Optimal Network Defense Strategy Selection Based on Markov Bayesian Game

Zengguang Wang[1], Yu Lu[1], Xi Li[1] and Wei Nie[2*]
[1] Shijiazhuang Campus of Army Engineering University
Shijiazhuang, Hebei 050003 - China
[2] Shenzhen University
Shenzhen, Guangdong 518060 - China
[e-mail: wzg6410@163.com]
*Corresponding author: Wei Nie

Received May 7, 2019; accepted May 26, 2019; published November 30, 2019

Abstract

The existing defense strategy selection methods based on game theory basically select the optimal defense strategy in the form of mixed strategy. However, it is hard for network managers to understand and implement the defense strategy in this way. To address this problem, we constructed the incomplete information stochastic game model for the dynamic analysis to predict multi-stage attack-defense process by combining Bayesian game theory and the Markov decision-making method. In addition, the payoffs are quantified from the impact value of attack-defense actions. Based on previous statements, we designed an optimal defense strategy selection method. The optimal defense strategy is selected, which regards defense effectiveness as the criterion. The proposed method is feasibly verified via a representative experiment. Compared to the classical strategy selection methods based on the game theory, the proposed method can select the optimal strategy of the multi-stage attack-defense process in the form of pure strategy, which has been proved more operable than the compared ones.

Keywords: Network security, Bayesian game, Markov decision-making, strategy selection, defense effectiveness

This research was supported by the National Natural Science Foundation of China [No. 61271152, No.51377170] and the Natural Youth Science Foundation of China [No. 61602505].

http://doi.org/10.3837/tiis.2019.11.020 **ISSN : 1976-7277**

1. Introduction

With the increasingly diversified means of network attack, the situation of network security is facing more and more serious challenges [1]. It is urgent to take effective defense to ensure network security. The most ideal network security defense system can defend against all attacks in the network. However, this "at all costs" defense idea can not be applicable to the realistic network security defense [2]. The essence of network attack and defense can be formulated as a game between attacker and defender, its main process is to make reasonable decisions under limited resources, which is accordance with the game theory [3].Therefore, the network security defense method based on game theory has gradually become a hot topic.

In recent years, some researches regarding network security defense based on game theory were done. By constructing the complete information game model, Li et al. [4] proposed a network security situation-aware method based on the game theory and Markov decision process, which determines the transformation of network security state by solving the Nash equilibrium. Zhang [5] proposed a network attack-defense game model based on the defense graph to describe the dynamic changes of network security state, which provides guidance for the selection of defense strategies. Afraa et al. [6] proposed a dynamic game theory framework, which selects the effective defense strategy via Nash equilibrium. By constructing the incomplete information game model, Liu et al. [7] proposed a performance evaluation model based on static Bayesian game to evaluate the performance of worm attack and defense strategies in confrontation situations. However, the information of defender is assumed to be known for the attacker and the impact of defense types on the attack-defense process is not considered. Zhang et al. [8] proposed a risk prediction model based on static Bayesian game to analyze the network security situation. The attack behaviors are predicted via the mixed strategy. However, the model is only applicable to specific attack and defense scenarios with risk attacker and cautious attacker, which make the model less practical. Yang et al. [9] designed a network attack prediction model based on Bayesian game to solve the problem of Advanced Persistent Threat (APT), which analyzes possible attack behaviors in the network via Nash equilibrium. However, the quantification method of attack and defense payoff is not given. Liu et al. [10] studied the optimal defense problem of dynamic targets based on the incomplete information dynamic game. The optimal defense strategy is selected via the analysis of perfect Bayesian equilibrium. However, the situation of multiple attacker types is not considered in the model. In order to analyze the changes of network security status in real time, differential game is introduced into the field of network security. Liang et al. [11] studied the optimal strategy selection based on the differential game model and analyzed the general laws of the network attack and defense process over time. Zhang et al. [12] proposed a differential game model and analyzed the changes in the security state of network system via the infectious disease dynamics. The optimal defense strategy is selected through the analysis of the dynamic continuous attack and defense process. The above research results provide guidance for the solution of network security issues. However, the selection of the optimal strategies in the above research results are given in the form of mixed strategy. In the realistic network security defense, it is hard for the network managers to understand the scheme, which leads to confusion about how to implement optimal defense.

The key of network security defense research based on game theory is the design of network attack-defense game model, which conforms the realistic situation of network operation [13]. The complete information game model requires that the information of both attackers and defenders are known to each other, but it is difficult to achieve in the realistic network [14].

The differential game model analyzes the process of real-time attack and defense confrontation through the number of nodes in the network security state. This requires an extremely large number of network nodes to satisfy the conditions that can be analyzed by differential equations. Therefore, this method is only applicable to ultra-large-scale complex networks, which limits its practicability [15]. The Bayesian game is a classic incomplete information game, which requires at least one participant's information to be unknown to other participants, but the participants can have an initial judgment on the probability of other participants' types [16]. This is basically consistent with the realistic situation of network operation. Therefore, we choose Bayesian game as the theoretical basis to solve the problem that it is difficult to understand and implement the optimal strategy in the form of mixed strategy in the existing strategy selection scheme based on game model. The actual network attack-defense confrontation consists of multiple attack-defense stages. To achieve optimal defense strategy selection in a multi-stage attack-defense scenario, we constructed the incomplete information stochastic game model by combining Bayesian game theory and the Markov decision-making method.

In this work, we propose a novel method that selects the optimal defense strategy based on the Markov Bayesian game. Specifically, the contributions of this paper are:

(1) The incomplete information stochastic game model on network security attack-defense is developed. Our model transforms the uncertainty of attack-defense preferences to the uncertainty of attack-defense players' types, which makes the proposed model more applicable to the realistic situation of network attack and defense.

(2) The quantitative method of attack-defense payoff is designed. The payoff is quantified from the perspective of the impact value of attack-defense actions, which makes the calculation of the payoff more accurate and practical.

(3) The optimal defense strategy selection algorithm is designed. The optimal defense strategy in each security state obtained by the defense effectiveness is the pure strategy rather than the mixed strategy, which has a better guiding significance for the network security defense.

The remainder of this paper is organized as follows: we start with the design of network attack-defense game model (*Section 2*). Second, the quantification of network attack-defense payoff is presented (*Section 3*). Third, a novel optimal defense strategy selection method is proposed based on defense effectiveness (*Section 4*). Then, the proposed method is compared to the classical ones (*Section 5*). Next, the feasibility of the proposed method is verified via a representative experiment (*Section 6*). Finally, we conclude this work by promising future direction (*Section 7*).

2. Network Attack-Defense Game Model

2.1 Basic assumptions

In the realistic network attack and defense process, each decision made by attacker or defender is to maximize the payoff. Both attacker and defender should not only consider the gain of the action but also the cost of the action. Due to the complexity of the network and the confidentiality of the information, the attacker and defender can not determine the opponent's strategy payoff in the process of confrontation. But the attacker and defender can analyze their opponents via historical statistics data and judge the type of opponent in the form of probability distribution. We can assign each player with a unique type related to its strategy

and payoff. The ultimate goal of both sides is to protect their own network system. Therefore, the attack-defense effect can be measured by the value of network system. Based on the above analysis of the realistic network attack and defense [8][17], the establishment of the attack-defense game model needs to satisfy three assumptions as follows:

Assumption 1. **Rational assumption**. Attacker and defender are rational decision-making subjects and will not take unprofitable actions. The obtained payoffs of attacker and defender depend on the adopted strategies.

Assumption 2. **Type assumption**. The uncertainty of strategy payoff can be converted to the uncertainty of type by using the Harsanyi transformation.

Assumption 3. **Payoff assumption**. The payoffs of attacker and defender can be quantified via the security value of the network system.

2.2 Game model

Attackers and defenders confront each other according to the network security. Both attacker and defender select appropriate attack-defense actions according to the attack-defense purpose in each network security state. The network security state changes according to attack-defense actions. Due to the complexity characteristics of networks, neither of them can determine the attack-defense preferences of the opponent in the confrontation process. Consequently, the problem of network attack-defense confrontation can be formulated as the incomplete information stochastic game model. In order to simplify the analysis, we only consider the case of two players in the game. The network attack-defense game model is represented by a 7-tuple NADG $=\{N,\Omega,T,S,P,\eta,U\}$.

(1) $N=\{N_A,N_D\}$ represents the set of game players. The participants are the two sides of network attack and defense, where N_A represents the network attacker, and N_D represents the network defender.

(2) $\Omega=\{\Omega_1,\Omega_2,...,\Omega_k\}$ represents the state set of stochastic game. Each game state represents the network security state. It is necessary to choose a specific network security state representation according to the actual situation of the network.

