Hindawi
Security and Communication Networks
Volume 2019, Article ID 1362964, 14 pages
https://doi.org/10.1155/2019/1362964

Research Article

Active Defense Strategy Selection Method Based on Two-Way Signaling Game

Xiaohu Liu[1,2] **Hengwei Zhang**[1,2] **Yuchen Zhang,**[1,2] **Lulu Shao,**[1] **and Jihong Han**[1,2]

[1]*Zhengzhou Information Science and Technology Institute, Zhengzhou 450001, China*
[2]*State Key Laboratory of Mathematical Engineering and Advanced Computing, Zhengzhou 450001, China*

Correspondence should be addressed to Hengwei Zhang; wlby_zzmy_henan@163.com

Received 23 June 2019; Revised 25 September 2019; Accepted 31 October 2019; Published 29 November 2019

Guest Editor: Akbar S. Namin

Most network security research studies based on signaling games assume that either the attacker or the defender is the sender of the signal and the other party is the receiver of the signal. The attack and defense process is commonly modeled and analyzed from the perspective of one-way signal transmission. Aiming at the reality of two-way signal transmission in network attack and defense confrontation, we propose a method of active defense strategy selection based on a two-way signaling game. In this paper, a two-way signaling game model is constructed to analyze the network attack and defense processes. Based on the solution of a perfect Bayesian equilibrium, a defense strategy selection algorithm is presented. The feasibility and effectiveness of the method are verified using examples from real-world applications. In addition, the mechanism of the deception signal is analyzed, and conclusions for guiding the selection of active defense strategies are provided.

1. Introduction

Network information technology is developing rapidly, and interconnected systems are on the rise [1]. However, network security incidents pose a major and perpetual problem [2]. Defense technologies represented by firewalls, intrusion detection, and antivirus software provide passive response defense based on a priori knowledge and attack characteristics, but they cannot respond to new types of complex network attacks in an effective and timely manner [3]. If the defending party can actively select a targeted defense strategy by predicting the attacker's actions and disrupt or block the attack process, while simultaneously maximizing its own benefits, then the defense may be called an active defense [4]. The essence of cybersecurity is a battle between the offense and defense. The effectiveness of the defense depends not only on its own strategic action, but also influenced and constrained by the attacker's action [5]. The key issue is how to select the optimal active defense strategy in an information-constrained confrontation environment.

The characteristics of opposite goals, strategic dependence, and noncooperative relationships in network attack and defense are in line with the core philosophy of game theory, namely, optimal decision in an environment of conflict. Some scholars, such as the authors of Refs. [6–11], have established network security models based on game theory, analyzed the offensive and defensive confrontation process, and solved the game equilibrium to determine the defense strategy and guide defense actions. We classified and analyzed the existing research results by combining the two factors of game information and action timing and came to the following conclusions:

(1) In a static game with complete information, there are many premise assumptions and the model is easy to establish, as demonstrated in Ref. [12].

(2) In a dynamic game with complete information, given the sustained nature of the offensive and defensive confrontation process, previous actions could be studied to affect the subsequent game process, as shown in Ref. [13].

(3) In a static game with incomplete information, the players may use the static Bayes' rule to infer the opponent's private information and break through

the complete information assumption, such as in Ref. [14].

(4) In a dynamic game with incomplete information, the late player observes the partial action of the early player, even without fully understanding the behavior type. However, since the behavior is type dependent, one can modify the a priori judgment of the behavior type of the early player by using the dynamic Bayes' rule, as depicted in Ref. [15]. Since neither the offense player nor the defense player can fully understand the opponent's information, influenced by the dynamic and persistent nature of the confrontation process, the dynamic game with incomplete information is more in line with the actual network attack and defense. Hence, this type of game is the focus of current network security game research.

A signaling game is a typical dynamic game with incomplete information, which provides a formal mathematical way to analyze how identity and deception are coupled in cyber-social systems. [16] It describes the strategic interplay of the game process through signal transmission [17], which is well-suited for studying the selection of active defense strategy. In Ref. [18], from the perspective of dynamic confrontation and limited information, a two-stage signaling game model is constructed to derive an optimal defense strategy. As demonstrated in Ref. [19], the signaling game model can be used to analyze the moving target defense. The defense side can alter the information asymmetry of the two sides by releasing the dynamically transformed signal and thereby expand its own benefits. In Ref. [20], the DDoS attack and defense process is modeled as a multistage signaling game, and an equilibrium solution is found. Moreover, the server port hopping defense strategy has been demonstrated to be effective. In Ref. [21], a multistage offensive and defensive signaling game model is constructed for modeling the multistage dynamic attack and defense process under incomplete information constraints. Also, the signal attenuation factor is used to quantify the influence of the defensive signal of the defending party. In Ref. [22], to address the spear-phishing attack of industrial control systems, a multistage offensive and defensive game model is established. Defense strategies are selected based on the comprehensive consideration of the benefits and costs. Finally, Ref. [23] analyzes the security issues of the Internet of Things through a multistage game model and provides specific defense strategies.

Despite their strengths, all the studies above assume that the network attack and defense process involve only one-way signal transmission, so the attack and defense process is modeled and analyzed by designating either the attacker or defender as the signal sender and the other party as the signal receiver. However, in an actual network attack and defense process, the attacker and the defender will have a series of strategic interactions. The attack and defense parties are generally both senders and receivers of signals. If the sender's transmitted signal is viewed as a stimulus, then the response chosen by the recipient is a reaction. In a two-way sustained stimulus-response process, the defender and the attacker are constantly adjusting and optimizing their respective strategies, thus dynamically propelling the attack and defense evolution [24]. Therefore, the game signal in network attack and defense should be a two-way send-and-receive mechanism.

To address the problem described above, we construct a two-way signaling game model to analyze the network attack and defense processes based on a two-way transmission mechanism of actual attack and defense signals. Based on the solution of the perfect Bayesian equilibrium, a defense strategy selection algorithm is presented. The main contributions of this work are as follows:

(1) Two-way signal transmission mechanism: both the offense and defense parties play a dual role of the sender and receiver. While affecting the other party's strategy selection by releasing the signal, they are also affected by the signal released by the other party.

(2) Game signal set containing both true and fake signals: in order to disrupt the cognitive decision-making process of the other party, both the offense and defense sides in the process of network confrontation use information countermeasures that release a mixture of true and false signals. Since the signal recipient has a certain discriminating ability against false signals, the deceptive effect of the false signal diminishes as the attack and defense game progresses.

(3) Dynamic multistage game process: the offensive and defensive confrontation continues in multiple stages as both sides continue to learn and evolve based on the interaction of signals, dynamically adjust the action strategy, and maximize their gains. Through a two-way signal transmission mechanism, the method proposed in this paper can more accurately characterize the offensive and defensive strategy confrontation process. Hence, this method more closely models an actual network attack and defense process. It also serves as a better theoretical reference, providing practical guidance in the selection of active defense strategies under dynamic conditions of incomplete information.

2. Construction of a Two-Way Attack and Defense Game Signal Model

2.1. Analysis of Attack and Defense Game Process

2.1.1. Basic Signaling Game Process. The basic signaling game consists of two players: the signal sender and the signal receiver. First, according to the Harsanyi conversion [25], the virtual player "Nature" selects the type of signal sender as θ and transforms the selection problem under the condition of incomplete information into a selection problem under the condition of uncertainty type. The signal sender knows that its type is θ, but the signal receiver only knows the a priori probability $P(\theta)$ that the sender belongs to type θ. The signal sender releases a signal H, and the signal receiver,

having observed signal H, uses Bayes' rule to deduce the posteriori probability $P(\theta \mid H)$ from the a priori probability $P(\theta)$ and subsequently selects an action strategy. The signal sender determines its own action strategy by predicting the signal receiver's action strategy, and both parties strive to maximize their respective gains. The process of the basic signaling game is shown in Figure 1.

2.1.2. Two-Way Attack and Defense Signaling Game Process. Network confrontations are dynamic and sustained. The attacker and the defender take sequential actions, and each party selects its own action strategy after observing the signal released by the other party. The two-way signaling game process is shown in Figure 2.

(1) Initial Configuration (ICN). The defender acts as the signal sender, and the attacker acts as the signal receiver. The defender deploys the network information system and configures the network topography, IP address, and network segmentation. Since the network must provide services to the outside world, it is characterized by open sharing, interconnection, and interoperability. The network must also have homologous, isomorphic, and homogenous characteristics of information network products. The attacker can gather information on the initial configuration of the defender through a variety of avenues, including infiltration by social engineering means, continuous scanning and detection, and public information acquisition [26]. Such information serves as the basis for the attacker to launch a network attack. In this work, the information is treated as a signal H_D released by the defender. The attacker observes the signal H_D, corrects the a priori judgment regarding the type of defender, and identifies its attack strategy. The game process is shown in the S_1 stage of Figure 2.

