SOCMTD: Selecting Optimal Countermeasure for Moving Target Defense Using Dynamic Game

Hao Hu*, Jing Liu, Jinglei Tan, Jiang Liu
State Key Laboratory of Mathematical Engineering and Advanced Computing, Zhengzhou, 450001 China
Zhengzhou Information Science and Technology Institute, Zhengzhou 450001, China
*Corresponding author: Hao Hu [e-mail: wjjhh_908@163.com]

Received February 18, 2020; revised June 20, 2020; accepted August 11, 2020; published October 31, 2020

Abstract

Moving target defense, as a 'game-changing' security technique for network warfare, realizes proactive defense by increasing network dynamics, uncertainty and redundancy. How to select the best countermeasure from the candidate countermeasures to maximize defense payoff becomes one of the core issues. In order to improve the dynamic analysis for existing decision-making, a novel approach of selecting the optimal countermeasure using game theory is proposed. Based on the signal game theory, a multi-stage adversary model for dynamic defense is established. Afterwards, the payoffs of candidate attack-defense strategies are quantified from the viewpoint of attack surface transfer. Then the perfect Bayesian equilibrium is calculated. The inference of attacker type is presented through signal reception and recognition. Finally the countermeasure for selecting optimal defense strategy is designed on the tradeoff between defense cost and benefit for dynamic network. A case study of attack-defense confrontation in small-scale LAN shows that the proposed approach is correct and efficient.

Keywords: moving target defense, dynamic defense, signal game, optimal countermeasure, cost and benefit

http://doi.org/10.3837/tiis.2020.10.013 **ISSN : 1976-7277**

1. Introduction

The global network security attack-defense competition has reached an unprecedented intensity[1]. All kinds of network attacks are becoming more and more rampant [2, 3]. Network attackers are constantly developing new attack strategies. Among them, the technology of moving target attack (MTA) is one of the most popular attack methods for attackers [4]. It uses various uncertain attack methods to hide the intention of attack, and tries to avoid the detection mechanism of traditional network defense [5]. Because the traditional network defense mechanism can't predict the attacker's next attack action accurately, the MTA technology gradually obtains the competitive advantage in the network attack-defense game, which not only poses a great security threat to the network space, but also produces a high defense cost[6].

At the same time, the network security strategy has undergone an evolution from passive defense to active defense, and the emerging moving target defense (MTD) technology [7] has become a new method to balance the competitive environment of network security. It protects the network space by introducing dynamics, randomness and heterogeneity. It aims to break the static characteristics of the network system by using the dynamic transformation of the attack surface and present an unpredictable network state to the attacker, so as to prevent the malicious behavior of the attacker and make it harder for attackers to attack successfully.

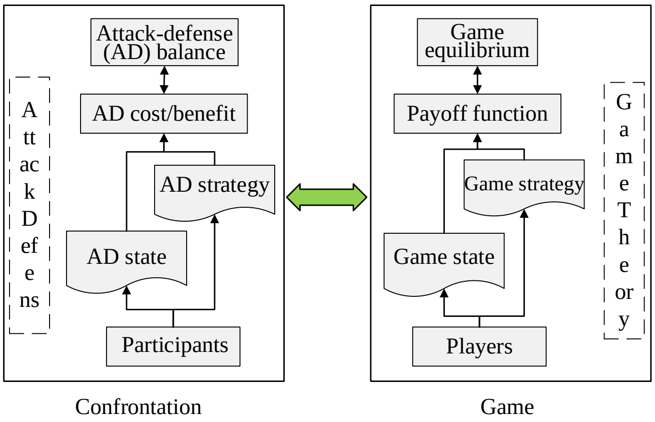

Fig. 1. Corresponding relationship between attack-defense and game theory

Due to the complex nonlinear and uncertain characteristics of network attack-defense [8], the traditional behavior modeling methods based on system identification or engineering experience cannot accurately model the entity game behavior of attack-defense players accurately. Game theory is a mathematical theoretical tool to study the interdependence and competition among decision-makers [9, 10]. It It fits well with the nature of cyberspace confrontation in terms of object opposition, strategic dependence, non-cooperative relationship and dynamic multi-stage as shown in **Fig. 1**.