(3) $T=\{T_A,T_D\}$ represents the type set of game players, where $T_A=\{T_1^A,T_2^A,...,T_m^A\}$ represents the type set of the attacker, and $T_D=\{T_1^D,T_2^D,...,T_n^D\}$ represents the type set of the defender.

(4) $S=\{S_A,S_D\}$ represents the set of attack-defense strategies of game participants with complete action plans selected. For the attacker, $S_A=\{S_1^A,S_2^A,...,S_k^A\}$ represents the set of attack strategies, and $S_k^A=\{A_1(S_k^A),A_2(S_k^A),...,A_g(S_k^A)\}$ represents the set of attack actions of the attack strategy S_k^A. For the defender, $S_D=\{S_1^D,S_2^D,...,S_l^D\}$ represents the set of defense strategies, and $S_l^D=\{D_1(S_l^D),D_2(S_l^D),...,D_r(S_l^D)\}$ represents the set of defense actions of the defense strategy S_l^D.

(5) $P=\{P_A,P_D\}$ represents a set of prior beliefs of game players. $P_A=\{P_A(T_i^D,\Omega_k)\,|\,1\le i\le m,1\le k\le h\}$ represents prior beliefs of attacker. $P_D=\{P_D(T_j^A,\Omega_k)\,|\,1\le j\le n,1\le k\le h\}$ represent prior beliefs of defender.

(6) $\eta(\Omega_j \mid \Omega_i)$ represents the state transition function, which denotes the probability that the network will transform from security state Ω_i to security state Ω_j.

(7) $U = \{U_A, U_D\}$ represents a set of payoff functions for game players. $U_A(T_j^A, S_k^A, S_l^D, \Omega_h)$ represents the profit gained by the attacker, when the security state is Ω_h and the type of attacker is T_j^A, the attacker takes attack action S_k^A to attack the defender, and the defender uses the action S_l^D to resist. $U_D(T_i^D, S_k^A, S_l^D, \Omega_h)$ represents the payoff gained by defender, when the security state is Ω_h and the type of defender is T_i^D, the defender takes defense action S_l^D to defend the attacker's attack S_k^A.

In order to select the optimal defense strategy, we design an objective criteria function *H* to judge the merits of attack-defense strategies. The network attack and defense is a sequential decision problem. The common objective criterion functions for such problems include the average return criterion function and the discount expectation criterion function [18]. Due to the relationship of the game payoff and time, we choose the discount expectation criterion function as the objective criteria function in this paper. The discount factor ε is introduced to indicate the attack-defense preference of current and future payoffs. The objective criteria function *H* can be described as follows:

$$\begin{cases} H_A(T_j^A, S_k^A, S_l^D, \Omega_m) = U_A(T_j^A, S_k^A, S_l^D, \Omega_m) + \varepsilon \sum_{\Omega_n \in \Omega} \eta(\Omega_n \mid \Omega_m) H_A(T_j^A, S_k^A, S_l^D, \Omega_n) \\ H_D(T_i^D, S_k^A, S_l^D, \Omega_m) = U_D(T_i^D, S_k^A, S_l^D, \Omega_m) + \varepsilon \sum_{\Omega_n \in \Omega} \eta(\Omega_n \mid \Omega_m) H_D(T_i^D, S_k^A, S_l^D, \Omega_n) \end{cases} \quad (1)$$

3. Quantification of Network Attack-Defense Payoff

In the process of network attack-defense game, the selection of both rational side is based on payoff [19]. Therefore, the selection methods of network optimal defense strategy are affected by the quantification rationality of the attack-defense payoff. Based on the realistic situation of network attack and defense, the payoffs of attacker and defender can be quantified from the perspective of the impact value of attack-defense strategies.

3.1 Basis for the quantification of attack-defense payoff

Quantification of attack-defense payoff is the basis for optimal defense strategy selection. Some definitions are introduced to indicate the quantitative process of attack- defense payoff.

Definition 1. Network System Value. This is the value of the network resource. The system value can be reflected in the security characteristics (*i.e.*, confidentiality, integrity and availability) of network device. Network device value R consists of $R(C_1)$, $R(C_2)$ and $R(C_3)$, which are the values of device in terms of confidentiality, integrity and availability respectively.

Definition 2. Attack Impact Factor. This is the impact of attack actions on the value of network system. The attack impact factor W consists of $W(C_1)$, $W(C_2)$ and $W(C_3)$, which are the weighting factors that represents attack actions on the confidentiality, integrity and availability of network devices.

Definition 3. Attack Success Rate. This is the success probability of the selected attack action which can break through the defense and obtain information resources. Whether the attack action is successful or not is closely related to the attack detection ability and the defense ability. Thus, the attack success rate θ can be quantified by the probability of the attack detected λ and the probability of the successful defense β. The attack will fail only if the detection and the defense are both successful. Therefore, the attack success rate is $\theta = 1 - \lambda\beta$.

Definition 4. Attack Payoff. This is the payoff that the attacker can achieve by implementing an attack action. If the attack action fails, the attacker can obtain the relevant defense information of the defender in the process of attack, but it will leave attack record in the defense system. The defender will change the defense deployment according to the attack record to focus on the vulnerability in the network system. It is difficult for attacker to achieve successful attack which implement the attack action based on the previous defense information. Therefore, the attacker can only gain payoff when the attack action is successful. The attack payoff can be calculated by the damage to the network system value, which is caused by the successful attack actions.

Definition 5. Defense Payoff. This is the payoff that the defender can achieve by implementing the defense action. The ultimate goal of the defender is to protect the value of the network system. The defense payoff can be calculated by the value of the network system, which is successfully protected. Whether the defense action is successful or not, the defender can get payoff.

- If the defender successfully implemented defense action, it can successfully protect the network system and gain direct defense payoff. Direct defense payoff can be calculated via the value of successful protection of network systems.
- If the defense action fails, the defender can not directly protect the value of network system. However, in the process of defense, the defender can obtain the relevant attack information of the attacker. The probability of successful defense can be increased in this way and the defender can gain the indirect payoff. Indirect defense payoff can be calculated via both discount factor and network system value that can be protected by successful defense action.

3.2 Quantification of attack-defense payoff

Note that Ω_m represents the network security state of attack-defense confrontation. T_j^A represents the type of the attacker, and the attack action $A_h(S_k^A)$ of attack strategy S_k^A is used to attack the target. T_i^D represents the type of defender, and the defense action $D_q(S_g^D)$ of defense strategy S_g^D is used to defend the target network. λ_h represents the probability of attack detected by defender, β_{hq} represents the probability of successful defense.

If the attacker implements the attack action successfully, the expected value $E_{A_h}(C_x)$ in the security attribute C_x can be mathematically described as:

$$E_{A_h}(C_x) = (1 - \lambda_h \beta_{hq}) W_{A_h}(C_x) R(C_x) \tag{2}$$

If the attack action $A_h(S_k^A)$ occurs, the payoff of the attacker U_{A_h} can be mathematically formulated as:

$$U_{A_h} = \sum_{x=1}^{3}(1-\lambda_h\beta_{hq})W_{A_h}(C_x)R(C_x) - B(A_h) \tag{3}$$

where $W_{A_h}(C_x)$ represents the damage weighting factor of the attack action $A_h(S_k^A)$ to the network device security attribute C_x, $R(C_x)$ represents the value of the attacked network device in terms of security attributes, and $B(A_h)$ represents the cost of implementing the attack action $A_h(S_k^A)$.

The value of security attributes successfully protected by the defense action is equal to the value of security attributes damaged by successful attack action. If the defender successfully implement the defense action, the expected value $E_{D_{q_1}}(C_x)$ of the defender in the security attribute C_x can be mathematically described as:

$$E_{D_{q_1}}(C_x) = \lambda_h\beta_{hq}W_{A_h}(C_x)R(C_x) \tag{4}$$

If the defense fails, the expected value $E_{D_{q2}}(C_x)$ of the defender in the security attribute C_x can be formulated as follows:

$$E_{D_{q_2}}(C_x) = \mu_q(1-\lambda_h\beta_{hq})W_{A_h}(C_x)R(C_x) \tag{5}$$

If the defense action $D_q(S_g^D)$ is taken, the payoff of the defender can be mathematically described as:

$$U_{D_q} = \sum_{x=1}^{3}[\mu_q(1-\lambda_h\beta_{hq}) + \lambda_h\beta_{hq}]W_{A_h}(C_x)R(C_x) - B(D_q) \tag{6}$$

where μ_q represents discount factor when the defense is failed and $B(D_q)$ represents the cost of launching the defense action $D_q(S_g^D)$.

