A Multistage Dynamic Defense Method for Evolutionary Games

Zhiyong Luo, Yutong Cao, Weiwei Song, and Jie Li
(Corresponding author: Luo Zhiyong)

School of Computer Science and Technology, Harbin University of Science and Technology
Harbin 150080,China
Email: luozhiyongemail@sina.com

(Received Jan. 27, 2022; Revised and Accepted Oct. 13, 2022; First Online Oct. 22, 2022)

Abstract

This paper proposes a multistage dynamic defense method for evolutionary games to address the challenge of accurately sensing unknown and homeopathic attacks on each node in the network and effectively accomplishing dynamic security. The method combines the replicator dynamic equations and the characteristics of the attack-defense game adversarial process to establish a discrete multistage offense-defense game model, and then quantifies the gains and equilibrium solutions for the model, simulates the multistage offense-defense game process under the high selection rate of the attack and defense strategy, and calculates the maximum objective function values of both sides. By analyzing these function values, the security situational awareness of the whole network nodes is completed to predict future security situations and system maintenance. Experimental comparisons show that the model approach has high operational efficiency and better defensive performance to ensure system integrity and other advantages.

Keywords: Dynamic Interaction; Multistage; Replicator Dynamic Equations; Situational Awareness

1 Introduction

With the widespread use of the Internet, people are becoming more and more inseparable from computers in their production and life. While we are currently improving network connectivity through constantly updated network technologies, we are also witnessing an era of unprecedented cyber attacks. Ensuring the confidentiality, integrity and availability of data, devices, networks and users has become critical. Most cybersecurity research has focused either on targeting specific vulnerabilities or proposing specific defense algorithms to defend against well-defined attack scenarios. Most of the defense techniques, which are static and passive, however, cannot effectively accomplish dynamic security in the face of unknown attacks, transient attacks in the network. Although such network security research is important, little attention has been paid to the dynamic interaction between attackers and defenders.

Game theory has the characteristics of goal opposition, relationship non-cooperation, and strategy dependence, which are consistent with the basic characteristics of network attack and defense [15]. Therefore, people apply game models to the field of network security to reason about intrusion intentions, targets and strategies. Traditional games are built on the premise of complete rationality of decision makers, which does not match with the actual attack and defense and reduces the effectiveness of models and methods. Considering the limited rationality of the attack and defense sides in the real network, evolutionary game theory is applied to the study of the attack and defense process [11].

Yao *et al.* proposed a multi-variant execution architecture-based CFI (MVX-CFI). MVX-CFI is an execution architecture-based, dynamic, and transparent CFI (control integrity) implementation that effectively captures the direction of control flow throughout the software runtime and detects illegal path shifts caused by malicious behaviors such as attacks [20]. Tian *et al.* modeled sequential attacks in complex networks as a partially observable Markov decision process (POMDP). Then a POMDP reinforcement learning method is proposed to analyze the dynamic robustness of complex networks under sequential attacks when the network information is incomplete [19]. Foschini *et al.* analyzed and identified the correct detection and mitigation strategies for DoS attacks in IT/OT networks. Provided DoS detection and mitigation strategies in business-centric IT/environment and production-centric operational technology (IT/OT) networks [6]. Hu *et al.* extend the game model to a novel two-way signaling game model and proposed an algorithm to identify the refined Bayesian equilibrium. Based on the calculated payoffs, the optimal strategy choices of the attacking and defending parties during the game are analyzed [9]. Hsieh *et al.* characterized the equilibrium of the underlying game and used the Bayesian dual Metropolis-

Hastings algorithm to estimate the model. And further extended the model to incorporate unobserved heterogeneity and showed that ignoring unobserved heterogeneity leads to biased estimation in simulation experimentsHsieh2020A. Zhang *et al.* proposed heterogeneity to improve the security of web servers with mimetic constructs and pointed out the importance of quantifying heterogeneity. A quantification method applicable to quantify heterogeneity is proposed, by which the factors affecting the heterogeneity of a mimetically constructed Web server are analyzed. A new method is provided for the quantitative assessment of mimetic defense [21]. Chen *et al.* innovatively used users as third-party participants in the moving target defense game and combined Stackelberg game and Markov model to construct a non-reciprocal three-party game to determine the optimal strategy for moving target defense. The proposed model can balance the cost and benefit of defenders and users, avoid excessive defense and inappropriate defense, and effectively achieve intelligent defense strategy decision [4]. Huanruo *et al.* conducted a comprehensive survey of the current state-of-the-art quantitative evaluation. MTD techniques based on the software stack model for classification. Then, a specific review and comparison of existing quantitative evaluations of MTD is presented [10]. Sengupta *et al.* provided a comprehensive survey of MTD and implementation strategies from the perspective of complete network system architecture. Discussed how various MTDs are implemented, analyzed MTD testbeds and case studies, and classified MTDs according to qualitative and quantitative metrics of security and performance effectivenessSengupta2020A. Zhang *et al.* proposed a user-oriented anti-censorship approach that significantly increases the cost to attackers. Representing Web services as mobile nodes forms a moving target defense strategy by using mobile IPv6 [22]. Sharma *et al.* proposed the random host and service multiplexing technique, RHSM. This technique uses shuffling of IP addresses and port numbers and aims to obfuscate the true identity of hosts and services at the network and transport layers to defend against network reconnaissance and scanning attacks [17].

The above studies have established different network security risk assessment models based on game theory, but they are too ideal for the establishment of attack and defense models, which cannot truly reflect the possibility of attackers' choice of target networks and attack methods, and do not quantify the probability situation of strategy selection for the intentions of both attackers and defenders. In this paper, a multistage dynamic attack and defense model based on evolutionary game theory is proposed under the premise of information asymmetry between offensive and defensive, and the main work and innovations are as follows.

1) Combining evolutionary game theory and Nash equilibrium, each round of attack and defense will adjust the strategy according to the newly acquired information and vested interests, and the game process goes through many iterations to finally reach the dynamic equilibrium state;

2) Based on the limited amount of information, in response to the attacker's attack intention, the defender releases false defense signals to cope with the complex network with changing security elements and improve the predictive capability of the defense situation;

3) Considering the complex factors affecting the attacker's attack behavior, the attack probability is calculated from three indicators: signal deception cost, attack cost and defense cost, which more realistically reflects the attacker's situation in the actual network.