MTA and MTD players select the optimal strategy according to the cost-benefit of attack-defense to maximize the attack-defense payoff, which has a non-cooperative relationship. In the process of defense confrontation, MTA attempts to control the attack surface of the system through various attack means, expanding the exposure scope of the attack surface, and preparing for the subsequent continuous attack [11]. On the contrary, MTD controls the attack surface of the system through dynamic, randomized and diversified methods, transferring or reducing the attack surface of the system, so as to reject the attack action of MTA. Therefore, both two sides of attacker and defender are opposite. The selection of the optimal strategy for both sides of the MTA and MTD depends not only on themselves but also on the adversaries. Both sides of the MTA and MTD have the dependent strategies. The relationship noncooperation, target opposition and strategy dependence in the process of confrontation are consistent with the game theory. Game theory plays an important role in the study of selecting the optimal strategy of MTA.

The rest of this paper is organized as follows. Section 2 gives an overview of some related works. Section 3 describes some preliminaries as well as the model of the attack-defense confrontation to demonstrate the signal game between attacker and defender. Section 4 specifies the attack-defense strategy generation and payoff quantification. Section 5 shows how to calculate the perfect Bayesian equilibrium between attack-defense in Section 4. The optimal strategy selecting algorithm is designed in Section 6. The corresponding experiments are presented to demonstrate the feasibility and flexibility of the proposed approaches in Section 7. Finally, Section 8 summarizes this paper and discusses some future works.

2. Related works

The optimal game refers to the selection and arrangement of candidate strategies. The existing methods of MTA decision-making are mainly divided into the following two categories.

(1) Attack graph based approaches

S. JHA et al., “[12] defined a certain number of sets of atomic attacks as attack critical sets. By removing the atomic attacks in the attack critical set, the attacker was prevented from reaching the target attack state node from the initial attack state node, and the minimum critical set was defined as the set with the least elements of all key sets. The minimum critical set was calculated by using state attack graph and greedy algorithm. Poolsappasit et al., “[13] took the cost and benefit factors of security control into account in the security risk assessment and introduced a Bayesian attack graph. He gave a multi-target defense strategy calculation method, which tries to find a balance between the cost of security control and the overall benefit. Chen et al., “[14] proposed a probability attack graph (MPAG) model of defense strategy for internal threats. He used greedy algorithm to solve the optimal defense strategy of internal threats.

(2) Game theory based approaches

The nature of cyberspace security confrontation is the interdependence and mutual restriction of attack-defense strategies. Whoever has a good strategy will be invincible in the confrontation. The effectiveness of the defense strategy depends not only on the behavior of the defender, but also on the attack strategy. Using game theory to study attack -defense confrontation and its optimal defense strategy has natural connectivity. Lye et al., “[15] first introduced game theory into the field of network security. Based on the game theory, [15] described and analyzed the network attack-defense confrontation in detail, established the network attack-defense game model, and obtained the optimal strategies of both sides by calculating the Nash equilibrium. Jiang et al., “[16] systematically gave the classification of attack-defense strategies and their cost / benefit analysis, and put forward an algorithm for selecting the optimal defense strategies based on the two-player non-cooperative game model. However, this method is based on the complete knowledge of each other’s strategies. It is a complete information game essentially, and cannot describe the difference of strategy set under different attack capabilities. At the same time, aiming at the problem that static game is difficult to reflect the deduction of attack-defense state. The researchers put forward a stochastic attack-defense game model [17], which describes the dynamic deduction of network security evolution. It employs matrix game model and Markov decision-making process. Based on this model, attack strategy prediction and optimal active defense decision are made. Tan et al., "[18] put forward a dynamic game model for complete information attack-defense, while it can't deal with the dynamic change of attack intention and behavior in static game. Tan introduced 'virtual node' to convert the network attack-defense map into attack-defense game tree, and gave an algorithm to solve the optimal attack-defense strategy set by using non-cooperative dynamic game. However, the assumption of complete information of both sides of attacker and defender is hard to satisfy the practical application.