When multiple defense actions of the defense strategy have defense effect on an attack action of the attack strategy, the payoff of the attack action is the minimum value of the payoffs, the payoff of the defense action is the maximum value of the payoffs. If attacker and defender implement attack-defense confrontation by selecting strategies (S_k^A, S_g^D), the payoffs of both sides can be formulated as follows:

$$\begin{aligned} U_A(T_j^A, S_k^A, S_l^D, \Omega_m) &= \sum_{A_h \in S_k^A} \min_{D_q \in S_g^D} U_{A_h} \\ &= \sum_{A_h \in S_k^A} \min_{D_q \in S_g^D} (\sum_{x=1}^{3}(1-\lambda_h\beta_{hq})W_{A_h}(C_x)R(C_x) - B(A_h)) \end{aligned} \tag{7}$$

$$\begin{aligned} U_D(T_i^D, S_k^A, S_l^D, \Omega_m) &= \sum_{D_q \in S_g^D} \max_{A_h \in S_k^A} U_{D_q} \\ &= \sum_{D_q \in S_g^D} \max_{A_h \in S_k^A} (\sum_{y=1}^{3}[\mu_q(1-\lambda_h\beta_{hq}) + \lambda_h\beta_q]W_{A_h}(C_x)R(C_y) - B(D_q)) \end{aligned} \tag{8}$$

To make it clear for readers to understand the computational complexity of payoffs in complex attack and defense scenarios, we take DDoS attack as an example to quantify the attack-defense payoff.

Example: Suppose that in the network attack and defense scenario, the attacker selects the DDoS to attack the web server. The attack strategy DDoS consists of attack actions, including network scanning, code injection and exception service request. The attacker scans the vulnerabilities in the target network through network scanning. The control privileges of the hosts are obtained by code injection. The attacker attacks the web server through exception service request. The defense strategy implemented by the defender consists of defense actions, including network sniffer and limit request frequency. According to the service and importance provided by the hosts and web server in the network, the attribute value of the hosts is set to (60, 60, 60) and the attribute value of the web server is set to (80, 90, 100). According to the historical data of the network attack and defense, the probability of the attack actions detected by the defender is (0.8, 0.7, 0.8). The probability of successful defense against the attack actions by network sniffer is (0.6, 0.7, 0.5). The probability of successful defense against the attack actions by limit request frequency is (0.7, 0.4, 0.8). The discount factor is set to (0.2, 0.2) according to the importance of attacker's information obtained by the defender. The attack impact factor of attack actions is set to (0.3, 0, 0), (0.2, 0.1, 0.4) and (0, 0, 0.8). The cost of attack-defense action is quantified according to the degree of difficulty in implementing attack-defense action. The cost of attack actions is set to (4, 5, 6) and the cost of defense actions is set to (5, 4). If attacker and defender select attack-defense actions (network scanning, network sniffer) to have attack-defense confrontation, the payoffs of attacker and defender can be mathematically described as:

$$\begin{aligned} U_{A_{11}} &= [(1-0.8\times0.6)\times0.3\times60]+[(1-0.8\times0.6)\times0\times60] \\ &\quad +[(1-0.8\times0.6)\times0\times60]-4 \\ &= 5.36 \end{aligned}$$

$$\begin{aligned} U_{D_{11}} &= \{[0.2\times(1-0.8\times0.7)+0.8\times0.7]\times0.3\times60\}+\{[0.2\times(1-0.8\times0.7)+0.8\times0.7]\times0\times60\} \\ &\quad +\{[0.2\times(1-0.8\times0.7)+0.8\times0.7]\times0\times60\}-5 \\ &= 5.512 \end{aligned}$$

If attacker and defender select attack-defense actions (network scanning, limit request frequency) to have attack-defense confrontation, the payoffs of attacker and defender can be mathematically described as:

$$\begin{aligned} U_{A_{12}} &= [(1-0.8\times0.7)\times0.3\times60]+[(1-0.8\times0.7)\times0\times60] \\ &\quad +[(1-0.8\times0.7)\times0\times60]-4 \\ &= 3.92 \end{aligned}$$

$$\begin{aligned} U_{D_{12}} &= \{[0.2\times(1-0.8\times0.6)+0.8\times0.6]\times0.3\times60\}+\{[0.2\times(1-0.8\times0.6)+0.8\times0.6]\times0\times60\} \\ &\quad +\{[0.2\times(1-0.8\times0.6)+0.8\times0.6]\times0\times60\}-4 \\ &= 7.664 \end{aligned}$$

The ultimate payoff of the attacker implementing network scanning is $U_{A_1}=\min\{U_{A_{11}},U_{A_{12}}\}$. In the same way, we can quantify the payoffs of other attack actions. The payoff of DDoS can be obtained by summing payoffs of the above attack actions.

4. Optimal Defense Strategy Selection

4.1 Game equilibrium analysis

In the network attack-defense game, attacker and defender expect to obtain the maximum payoff based on both the given type set and the prior belief set in each security state. Under the

guidance of the above principle, both attacker and defender will achieve a balance. The pure strategy can be considered as a specific case of the mixed strategy. Consequently, the mixed strategy can be used to analyze the equilibrium of the NADG model.

Definition 6. Mixed strategy. The attacker selects a pure attack strategy with the probability $f_k^A(T_j^A,\Omega_m)$. Based on the attack strategy selecting, we need to consider the constraints $(i.e., 0 \le f_k^A(T_j^A,\Omega_m) \le 1, \sum_{k=1}^{k_1} f_k^A(T_j^A,\Omega_m) = 1)$. $F_A(T_j^A,\Omega_m) = \{f_1^A(T_j^A,\Omega_m), f_2^A(T_j^A,\Omega_m),..., f_{k_1}^A(T_j^A,\Omega_m)\}$ is a mixed strategy of the attacker under the type T_j^A in security state Ω_m .Similarly, $F_D(T_i^D,\Omega_m) = \{f_1^D(T_i^D,\Omega_m), f_2^D(T_i^D,\Omega_m),..., f_{k_2}^D(T_i^D,\Omega_m)\}$ is also a mixed strategy of the defender under type T_i^D in security state Ω_m.

Definition 7. Mixed strategy Bayesian Nash equilibrium. According to Bayesian game, the mixed strategy $(F_A^*(T_j^A,\Omega_m), F_D^*(T_i^D,\Omega_m))$ can be considered as the Bayesian Nash Equilibrium in the security state Ω_m, if the mixed strategy satisfies the constraints as follows:

$$\begin{cases} \forall T_j^A \in T_A, F_A^*(T_j^A,\Omega_m) \in \arg\max \sum_{i=1}^{k} P_A(T_i^D,\Omega_m) H_A(T_j^A, F_A(T_j^A,\Omega_m), F_D^*(T_i^D,\Omega_m),\Omega_m),\Omega_m) \\ \forall T_i^D \in T_D, F_D^*(T_i^D,\Omega_m) \in \arg\max \sum_{j=1}^{n} P_D(T_j^A,\Omega_m) H_D(T_i^D, F_A^*(T_j^A,\Omega_m), F_D(T_i^D,\Omega_m),\Omega_m),\Omega_m) \end{cases} \tag{9}$$

Theorem 1. The mixed strategy Bayesian Nash equilibrium of the network attack-defense game $\mathrm{NADG} = \{N, \Omega, T, S, P, T, U\}$ exists.

Proof. Above all, the NADG consists of several independent and similar Bayesian games. Since each independent Bayesian game belongs to the finite game, the basic theorem of Bayesian game [20] shows that the mixed strategy Nash equilibrium exist. Moreover, according to the definition of network attack-defense game, its payoff function is a convex function based on the transition probability and payoff functions. According to the existence theorem of equilibrium strategy in finite stochastic game, we can prove that the mixed strategy Bayesian Nash equilibrium of NADG exists.

Chatterjee B [20], Cheng L et al. [21], Jiang W et al. [22], and Li C et al. [23] reported that the solution of the mixed strategy Bayesian Nash equilibrium can be formulated as the standard nonlinear programming problem. Therefore, the equilibrium strategy solution of NADG can be equally converted into the problem of solving the optimal value of nonlinear programming. The mixed strategy Bayesian Nash equilibrium can be obtained by solving the nonlinear programming model. According to the game theory, the mixed strategy $(F_A^*(T_j^A,\Omega_m), F_D^*(T_i^D,\Omega_m))$ is the selected strategy when both sides reach equilibrium state in security state Ω_m. Therefore, the mixed strategy $(F_A^*(T_j^A,\Omega_m), F_D^*(T_i^D,\Omega_m))$ can be solved by forming a quadratic programming which combines the objective function of A with that of D, under the condition that the constraints of objective functions are satisfied. The procedure of solving mixed strategy Bayesian Nash equilibrium can be mathematically described as:

$$\begin{cases} \arg\max z = \sum_{i=1}^{k} P_A(T_i^D, \Omega_m) H_A(T_j^A, F_A^*(T_j^A, \Omega_m), F_D^*(T_i^D, \Omega_m), \Omega_m), \Omega_m) + \\ \quad \sum_{j=1}^{n} P_D(T_j^A, \Omega_m) H_D(T_i^D, F_A^*(T_j^A, \Omega_m), F_D^*(T_i^D, \Omega_m), \Omega_m), \Omega_m) - v_1 - v_2; \\ \quad \sum_{i=1}^{k} P_A(T_i^D, \Omega_m) H_A(T_j^A, F_A(T_j^A, \Omega_m), F_D^*(T_i^D, \Omega_m), \Omega_m), \Omega_m) \le v_1; \\ \quad \sum_{j=1}^{n} P_D(T_j^A, \Omega_m) H_D(T_i^D, F_A^*(T_j^A, \Omega_m), F_D(T_i^D, \Omega_m), \Omega_m), \Omega_m) \le v_2; \\ \quad \forall \Omega_m \in \Omega, \sum_{i=1}^{m} P_A(T_i^D, \Omega_m) \in [0,1], \sum_{j=1}^{n} P_D(T_j^A, \Omega_m) \in [0,1]; \\ \quad \forall \Omega_m \in \Omega, \sum_{i=1}^{m} P_A(T_i^D, \Omega_m) = 1, \sum_{j=1}^{n} P_D(T_j^A, \Omega_m) = 1. \end{cases} \tag{10}$$

