarrow s_2 \rightarrow s_3 \rightarrow s_4$ - a : $s_1 \rightarrow s_2 \rightarrow s_3 \rightarrow s_4$

The defender's defense strategies are: - d : $\{v_1, v_2, v_3, v_4\}$ - d : $\{v_1, v_2, v_3\}$ - d : $\{v_1, v_2, v_3, v_4\}$

Using the quantification method from reference [6], the payoff matrices are obtained:

$$M_h = \begin{bmatrix} (12, 5) & (19, 7) & (8, 12) \\ (23, 31) & (14, 17) & (11, 6) \\ (3, 9) & (8, 13) & (13, 21) \end{bmatrix}$$

$$M_m = \begin{bmatrix} (13, 21) & (15, 23) & (6, 9) \\ (11, 23) & (15, 17) & (11, 9) \\ (2, 11) & (7, 15) & (13, 19) \end{bmatrix}$$

$$M_l = \begin{bmatrix} (8, 9) & (12, 19) & (7, 21) \\ (7, 12) & (11, 21) & (13, 14) \\ (1, 11) & (7, 17) & (14, 22) \end{bmatrix}$$

Inputting the payoff matrices into game tool Gambit for equilibrium solving yields mixed strategy Nash equilibria: - * = (0, 0.324, 0.676) - * = (0.75, 0.25, 0) - * = (0, 0, 1)

Using Equation (1) to calculate transfer probability:

$$t_{01} = 0.3 \times (0.324 + 0) + 0.4 \times (0.75 + 0.25) + 0.3 \times (0 + 0) = 0.497$$

$$t_{02} = 0.3 \times 0.676 + 0.4 \times 0 + 0.3 \times 1 = 0.503$$

Using BFS to select the next vulnerability node s and update k . The exploitation probability of node s is $p = t = 0.497$.

Using Equations (4)-(6) to calculate prior probabilities:

$$p_h^1 = \frac{1 \times 0.3 \times (0.324 + 0)}{0.497} = 0.195$$

$$p_m^1 = \frac{1 \times 0.4 \times (0.75 + 0.25)}{0.497} = 0.805$$

$$p_l^1 = \frac{1 \times 0.3 \times (0 + 0)}{0.497} = 0$$

Since there is no vulnerability change at node s , the Nash equilibrium is not broken. Use Equation (2) to calculate transfer probability:

$$t_{13} = \frac{0.3 \times 0 + 0.4 \times 0.75 + 0.3 \times 0}{0.604} = 0.497$$

$$t_{14} = \frac{0.3 \times 0.324 + 0.4 \times 0.25 + 0.3 \times 0}{0.396} = 0.497$$

Following BFS to select the next vulnerability node s and perform vulnerability update. According to the vulnerability update template and target environment, after the intruder exploits CVE-2016-3387 to intrude the administrator host, CVE-2016-2555 vulnerability in the database server will be discovered. At this

point, the intruder's vulnerability set needs to be updated to $N = \{s, s, s, s, s, s, s, s, s, s\}$, and corresponding intrusion and defense strategies change, breaking the Nash equilibrium and requiring strategy adjustment.

From the updated dynamic defense graph, the intruder's strategies are: - a : $s \rightarrow s \rightarrow s \rightarrow s$ - a : $s \rightarrow s \rightarrow s \rightarrow s$

The defender's strategies are: - d : $\{v, v, v, v, v\}$ - d : $\{v, v, v, v\}$

Using reference [6] for quantification and Nash equilibrium solving yields: - * = (0.1, 0.9) - * = (0.3, 0.7) - * = (0.2, 0.8)

Using Equation (1) to calculate transfer probability:

$$t_{23} = 0.403 \times 0.1 + 0 \times 0.3 + 0.597 \times 0.2 = 0.16$$

$$t_{26} = 0.403 \times 0.9 + 0 \times 0.7 + 0.597 \times 0.8 = 0.84$$

Continue calculations according to Algorithm 1 until the algorithm terminates. The final results are shown in [Figure 6: see original paper].

4.3 Results Analysis

The results show that the most likely intrusion path is $s \rightarrow s \rightarrow s \rightarrow s$ with probability 0.423, where node s has exploitation probability 0.503 and node s has exploitation probability 0.423, representing significant security risks. The probabilities of the two intrusion paths $s \rightarrow s \rightarrow s \rightarrow s$ and $s \rightarrow s \rightarrow s \rightarrow s$ are smaller at 0.3 and 0.197 respectively. However, since both paths pass through node s , s also has relatively high exploitation probability at 0.497. Defenders cannot ignore vulnerability node s when taking defense measures.

Using methods from references [3,20] for analysis: If the initial vulnerability set does not contain node s , since these methods do not perform vulnerability updates and only conduct one game, the final results would ignore node s . If the initial vulnerability set contains node s , the high payoff of the intrusion path $s \rightarrow s \rightarrow s \rightarrow s$ would bias the game results toward this path, causing the exploitation probability of node s to be lower than the actual probability. In practice, due to missing initial state information, intruders cannot directly and accurately determine the objectively best intrusion path, resulting in node s also having relatively high utilization. Through the above analysis, we can see that in this scenario, the proposed method is more realistic, reasonable, and accurate than methods in references [3,20].