2 Construction of Multistage Dynamic Game Model

2.1 Multistage Dynamic Game Process Analysis

In the traditional network attack and defense process, attackers mainly uses network attacks or detection methods to obtain information about the target network, so as to achieve the analysis and penetration of the vulnerability of the target system and finally find the most appropriate network attack strategy to make the optimal network attack benefit [7]. Due to the natural asymmetry of the network attack process, the attacker is able to actively obtain the information about the target network and carry out the network attack at any time, while the defender is often in a passive defense state [5]. In order to change the passive defense situation, the defender takes the initiative to release defense signals so that the network attacker cannot judge the authenticity of the information, thus influencing the attacker's choice of attack strategy and making the network defense passive to active. The attack and defensive game process analysis of defense signals is a key factor in the attacker's analysis to determine the type of defender and the choice of action decisions [13].

In the initial stage of the game, the defender deceives and restricts the attacker by releasing a false defense signal so that the attacker cannot obtain the real state of the target system; the attacker forms an initial a priori probability judgment of the defender by combining the detection behavior and intelligence collection of the target in the early stage, and then forms a posteriori probability of the defense type based on the defense signal released by the defender, inspired by Bayes' law. Define the replicator dynamic equation to select the optimal network attack strategy, thus completing the initial process of the game. After the initial stage of the game, the defender releases the deception signal suitable for this stage again and selects the corresponding optimal defense strategy; the attacker takes the posterior probability of the defense type obtained in the previous stage as the prior probability of this stage, and combines the defense signal received in this stage to derive the posterior probability of the defender type in this stage and selects the corresponding optimal attack strategy. The specific process is shown in Figure 1.

Analyzed from the perspective of the discount factor, the defense signal is strongest in the first phase with a discount factor expressed as $\delta_T = 1$. With the progress of the attack and defense process, the attacker absorbs the learning experience so that the effectiveness of the defense deception signal decreases, with$0 < \delta_i < 1$.

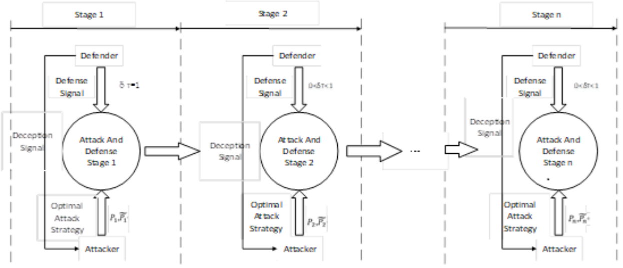


Figure 1: Multi-stage network attack and defense game model

2.2 Construction of Multistage Dynamic Game Model

Definition 1. *Multistage Dynamic Attack And Defense Game Model MDADGM as an octet(N,T,S,θ,M,δ, P,U). The details are as follows:*

1) *$N = (N_A, N_D)$ denotes the set of game participants, where N_A is the defender and N_D is the attacker;*

2) *T denotes the number of stages in the multi-stage game process, i.e., T = 1..., n;*

3) *$S = (S_A, S_D)$ denotes the set of game strategies, where S_D represents the set of strategies of the defender, $S_D = \{S_{D_j} | j = 1, 2..., n\}$;$S_A$ represents the set of strategies of the attacker, $S_A = \{S_{A_i} | i = 1, 2..., m\}$;*

4) *θ denotes the set of game participant types, and $\theta_D = \{\theta_{D_i} | i = 1, 2..., n\}$ indicates the set of defender types, and the defense type is the defender's private information, There is only one type of attacker, namely $\theta_A = (\eta)$ [12];*

5) *M denotes the defense signal space, satisfying $M \neq \oslash$,$M = \{m_k | k = 1, 2, ...\}$. Defenders for the purpose of deterring, deceiving and luring attackers, the network defense signal and the defender's true type may not match;*

6) *δ is the discount factor, which indicates the degree of discount of the defense signal in the game stage T compared with the previous game stage, and satisfies $0 < \delta < 1$;*

7) *P represents the set of game beliefs, where P_A =$\{P_A(\theta_{D_1}), P_A(\theta_{D_2})...P_A(\theta_{D_n})\}$ denotes the attacker's prior probability for different defense types. $\widetilde{P_A}$ =$\widetilde{P_A}(\theta_{D_i} | m_k)$ denotes the posterior probability calculated by the attacker combining the defense signal and the prior probability [2];*

8) *U denotes the set of gain functions of both attackers and defenders, where, $U_A$denotes the attacker's gain function and U_A denotes the defender's gain function.*

2.3 Quantification of Game Gains

The game theory is applied to the analysis of network attack and defense, and the quantification of game gain is the key to determine the accuracy of the final game result.

Definition 2. *Attack cost (AC) represents the economic, time, hardware and software, and human resources spent by the attacker due to the choice of attack strategy.*

Definition 3. *Defense cost (DC) indicates resources such as economy, time, hardware and software equipment, and labor, as well as the impact of degradation of service quality due to the defender's choice of a defense strategy.*

Definition 4. *System damage cost (SDC) represents the damage caused to the system after an attacker initiates an attack.*

Definition 5. *Defense effectiveness ε indicates the effectiveness of defense policy d against attack a. When the attack can be completely blocked, $\varepsilon(a, b) = 1$; when the defense strategy is ineffective, $\varepsilon(a_i, b_j) = 0$; in other cases, $0 < \varepsilon(a_i, b_j) < 1$ [**?**].*

Definition 6. *Signal deception explore (SDE) represents the cost of the defender to deceive the attacker by releasing false signals. If the signal matches the true type of the defender, the SDE is zero. The SDE is relatively quantified according to the gap between the true and false defense information and is expressed as an integer value within the interval $[0, 100]$ [18]. In general, the greater the gap between defender and defensive signal, the more difficult and costly to camouflage. The classification and quantification of SDE are shown in Table 1.*