4.2 Optimal defense strategy selection method

The optimal defense strategy selection is a complicated process, which fully consider the payoffs of both sides with incomplete information. In the realistic network security defense, the defender can only choose one defense strategy at a time with limited resources. However, the traditional selection methods based on the game theory basically gives the optimal defense strategy in the form of mixed strategy, which is hard for network managers to select optimal defense strategy. To solve this problem, the concept of defense effectiveness is proposed that quantifies the effect of defense strategy against attack strategy when both sides reach the balanced state in each security state. The optimal defense strategy is selected, which takes the defense effectiveness as the criterion. Compared to the related works based on the game theory, the proposed method gives the optimal defense strategy in the form of pure strategy, which can perform better in the realistic network security defense.

Considering the possibility of selected defense strategy in each security state, the defense effectiveness is quantified based on the payoff function and the prior probability of defender. In the network attack-defense game scenario, $F_D(T_i^D, \Omega_m) = \{ f_1^D(T_i^D, \Omega_m), f_2^D(T_i^D, \Omega_m), ..., f_{k_2}^D(T_i^D, \Omega_m)\}$ represents the probability that the defender of type T_i^D selects the defense strategy in the equilibrium state, $P_D(T_j^A, \Omega_m)$ represents a prior belief of defender, and $H_D(T_i^D, S_k^A(T_j^A), S_l^D, \Omega_m)$ represents the payoff of the selected defense strategy. The defense effectiveness of defense strategy S_l^D in the security state Ω_m can be quantified as follows:

$$E(S_l^D, \Omega_m) = \sum_{j=1}^{n} P_D(T_j^A, \Omega_m) \sum_{k=1}^{k_2} H_D(T_i^D, S_k^A(T_j^A), S_l^D, \Omega_m) F_D^*(T_i^D, \Omega_m) \tag{11}$$

After calculating the defense effectiveness of all defense strategies in the security Ω_m, the effect of the defense strategy against the attack action in the equilibrium state can be obtained via the defense effectiveness, and the defense strategies can be sorted based on the defense effectiveness. With the permission of network resources, the defender can select the defense strategy by running the defense ranking strategies, which prioritizes the optimal defense

strategy with the maximum defense effectiveness. In this way, the optimal defense strategy can be selected in the form of pure strategy.

4. 3 Algorithm description

Based on the quantification of the attack-defense payoffs, the possibility of the defender selection strategy is obtained by solving the mixed strategy Bayesian Nash equilibrium. The defense effectiveness is quantified based on the prior belief and payoff of defender. The strategy with the maximum defense effectiveness is selected as the optimal defense strategy. Compared to the current classical algorithms that select the optimal strategy in the form of mixed strategy, the proposed model can select the optimal strategy in the form of pure strategy, which has a stronger operability. Therefore, network defense can be implemented via an optimal defense strategy before the network security threats occur, thus implementing active defense. The main steps of our novel defense strategy can be described as follows:

Optimal defense strategy selection algorithm based on attack defense game

Input：Network attack-defense game model NADG

Output：Optimal defense strategy

1) Initialize the model parameter $\text{NADG} = \{N, \Omega, T, S, P, \eta, U\}$;

2) Construct the security states $\{\Omega_1, \Omega_2, ..., \Omega_k\}$

3) Construct attack and defense type set T_A and T_D;

4) Construct attack and defense strategy set S_A and S_D;

5) Obtain attack and defense prior belief set P_A and P_D;

6) Obtain security state transition probability $\eta(\Omega_j \mid \Omega_i)$

7) **while** $\Omega_m \in \Omega$ **do**

{//Select the optimal defense strategy in each security state

8) **while** $S_k^A \in S_A, S_l^D \in S_D$ **do**

{//Calculate the payoffs of attack-defense strategies

9) Calculate the payoff of attack strategy $U_A(T_j^A, S_k^A, S_l^D, \Omega_m)$ by formula (7);

10) Calculate the payoff of attack strategy $U_D(T_i^D, S_k^A, S_l^D, \Omega_m)$ by formula (8);}

11) **end while**

12) Calculate the discount payoff of attack-defense strategies $\varepsilon \sum_{\Omega_n \in \Omega} \eta(\Omega_n \mid \Omega_m) H_A(T_j^A, S_k^A, S_l^D, \Omega_n)$ and $\varepsilon \sum_{\Omega_n \in \Omega} \eta(\Omega_n \mid \Omega_m) H_D(T_i^D, S_k^A, S_l^D, \Omega_n)$ by discount factor ε and formula (1);

13) The mixed strategy $(F_A^*(T_j^A, \Omega_m), F_D^*(T_i^D, \Omega_m))$ can be obtained by solving the formula (10);

14) The defense effectiveness $E(S_l^D, \Omega_m)$ is obtained by formula (11);

15) Return (argmax($E(S_l^D, \Omega_m)$));//output the optimal defense strategies of each security state}

16) **end while**

5. Comparison of Related Works

In this section, we compare the method proposed in this paper to the other classical methods. According to [24], the comparison can be evaluated in terms of game assumption, types of attacker and defender, payoff quantification and operability. Game assumption is an important indicator to measure the availability of game model. It is difficult to apply the model to the realistic network security defense, if the game assumption is too idealized. Attacker type refers to whether the attacker information is assumed to be known in the model design process. In a similar way, defender type refers to whether the defender information is assumed to be known in the model design process. Whether the types of attacker and defender can be extended to N is an important indicator to measure the generality of game model. The game model is only suitable for special cases, and the possibility of promotion is poor, if the types of attacker or defender can not be extended to N. The payoff quantification refers to whether the payoff of both sides are quantified according to the realistic network, and the steps are detailed and feasible. Operability refers to the strong practicability of the models given in the literature for the selection of the defense strategy. The results of comparison are shown in **Table 1**.

Table 1. Results of comparison

Game model	Game assumption	Attacker type	Defender type	Payoff quantification	Operability
MGM	Complete information	1	1	—	Poor
ADRG	Complete information	1	1	Simple	General
APM-SBG	Incomplete information	2	1	Simple	General
MS^2GM	Incomplete information	N	1	Detailed	General
Ours	Incomplete information	N	N	Detailed	Good

The MGM model mentioned in [4] and the ADRG model mentioned in [5] are based on the complete information without considering the impact of both types of attacker and defender on the network attack-defense game. The game models assume that the information of both attacker and defender is known, which is quite different from the realistic network environment. The MGM model does not give a quantitative method for the payoffs of attacker and defender. The payoffs are given by hypothesis, which leads to the inaccurate selection of the optimal defense strategy. The general method of equilibrium solution is not given and the guidance to realistic network security defense is weak. The operability of the game model is poor. The ADRG model quantifies the payoffs of attacker and defender from the perspective of atomic attack. However, in the process of quantification, the payoff of the attacker is ideally equivalent to that of the defender, thus resulting in a certain error in the calculation of the utility functions for game participants. The Nash equilibrium of this model is solved by value iteration method, which has high complexity and low efficiency. The operability of the game model is general.

The APM-SBG model mentioned in [8] and the MS^{2}GM model mentioned in [10] are based on the incomplete information without considering the impact of defender types on the network attack-defense game. However, in the design of the game model, only the incompleteness of the attacker information is considered, and the information of the defender is assumed to be to be complete information. The APM-SBG model quantifies the payoffs of attacker and defender from the perspective of gain and cost. However, it can only quantify the special attack-defense players and does not give a general method of payoff quantification. The number of attacker type in the game model is two, which is only applicable to specific attack-defense scenarios with risk attacker and cautious attacker. The versatility of the game model is limited and the operability of the game model is general. The MS^{2}GM model quantifies the payoffs of attacker and defender via attack-defense costs and system damage. However, in the process of quantification, the impact of attack success rate on the payoff is not considered. The optimal defense is not accurate, which is based on the payoffs of attacker and defender. The steps of the optimal defense strategy selection are not given in the model. The operability of the game model is general.