Colbaugh et al., "[19] used game theory and machine learning to analyze the MTD strategies for attackers with self-learning ability, and pointed out that uniform randomization can reduce the probability of attackers to predict defense strategies and achieve the best defense effect. However, the model focuses on the attackers with prediction ability, ignoring the diversity and dynamics of MTD strategy itself. At the same time, the model is a single-stage game model, which is difficult to describe the dynamic attack-defense deduction process. Manadhata [20] first presented the concept of attack surface transfer and made a quantitative analysis of it. The attack-defense interaction in MTD environment is modeled as a complete information static game, and the optimal MTD strategy is regarded as a balance between system security and availability. However, the assumption of complete information of both sides of attacker and defender is not consistent with the actual attack-defense confrontation, and it is a single-stage static game actually. Zhu et al., "[21] proposed a multi-stage defense mechanism based on feedback information architecture, and constructed a two person zero sum game model, focusing on the analysis of the cost of the defender changing the attack surface and the cost of the attacker mining system vulnerabilities and changing the attack vector. However, the quantification of the benefits of the attack-defense strategies is not comprehensive enough to reflect the characteristics of the MTD strategy. Carter et al., "[22] focused on analyzing the dynamic platform migration strategy, modeling it as a 'leader follower' incomplete information game model. He constructed the threat model as an attacker to continuously break *k* platforms with vulnerabilities, and analyzed the optimal dynamic platform switching strategy under the static and adaptive attack conditions respectively. Among them, the optimal strategy of the former is to maximize platform differences, while the optimal strategy of the latter is to maximize platform differences and minimize the balance between attackers mining new vulnerabilities. However, the threat model is too rigorous and does not consider the cost of attack-defense. Winterrose et al., "[23] put forward a multi-stage game model based on the dynamic migration of operating system. He pointed out that when the defender uses diversity strategy, the attacker should select the operating system with the least similarity for vulnerability mining. On the contrary, when the defender uses randomness strategy, the attacker should select the operating system with the most similarity for vulnerability mining from an attacker's perspective. Prakash et al., "[24] used the improved Flipit game model to analyze a variety of different combinations of attack-defense strategies. He pointed out that the effectiveness of MTD defense depends on the ability of attack detection, but his game process only considers probe attack and reimage defense, which is difficult to reflect the general law of MTD defense. Jones et al., "[25] proposed a novel MTD game model PLADD (probabilistic learning attacker, dynamic defender) to improve Flipit. He focused on the analysis of the probability density of time to success and the impact of different types of defense strategies on game payoff. However, the assumption of the attacker's complete information and the defender's zero cognitive ability is not reasonable. Vadlamudi et al., “[26] modeled the MTD defense in the web application environment as Bayes Steinberg game model. Based on the game equilibrium solution, he analyzed the impact of the importance of the system configuration vulnerability. He further studied the sensitivity of different attacker types on the attack-defense benefits so as to find the optimal transfer strategy. However, this model is a single-stage game model and does not quantify the attack-defense benefits. Moreover, it is only suitable for the web software stack environment and has poor universality. Maleki et al., “[27] put forward the MTD game model based on Markov decision process, which analyzed the single target IP hop and multi-target IP hop strategies, and showed that multi-element hop can effectively improve the defense benefits. To sum up, the existing model construction only aims at the specific scene of IP hop, which has poor universality. Meanwhile, no specific strategy selection algorithm is given, which is not very instructive.