During network attack and defense, the attacker aims to minimize the cost of attack while maximizing the cost of system loss, while the defender minimizes the cost of defense, the cost of network deception, and the cost of system loss. Using different strategies for offensive and defensive confrontation generates different offensive and defensive payoffs [3]. At each stage, the reward expectations for attackers and defenders are as shown in Equations (1) and (2).

$$U_A = (1 - \varepsilon)SDE + DE - AC \qquad (1)$$

Table 1: Quantitative table of signal spoofing costs

Defender Real Type	Defender Spoofing Signals	SDE Level	Quantitative Allocation
High level defender	High level defender	SDE0	10
	Medium level defender	SDE1	40
	Low level defender	SDE2	70
Medium level defender	High level defender	SDE1	10
	Medium level defender	SDE0	40
	Low level defender	SDE2	70
Low level defender	High level defender	SDE2	70
	Medium level defender	SDE1	40
	Low level defender	SDE0	10

$$U_D = AC - (1-\varepsilon)SDE - DC \tag{2}$$

3 Equilibrium Solution of the Game Process and Optimal Strategy Selection

3.1 Nash Equilibrium

Definition 7. *(Nash Equilibrium) Given a network attack and defense game model MDADGM (N,T,S,θ,M,δ, P,U), S_{A_i} is the attacker strategy and S_{D_j} is the defense system strategy. The strategy (S_A^*, S_D^*) is a Nash equilibrium [14] when and only when the strategy is optimal for both the attacker and the defender, that is, it satisfies:*

$$\forall i, U_A(S_A^*, S_D^*) \geq U_A(S_{A_i}^*, S_D^*) \tag{3}$$

$$\forall j, U_D(S_A^*, S_D^*) \geq U_D(S_A^*, S_{D_j}^*) \tag{4}$$

Theorem 1. *(Nash Equilibrium Existence) Given a network attack and defense game model MDADGM (N,T,S,θ,M,δ, P,U), there exists at least one Nash equilibrium.*

The network attack and defense game model MDADGM (N,T,S,θ,M,δ, P,U) is a matrix-type game whose set of attack and defense strategies S, and gain functions are finite, so the network attack and defense game model MDADGM is a finite game. Nash proves that every finite game has a Nash equilibrium using the immobility theorem, so the network attack and defense game model MDADGM has a stable Nash equilibrium, that is, given a network attack and defense game model, it must be possible to solve its optimal dynamic equilibrium attack and defense strategy.

3.2 Replicator Dynamic Equation

In the set of game participants, when the attacker receives the defense deception signal $m_k \in M$ the posterior probability $P(\theta_{D_i}|m_k)$ is calculated in conjunction with the prior probability, and at the same time, the defender is able to anticipate that the attacker will pick the inferentially dependent optimal attack strategy $S_A^*(m_k)$ based on the network deception signal m_k released by itself, so the defender picks the optimal network defense policy $S_D^*(m_k)$ that maximizes the expected gain of defense. At this point, we can interpret that the probability of the chosen strategy changes for both sides through learning from the previous round of the game.

Theorem 2. *(Replicator Dynamics Equation) The growth rate of the number of individuals choosing strategy S will be less than 0 if the gain obtained by individuals choosing strategy S is less than the average gain of the population and vice versa. The replicator dynamic equation is a dynamic differential equation that reflects the frequency when a strategy is adopted in the set [23]. It is usually expressed by equation 5:*

$$\frac{\mathrm{d}x_i}{\mathrm{d}i} = x_i(U_{S_{A_i}} - \overline{U_A}) \tag{5}$$

where x_i indicates the proportion of strategy S adopted in the set,$U_{S_{A_i}}$ indicates the return when strategy S is adopted, and $\overline{U_A}$ represents the average return.

3.3 Optimal Strategy Selection

In choosing the optimal defense strategy, the equilibrium state of the multistage game is analyzed and the evolutionary equilibrium is solved by using the replicator dynamic equation. The equilibrium solution steps of the evolutionary game can be obtained as follows:

1) Based on the defense signal released by the defender, the probability of the strategy selected by the attacker on the strategy set is p, and the attacker's replicator dynamic equation is:

$$A(P) = \frac{\mathrm{d}p_i(t)}{\mathrm{d}x} = p(U_{S_{A_i}} - \overline{U_A}) \tag{6}$$

Among them, $U_{S_{A_i}} = \sum_{j=1}^{n} q_j a_{ij}$, $\overline{U_A} = \sum_{i=1}^{m} p_i U_{S_{A_i}}$, $U_{S_{A_i}}$ represents the gain function when the attacker chooses the attack strategy a_{ij}.

2) The probability that a defender chooses a strategy on the strategy set is q. The defender's replicator dynamics equation is:

$$D(q) = \frac{\mathrm{d}q_j(t)}{\mathrm{d}t} = q(U_{S_{D_j}} - \overline{U_D}) \tag{7}$$

3) Quantification of Utility Functions
In a multistage game, as the attacker gradually determines the type of the defender, the gain obtained gradually decreases based on the initial gain. Therefore, in this paper, a discount factor δ_T is introduced to calculate the future gain based on the original gain function U. The calculation is as follows:

$$\begin{cases} U_D^k = U_D^k + \sum_{h-1}^{k-1} \delta_T U_D^h \\ U_A^k = U_A^k + \sum_{h-1}^{k-1} \delta_T U_A^h \end{cases} \tag{8}$$

k=$\{1, 2...T\}$.

4) Equilibrium Solution
According to the evolutionary game equilibrium state, the replicator dynamic equations of the attacker and defender should be equal to 0. The game equilibrium solution should satisfy the following equation:

$$\gamma = \begin{bmatrix} maxU_A^k \\ maxU_D^k \\ A(p) = 0 \\ D(p) = 0 \end{bmatrix} \tag{9}$$

k=$\{1, 2...T\}$.

By solving the above equations together, the set of strategy choices under the evolutionary equilibrium state $(S_{D_j}^k, S_{A_i}^k)$ can be obtained. According to the evolutionary game theory, the offensive and defensive strategies at this time are the best choices for both attackers and defenders.