Compared to the related game models, the game model designed in this paper is based on incomplete information. The realistic situation of network attack and defense is fully considered in the designed model, the impact of both types of attacker and defender on the network attack-defense game is considered. Both types of attacker and defender can be extended to N, and the versatility of the game model is better. The payoffs of the strategy are quantified from the impact of attack-defense actions on the security value of network device. In addition, the steps of payoff quantification are detailed and feasible. The optimal defense strategy is selected according to defense effectiveness on the basis of Bayesian Nash equilibrium, and the detailed steps of the optimal network defense strategy selection are given. The optimal defense strategy obtained in this way is a pure strategy rather than a mixed strategy, which is easier for network managers to understand and implement defense strategy. Therefore, the operability of the game model is good.

6. Experiment and Analysis

6. 1 Experiment environment

In order to verify the effectiveness of the proposed optimal defense strategy selection method, the experimental network system was deployed according to the typical experimental network environment proposed by Li et al. [4] and Hu et al [17]. The structure of the experimental network system is shown in **Fig. 1**, where the network security protection devices are consist of firewall, virus detection system and intrusion detection system. The DMZ zone is consist of web server and host groups. The trusted zone is consist of database server group and file server group. The network system is scanned by Nessus, and the vulnerability information of the network system is shown in **Table 2**. The attacker is in the external network. The firewall prohibits external hosts from accessing the hosts and servers in the trusted zone. External hosts can only access web server and host group in the DMZ zone. The web server and the host group can access the database server group and the file server group according to the access control rules. The attacker can not directly access the database server at the initial moment. The ultimate goal of attacker is to gain root access to the database server through a series of attack actions.

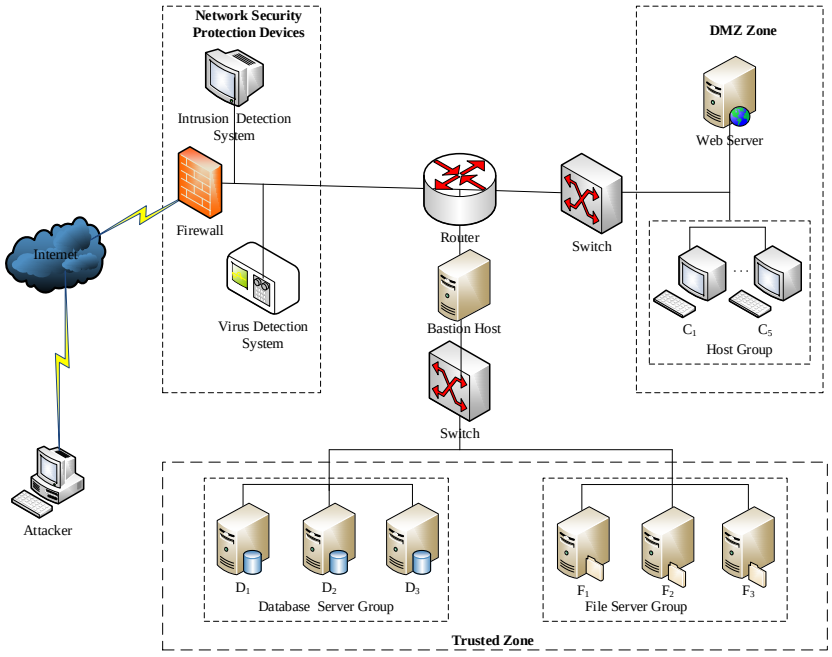


Fig. 1. Topology of the experimental network system

Table 2. Vulnerability information of experimental network system

Host	OS	Vulnerability-ID	Vulnerability Description	Result
Web Server	Linux 16.04	CVE-2015-1635	HTTP.sys	root
		CVE-2018-0098	Apache Chunked-Enc	root
		CVE-2018-4800	Write$home/.rhost	user
Host Group	Windows 10	CVE-2017-7269	Code injection	root
Bastion Host	Linux 16.04	CVE-2018-8517	Remote buffer overflow	root
		CVE-2018-4800	Write$home/.rhost	user
Database Server Group	Linux 16.04	CVE-2016-6662	Local buffer overflow	root
		CVE-2018-4326	Oracle TNS Listener	user
		CVE-2015-0104	Oracle Teloperation	root
File Server Group	Linux 16.04	CVE-2016-1143	Ftp.rhost	user
		CVE-2018-4800	Write$home/.rhost	user

The number of attack-defense types is an important measure of the generality of the game model. With the increase of attack-defense types, the branches of attack-defense confrontation will increase. However, the principle and the calculation process of optimal defense selection are similar. To illustrate the problem, the types of attackers can be divided into three categories (*i.e.*, T_A={ T_1^A, T_2^A, T_3^A }, which represent risky attacker, balanced attacker and conservative attacker)based on the analysis and summary of historical attack data. The specific characteristic of the risky attacker is that it is willing to adopt high-cost attack actions, thus achieving stronger attack capability. The specific characteristic of the conservative attacker is that it is willing to adopt low-cost attack actions, thus achieving its goals. The balanced attacker is between the risky attacker and the conservative attacker. The attack cost and attack success rate are both moderate. The types of defenders can be divided into two categories (*i.e.*, T_D={ T_1^D, T_2^D }, which represent senior defender and primary defender) according to the defense

effect and defense cost. The specific characteristic of the senior defender is that it is willing to adopt high-cost defense actions, thus achieving stronger defense effect. The specific characteristic of the primary defender is that it is willing to adopt low-cost defense actions, thus achieving its goals.

In order to clarify the difference in the actual network security state, the network security states are divided into seven types as shown in **Table 3** depending on host permissions. For the state transition between different security states, according to the literature [18] and [20], we assume that the state transition probabilities are fixed. Furthermore, the values of security state transition probabilities between security states and the prior beliefs are determined by the historical data and experts' experience, as shown in **Table 4** and **Table 5**.

Table 3. Network security states

Security state	Description of security state
Ω_0	Root privilege on Attacker Host
Ω_1	Root privilege on Web Server
Ω_2	Root privilege on Host
Ω_3	Root privilege on Bastion Host
Ω_4	Root privilege on File Server
Ω_5	User privilege on Database Server
Ω_6	Root privilege on Database Server

Table 4. Probabilities of security state transition

Security state transition	Transition probabilities	Security state transition	Transition probabilities	Security state transition	Transition probabilities
$\Omega_0 \to \Omega_0$	0.2	$\Omega_2 \to \Omega_2$	0.4	$\Omega_4 \to \Omega_4$	0.2
$\Omega_0 \to \Omega_1$	0.4	$\Omega_2 \to \Omega_3$	0.6	$\Omega_4 \to \Omega_5$	0.6
$\Omega_0 \to \Omega_2$	0.4	$\Omega_3 \to \Omega_3$	0.3	$\Omega_4 \to \Omega_6$	0.2
$\Omega_1 \to \Omega_1$	0.3	$\Omega_3 \to \Omega_4$	0.4	$\Omega_5 \to \Omega_5$	0.5
$\Omega_1 \to \Omega_3$	0.7	$\Omega_3 \to \Omega_5$	0.3	$\Omega_5 \to \Omega_6$	0.5

Table 5. Prior beliefs of attacker and defender

Security state	Prior beliefs of attacker		Prior beliefs of defender		
	$P_A(T_1^D, \Omega)$	$P_A(T_2^D, \Omega)$	$P_D(T_1^A, \Omega)$	$P_D(T_2^A, \Omega)$	$P_D(T_3^A, \Omega)$
Ω_0	0.6	0.4	0.3	0.5	0.2
Ω_1	0.5	0.5	0.2	0.1	0.7
Ω_2	0.5	0.5	0.1	0.6	0.3
Ω_3	0.3	0.7	0.4	0.2	0.4
Ω_4	0.3	0.7	0.1	0.6	0.3
Ω_5	0.2	0.8	0.6	0.2	0.2

The attack and defense strategies in the network consist of a series of attack and defense actions. Based on the National Information Security Vulnerability Database (CNNVD) [25], we analyzed the vulnerability information and constructed the attack and defense action set as shown in **Table 6** and **Table 7**. Then referring to the classification of network attack and defense conducted by MIT Lincoln Laboratory [26], we constructed the attack and defense strategy set as shown in **Table 8** and **Table 9**. We set the discount factor $\varepsilon = 0.5$ referring to the literature [27]. According to the service and importance provided by the network device, based on the quantitative method of attribute value proposed in the literature [28], the attribute value of the hosts is set to (150, 150, 150), and the attribute value of the web server is set to

(180, 210, 240), and the attribute value of the bastion host is set to (250, 220, 260)， and the attribute value of the file server is set to (260, 280, 290) and the attribute value of the database server is set to (380, 360, 380). The probability of the attack detection λ and the probability of the successful defense β can be determined by historical attack and defense data. On the basis of the above, the payoffs of attack and defense strategy are quantified.