(2) Dynamic Confrontation (DCN). Both the offense and defense sides are constantly switching between the role of the signal sender and the signal receiver. Each stage of the game consists of a basic signaling game, as shown in the S2, S3, and S*i* stages in Figure 2. In the S_2 phase, the attacker selects the attack strategy and releases the signal H_A. The defender receives the signal H_A, corrects the a priori judgment about the type of the attacker, and selects the defense strategy accordingly. In the S_3 stage, the defender releases the signal H_D and the attacker receives the signal H_D and again corrects the a priori assessment regarding the type of the defender to determine the attack strategy. In the process of dynamic confrontation, the signal is transmitted in both directions, and both the offense and defense sides use Bayes' rule to incrementally correct their estimate of the true type of the other party. From the perspective of the defender, the termination condition of the game is when the attacker stops the attack and no longer releases signals. The game process is shown in the S_n phase of Figure 2.

2.2. Definition of Two-Way Attack-Defense Signaling Game Model. The signal plays a role in the strategic interaction between the sender and receiver. The sender of the signal determines the content of the signal and influences the recipient's action strategy through the signal. According to the Cyber Kill Chain model [27], the first stage of network reconnaissance is an intelligence gathering activity, such as detection and scanning, which is conducted by the attacker on the defender. This may be regarded as receiving the signal released by the defender. In the course of the confrontation, the sender of the signal may adopt the idea of deception by releasing signals that do not match its own type for the purpose of misleading the other party's judgment and expanding its own gain [28]. Therefore, the signals transmitted by both the offense and defense parties can be divided into two types: real signals and deception signals.

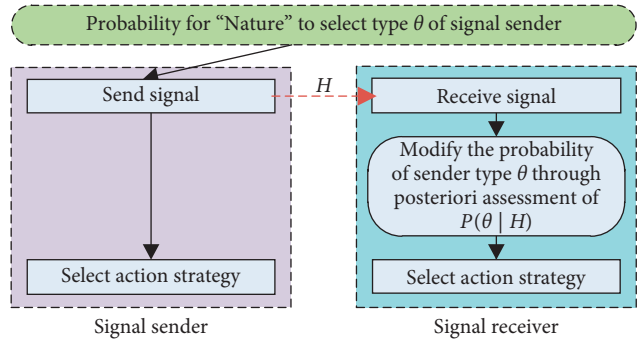

Figure 1: The basic signaling game process.

Definition 1 (real signal (RS)). A real signal is a signal that reflects the true type of the player. The player chooses the action strategy according to its own type. In the process of implementing its strategy, some private information is inevitably exposed; this information is transmitted to the receiver as a real signal. A real signal is accompanied by an action strategy, and the release of a real signal does not require additional cost.

Definition 2 (deception signal (DS)). A deception signal is a signal that does not match the true type of the player. In order to conceal its real type, the player induces the signal receiver to establish a wrong correction to the a priori probability by sending a signal that does not match its type, thereby rendering the receiver into a passive state. Since a signal will not be generated for no reason, the deceptive player must pay an extra cost to release the deceptive signal [29]. For example, if a low-defense user wishes to spoof as a high-defense user, it must deploy some camouflage facility and pay a certain defense cost to release the spoofing signal. The release of defensive signals by the defense player is a concrete manifestation of the active defense philosophy [30], in line with the deceptive concept that "when we are able to attack, we must seem unable; when using our forces, we must seem inactive" in Sun Tzu's *The Art of War*.

Based on the above analysis, a two-way signaling game (TWSG) model is constructed for the two-way transmission mechanism in the actual network attack and defense confrontation process.

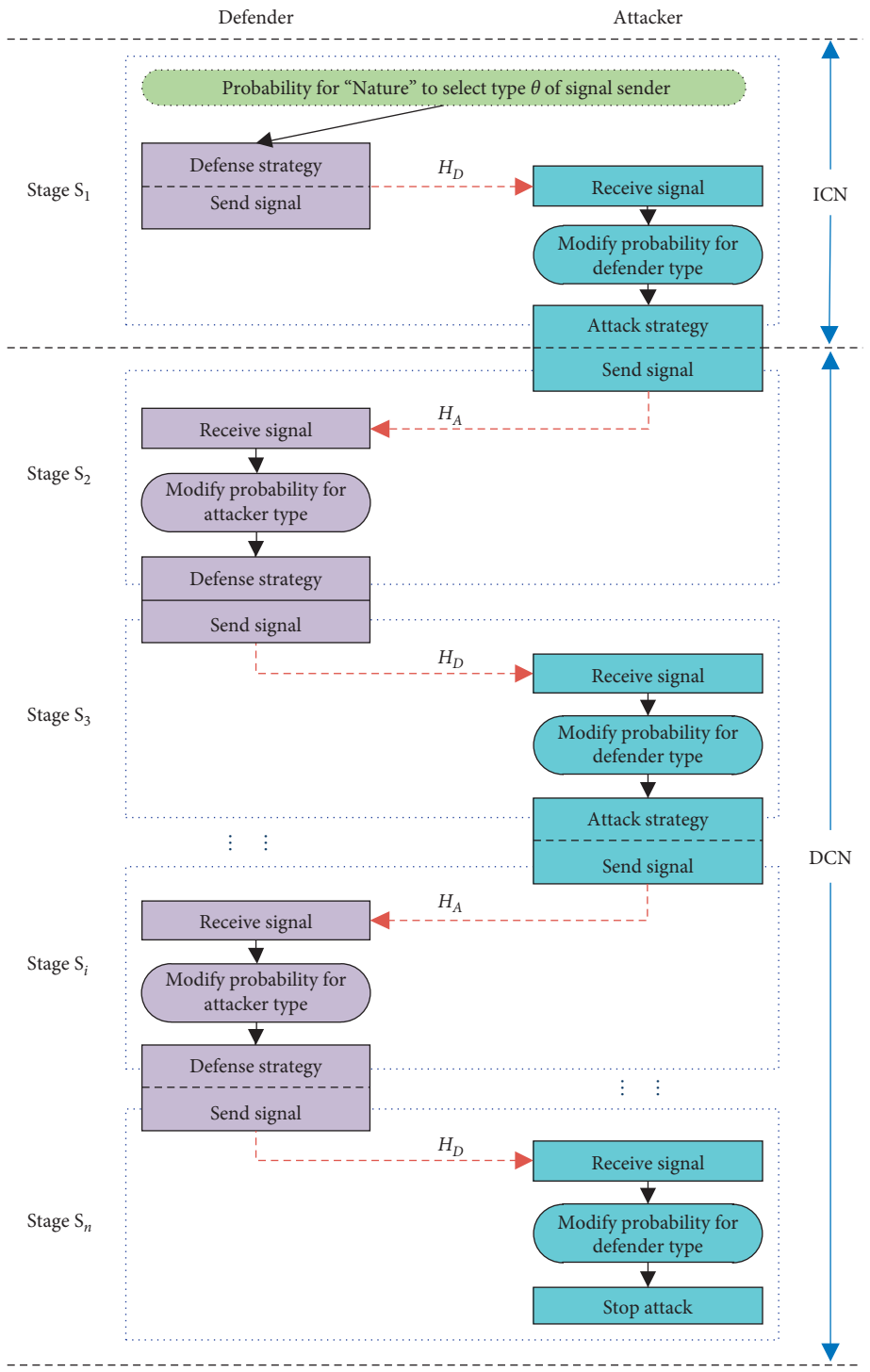

Figure 2: Two-way signaling game process.

Definition 3. The *TWSG* model has ten elements, where TWSG = $(N, \Theta, H, T, \sigma, \xi, S, P, \tilde{P}, U)$.

① $N = (N_D, N_A)$ is the player space of the game. It includes two players: the defender N_D and the attacker N_A.

② $\Theta = (\theta_D, \theta_A)$ is the type space. θ_D is the type of the defender, $\theta_D = (\phi_i \mid i = 1, 2, \ldots n)$, $n \geq 2$, and θ_A is the type of the attacker, $\theta_A = (\varphi_j \mid j = 1, 2, \ldots m)$, $m \geq 2$. The type of the player is private information, determined by the action strategy, and the player type can affect the game return of both parties.