3.4 The Equilibrium Solution Process of Multistage Attack and Defense Game

Suppose the defender type θ_D is divided into high level defender θ_h, medium level defender θ_m, and low level defender θ_l, the corresponding defense signal space M has m_h, m_m,m_l,the defense strategy space is $m^*(\theta)$, the attacker type θ_A =(η), the attack strategy space is $\{S_{A_1}, S_{A_2}, S_{A_3}\}$, the defense type prior probability is P_A, and the attack and defense gains are(U_A, U_D).

1) When T=1, enter the first stage of the attack and defense game
For the initial stage of the attack and defense game, nature selects the defender type with probabilities $P_A(m_h)$, $P_A(m_h)$, $P_A(m_h)$. If a network defender releases a false defense signal m_1, when the attacker receives the signal m_1, it will probabilistically correct the defender type with the posterior probability $\{\widetilde{P_A}(\theta_{D_1}|m_1), \widetilde{P_A}(\theta_{D_2}|m_1), \widetilde{P_A}(\theta_{D_n}|m_1)\}$ to discriminate the defender type as $\{\theta_{D_1}, \theta_{D_2}, \cdots, \theta_{D_n}\}$. Similarly, when the attacker receives the defense signal m_i, the defender type is determined with the posterior probability of $\{\widetilde{P_A}(\theta_{D_1}|m_i), \widetilde{P_A}(\theta_{D_2}|m_i), \widetilde{P_A}(\theta_{D_n}|m_i)\}$. The attack and defense signal game tree is shown in Figure 2.

According to the Nash equilibrium existence theorem, given the participant types and limited attack and defense strategies, both attackers and defenders are maximizing the expected gain as much as possible, and combining Theorem 2, we obtain the replicator dynamic equations $A(p)$ and $D(p)$ for both attackers and defenders. The set of equations is constructed by associating the gain functions to obtain the optimal set of attack and defense strategies $(S_A^*(m), S_D^*(m))$ at this stage. In the initial stage of the game, the attacker cannot analyze the actual type of the defender from the pre-confrontation process, and there is no signal attenuation effect of the false defense signal released by the defender. At this time, $\delta_1 = 1$.

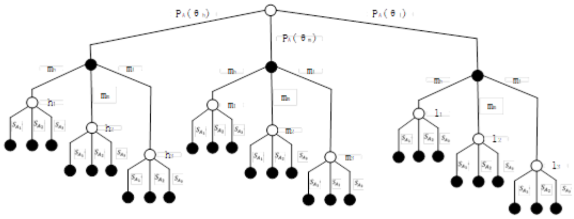


Figure 2: Attack and defense signal game tree

2) When T = 2, enter the second stage of the attack and defense game
By analyzing and comparing the game process and outcome information of the previous stage, the attacker enhances the analysis and screening ability of the defense signals, so from the second stage, $0 < \delta_2 < 1$,the signal discounting effect starts to appear. The attacker takes the posterior probability obtained in the previous stage as the prior probability in this stage, and then combines the false defense signal released in this stage to obtain $\delta_2 \widetilde{P_A}(\theta_{D_2}|m_2))$. The set of most attack and defense strategies for the second stage is again obtained by Nash equilibrium and replicator dynamic equations.

3) When T = n, enter the n stage of the attack and defense game
When the number of game phases T tends to be large or even infinite and the defender releases false signals

more often, $\delta^{T-r-1} \approx 0$, where r represents the number of phases in which the defender releases real defense signals. According to the basic theory of signal game, the game stage becomes a static game with incomplete information, as shown in the game tree in Figure 3. The method of solving the incomplete information static game can be found in the literature [1].

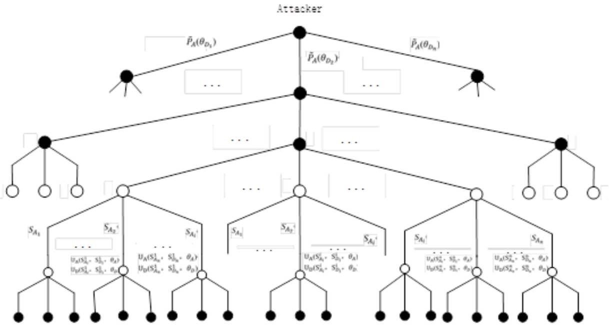

Attacker
$\tilde{P}_A(\theta_{D_1})$
$\tilde{P}_A(\theta_{D_2})$
$\tilde{P}_A(\theta_{D_n})$
...
S_{A_1}
S_{A_2}
S_{A_i}
S_{A_n}
$U_A(S^1_{A_1}, S^1_{D_1}, \theta_A)$
$U_D(S^1_{A_1}, S^1_{D_1}, \theta_D)$

Figure 3: Incomplete information static attack and defense game tree

3.5 Algorithm Design

Based on the analysis process of the above multistage network attack and defense game, the optimal defense strategy selection algorithm is designed as follows.