Table 6. Attack action set

Serial number	Attack action description
a_1	Network scanning
a_2	Code injection
a_3	Send abnormal data
a_4	Remote buffer overflow
a_5	Steal or tamper with data
a_6	Steal account and crack it
a_7	Oracle TNS listener
a_8	Web-rhost attack
a_9	Local buffer overflow
a_{10}	SMTP sniffer
a_{11}	Homepage attack
a_{12}	Install Trojan

Table 7. Defense action set

Serial number	Defense action description
d_1	Limit packets from port
d_2	Reinstall listener program
d_3	Uninstall delete Trojan
d_4	Filtrate malicious packets
d_5	Reset Oracle access authority
d_6	Limit SYN/ICMP packets
d_7	Delete suspicious account
d_8	Repair database
d_9	Correct homepage
d_{10}	Address blacklist
d_{11}	Install SGD on web server
d_{12}	Alter data read-write rule

Table 8. Attack strategy set in different security state

Attacker types	Ω_0	Ω_1	Ω_2	Ω_3	Ω_4	Ω_5
T_1^A	$S_1^A\{a_1,a_8,a_{11}\}$	$S_7^A\{a_1,a_2,a_5\}$	$S_{13}^A\{a_1,a_3,a_4\}$	$S_{19}^A\{a_2,a_7,a_9\}$	$S_{25}^A\{a_6,a_7,a_9\}$	$S_{31}^A\{a_3,a_4,a_7\}$
	$S_2^A\{a_1,a_2,a_6\}$	$S_8^A\{a_2,a_3,a_4\}$	$S_{14}^A\{a_1,a_5,a_6\}$	$S_{20}^A\{a_1,a_3,a_9\}$	$S_{26}^A\{a_2,a_5,a_7\}$	$S_{32}^A\{a_2,a_7,a_{12}\}$
T_2^A	$S_3^A\{a_1,a_3\}$	$S_9^A\{a_1a_5\}$	$S_{15}^A\{a_2,a_6\}$	$S_{21}^A\{a_1,a_6\}$	$S_{27}^A\{a_3,a_4\}$	$S_{33}^A\{a_3,a_7\}$
	$S_4^A\{a_2,a_8\}$	$S_{10}^A\{a_1,a_2\}$	$S_{16}^A\{a_1,a_{12}\}$	$S_{22}^A\{a_2,a_9\}$	$S_{28}^A\{a_4,a_7\}$	$S_{34}^A\{a_4,a_{12}\}$
T_3^A	$S_5^A\{a_3,a_{11}\}$	$S_{11}^A\{a_3,a_9\}$	$S_{17}^A\{a_2,a_9\}$	$S_{23}^A\{a_2,a_3\}$	$S_{29}^A\{a_2,a_3\}$	$S_{35}^A\{a_2,a_{10}\}$
	$S_6^A\{a_{10},a_{11}\}$	$S_{12}^A\{a_1,a_4\}$	$S_{18}^A\{a_3,a_6\}$	$S_{24}^A\{a_2,a_7\}$	$S_{30}^A\{a_3,a_9\}$	$S_{36}^A\{a_4,a_9\}$

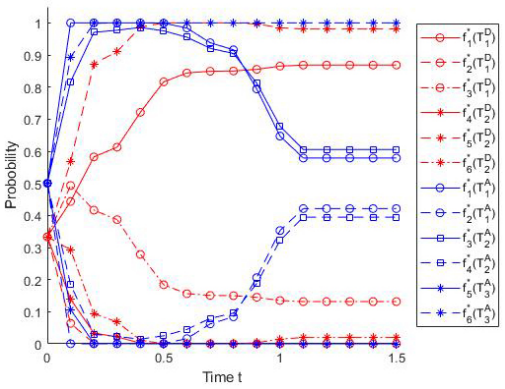
Table 9. Defense strategy set in different security state

Defender types	Ω_0	Ω_1	Ω_2	Ω_3	Ω_4	Ω_5
T_1^D	$S_1^D\{d_1,d_2,d_{10}\}$	$S_7^D\{d_1,d_4,d_7\}$	$S_{13}^D\{d_2,d_3,d_4\}$	$S_{19}^D\{d_4,d_5,d_6\}$	$S_{25}^D\{d_1,d_4,d_5\}$	$S_{31}^D\{d_1,d_5,d_8\}$
	$S_2^D\{d_3,d_4,d_{11}\}$	$S_8^D\{d_3,d_7,d_{10}\}$	$S_{14}^D\{d_1,d_6,d_7\}$	$S_{20}^D\{d_2,d_5,d_8\}$	$S_{26}^D\{d_2,d_5,d_6\}$	$S_{32}^D\{d_2,d_6,d_9\}$
	$S_3^D\{d_2,d_3,d_6\}$	$S_9^D\{d_1,d_2,d_4\}$	$S_{15}^D\{d_2,d_4,d_9\}$	$S_{21}^D\{d_2,d_3,d_4\}$	$S_{27}^D\{d_4,d_5,d_8\}$	$S_{33}^D\{d_4,d_6,d_8\}$
T_2^D	$S_4^D\{d_3,d_7\}$	$S_{10}^D\{d_1,d_3\}$	$S_{16}^D\{d_3,d_{12}\}$	$S_{22}^D\{d_5,d_6\}$	$S_{28}^D\{d_6,d_{12}\}$	$S_{34}^D\{d_5,d_{12}\}$
	$S_5^D\{d_1,d_9\}$	$S_{11}^D\{d_2,d_7\}$	$S_{17}^D\{d_1,d_3\}$	$S_{23}^D\{d_4,d_{12}\}$	$S_{29}^D\{d_8,d_{10}\}$	$S_{35}^D\{d_2,d_5\}$
	$S_6^D\{d_4,d_{10}\}$	$S_{12}^D\{d_3,d_4\}$	$S_{18}^D\{d_3,d_{10}\}$	$S_{24}^D\{d_5,d_{10}\}$	$S_{30}^D\{d_5,d_8\}$	$S_{36}^D\{d_8,d_{12}\}$

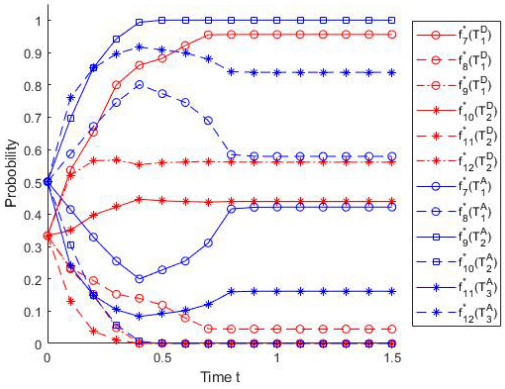
6. 2 Experimental analysis

The attacker implements multi-step attacks with the root authority of the database as the target, and the defender selects appropriate defense strategies to defend the target network. The attack and defense confrontation process can cause changes in the network security states. In this experiment, the attack and defense states of network are divided into seven types, and the attacker and defender rely on their different strategy sets (see **Table 8** and **9**) for confrontation. In the security state Ω_6, the attacker achieves the attack purpose and stops the attack-defense

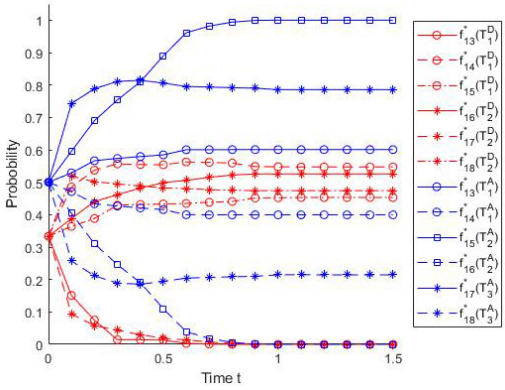
confrontation. The probability change process of attack and defense strategy in each security state is shown in **Fig. 2**. We observe that the probability of each strategy selected by attacker and defender gradually converges and finally stabilizes. The attacker and defender eventually reach equilibrium, and the mixed strategy Nash equilibrium is shown in **Table 10**.