③ $H = (H_D, H_A)$ is the signal space. H_D is the defense signal, $H_D = (h_{Dk} \mid k = 1, 2, \ldots v)$, $v \geq 2$, and H_A is the attack signal, $H_A = (h_{Al} \mid l = 1, 2, \ldots w)$, $w \geq 2$. The signal receiver can estimate the type of sender according to the signal received, and the signal space

logically corresponds to the type space. However, due to the existence of the spoofing signal, a specific signal does not have a strict correspondence relationship with the specific type of the attacker or defender.

④ T is the number of game stages, and $T = (1, 2, 3, \ldots, t)$, $t \geq 3$. The two-way signaling game continues in multiple stages, and the tth stage of the game is represented as TWSG(t).

⑤ σ is the spoofing signal attenuation factor. After multiple strategic interactions between the attacker and defender, the two sides become more familiar with each other, and the influence of deception signals is gradually attenuated. The posteriori probability generated in the tth stage of the game is modified by the factor σ_t to make it more realistic, where $0 \leq \sigma_t \leq 1$. The initial stage deception signal is not attenuated. The degree of attenuation of the deception signal at the TWSG(t) stage is expressed as $\sigma_t = \sigma^{t-1}$. For a sufficiently large T, $\sigma_T = \sigma^{T-1} \approx 0$, and the influence of the spoofing signal disappears completely. The signal and type constitute a corresponding relationship, and the two-way signaling game degenerates into a static game of incomplete information.

⑥ ξ is the gain discount factor and ξ represents the discount ratio of the gain in the $t+1$ stage as well as the gain in the t-stage. The discount ratio is used to convert the gain of a future stage into the present value.

⑦ $S = (S_D, S_A)$ is the strategy space. S_D is a defensive party strategy, $S_D = \{d_g \mid g = 1, 2, \ldots\}$ and S_A is an attacker party strategy, $S_A = \{a_h \mid h = 1, 2, \ldots\}$.

⑧ $P = (P_D, P_A)$ is the a priori probability space. P_D is the set of a priori probability of the defender, and it represents the a priori probability of the attacker's type known to the defender, where $P_D \neq \varnothing$, $P_D = [p_{D1}, p_{D2}, \ldots, p_{DT}]$. P_A is the a priori probability of the attacker, and it represents the a priori probability of the defender's type known to the attacker, where $P_A \neq \varnothing, P_A = [p_{A1}, p_{A2}, \ldots, p_{AT}]$.

⑨ $\widetilde{P} = (\widetilde{P}_D, \widetilde{P}_A)$ is the posteriori probability space. $\widetilde{P}_D$ is a set of posteriori probability of the defender, meaning the defender's posteriori assessment of the attacker's type, where $\widetilde{P}_D(\varphi_j \mid h_l) = (\varepsilon_{D1}, \varepsilon_{D2}, \ldots, \varepsilon_{\mathrm{DT}})$. $\widetilde{P}_A$ is the attacker's posteriori probability set, meaning the attacker's posteriori assessment of the defender's type, where $\widetilde{P}_A(\phi_i \mid h_k) = (\varepsilon_{A1}, \varepsilon_{A2}, \ldots, \varepsilon_{\mathrm{AT}})$.

⑩ $U = (U_D, U_A)$ is the gain space. U_D and U_A represent the defender's gain and the attacker's gain, respectively.

2.3. Gain Calculation. Based on the characteristics of the two-way signaling game model, we provide the following definition and calculation method for the game return.

Definition 4. The system damage cost (SDC), attack cost (AC), defense cost (DC), and related definitions and calculation methods can be found in Refs. [23, 31, 32]. Among them, SDC is affected by the combination of attack and defense strategies and is often recorded as $\mathrm{SDC}(d_g, a_h)$, which represents the value that the system suffers when the defense strategy is d_g and the attack strategy is a_h.

Definition 5 (deception cost). The deception defense cost (DDC) is the cost incurred to the defense party for actively releasing a spoofing signal to confuse the attacker. The deception attack cost (DAC) is the cost incurred to the attacking party for actively releasing a spoofing signal to confuse the defender.

According to the cost/reward calculation method, the returns of the attacker are the SDC and the total cost is the sum of the AC and DAC. The defender's cost is the sum of the SDC, DC, and DDC.

The discount factor ξ is used to convert future earnings into current gain. The gain target functions of the offensive and defensive parties can be expressed, respectively, as follows:

$$U_A(d_g, a_h, t) = \sum_{g,h,t} \xi^{t-1}\left[\mathrm{SDC}(d_g, a_h) - \mathrm{AC} - \mathrm{DAC}\right],$$
$$U_D(d_g, a_h, t) = -\sum_{g,h,t} \xi^{t-1}\left[\mathrm{SDC}(d_g, a_h) + \mathrm{DC} + \mathrm{DDC}\right]. \tag{1}$$

According to the attack-defense types of θ_A and θ_D, the attack-defense strategies can be divided into different levels, such as enhanced type and regular type. The costs and returns of the strategies at the same level are basically the same. For example, if an attack level contains a total of h attack policies, then the probability that the attacker selects the strategy a_h is *1/h*. The gain from this attack level can be expressed as an average of $\overline{U}_A(d_g, a_h, t) = \sum_h U_A(d_g, a_h, t)/h$. Similarly, if a defense level has a total of g defensive strategies, the gain of the defense level is $\overline{U}_D(d_g, a_h, t) = \sum_g U_D((d_g, a_h, t))/g$.

3. Two-Way Signaling Game Equilibrium Solution and Defense Strategy Selection

A two-way signaling game is a finite game consisting of several basic signaling games. In the game, the attacker and defender alternately act as signal senders and receivers and the single role equilibrium solution is no longer applicable. In this paper, we first present the solution process for a one-stage game equilibrium and then apply it to a multistage equilibrium solution.

We carry out the calculation and analysis for the single-stage game equilibrium solution by referring to the signal sender as the Leader and the signal receiver as the Follower. The relevant parameters are set as follows:

① Signal sender action strategy $\{l_1, l_2, \ldots, l_n\}$

② Signal receiver action strategy $\{f_1, f_2, \ldots, f_m\}$

③ Defender type space $\theta_D = (\phi_{\mathrm{DH}}, \phi_{\mathrm{DM}})$ = (enhanced type defense, regular type defense)

④ Defender's signal space $H_D = (h_{\mathrm{DH}}, h_{\mathrm{DM}})$ = (enhanced defense signal, regular defense signal)

⑤ Attacker type space $\theta_A = (\varphi_{\mathrm{AH}}, \varphi_{\mathrm{AM}})$ = (enhanced attack, regular attack)

⑥ Attacker signal space $H_A = (h_{\mathrm{AH}}, h_{\mathrm{AM}})$ = (enhanced attack signal, regular attack signal)

3.1. Single-Stage Game Equilibrium Solution

Definition 6. The TWSG(t) game equilibrium solution is $\mathrm{EQ}_t = (h^*(l^*, \Theta), f^*(h), \widetilde{P}_F(\Theta \mid h))$, where $h^*(l^*, \Theta)$ is the Leader's signal strategy, abbreviated as $h^*(\Theta)$, $f^*(h)$ is the Follower's strategy, abbreviated as $f^*(h)$, and $\widetilde{P}_F(\Theta \mid h)$ is the Follower's posteriori probability of the Leader type, where the parameter $F \in \{A, D\}$ indicates that the Follower can be an attacker or defender in different game stages, abbreviated as $\widetilde{P}_F(\Theta)$. According to game theory, the equilibrium should satisfy two conditions:

(i) $f^*(h) \in \arg\max_{f \in F} \sum \widetilde{P}_F(\Theta \mid h) U_F(h^*(\Theta), f, \Theta)$, indicating that under the condition of posteriori probability $\widetilde{P}_F(\Theta \mid h)$, the Follower is the optimal strategy for the Leader

(ii) $h^*(\Theta) \in \arg\max_{h \in H} U_L(h, f^*(h), \Theta)$, indicating that the Leader is the optimal strategy for the Follower

Here, $\widetilde{P}_F(\Theta \mid h)$ represents the posteriori probability of the Leader type calculated for the Follower based on a priori probability P, observed signal h, and its own strategy $f^*(h)$.

The steps for solving the perfect Bayesian equilibrium is more complex, and the entire process may be divided into the following three steps:

(1) *Step 1*. Calculate optimal strategy $f^*(h)$ based on the signal received by the Follower

(2) *Step 2*. Leader reduces the optimal strategy $h^*(\Theta)$

(3) *Step 3*. Select the perfect equilibrium solution $\mathrm{EQ}_t = (h^*(\Theta), f^*(h), \widetilde{P}_F(\Theta))$

The detailed process is shown in the Appendix.