Algorithm Optimal defense strategy selection algorithm for multi-stage network attack and defense games.

```
Input (N,T,S,$\theta$,M,$\delta$, P,U)
Output Optimal defense strategy $S^K_{D_j}$
1) Initialize MDADGM=(N,T,S,$\theta$,M,$\delta$, P,U), $S_{A_i}$;
2) Build a defensive operational space $S_D$= $\{S^k_{D_j}|1 \le k \le T, 1 \le j \le n\}$, attack Space $S_A$= $\{S^k_{A_i}|1 \le k \le T, 1 \le i \le m\}$ and k denotes the number of game stages;
3) Build a defender type space $\theta_D = (\theta_h, \theta_m, \theta_l)$ and attacker type space$\theta_A = (\eta)$ ;
4) Initialize defender defense signal space $M \neq \oslash$, $M = (m_h, m_m, m_l)$;
5) while ($\exists \varepsilon \&\& S_{D_j} \in S_D \&\& S_{A_i} \in S_A$) { //Calculate earnings
   $$U^k_D = U^k_D + \sum_{h-1}^{k-1} \delta_T U^h_D$$
   $$U^k_A = U^k_A + \sum_{h-1}^{k-1} \delta_T U^h_A$$
   for ($i = 1; i \le m; i++$)// Learn the set of offensive and defensive strategies to construct replicator dynamic equations
   $$A(p) = \frac{\mathrm{d}p_i(t)}{\mathrm{d}t} = p(U_{S_{A_i}} - \overline{U_A});$$
   $$D(q) = \frac{\mathrm{d}q_j(t)}{\mathrm{d}t} = q(U_{S_{D_j}} - \overline{U_D}).$$
6) for($k = 1; k \le T; k++$)// Building a network attack and defense game at different stages
   {if($\delta_T > 0$) // Discount factor to solve the replicator dynamic equation solution
   { $U^k_D = U^k_D + \sum_{h=1}^{k-1} \delta_T U^h_D$
   $U^k_A = U^k_A + \sum_{h=1}^{k-1} \delta_T U^h_A$;}
   while($\exists max U^k_A \&\& max U^k_D$)
   { $A(p) = 0$;
   $D(p) = 0$;}
   if($\delta_T = 0$)
   {// Incomplete information game solve
   }}
7) end for;
8) return $\{S^1_{D_j}, S^2_{D_j}, S^3_{D_j} ... S^n_{D_j}\}$.
```

4 Experiment and Analysis

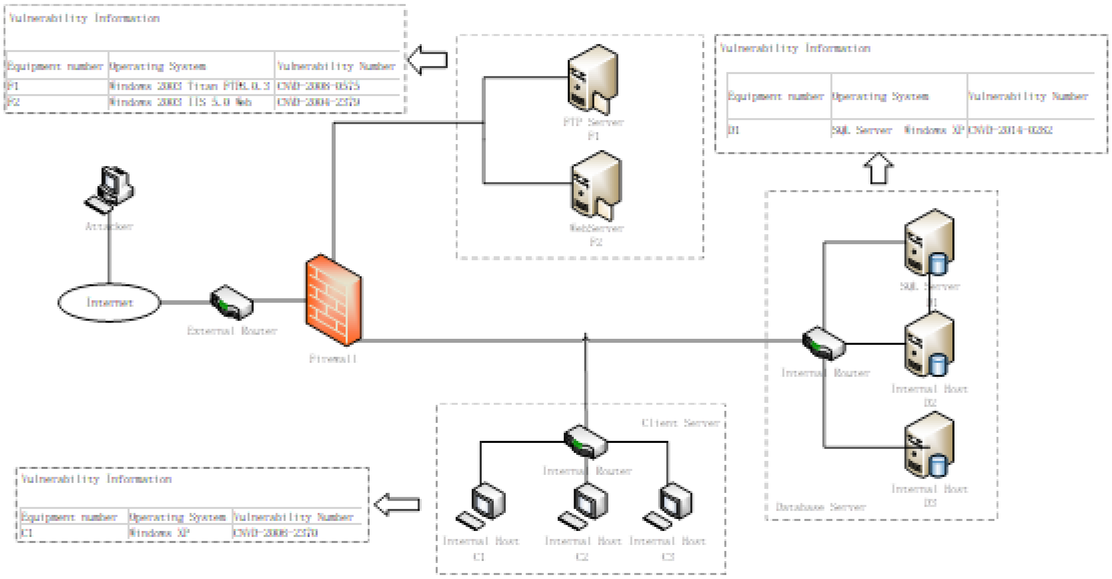
4.1 Description of Experimental Environment

In order to further illustrate the effectiveness of the proposed active defense model and its related algorithms, a simulation experiment is carried out by deploying the experimental scenario shown in Figure 4. The experimental environment mainly consists of network defense devices, network servers, file servers, database servers, client servers, etc., and mainly Windows and Linux operating systems are installed. The security defense rules are to restrict the access requests from hosts outside the system (including attackers), and stipulates that they can only access the network server; the file server and the network server are allowed to access the database server. However, with the help of a multi-step attack process, the attacker is able to gain access to the application server and the database server.

4.2 Calculation of Game Profit

The method for analyzing routing files, vulnerability databases and defense strategies in literature [7] is combined with the information on atomic attacks given in literature [16], as shown in Table 2.

In this network, the SQL database server exists with important data, and a_3 can be considered as the attacker's intrusion intent to set the attack strategy using the scanned vulnerability information, the relationship between vulnerabilities, host and server information, network configuration and other data. To simplify the calculation, we only consider high-level defenders and low-level defenders. The descriptions of the attack and defense


Vulnerability Information

Equipment number	Operating System	Vulnerability Number
F1	Windows 2003 Titan FTP6.0.3	CNVD-2008-0575
F2	Windows 2003 IIS 5.0 Web	CNVD-2004-2379

Vulnerability Information

Equipment number	Operating System	Vulnerability Number
D1	SQL Server Windows XP	CNVD-2014-0262

Vulnerability Information

Equipment number	Operating System	Vulnerability Number
C1	Windows XP	CNVD-2008-2370

Figure 4: Experimental environment network topology

Table 2: Atomic attack strategy

Serial Number	**Description of the Atomic Attack Strategy**	**Permission**	**AL**
a1	Attack SSH on FTP sever	root	10
a2	Unsigned firmware update	root	10
a3	Database rights	root	11
a4	LPC to LSASS process	root	11
a5	Remote code execution	uesr	8
a6	Attack address blacklist	user	9
a7	Oracle TNS listener	user	7
a8	Remote buffer overflow	user	8
a9	Install SQL listener program	access	6

atomic strategies are given by using the Nessus scanning experimental information system with reference to the attack and defense classification method of MIT Lincoln Laboratory [8] and the National Information Security Vulnerability Database are shown in Table 3.