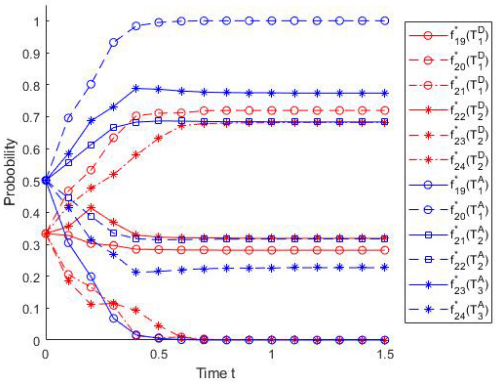
(a) Probability change of attack-defense strategy in security state Ω_0



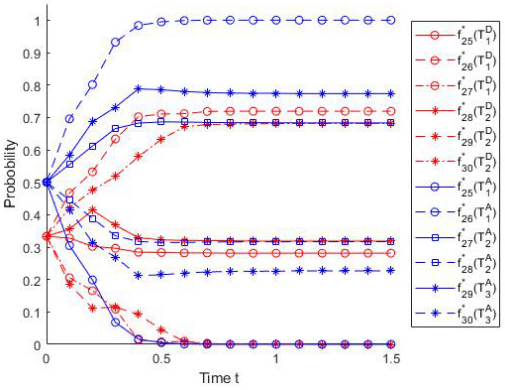
(b) Probability change of attack-defense strategy in security state Ω_1



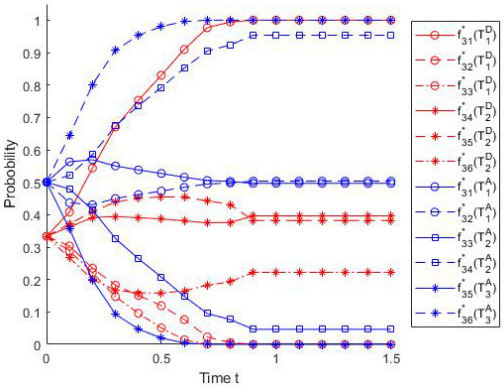
(c) Probability change of attack-defense strategy in security state Ω_2



(d) Probability change of attack-defense strategy in security state Ω_3



(e) Probability change of attack-defense strategy in security state Ω_4



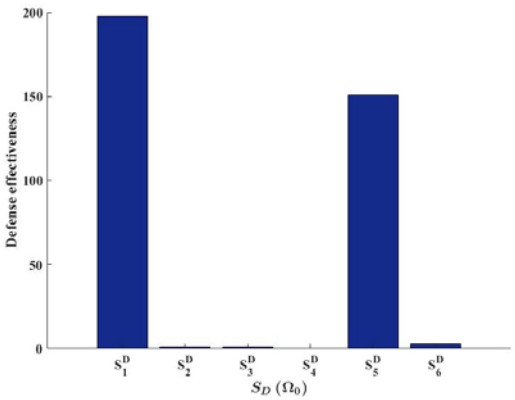
(f) Probability change of attack-defense strategy in security state Ω_5

Fig. 2. Probability change of attack-defense strategy

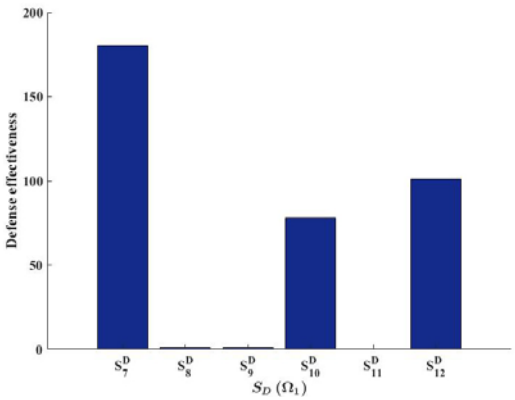
Table 10. Mixed strategy Nash equilibrium

Security state	$F_{\mathrm{A}}^{*}(T_1^{\mathrm{A}})$	$F_{\mathrm{A}}^{*}(T_2^{\mathrm{A}})$	$F_{\mathrm{A}}^{*}(T_3^{\mathrm{A}})$	$F_{\mathrm{D}}^{*}(T_1^{\mathrm{D}})$	$F_{\mathrm{D}}^{*}(T_2^{\mathrm{D}})$
Ω_0	[0.579,0.421]	[0.605,0.395]	[0,1]	[0.868,0,0.132]	[0,0.981,0.019]
Ω_1	[0.422,0.578]	[1,0]	[0.162,0.838]	[0.956,0,0.044]	[0.439,0,0.561]
Ω_2	[0.600,0.400]	[1,0]	[0.786,0.214]	[0,0.547,0.453]	[0.526,0,0.474]
Ω_3	[0,1]	[0.683,0.317]	[0.733,0.227]	[0.281,0.719,0]	[0.318,0,0.682]
Ω_4	[1,0]	[0.042,0.958]	[0.859,0.141]	[1,0,0]	[0.143,0.725,0.132]
Ω_5	[0.497,0.5033]	[0.046,0.954]	[0,1]	[1,0,0]	[0.396,0.382,0.222]

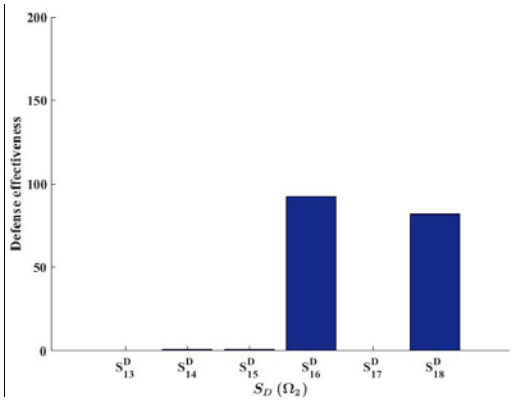
We observe that different types of attackers and defenders can select the attack-defense strategies with a certain probability, which makes both sides reach a balanced state in each security state. The traditional defense strategy selection methods take the mixed strategy of defender as the optimal defense strategy. However, it is hard for the network manager to select the optimal defense strategy in the form of mixed strategy. To solve this problem, we propose an optimal defense strategy selection method, which quantifies the defense effectiveness based on the Nash equilibrium. The optimal defense strategies in each security state are selected, which take the defense effectiveness as the criterion. The defense effectiveness of defense strategies in different security state are shown in **Fig. 3**.



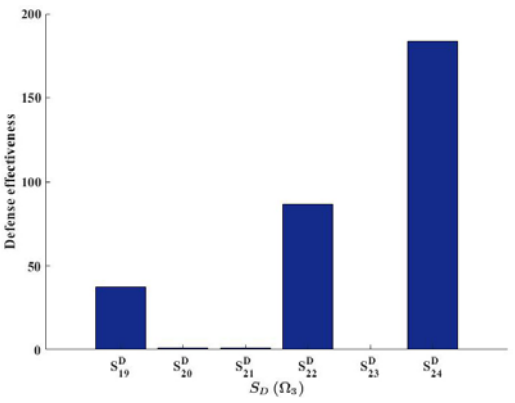
(a) Defense effectiveness of defense strategies in security state Ω_0



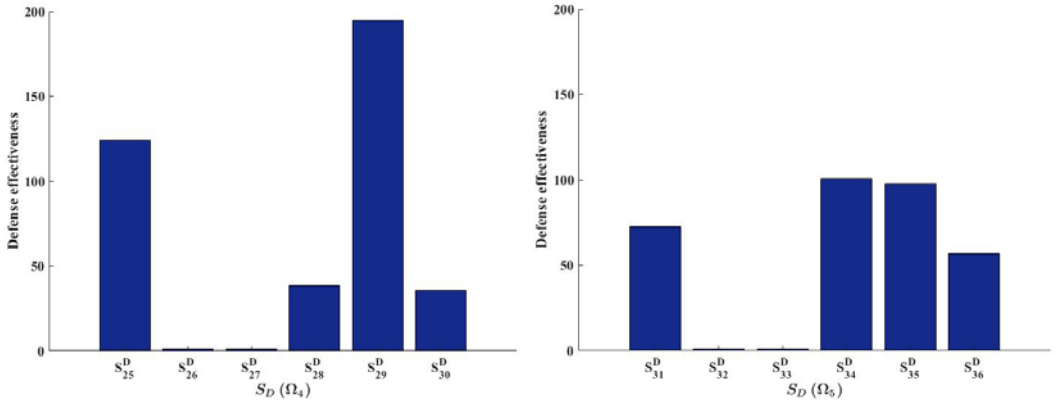
(b) Defense effectiveness of defense strategies in security state Ω_1



(c) Defense effectiveness of defense strategies in security state Ω_2



(d) Defense effectiveness of defense strategies in security state Ω_3



(e) Defense effectiveness of defense strategies in security state Ω_4

(f) Defense effectiveness of defense strategies in security state Ω_5

Fig. 3. Defense effectiveness of defense strategies

The network security defense under multi-step attacks can be realized by selecting the optimal defense strategy in different security states. We take attack path $\Omega_0 \to \Omega_1 \to \Omega_3 \to \Omega_5 \to \Omega_6$ as an example for analysis. In the security state Ω_0, the result$\{ E(S_1^D), E(S_5^D), E(S_6^D), E(S_3^D), E(S_2^D), E(S_4^D) \}$ can be obtained in a descending manner based on the defense effectiveness as shown in **Fig. 3(a)**. We observe that the defense strategy S_1^D has the greatest defense effectiveness, that is, S_1^D represents the optimal defense strategy in the security state Ω_0. When network resources are limited, the defender should preferably select the defense strategy S_1^D to protect the network in the security state Ω_0, thus achieving the optimal defense effect. In the same way, the optimal defense strategies in security state Ω_1, Ω_3 and Ω_5 are S_7^D, S_{24}^D and S_{34}^D. Similarly, defense strategies against other attack paths can be obtained. Thus, the defense strategy selection method proposed in this paper can achieve the optimal defense strategy selection under multi-step attacks. Compared to the optimal defense strategy given in the form of mixed strategy, the novel method proposed in this paper selects the optimal strategy in the form of the pure strategy, which is more applicable to the network security protection.