Based on game theory, the perfect Bayesian equilibrium solution is the optimal strategy for the player [33]. Therefore, the defender should determine the active defense strategy based on its role and game equilibrium EQ_t.

3.2. Multistage Game Equilibrium Solution. In the multistage continuous confrontation process, the defense party may incrementally modify the attacker's motivation and behavioral preference using the stimulus-response learning mechanism, reduce the impact of the attacker's deception signal, and implement a targeted active defense strategy to maximize the expected return.

(1) In the first stage of the game TWSG(1), the Leader is the defender and the Follower is the attacker.

Based on the Harsanyi conversion, the viral player "Nature" selects the type of the defender. Type ϕ_{DH} is selected with a priori probability p_1, and type ϕ_{DM} is selected with probability $1 - p_1$. The defender releases the signals h_{DH} and h_{DM}. Based on the observed signals, the attacker selects strategy types φ_{AH} and φ_{AM} and corrects its a priori assessment of the defender type. According to the single-stage game equilibrium solution process in Section 3.1, the game equilibrium $\mathrm{EQ}_1 = (h^*(\Theta), f^*(h), \widetilde{P}_A(\Theta))$ can be obtained for TWSG(1). The TWSG(1) game tree is shown in Figure 3.

(2) In the second stage of the game TWSG(2), the Leader is the attacker and the Follower is the defender.

The attacker selects the attack strategy according to EQ_1 and sends a signal to the defender. The offense and defense sides have interchanged their role as the sender and receiver of the signal. Through the TWSG(1) game, both the offensive and defensive sides have gained some mutual understanding and the decay phenomenon of the deception signal begins to emerge. At this point, the attacker no longer relies on "Nature" to select the type. Instead, the selection is determined by the signal attenuation factor σ of the deception signal and the posteriori probability $\mathrm{EQ}_1(\widetilde{P}_A(\Theta))$ in EQ_1, as expressed by $\sigma\mathrm{EQ}_1(\widetilde{P}_A(\Theta))$. The attacker chooses φ_{AH} with probability $\sigma\mathrm{EQ}_1(\widetilde{P}_A(\Theta))$ and chooses φ_{AM} with probability $1 - \sigma\mathrm{EQ}_1(\widetilde{P}_A(\Theta))$. The TWSG(2) game tree is shown in Figure 4.

(3) In the third stage game TWSG(3), the Leader is the defender and the Follower is the attacker. The TWSG(3) game tree is shown in Figure 5.

The defender selects the defense strategy according to EQ_2 and sends a signal to the attacker. The attack and defense roles are interchanged again. After the first two stages of the game, the attenuation effect of the deception signal is more pronounced, as represented by the expression $\sigma^2\mathrm{EQ}_2(\widetilde{P}_D(\Theta))$. The defender chooses ϕ_{DH} with probability $\sigma^2\mathrm{EQ}_2(\widetilde{P}_D(\Theta))$ and selects ϕ_{DM} with probability $1 - \sigma^2\mathrm{EQ}_2(\widetilde{P}_D(\Theta))$.

(4) In the T-stage of the game TWSG(T), the Leader is the defender and the Follower is the attacker.

As described in Section 2.1.2, both the attacker and the defender continuously interchange their roles as the sender and receiver of the signal during the ongoing confrontation, which dynamically adjusts the strategy and moves the game process forward. When the game stage T is large enough, the spoofing signal will be screened by the other party and its influence will completely disappear. The two-way signaling game will degenerate into a static game of incomplete information. The defender will continue to use defensive measures as the Leader releases signals to the outside world. The attacker will terminate the confrontational behavior and act only as the Follower to receive the signals sent by the defender. The TWSG(T) game tree is shown in Figure 6.

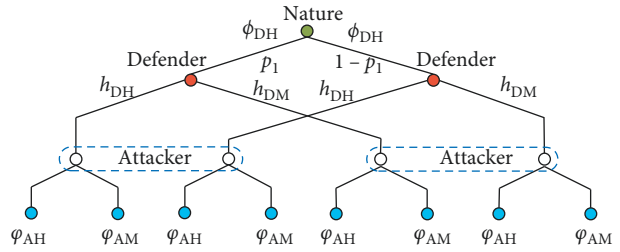

Figure 3: TWSG(1) game tree.

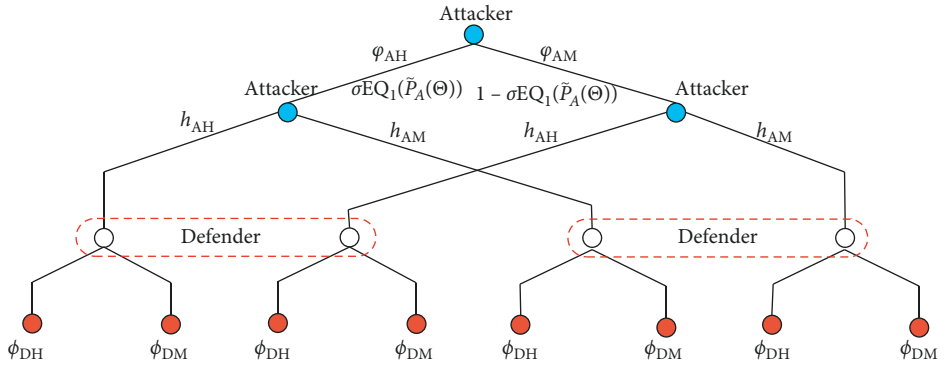

Figure 4: TWSG(2) game tree.

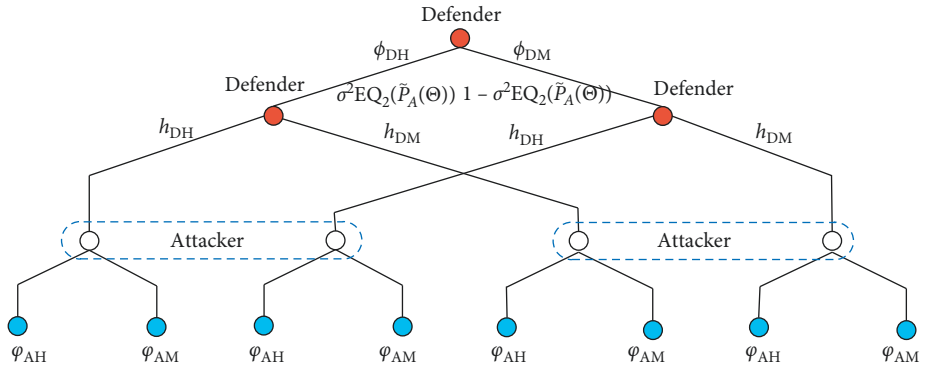

Figure 5: TWSG(3) game tree.

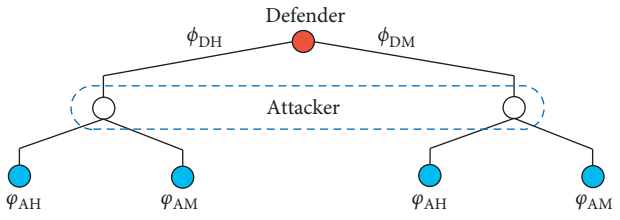

Figure 6: TWSG(T) game tree.

3.3. Defense Strategy Selection Algorithm and Comparison with Results. The algorithm for designing the active defense strategy is shown in Algorithm 1.

If the number of types on the defense side is n, the number of types on the attacker side is m, the number of game stages is t, the number of defense strategies is g, and the number of attack strategies is h, then according to Refs. [17, 21], the time complexity of the active defense strategy selection algorithm is $O(2t(mn + \max(g,h)^3))$ and the space complexity is $O(mn\max(g,h))$.

The results of our method are compared with available research on signaling games in Table 1.

The signal transmission mechanism refers to whether the signal transmission direction is one-way or two-way in the model. The attenuation of the deception signal indicates whether the model characterizes the deception signal attenuation phenomenon. The game process is used to distinguish whether the model has single-stage analysis capability or multistage analysis capability. The model expansion indicates whether the type and strategy of attack and defense in the model can be expanded. The better the

Input: Two-way signaling game model
Output: Active defense strategy
Initialize TWSG = $(N, \Theta, H, T, \sigma, \xi, S, P, \widetilde{P}, U)$
Calculate attack gain $U_A(d_g, a_h, t)$;
Calculate defense gain $U_D(d_g, a_h, t)$;
for $(t = 1, t \le T, t++)$
{
Initialize $P(\Theta \mid h)$;
Leader releases signal H;
Calculate {Inferred optimal dependence strategy $f^*(h)$ for Follower};
Calculate {Inferred optimal dependence strategy $h^*(\Theta)$ for Leader};
Generate posteriori inference of $\widetilde{P}_F(\Theta)$ for Follower based on Bayes' rule;
If $\widetilde{P}_F(\Theta)$ and $P(\Theta \mid h)$ not in conflict;
Then, Create $\mathrm{EQ}_t = (h^*(\Theta), f^*(h), \widetilde{P}_F(\Theta))$;
Return S_D^*;
$\widetilde{P}_F(\Theta) = \sigma^{t-1}\mathrm{EQ}_t(\widetilde{P}_F(\Theta))$;
}
End

Algorithm 1: Active defense strategy selection algorithm.