The real defender level is set to low defender and the defensive signal is set to high defender, that is, SDE = 70. The set of attack strategies in this system includes a1 and a2, and the set of defensive strategies includes d1 and d2. According to the calculation formula given in Section 2.3, the quantization of the attack and defensive strategies in the experiment are:

$$\begin{aligned}(a_{11}, d_{11}) &= (45, -45)\\ (a_{12}, d_{12}) &= (68, -68)\\ (a_{21}, d_{21}) &= (22, -22)\\ (a_{22}, d_{22}) &= (45, -45).\end{aligned}$$

When the defender uses strategies d1 and d2, respectively, the expected gain are:

$$\begin{aligned}U_{DS_1} &= pd_{11} + (1-p)d_{21} = -45p - 22(1-p)\\ U_{DS_2} &= pd_{12} + (1-p)d_{22} = -68p - 45(1-p).\end{aligned}$$

The average gain for the defenders is:

$$\begin{aligned}\overline{U_D} &= qU_{DS_1} + (1-q)U_{DS_2}\\ &= q[45p - 22(1-p)] + (1-q)[-68p - 45(1-p)].\end{aligned}$$

For the defensive strategy DS_1, the probability that the defender chooses this strategy is a function of time and its dynamic rate of change can be expressed as:

$$\begin{aligned}D(q) &= \frac{\mathrm{d}q(t)}{\mathrm{d}t}\\ &= q[-45p - 22(1-p) - 45pq + 22q(1-p)\\ &\quad +68p(1-q) + 45(1-q)(1-p)].\end{aligned}$$

Similarly, the expected gains obtained by the attacker using strategies a1 and a2 are:

$$U_{AS_1} = qa_{11} + (1-q)a_{12} = 45q + 68(1-q) \qquad (10)$$

Table 3: Atomic defense strategy

Description of the Atomic Attack Strategy	θ_{D_H}			θ_{D_L}		
	S_{D_1}	S_{D_2}	S_{D_3}	S_{D_4}	S_{D_5}	S_{D_6}
Limit packets from ports	√	×	√	×	√	√
Install Oracle patche	×	√	√	√	×	×
Reinstall listener program	√	√	×	×	×	√
Uninstall delete Trojan	√	√	√	√	√	×
Renew data(root)	√	×	×	×	×	√
Restart database sever	×	√	√	√	×	×
Limit SYN/ICMP packets	√	√	√	×	√	√
Add physical resourse	√	×	√	√	√	×
Repair database	√	√	×	×	×	×

$$U_{AS_2} = qa_{21} + (1-q)a_{22} = 22q + 45(1-q) \qquad (11)$$

The average gain for attackers is:

$$\begin{aligned}\overline{U_A} &= pU_{AS_1} + (1-p)U_{AS_2} p[45q + 68(1-q)] \\ &\quad +(1-p)[22q + 45(1-q)].\end{aligned}$$

The dynamic rate of change of the selection strategy AS_1 is:

$$\begin{aligned}A(p) &= \frac{\mathrm{d}p(t)}{\mathrm{d}t} = p(U_{AS_1} - \overline{U_A}) \\ &= p[45q + 68(1-q) - 45pq - 68p(1-q) \\ &\quad -22q(1-p) + 45(1-q)(1-p)].\end{aligned}$$

Since the optimal defense strategy is generated in equilibrium, this experiment analyzes and solves the equilibrium of the evolutionary game in the final stage. At this point, the effect of the false defense signal released by the defender on the game outcome completely disappears, $\delta_T = 0$, and the objective function is equal to the gain function, that is, $\gamma = [A(p), D(q)] = 0$. Then, the data are re-substituted into the algorithm for validation, and the evolutionary stable strategy is obtained from the image.

4.3 Example Analysis

The attacker's attack and control of the network is reflected in the control of each node component of the network. This example expresses the attacker's invasion status of the network as the attacker's access authority to each node component of the network, and the level of authority is divided into: no access permissions, remote access permissions, local user access permissions, and root access permissions. At the beginning of the game phase, because in the actual network application to define the specific start and end nodes is a rather difficult thing, this paper starts the attacker to launch an attack on different nodes to invade the test, and the defender releases the deception factor after discovering it. Taking the atomic attack strategy a_3 as an example, the attack paths are $S_{A_1} = \{a_1, a_2, a_5\}$, $S_{A_2} = \{a_3, a_5, a_7\}$, $S_{A_3} = \{a_4, a_6, a_9\}$. Assuming that a total of three stages of the games are required, the first and second stages are both offensive and defensive signal games, where the role of false defense signals decreases continuously and the probability of the attacker inferring that the defender is a low defense level increases continuously, and in the third stage, the attacker is able to completely screen false signals, the role of defensive signals disappears, and the attack and defensive games degenerate into incomplete information static games. The game tree is formed as shown in Figure 5.

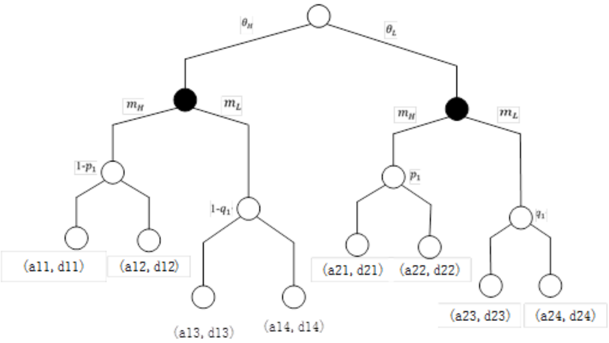


Figure 5: Game tree

Figure 6 shows the analysis of the evolutionary stability strategy of the defender with different probabilities of the attacker's choice of strategy. According to the defender's replicator dynamic equation, the defender's evolutionary stability strategy choice has the following cases.