3. Construction of attack-defense dynamic game model

The network attack-defense process based on dynamic network defense is a non-cooperative, incomplete information, multistage, dynamic game process. The signal game is a dynamic game model of incomplete information with information transmission mechanism. The behavior of the first action signal sender has the function of transmitting information to the second action signal receiver. Therefore, this paper uses the theory of signal game to describe and analyze the process of attack-defense confrontation in the dynamic defense environment of the network, in which the attacker is the signal sender, the defender is the signal receiver, and the attack strategy can be regarded as the signal sent by the attacker. The defender can infer the attacker type and gradually modify the inference of the attacker type through the analysis of the attack behavior, and then select the attacker type to take the best defense strategy.

Definition 1 Dynamic attack-defense signal game model can be expressed as 9-tuples, in which

① $N=(N_D,N_A)$ is the player set, N_D is the defender and the signal sender, N_A is the attacker and the signal receiver.

② $\Theta=(\Theta_D,\Theta_A)$ is the player type set, the types of defenders belong to private information, which can be divided into several types according to their defense capabilities, namely $\Theta_D=\left(\theta_i \middle| i=1,2,...,n\right)$, attacker type is $\Theta_A=(\eta)$.

③ M is the set of defense signal, $M\neq\varnothing$ and $M=(m_j \mid j=1,2,...)$. The defender selects and releases signals according to the set signal release mechanism. For the convenience of expression, the name of the signal is consistent with the name of the defender type. For the purpose of deterring, cheating and trapping attackers, the defense signals and defender types may be different.

④ T is the number of stages in a multi-stage game, $T=\{1,2,...,n\}$, the game process in the current stage is represented by $G(T)$.

⑤ δ_T is a signal attenuating factor, which describes the attenuation degree of defense signal effect with the increase of game stages, $0\leq\delta_T\leq 1$.

⑥ $S=(D,A)$ is the strategy space of attacker and defender, $D=\{d_g \mid g=1,2,...\}$ and $A=\{a_h \mid h=1,2,....\}$ denote the defense and attack strategies respectively.

⑦ P_A is the prior probability judgment of the attacker on the defender type, where $P_A = (p_A(\theta_1), p_A(\theta_2),..., p_A(\theta_n)) = (\gamma_1,...,\gamma_n)$.

⑧ P'_A is the posterior probability set of attackers, $P'_A = P'_A(\theta_i \mid m_j) = (\mu_1,...,\mu_n)$ is the posterior probability of defender type calculated by Bayes rule after the attacker observes the defense signal m_j.

⑨ $U = (U_D, U_A)$ is the payoff function of defender and attacker.

For the fake defense signal, when $T = 1$ and $\delta_1 = 1$, in the first stage of attack-defense game, the signal does not decay. The deterrent, deception and inducement effect of the false defense signal works best. When $1 < T < n$, we can derive $0 < \delta_T < 1$. In particular, if $T < T'$, we can derive $0 < \delta_{T'} < \delta_T < 1$. As the game carrying forward, the signal decays gradually and the degree of attenuation is increasing, the effect of deterrent, deception and inducement decreases as well. When $T = n$ and $\delta_n = 0$, the impact of fake defense signals on the attack-defense game disappears. At this time, the influence of false defense signal on attack-defense game disappears and degenerates into incomplete information static game $G(T)$. The signal attenuation factor directly affects the posterior probability of selecting different candidate strategies when calculating the game equilibrium.