Table 1: Comparison of research methods.

Reference	Signal transmission mechanism	Deception signal attenuation	Game process	Model expandability	Equilibrium solution	Operating costs	Performances
Ref. [16]	One-way	No	Single stage	Average	Detailed	Low	Poor
Ref. [18]	One-way	No	Single stage	Better	Simple	Low	Poor
Ref. [19]	One-way	No	Multistage	Average	Simple	High	Medium
Ref. [20]	One-way	No	Multistage	Average	Simple	High	Medium
Ref. [21]	One-way	Yes	Multistage	Good	Detailed	High	Medium
Ref. [22]	One-way	Yes	Multistage	Good	Detailed	High	Medium
This study	Two-way	Yes	Multistage	Good	Detailed	High	Good

expansion ability, the wider the scope of application of the model. The equilibrium solution of the model represents the degree of detail of the game equilibrium solution process. The more detailed the solution process is, the more practical it is. In terms of operating costs, it means time complexity and space complexity of the defense strategy selection algorithm. The lower the operation cost, the better; the better the performance, the better. Most previous studies use the one-way signal transmission mechanism to model the attack and defense process, and less consideration is given to the phenomenon of deception signal attenuation in the confrontation. Additionally, some studies are limited to single-stage game analysis. In this paper, we conduct an in-depth analysis of the two-way signal transmission mechanism, establish a two-way signaling game model, provide a detailed game equilibrium solution process, and design a defense strategy selection algorithm. In terms of signal transmission mechanisms, deception signal attenuation, and game process, this work comes closer to actual network attack and defense, and the model has better scalability and practicability. By sending deception signals from both the offense and defense sides, the parties seek to control the other party's strategy selection as well as maximize their own expected returns. This process embodies the confrontational philosophy under the condition of limited information.

Zhu et al. [34] propose two iterative reinforcement learning algorithms which allow the defender to identify optimal defenses. Reinforcement learning and signaling game model have their own advantages and disadvantages, and they should be adapted to different application scenarios. The purpose of this paper is to analyze process of network attack and defense. Reinforcement learning is a black box. Although the optimal defenses can be obtained, the analysis process and principles cannot be visualized. Using the two-way signaling game model to conduct the network attack-defense confrontation analysis, the analysis process and principles can be visulized more cleraly.

4. Real Case Application and Results Analysis

4.1. Experimental Environment and Parameter Configuration. In order to verify the feasibility and effectiveness of the proposed method, an experimental network environment was set up to carry out a simulation experiment. The

experimental network was a typical business network, which was divided into three areas: external network, internal network, and DMZ. The attack and defense scenario are set as follows: the attacker located in the external network area and attempted to remotely attack the internal network zone of the enterprise intranet. The defender was the network security administrator of the enterprise and selected the active defense strategy according to the method in the paper. The topography of the experimental network is shown in Figure 7.

To ensure the availability and security of the enterprise network, a set of access control rules were set up between the network partitions as shown in Table 2. Among them, ⊕ indicates that access was allowed; × indicates that access was not allowed; and ∅ indicates that access requires certain permissions.

In general, the database server (databaseserver) stores a large amount of confidential data of the enterprise, so it was set as the target of attack in the experiment. According to the access control rules in Table 2, the attacker cannot directly access the databaseserver; however, through multiple steps, the vulnerability of the bastion server in the DMZ area can be used to obtain access to the internal network area, thereby achieving the goal of the attack.

Combined with the description of Common Vulnerabilities and Exposures (CVE) information in the information security vulnerability library [35], the vulnerability scanning tool Nessus was used to detect and discover the security vulnerabilities that existed in the experimental network. The security vulnerability of the experimental network is given in Table 3.

The attacker used the security vulnerabilities and defects that existed in the enterprise network to select an attack strategy consisting of several atomic attack actions. The defender selected a defense strategy containing different atomic defense actions in a targeted manner [36]. According to the attack and defense classification of the Lincoln Laboratory [37], we obtained the attack and defense strategies and their operating costs, as shown in Table 4.

In Refs. [17, 28], historical statistical data and expert experience were combined to provide the SDC values for different combinations of attack and defense strategies, as shown in Table 5, and to set $\xi = 0.5$ and $\sigma = 0.6$. In the ninth stage, $\xi^{t-1} = 0.5^8 \approx 0.0039$, which shows that after this stage, the gain has very less influence on the total return calculation; thus, the number of game stages was set to $T = 9$.

4.2. Equilibrium Solution and Strategy Selection

4.2.1. TWSG(1) Game Equilibrium and Defense Strategy. "Nature" selects the type of defense strategy with a probability of (0.4, 0.6). When the strategy type of the defender is φ_{DH}, the signal h_{DH} is sent out. When the type of the attack strategy is φ_{AH}, there are a total of four strategy combinations: (d_1, a_1), (d_1, a_2), (d_2, a_1), and (d_2, a_2). The SDC values for different combinations of attack-defense strategies are given in Table 5.

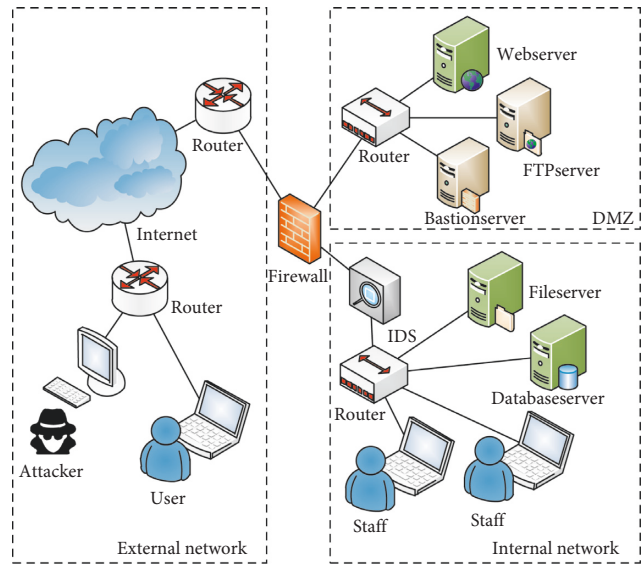

Figure 7: Topography of the experimental network.

Under the first strategy combination (d_1, a_1), the spoof signal of the attacker is DAC = 0. Thus,

$$\begin{aligned} U_A(d_1, a_1, 1) &= \mathrm{SDC}(d_1, a_1) - \mathrm{AC} - \mathrm{DAC} \\ &= 2320 - 480 - 0 = 1840. \end{aligned} \tag{2}$$

The gains for the other three strategy combinations can be calculated in the same way:

$U_A(d_1, a_2, 1) = 1810$, $U_A(d_2, a_1, 1) = 1900$, and $U_A(d_2, a_2, 1) = 1770$.

Since the probability for selecting different strategies at the same attack and defense level is the same, the probability for each strategy combination is 0.25, and therefore the average gain u_{12} of the attacker under strategy type φ_{AH} is

$$\begin{aligned} u_{12} &= \overline{U}_A(\phi_{\mathrm{DH}}, \varphi_{\mathrm{AH}}, 1) \\ &= 0.25U_A(d_1, a_1, 1) + 0.25U_A(d_1, a_2, 1) \\ &\quad + 0.25U_A(d_2, a_1, 1) + 0.25U_A(d_2, a_2, 1) \\ &= 1830. \end{aligned} \tag{3}$$

Similarly, we have.

$U_D(d_1, a_1, 1) = -[\mathrm{SDC}(d_1, a_1) + \mathrm{DC} + \mathrm{DDC}] = -3000$, $U_D(d_1, a_2, 1) = -2950$, $U_D(d_2, a_1, 1) = -3020$, and $U_D(d_2, a_2, 1) = -2870$.

$$\begin{aligned} u_{11} &= \overline{U}_D(\phi_{\mathrm{DH}}, \varphi_{\mathrm{AH}}, 1) \\ &= 0.25U_D(d_1, a_1, 1) + 0.25U_D(d_1, a_2, 1) \\ &\quad + 0.25U_D(d_2, a_1, 1) + 0.25U_D(d_2, a_2, 1) \\ &= -2960. \end{aligned} \tag{4}$$

Similarly, the above method can be used to obtain the offensive and defensive gains under different combinations of strategy types.