When p=0, there exists $D(p) = 0$ for any defensive strategy selection probability. when $p \neq 0$, $D(q)$ changes significantly, as shown in Figures 6 and 7. When the slope of the tangent line of the curve is positive, the evolutionary stable strategy of the defender is obtained. Therefore, in Figure 6, when p¿0, $d_1(q = 1)$ is the defender's evolutionary stable strategy. In Figure 7, when p¡0, $d_2(q = 0)$ is the defender evolutionary stable strategy. And since p cannot be less than 0, d_2 is the best defense strategy

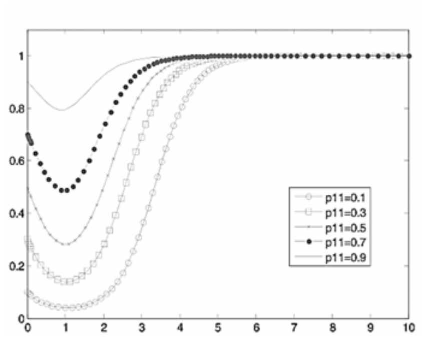


Figure 6: Evolutionary stability strategy selection diagram

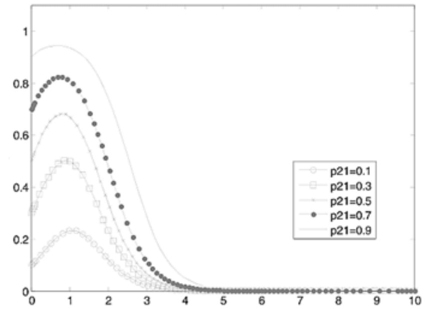


Figure 7: Evolutionary stability strategy selection diagram

at this point. To demonstrate that the analytical results are consistent with the results in the real scenario, we use Matlab to perform simulations to obtain the evolutionary game equilibrium. The results are shown in Figure 8. The horizontal axis represents time and the vertical axis represents the initial value of q. From the figure, we can see that the system finally reaches stability at q=1 regardless of the initial probability p. It is proved that d_1 is the optimal defense strategy solution and the proposed MDADGM model is feasible and effective.

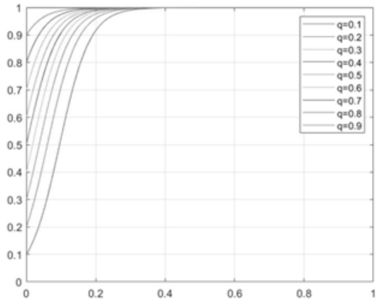


Figure 8: Curve simulation diagram

4.4 Comparison of Methods

The reachability probability of each attribute node in the attack and defense process is the main indicator for network security risk assessment, and the prediction of attack paths can provide network administrators with a basis for intrusion defense. In order to verify the superiority of the model in this paper, in the same network environment, it is necessary to conduct simulation research on different test models. The specific operation process is as follows: Five models were selected for testing: model A in this paper, model B in literature 5, model C in literature 7, model D in literature 16, and model E in literature 23.

Figure 9 shows the running time of the five algorithms in a single stage, three stages, six stages, and 12 stages, respectively, under the same network environment. Models B and C also use the signaling game model to describe the causal relationship between network attack behaviors. However, because their evaluation indicators of vulnerabilities are too single and do not take into account the costs and benefits of attacks, the vulnerability utilization of both does not truly reflect the exploited situation of vulnerabilities in the network. Model D and model E are solved by refined Bayesian equilibrium, but they also lack multiple metrics to evaluate the calculation. As can be seen from Figure 9, the accuracy of the evaluation models in this paper is significantly better than the other four models, because this paper calculates the probability of atomic attacks from multiple indicators and evaluates them more accurately.

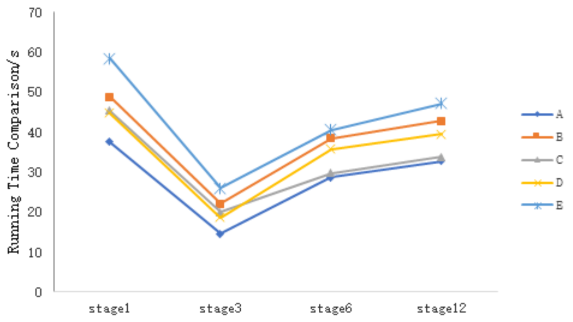


Figure 9: Running time comparison

The total plenum value E(Q) of the minimum critical strategy set of different defense methods is used as an important index to evaluate the defense performance of different methods, where the larger the value of E(Q) is, the better the defense performance of the algorithm is. The specific experimental comparison results are given below.

As can be seen in Figure 10, the total weight of the minimum critical policy set of model A in this paper is higher than that of model B in reference 5, model C in reference 7, model D in reference 16, and model E in reference 23. The E(Q) values of model B in reference 5 and model C in reference 7 are about 65% and 55%, while the E(Q) value of the method in this paper is as high as about 80%. The attack and defense revenue strategy is fully considered in the atomic attack probability, and a discount factor is introduced to balance the error caused by deception signals in the multistage attack and defense process. Experimental research shows that the defense method proposed in this paper has better defense performance.

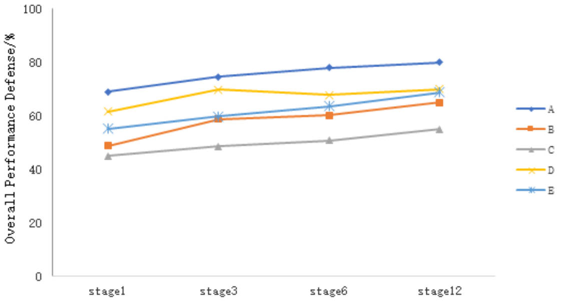


Figure 10: Comparison of overall defense performance results

5 Conclusion

For the traditional intrusion detection, firewall and other passive defense technologies can not cope with the increasingly prominent network security problems. This paper applies noncooperative signal game theory to network attack and defense analysis, makes full use of defense signals to confuse and deter attackers, constructs a multistage network deception game model, and conducts in-depth research on the signal deception mechanism existing in the process of network attack and defense. The research results can provide an effective modeling method for network security situation prediction and theoretical guidance for the application of defense deception in the field of network security. In response to the strong learning ability of attackers and the endless attacks, the next step needs to further consider how to adjust the defense strategy to make the network deception signal achieve better deception effect. And consider more influencing factors to optimize the gain function to make it more suitable for the actual situation.

Acknowledgments

This study was supported by the Natural Science Foundation of Heilongjiang Province:LH2021F030. The authors gratefully acknowledge the anonymous reviewers for their valuable comments.