To deeply analyze the relationship between information and intrusion path selection, this experiment selected a typical scenario. As the number of network nodes in the scenario increases, the fundamental impact of information on intruder path selection does not change essentially. Therefore, the rationality and accuracy of this algorithm for intrusion path prediction will not change when the number of nodes increases. Increasing network nodes will increase the number

of nodes in the defense graph, thereby increasing computation during defense graph generation. This paper uses the method from reference [21] to generate defense graphs, which has been experimentally verified to be applicable to large-scale networks. Therefore, this algorithm remains applicable when the number of network nodes increases.

5 Conclusion

To predict intrusion paths of intruders with limited information collection capabilities, we must first predict their understanding of the target network at different stages, and then predict their strategy adjustments at different stages. To address the first problem, this paper constructs a dynamic defense graph model based on hypergraph theory and proposes an update method for dynamic defense graphs. For the second problem, this paper establishes an incomplete information multi-stage game model on the dynamic defense graph to predict intruder decisions at different stages, and designs a dynamic defense graph path prediction algorithm based on incomplete information multi-stage games. Finally, a typical network example demonstrates the specific application of the proposed method in intrusion path prediction. Analysis of the results illustrates the rationality and accuracy of the method.

To further improve model accuracy, the next step requires refinement of strategy quantification methods.

References

- [1] Liang X, Xiao Y. Game theory for network security [J]. IEEE Communications Surveys & Tutorials, 2013, 15(1): 472-486.
- [2] Fallah M. A puzzle-based defense strategy against flooding attacks using game theory [J]. IEEE Trans on Dependable & Secure Computing, 2008, 7(1): 5-19.
- [3] 姜伟, 方滨兴, 田志宏, 等. 基于攻防随机博弈模型的防御策略选取研究 [J]. 计算机研究与发展, 2010, 47(10): 1714-1723.
- [4] Lye K W, Wing J M. Game strategies in network security [J]. International Journal of Information Security, 2005, 4(1-2): 71-86.
- [5] 林旺群, 王慧, 刘家红, 等. 基于非合作动态博弈的网络安全主动防御技术研究 [J]. 计算机研究与发展, 2011, 48(2): 306-316.
- [6] 张恒巍, 张健, 韩继红, 等. 基于博弈模型和风险矩阵的漏洞风险分析方法 [J]. 计算机工程与设计, 2016, 37(6): 1421-1427.
- [7] 张恒巍, 余定坤, 韩继红, 等. 基于攻防信号博弈模型的防御策略选取方法 [J]. 通信学报, 2016, 37(5): 51-61.

- [8] Sheyner O M. Scenario graphs and attack graphs [M]. 2004.
- [9] Rambo S I, Anka I M. Attack graph-based approach for enterprise networks security analysis [J]. International Journal of Advanced Trends in Computer Science & Engineering, 2016, 5(5): 16532-16538.
- [10] Kumar S, Negi A, Prasad K, et al. Evaluation of network risk using attack graph based security metrics [C]// Proc of IEEE International Conference on Dependable, Autonomic and Secure Computing, International Conference on Pervasive Intelligence and Computing, International Conference on Big Data Intelligence and Computing and Cyber Science and Technology Congress. 2016: 91-93.
- [11] 姜伟, 方滨兴, 田志宏, 等. 基于攻防博弈模型的网络安全测评和最优主动防御 [J]. 计算机学报, 2009, 32(4): 817-827.
- [12] Andersen J L, Flamm C, Merkle D, et al. Maximizing output and recognizing autocatalysis in chemical reaction networks is NP-complete [J]. Journal of Systems Chemistry, 2012, 3(1): 2012.
- [13] Berge C. Graphs and hypergraphs [M]. Amsterdam: North-Holland Publishing Company, 1973.
- [14] Berge C. Hypergraphs: combinatorics of finite sets [M]. 1984.
- [15] Li Wei. An approach to graph-based modeling of network exploitations [S.l.]: Mississippi State University, 2005.
- [16] 陈锋. 基于多目标攻击图的层次化网络安全风险评估方法研究 [D]. 长沙: 国防科学技术大学, 2009.
- [17] Common vulnerabilities and exposures [EB/OL]. [2017-7-10]. <http://cve.scap.org.cn>.
- [18] 叶云. 基于攻击图的网络安全风险计算研究 [D]. 长沙: 国防科学技术大学, 2012.
- [19] 张维迎. 博弈论与信息经济学 [M]. 上海: 格致出版社, 2012.
- [20] 张健, 王晋东, 张恒巍, 等. 基于节点博弈漏洞攻击图的网络风险分析方法 [J]. 计算机科学, 2014, 41(9): 169-173.
- [21] 叶云, 徐锡山, 齐治昌, 等. 大规模网络中攻击图自动构建算法研究 [J]. 计算机研究与发展, 2013, 50(10): 2133-2139.

Note: Figure translations are in progress. See original paper for figures.

Source: ChinaXiv – Machine translation. Verify with original.