Using the equilibrium solution algorithm of Section 3.3, a pooling equilibrium solution is obtained for TWSG(1). There are two possible combinations of strategy types:

Table 2: Access control rules.

Network region	External network	Internal network	DMZ
External network	⊕	×	⊕
Internal network	×	⊕	∅
DMZ	⊕	∅	⊕

Table 3: Security vulnerability of the experimental network.

No.	Object of action	CVE code	Threat type	Threat level
1	Webserver	CVE-2015-1635	Code injection	Extreme risk
2	Webserver	CVE-2017-7269	Buffer zone overflow	Extreme risk
3	FTPserver	CVE-2014-8517	Operating system command injection	High risk
4	Bastionserver	CVE-2014-3556	Operating system command injection	High risk
5	Fileserver	CVE-2013-4730	Buffer zone overflow	Extreme risk
6	Databaseserver	CVE-2016-6662	Authorization and access control	Extreme risk

Table 4: Attack-defense strategy and operating cost.

Atomic attack action	φ_{AH}		φ_{AM}		Atomic defense action	ϕ_{DH}		ϕ_{DM}	
	a_1	a_2	a_3	a_4		d_1	d_2	d_3	d_4
Install listener program	√	√	√	√	Uninstall listener program	√	√	√	
Remote buffer overflow	√	√	√		Buffer overflow protection	√	√		
Install delete Trojan	√				Uninstall delete Trojan	√	√		
Attack SSH on FTPServer	√	√			Restart FTPserver	√		√	√
Steal account and password	√		√	√	Change account and password		√	√	√
Raise authority	√	√			Delete suspicious account	√	√	√	
Remote code injection	√	√			Identify code injection	√			
Violent crack password			√	√	Increase password complexity	√	√	√	√
AC	480	460	240	220	DC	680	640	440	410
DAC	80	70	30	20	DDC	100	80	40	30

Table 5: SDC values for different combinations of attack-defense strategies.

d / a	d_1	d_2	d_3	d_4
a_1	$SDC(d_1, a_1) = 2320$	$SDC(d_2, a_1) = 2380$	$SDC(d_3, a_1) = 2640$	$SDC(d_4, a_1) = 2680$
a_2	$SDC(d_1, a_2) = 2270$	$SDC(d_2, a_2) = 2230$	$SDC(d_3, a_2) = 2520$	$SDC(d_4, a_2) = 2570$
a_3	$SDC(d_1, a_3) = 2180$	$SDC(d_2, a_3) = 2120$	$SDC(d_3, a_3) = 2280$	$SDC(d_4, a_3) = 2320$
a_4	$SDC(d_1, a_4) = 2120$	$SDC(d_2, a_4) = 2080$	$SDC(d_3, a_4) = 2210$	$SDC(d_4, a_4) = 2260$

Option 1: the defender selects strategy type ϕ_{DH} and releases signal h_{DH}, and the attacker selects strategy type φ_{AM}. This time, $U_{11} = -2960$ and $U_{12} = 1830$.

Option 2: the defender selects strategy type ϕ_{DM} and releases signal h_{DH}, and the attacker selects strategy type φ_{AM}. At this time, $U_{11} = -2727.5$ and $U_{12} = 2037.5$.

Therefore, the defender selects option 2 as the defense strategy, designated as (ϕ_{DM}, h_{DH}). The game tree of attack and defense is shown in Figure 8.

4.2.2. TWSG(2) Game Equilibrium and Defense Strategy. In the TWSG(1) equilibrium solution process, the attacker may choose either the strategy type φ_{AH} or φ_{AM}, and therefore the defender's posteriori probability of the attacker is modified to (0.5, 0.5). Using the equalization solution algorithm described in Section 3.3, the solution of TWSG(2) remains a pooling equilibrium. There are two possible combinations of strategies:

(i) The attacker selects the strategy type φ_{AH} and releases signal h_{AM}, and the defender chooses strategy type ϕ_{DM}

(ii) The attacker selects strategy type φ_{AM} and releases signal h_{AM}, and the defender selects strategy type ϕ_{DM}

Therefore, the defender selects the regular type strategy, designated as ϕ_{DM}.

4.2.3. Game Equilibrium and Defense Strategy for Stages Three through Nine. Using the above method, the game equilibrium for each stage is solved sequentially.

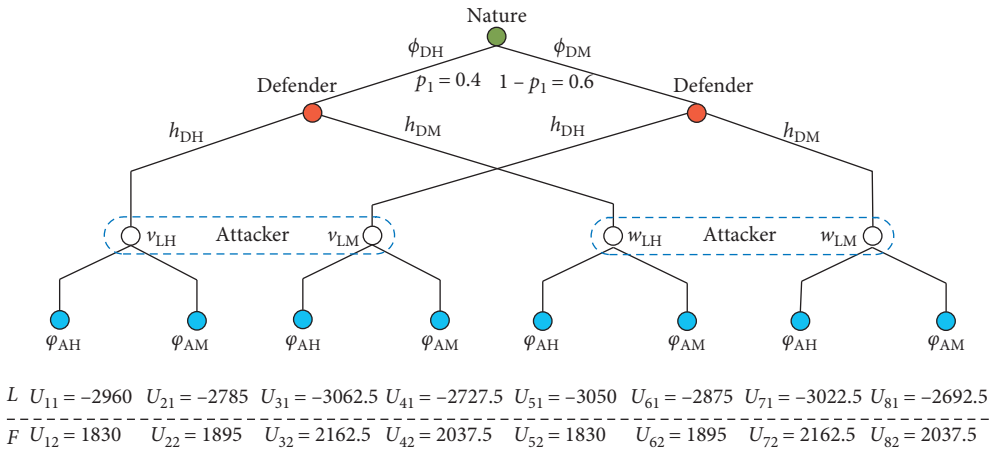

Figure 8: Game tree of attack and defense.

For stages three through six, as shown in Table 6, the game equilibrium solution remains a pooling equilibrium, but the deceptive signal is gradually attenuated. In stages seven through nine, the deception signal is completely attenuated, the game evolves into an incomplete information static game, and the pooling equilibrium solution becomes a separating equilibrium solution. At this point, the defender selects the enhanced ϕ_{DH} as the strategy type and releases an enhanced signal h_{DH}, designated as (ϕ_{DH}, h_{DH}).

4.3. Experimental Analysis. Based on the above experiments and data analysis, the following conclusions can be drawn from the general analysis of the offensive and defensive game equilibrium and the gain without considering specific parameter values.

(1) Deception signals can improve attack and defense performance.

The game equilibrium solutions for stages one through six are pooling equilibrium solutions, indicating that, in the initial stage of the offensive and defensive game, the defender may adopt the regular type of defense strategy ϕ_{DM} and confuse and mislead the attacker by releasing the spoofing signal h_{DH}. By disrupting the cognition of the attacker, the defender's own gain can be maximized at a small cost. The effectiveness of the spoofing signal should therefore be fully utilized to actively release the spoofing signal. At the same time, the ability to identify the attacking party's spoofing signals should be enhanced so that the motivation and preference of the attacker can be recognized as early as possible and a targeted active defense strategy can be implemented.

(2) The role of the spoofing signal is limited and attenuated.

As the game progresses, the spoofing signal becomes gradually attenuated. In the seventh through ninth stages of the game, the game equilibrium solution becomes a separating equilibrium solution, indicating that the function of the deception signal has completely disappeared. The defender no longer releases spoofing signals but instead increases the defensive input and adopts an enhanced defense strategy ϕ_{DH} to fight against network attacks. Therefore, when selecting the strategy, one should avoid the limitations of the spoofing signal and the attenuation process should be delayed by improving the quality of the spoofing signal. At the same time, attention should be given to collecting threat information and amplifying the limitations of the attacker's spoofing signal.

(3) Spoofing signals can delay the attack speed and reduce the suddenness of the attack.

An analysis of the first through ninth stages of the game shows that the deception signal released by the defender can delay the formation of the network kill chain and gain some reaction time for the defender. The deception signal can partially offset the time asymmetry advantage and the first-move advantage possessed by the attacker. However, due to the limitations of the spoofing signal, relying solely on the spoofing signal itself cannot completely resist network attacks. Therefore, the defending party should evolve according to the game process and use other means of defense to dynamically adjust the defense strategy to maximize its own return.