References

[1] M. Amin, S. Shetty, L. Njilla, D. K. Tosh, C. Kamhoua, "Hidden markov model and cyber deception for the prevention of adversarial lateral movement," *IEEE Access*, vol. PP, no. 99, pp. 1–1, 2021.

[2] J. U. Arshed, M. Ahmed, "Race: Resource aware cost-efficient scheduler for cloud fog environment," *IEEE Access*, vol. PP, no. 99, pp. 1–1, 2021.

[3] X. Chen, X. Liu, L. Zhang, C. Tang, "Optimal defense strategy selection for spear-phishing attack based on a multistage signaling game," *IEEE Access*, pp. 1–1, 2019.

[4] Z. H. Chen, G. Cheng, "Moving target defense technology based on stackelberg-markov non-peer-to-peer three-way game model," *Computer Science*, vol. 43, no. 3, p. 14, 2020.

[5] M. A. El-Zawawy, E. Losiouk, M. Conti, "Vulnerabilities in android webview objects: Still not the end!," *Computers & Security*, vol. 109, p. 102395, 2021.

[6] L. Foschini, V. Mignardi, R. Montanari, D. Scotece, "An SDN-enabled architecture for it/ot converged networks: A. proposal and qualitative analysis under ddos attacks," *Future Internet*, vol. 13, 2021.

[7] X. Hai, Z. Wang, Q. Feng, Y. Ren, H. Duan, "Mobile robot adrc with an automatic parameter tuning mechanism via modified pigeon-inspired optimization," *IEEE/ASME Transactions on Mechatronics*, vol. PP, no. 99, pp. 1–1, 2019.

[8] C. S. Hsieh, M. D. Knig, X. Liu, "A structural model for the coevolution of networks and behavior," *Review of Economics and Statistics*, pp. 1–14, 2020.

[9] Y. Hu, H. Zhang, Y. Guo, T. Li, J. Ma, "A novel attack-and-defense signaling game for optimal deceptive defense strategy choice," *Wireless Communications and Mobile Computing*, vol. 2020, no. 2, pp. 1–10, 2020.

[10] L. I. Huanruo, Y. Guo, S. Huo, G. Cheng, W. Liu, "Survey on quantitative evaluations of moving target defense," *Chinese Journal of Network and Information Security*, 2018.

[11] J. M. Huang, H. W. Zhang, "Optimal defense strategy selection based on improved replication dynamic evolutionary game model," *Communications Journal*, vol. 39, no. 1, pp. 170–182, 2018.

[12] M. F. Hyder, M. A. Ismail, "Securing control and data planes from reconnaissance attacks using distributed shadow controllers, reactive and proactive approaches," *IEEE Access*, vol. PP, no. 99, pp. 1–1, 2021.

[13] B. Liu and H. Wu, "Optimal d-facts placement in moving target defense against false data injection attacksn," *IEEE Transactions on Smart Grid*, vol. PP, no. 99, pp. 1–1, 2020.

[14] J. W. Liu, J. J. Liu, Y. L. Lu, B. Yang, K. L. Zhu, "Optimal defense strategy selection method based on network attack and defense game model," *Computer Science*, vol. 45, no. 6, pp. 117–123, 2018.

[15] Z. Y. Luo, X. Yang, J. H. Liu, R. Xu, "Network intrusion intention analysis model based on bayesian attack graph," *communications journal*, vol. 41, no. 9, pp. 160–169, 2020.

[16] D. Rotondo, H. S. Sánchez, V. Puig, T. Escobet, J. Quevedo, "A virtual actuator approach for the secure control of networked LPV systems under pulse-width modulated DoS attacks," *Neurocomputing*, vol. 365, no. 99, pp. 21–30, 2019.

[17] D. P. Sharma, J. H. Cho, T. J. Moore, F. F. Nelson, H. Lim, S. K. Dong, "Random host and service multiplexing for moving target defense in software-defined networks," in *ICC 2019 - 2019 IEEE International Conference on Communications (ICC)*, 2019.

[18] J. L. Tan, H. W. Zhang, H. Q. Zhang, H. Jin, C. Lei, "Optimal strategy selection method for moving target defense based on markov time game," *Journal of Communications*, vol. 41, no. 1, pp. 42–52, 2020.

[19] M. Tian, Z. Dong, X. Wang, "Reinforcement learning approach for robustness analysis of complex networks with incomplete information," *Chaos Solitons & Fractals*, vol. 144, no. 4-5, p. 110643, 2021.

[20] D. Yao, Z. Zhang, G. F. Zhang, J. X. Wu, "Mvx-cfi: a practical active defense architecture for software security," *Journal of Information Security*, vol. 5, no. 4, p. 11, 2020.

[21] J. X. Zhang, J. M. Pang, Z. Zhang, "Quantification method for heterogeneous web servers with mimicry construction," *software journal*, no. 2, p. 14, 2020.

[22] S. Zhang, "A moving target defense anti-censorship method based on mipv6," *Computer Applications and Software*, vol. 48, no. 6, pp. 2874–2883, 2019.

[23] Z. Zhang, R. Deng, P. Cheng, D. Yau, "Zero-parameter-information data integrity attacks and countermeasures in iot-based smart grid," *IEEE Internet of Things Journal*, vol. PP, no. 99, pp. 2327–4662, 2021.

Biography

Luo Zhiyong biography.Luo Zhiyong, male, was born in Shandong, China in July 1978.He is a professor at the School of Computer Science and Technology, Harbin University of Science and Technology. Research direction: computer network and information security, network optimization. words.

Cao Yutong biography. Cao Yutong, female, was born in Shandong, China in October 1997. She is a postgraduate student in the School of Computer Science and Technology, Harbin University of Science and Technology. Research direction: computer network and information security, network optimization, offensive and defensive games. optimization.

Song Weiwei biography.Song Weiwei, male, was born in May 1998 in Shanxi, China. He is a postgraduate student in the School of Computer Science and Technology, Harbin University of Science and Technology, Research direction: network security.

Li Jie biography.Li Jie is a lecturer in the School of Computer Science and Technology, Harbin University of Science and Technology. Research direction: network security.