7. Conclusion

In this paper, we combined the Bayesian game with the Markov decision method and constructed an incomplete information stochastic game model to analyze the multi-stage attack-defense process. To solve the problem that it is hard for network managers to understand and implement the optimal defense strategy in the form of mixed strategy, the proposed method takes the defense effectiveness as the basis of the optimal strategy selection and the optimal defense strategy can be selected in the form of pure strategy. In the process of designing the game model and quantifying the payoffs, the realistic situation of network attack and defense is fully considered. Furthermore, we proposed the equilibrium solution method based on the nonlinear programming and designed the optimal defense strategy selection algorithm. The advantage of the proposed method is verified by comparing to the other classical optimal strategy models. The effectiveness and feasibility of the proposed method is

demonstrated via a representative experiment. In our future work, it will be important to explore the probability of security state transition through reinforcement learning to increase the calculation accuracy.

References

[1] Moura J, Hutchison D, "Game theory for multi-access edge computing: survey, use case, and future trends," *IEEE Communications Surveys and Tutorials*, vol.21, no.1, pp.260-288, 2019. Article (CrossRef Link)

[2] Kang B G, Seo K M and Kim T G, "Communication analysis of network-centric warefare via transformation of system of systems model into integrated system model using neural network," *Complexity*, vol.2018, pp.1-16, 2018. Article (CrossRef Link)

[3] Do C, Tran N H and Hong C, "Game theory for cyber security and privacy," *ACM Computing Surveys*, vol.50, no.2, pp.1-37, 2017. Article (CrossRef Link)

[4] Li X, Lu Y and Liu S, "Network security situation assessment method based on Markov game model," *Ksii Transactions on Internet and Information Systems*, vol.12, no.5, pp.2414-2428, 2018. Article (CrossRef Link)

[5] Zhang N B, "Defensive strategy selection based on attack-defense game model in network security," *International Journal of Performability Engineering*, vol.14, no.11, pp.2633-2642, 2018. Article (CrossRef Link)

[6] Afraa A, Mainak C and Cliff C Z, "A game theoretic approach to model cyber attack and defense strategies," in *Proc. of IEEE International Conference on Communications*, pp.246-253, 2018. Article (CrossRef Link)

[7] Liu Y L, Fang D G and Wu L H, "Performance evaluation of worm attack and defense strategies based on static Bayesian game," *Journal of Software*, vol.23, no.3, pp.712-723, 2012. Article (CrossRef Link)

[8] Zhang J, Wang J D and Zhang H W, "Application of static Bayesian game in information system risk analysis," *Computer Engineering and Applications*, vol.51, no.11, pp.76-82, 2015.

[9] Yang H P, "Method for behavior-prediction of APT attack based on dynamic Bayesian game," in *Proc. of IEEE International Conference on Cloud Computing and Big Data Analysis*, pp.177-182, 2016. Article (CrossRef Link)

[10] Liu J, Zhang H Q and Liu Y, "Research on optimal selection of moving target defense policy based on dynamic game with incomplete information," *Acta Electronica Sinica*, vol.46, no.1, pp.82-89, 2018. Article (CrossRef Link)

[11] Liang L, Deng F and Peng Z H, "A differential game for cooperative target defense," *Automatica*, vol. 102, pp. 58-71, 2019. Article (CrossRef Link)

[12] Zhang H W, Li T and Huang S R, "Network defense decision-making method based on attack-defense differential game," *Acta Electronica Sinica*, vol.46, no.6, pp.1428-1435, 2018. Article (CrossRef Link)

[13] Ni Z, Li Q M, Liu G, "Game-model-based network security risk control," *Computer*, vol.51, no.4, pp.28-38, 2018. Article (CrossRef Link)

[14] Lee S, Kim S and Choi K B, "Game theory-based security vulnerability quantification for social internet of things," *Future Generation Computer Systems-the International of Science*, vol.82, pp.752-760, 2018. Article (CrossRef Link)

[15] Miao L, Li S and Wang Z Q, "Optimal dissemination strategy of security patch based on differential game in social network," *Wireless Personal Communications*, vol.98, no.1, pp.237-249, 2018. Article (CrossRef Link)

[16] Abuzainab N and Saad W, "A graphical Bayesian game for secure sensor activation in internet of battlefield things," *Ad Hoc Networks*, vol.85, pp.103-109, 2019. Article (CrossRef Link)

[17] Hu H, Liu Y L and Zhang H Q, "Optimal network defense strategy selection based on incomplete information evolutionary game," *IEEE Access*, vol.6, pp.29806-29821, 2018. Article (CrossRef Link)

[18] Huang S R, Zhang H W and Wang J D, "Markov differential game for network defense decision-making method," *IEEE Access*, vol.6, pp.39612-39634, 2018. Article (CrossRef Link)
[19] Wei J X, Zhang R, and Liu J Y, "Defense strategy of network security based on dynamic classification," *Ksii Transactions on Internet and Information Systems*, vol.9, no.12, pp.5116-5134, 2015. Article (CrossRef Link)
[20] Chatterjee B, "An optimization formulation to compute Nash equilibrium in finite games," in *Proc. of International Conference on Methods and Models in Computer Science*, pp.1-5, 2009. Article (CrossRef Link)
[21] Lei C, Zhang H Q and Wang L M, "Incomplete Information Markov game theoretic approach to strategy generation for moving target defense," *Computer Communications*, vol.116, pp.184-199, 2018. Article (CrossRef Link)
[22] Jiang W, Fang B X and Tian Z H, "Research on defense strategies based on attack-defense stochastic game model," *Journal of Computer Research and Development*, vol.47, no.10, pp.1714-1723, 2010.
[23] Li C B, Cao H Y and Du M K, "A novel method to compute Nash equilibrium in non-cooperative n-person games based on differential evolutionary algorithm," *Intelligent Decision Technologies*, vol.8, no.3, pp.207-213, 2014. Article (CrossRef Link)
[24] Zhang H W, Yu D K and Hang J H, "Defense policies selection method based on attack-defense signaling game model," *Journal on Communications*, vol.37, no.5, pp.51-61, 2016.
[25] "China National Vulnerability Database of Information Security," Aug.23, 2017. [Online]. Available: http://www.cnnvd.org.cn
[26] Aviad R, "Settling the complexity of computing approximate two-player Nash equilibria," in *Proc. of the 57th IEEE Annual Symposium on Foundations of Computer Science (FOCS)*, pp.258-265, 2016. Article (CrossRef Link)
[27] Huang J M, Zhang H W and Wang J D, "Markov evolutionary games for network defense strategy selection," *IEEE Access*, vol.5, pp.19505-19516, 2017. Article (CrossRef Link)
[28] Fu Y, Chen Y Q and Wu X P, "Network attack-defense strategies selection based on stochastic game model," *Journal of Beijing University of Posts and Telecommunications*, vol.37, pp.35-39, 2014.

Zengguang Wang received his Master's degree from Ordnance Engineering College. Now, he is a Ph. D candidate, with Equipment Command and Administration Department, Shijiazhuang Campus of Army Engineering University, Shijiazhuang, China. His current interests are network security and defense, equipment support informatization.

Yu Lu received his Ph. D degree from the Beijing University of Aeronautics and Astronautics, Beijing, China. He is currently a full professor, with Information Engineering Department, Shijiazhuang Campus of Army Engineering University. His current interests are in the area of network security control, software-defined security.

Xi Li received his Ph. D degree from Shijiazhuang Campus of Army Engineering University. He is currently a lecturer, with Equipment Simulation Training Center, Shijiazhuang Campus of Army Engineering University. His current interest is network security and information technology of equipment support.

Wei Nie received the Ph.D. from the University of Electronic Science and Technology. He is a Lecturer of Shenzhen University. His research interests include network security and Software Defined Network.