(4) Reduce security losses by enhancing defense capabilities.

We analyze the gamer's return when different strategy types are adopted. In the first through sixth stages, the defender adopts the regular type of defense strategy and the average return is −2853. In the seventh through ninth stages, the defender chooses the enhanced defense strategy type and the defender's average return is −2496. This shows that when faced with continuous high-intensity network attacks, the defending party should increase its security investment, enhance its defense capabilities, and reduce its security losses.

Table 6: Defense strategies of different stages and attack-defense returns.

Game stage	Defense role	Equilibrium type	Defense strategy	Attacker return	Defender return
TWSG (1)	Leader	Pooling equilibrium	(ϕ_{DM}, h_{DH})	2037.5	−2727.5
TWSG (2)	Follower	Pooling equilibrium	ϕ_{DM}	2053.5	−2785.5
TWSG (3)	Leader	Pooling equilibrium	(ϕ_{DM}, h_{DH})	2079.5	−2833.5
TWSG (4)	Follower	Pooling equilibrium	ϕ_{DM}	2112.5	−2894.5
TWSG (5)	Leader	Pooling equilibrium	(ϕ_{DM}, h_{DH})	2145.5	−2920.5
TWSG (6)	Follower	Pooling equilibrium	ϕ_{DM}	2069.5	−2956.5
TWSG (7)	Leader	Separating equilibrium	(ϕ_{DH}, h_{DH})	2011	−2460
TWSG (8)	Follower	Separating equilibrium	ϕ_{DH}	2038	−2492
TWSG (9)	Leader	Separating equilibrium	(ϕ_{DH}, h_{DH})	2089	−2536

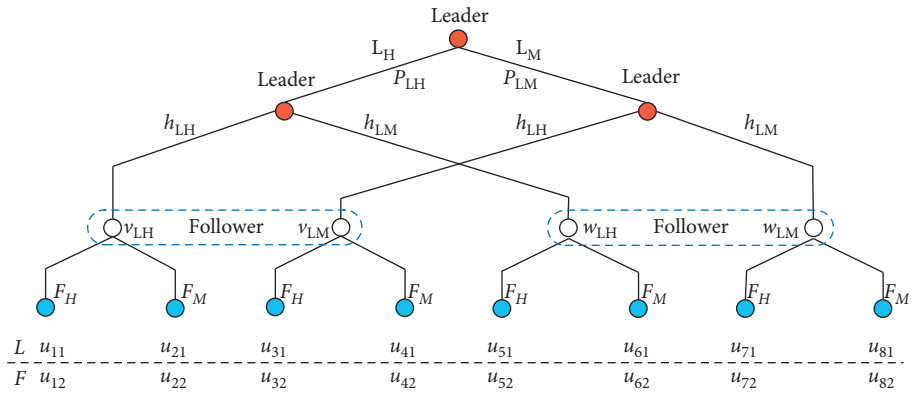

Figure 9: Single-stage signaling game tree.

5. Conclusion

Active defense is a topic at the forefront of research in the field of network security. Strategy selection is the key to defense effectiveness. Under the conditions of attack-defense confrontation and limited information, the defense party's optimal strategy is difficult to determine; however, a signaling game model is an effective way to solve this problem. To address the problem that one-way signal transmission does not conform to the actual problem of network attack and defense, we analyzed the two-way signal transmission process, constructed a two-way signaling game model, provided a multistage perfect Bayesian equilibrium solution process, and designed an active defense strategy selection algorithm in this paper. The feasibility and effectiveness of the method was verified through example applications and analysis. By analyzing the experimental results, we identified the mechanism driving the effectiveness and limitations of the deceptive signal and summarized four conclusions that guide the selection of active defense strategies. Compared with existing research, the two-way signaling game model proposed in this paper more accurately represents the offensive and defensive strategy confrontation process and more closely resembles an actual network attack and defense process. Thus, our work serves as the basis of, and provides reference to, the active defense strategy selection process under dynamic incomplete information conditions.

Appendix

Example Solution of Perfect Bayesian Equilibrium

Based on the parameter settings in this paper, the attacking party and defending party each have two strategy types and release two types of signals. The Leader type is represented by the symbols L_H and L_M, the signal space is represented by H_{LH} and H_{LM}, the Follower type is represented by the symbols F_H and F_M, $\{u_{11}, u_{21}, u_{31}, \ldots, u_{81}\}$ is the gain of the Leader, and $\{u_{12}, u_{22}, u_{32}, \ldots, u_{82}\}$ is the gain of the Follower. The single-stage signaling game tree is shown in Figure 9.

Step 1. Follower strategy calculation.

First, we assume that the posteriori inference of different signal sets on the single-stage game tree to be $P_F(\Theta \mid h)$. We then calculate the maximum return $\max_{f \in F} \sum P_F(\Theta \mid h) U_F(h^*(\Theta), f, \Theta)$.

When $H = h_{LH}$,

$$\begin{aligned}
&\max_{f \in F} \sum P_F(\Theta \mid h) U_F(h^*(\Theta), f, \Theta) \\
&= \max\{U_F(h_{LH}, F_H, L_H) \times \tilde{P}(L_H \mid h_{LH}) \\
&\quad + U_F(h_{LH}, F_H, L_M) \times \tilde{P}(L_M \mid h_{LH}), U_F(h_{LH}, F_M, L_H) \\
&\quad \times \tilde{P}(L_H \mid h_{LH}) + U_F(h_{LH}, F_M, L_M) \times \tilde{P}(L_M \mid h_{LH})\} \\
&= \max\{u_{12} \cdot v_{LH} + u_{32} \cdot v_{LM}, u_{22} \cdot v_{LH} + u_{42} \cdot v_{LM}\},
\end{aligned} \tag{A.1}$$

and the condition $v_{\mathrm{LH}} + v_{\mathrm{LM}} = 1$ is satisfied.

Assuming that $u_{12} \cdot v_{\mathrm{LH}} + u_{32} \cdot v_{\mathrm{LM}} = u_{22} \cdot v_{\mathrm{LH}} + u_{42} \cdot v_{\mathrm{LM}}$,

we solve and obtain $v^*_{\mathrm{LH}} = (u_{42} - u_{32}/u_{12} - u_{22} - u_{32} + u_{42})$, and $v^*_{\mathrm{LH}} \in [0, 1]$.

For $0 \le v_{\mathrm{LH}} \le v^*_{\mathrm{LH}}$, $(3) = u_{12} \cdot v_{\mathrm{LH}} + u_{32} \cdot v_{\mathrm{LM}}$ and $f^*(h) = F_H$.

For $v^*_{\mathrm{LH}} \le v_{\mathrm{LH}} \le 1$, $(3) = u_{22} \cdot v_{\mathrm{LH}} + u_{42} \cdot v_{\mathrm{LM}}$ and $f^*(h) = F_L$.

Similarly, we obtain $w^*_{\mathrm{LH}} = u_{82} - u_{72}/u_{52} - u_{62} - u_{72} + u_{82}$.

For $0 \le w_{\mathrm{LH}} \le w^*_{\mathrm{LH}}$, $f^*(h) = F_H$.

For $w^*_{\mathrm{LH}} \le w_{\mathrm{LH}} \le 1$, $f^*(h) = F_L$.

By repeating the above process, we calculate $f^*(h)$ for $H = h_{\mathrm{LM}}$.

Step 2. Leader strategy calculation.

$$\max_{h \in H} U_L(h, f^*(h), \Theta). \tag{A.2}$$

For $\Theta = L_H$, when $0 \le v_{\mathrm{LH}} \le v^*_{\mathrm{LH}}$ and $0 \le w_{\mathrm{LH}} \le w^*_{\mathrm{LH}}$,

$$\begin{aligned} &\max_{h \in H} U_L(h, f^*(h), \Theta) \\ &\quad = \max\{U_L(h_{\mathrm{LH}}, F_H, L_H), U_L(h_{\mathrm{LM}}, F_H, L_H)\} \\ &\quad = \max\{u_{11}, u_{51}\}, \end{aligned} \tag{A.3}$$

and we obtain $h^*(L_H)$.

Similarly, we obtain $h^*(L_H)$ for different sections of v_{LH} and w_{LH}.

By repeating the above process, we calculate $h^*(L_H)$ for $\Theta = L_M$.

Step 3. Calculate equilibrium solution.

We obtain $f^*(h)$ and $h^*(\Theta)$ in Step 1 and Step 2, respectively, by combining this with a priori probability PL and obtain the posteriori probability $\widetilde{P}_F(\Theta)$. If the calculated value of $\widetilde{P}_F(\Theta)$ is not in conflict with the premise hypothesis $P(\Theta \mid h)$, then the equilibrium solution is $\mathrm{EQ} = (h^*(\Theta), f^*(h), \widetilde{P}_F(\Theta))$.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest regarding the publication of this paper.

Authors' Contributions

Xiaohu Liu and Hengwei Zhang contributed equally to this work.

Acknowledgments

This work was supported in part by the National Natural Science Foundation of China under Grant nos. 61521003 and 61572517 and in part by the Henan Science and Technology Research Project under Grant no. 182102210144.

References

[1] Z. Tian, W. Shi, Y. Wang et al., "Real-time lateral movement detection based on evidence reasoning network for edge computing environment," *IEEE Transactions on Industrial Informatics*, vol. 15, no. 7, pp. 4285–4294, 2019.

[2] Z. Tian, M. Li, M. Qiu, Y. Sun, and S. Su, "Block-DEF: a secure digital evidence framework using blockchain," *Information Sciences*, vol. 491, pp. 151–165, 2019.

[3] R. K. Sharma, B. Issac, and H. K. Kalita, "Intrusion detection and response system inspired by the defense mechanism of plants," *IEEE Access*, vol. 7, pp. 52427–52439, 2019.

[4] D. E. Denning and E. Dorothy, "Framework and principles for active cyber defense," *Computers & Security*, vol. 40, pp. 108–113, 2014.

[5] S. Huang, H. Zhang, J. Wang, and J. Huang, "Markov differential game for network defense decision making method," *IEEE Access*, vol. 6, pp. 39621–39634, 2018.

[6] C. T. Do, N. H. Tran, C. Hong et al., "Game theory for cyber security and privacy," *ACM Computing Surveys*, vol. 50, no. 2, pp. 1–37, 2017.

[7] A. Farraj, E. Hammad, A. A. Daoud, and D. Kundur, "A game theoretic analysis of cyber switching attacks and mitigation in smart grid systems," *IEEE Transactions on Smart Grid*, vol. 7, no. 4, pp. 1846–1855, 2016.

[8] H. Hu, Y. Liu, H. Zhang, and R. Pan, "Optimal network defense strategy selection based on incomplete information evolutionary game," *IEEE Access*, vol. 6, pp. 29806–29821, 2018.

[9] J. Huang, H. Zhang, and J. Wang, "Markov evolutionary games for network defense strategy selection," *IEEE Access*, vol. 5, pp. 19505–19516, 2017.

[10] C. Lei, D.-H. Ma, and H.-Q. Zhang, "Optimal strategy selection for moving target defense based on Markov game," *IEEE Access*, vol. 5, pp. 156–169, 2017.

[11] Z. Tian, X. Gao, S. Su, J. Qiu, X. Du, and M. Guizani, "Evaluating reputation management schemes of internet of vehicles based on evolutionary game theory," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 6, pp. 5971–5980, 2019.

[12] Y.-L. Liu, D.-G. Feng, L.-H. Wu, and Y.-F. Lian, "Performance evaluation of worm attack and defense strategies based on static Bayesian game," *Journal of Software*, vol. 23, no. 3, pp. 712–723, 2012.

[13] L. Wangqun, H. Wang, and L. Jiahong, "Research on active defense technology in network security based on non-cooperative dynamic game theory," *Journal of Computer Research and Development*, vol. 48, no. 2, pp. 306–316, 2011.

[14] W. Jin-Dong, "Active defense strategy selection based on the static Bayesian game," *Journal of Xidian University*, vol. 43, no. 1, pp. 144–150, 2016.

[15] C. Lei, H.-Q. Zhang, L.-M. Wan, L. Liu, and D.-H. Ma, "Incomplete information Markov game theoretic approach to strategy generation for moving target defense," *Computer Communications*, vol. 116, pp. 184–199, 2018.

[16] W. Casey, A. Kellner, P. Memarmoshrefi, J. A. Morales, and B. Mishra, "Deception, identity, and security," *Communications of the ACM*, vol. 62, no. 1, pp. 85–93, 2018.

[17] H. Xu, R. Freeman, V. Conitzer, S. Dughmi, and M. Tambe, "Signaling in Bayesian stackelberg games," in *Proceedings of the International Joint Conference on Autonomous Agents and Multiagent Systems, AAMAS*, pp. 150–158, Singapore, January 2016.

[18] H. Jihong and Y. Dingkun, "Defense policies selection method based on attack-defense signaling game model," *Journal on Communications*, vol. 36, no. 4, pp. 121–132, 2016.

[19] X. Feng, Z. Zheng, D. Cansever, A. Swami, and P. Mohapatra, "A signaling game model for moving target defense," in *Proceedings of the 2017 IEEE Conference on Computer Communications, INFOCOM 2017*, pp. 1–9, Atlanta, GA, USA, May 2017.

[20] X. Gao and Y. Zhu, "DDoS defense mechanism analysis based on signaling game model," in *Proceedings of the 2013 5th International Conference on Intelligent Human-Machine Systems and Cybernetics*, pp. 414–417, Hangzhou, China, August 2013.

[21] Z. Hengwei, L. Tao, W. Jindong, and H. Jihong, "Optimal active defense using dynamic multi-stage signaling game," *China Communications*, vol. 12, no. 2, pp. 114–122, 2015.

[22] X. Chen, X. Liu, L. Zhang, and C. Tang, "Optimal defense strategy selection for spear-phishing attack based on a multistage signaling game," *IEEE Access*, vol. 7, pp. 19907–19921, 2019.

[23] Y. Yang, B. Che, Y. Zeng, Y. Cheng, and C. Li, "MAIAD: a multistage asymmetric information attack and defense model based on evolutionary game theory," *Symmetry*, vol. 11, no. 2, pp. 215–229, 2019.

[24] M. O. Sayin and T. basar, "Deception as defense framework for cyber-physical systems," 2019, https://arxiv.org/abs/1902.01364.

[25] J. C. Harsanyi, "Games with incomplete information played by "Bayesian" players," in *Game Theory*, pp. 154–170, Springer, Dordrecht, Netherlands, 1982.

[26] Q. Zhu, "Game theory for cyber deception: a tutorial," 2019, https://arxiv.org/abs/1903.01442.

[27] S. Clio, "Cyber kill chain based on threat taxonomy and its application on cyber common operational picture," in *Proceedings of the 2018 International Conference on Cyber Situational Awareness*, pp. 1–8, Data Analytics and Assessment (Cyber SA), Glasgow, Scotland, June 2018.

[28] J. Pawlick, E. Colbert, and Q. Zhu, "Modeling and analysis of leaky deception using signaling games with evidence," *IEEE Transactions on Information Forensics and Security*, vol. 14, no. 7, pp. 1871–1886, 2019.

[29] E. Al-Shaer, J. Wei, W. Kevin, Hamlen, and C. Wang, *Autonomous Cyber Deception*, Springer International Publishing, Berlin, Germany, 2019.

[30] T. Zhang, L. Huang, J. Pawlick, and Q. Zhu, "Game-theoretic analysis of cyber deception: evidence-based strategies and dynamic risk mitigation," 2019, https://arxiv.org/abs/1902.03925.

[31] W. Yuzhuo, Y. Jianye, and Q. Wen, "Evolutionary game model and analysis methods for network group behavior," *Chinese Journal of Computers*, vol. 38, no. 2, pp. 282–300, 2015.

[32] J. Wei and F. Bing-Xing, "Defense strategies selection based on attack-defense game model," *Journal of Computer Research and Development*, vol. 47, no. 12, pp. 714–723, 2014.

[33] D. Fudenberg and J. Tirole, "Perfect Bayesian equilibrium and sequential equilibrium," *Journal of Economic Theory*, vol. 53, no. 2, pp. 236–260, 1991.

[34] M. Zhu, Z. Hu, and P. Liu, "Reinforcement learning algorithms for adaptive cyber defense against heartbleed," in *Proceedings of the 2014 in Proceedings of the First ACM Workshop on Moving Target Defense—MTD'14*, Scottsdale, AR, USA, November 2014.

[35] National Vulnerability Database of Information Security, https://nvd.nist.gov/.

[36] Q. Tan, Y. Gao, J. Shi, X. Wang, B. Fang, and Z. Tian, "Toward a comprehensive insight into the eclipse attacks of tor hidden services," *IEEE Internet of Things Journal*, vol. 6, no. 2, pp. 1584–1593, 2019.

[37] L. Gordon, M. Loeb, W. Lucyshyn, and R. Richardson, "CSI/FBI computer crime and security survey," in *Proceedings of the Computer Security Institute*, pp. 48–64, San Francisco, CA, USA, 2015.