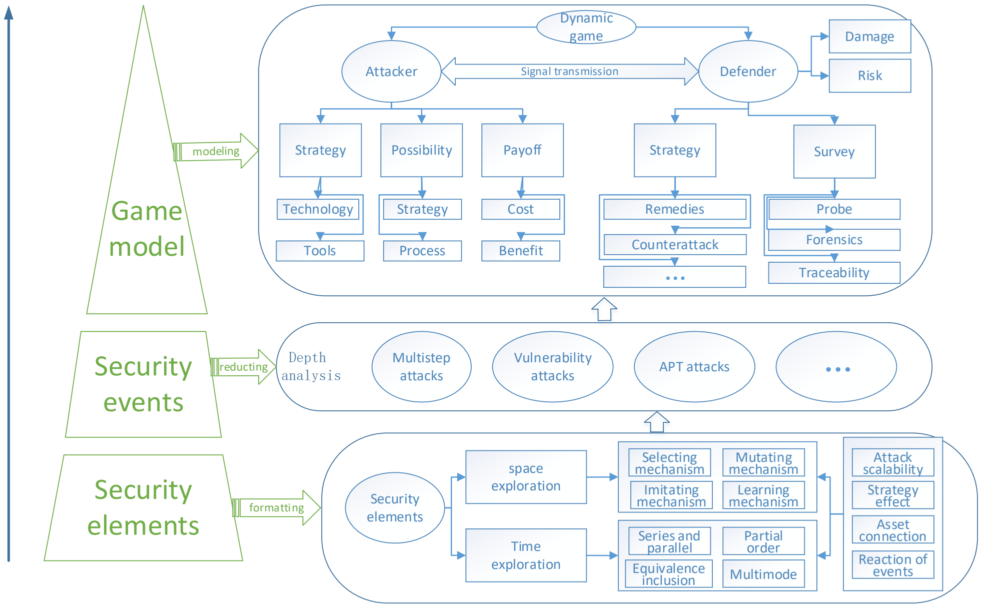

Fig. 2. Attack-defense modeling

The ideal of attack-defense modeling is illustrated in **Fig. 2**. Based on the attack chain or fragment from the two dimensions of time and space, we use honeypot data, log information and other quantitative measures to measure the attack capability, opportunity and risk reaction in the actual system. We infer and identify the attack theme under multiple correlation factors. Through describing the causal relationship of vulnerability exploitation in the two dimensions of time and space in the multistage attack- defense process, we can get obtain candidate strategy analysis and extraction.

4. Attack-defense strategy analysis and payoff quantification

Network state and attack-defense action are important parts of the game model. One of the key points to build the dynamic game model is the extraction of network states and attack-defense actions. For typical attack-defense scenarios, we select protégé knowledge map tool to construct attack-defense knowledge map and use its powerful visual analysis function for information extraction and knowledge reasoning. Using the powerful dependency relationship between graph nodes, we model the game between attacker and defender. Combining with necessary manual assistance, we can infer network states and extract possible attack-defense actions shown in **Fig. 3**, wherein the state set of attack-defense game model is composed of attack-defense knowledge map nodes extraction, the action set of attack-defense is extracted from the edge of attack-defense knowledge map to realize deduction and visualization of decision-making.

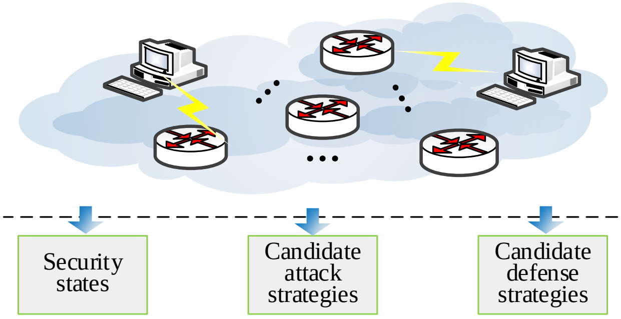

Fig. 3. Inference of candidate strategies and security states by attack graph

4.1. MTA strategy

MTA system has been gradually developed and continuously improved, and common strategies of MTA are shown in **Table 1**.

Table 1. MTA strategy classification

MTA strategy	Details
Multimode MTA	Transform malware signature
self-modified MTA	Dynamic transformation of malware code
Confused MTA	Hide malicious activity
Self-encrypted MTA	Transform malware signatures and hide malicious code and data
Anti-virtual machine / anti sandbox MTA	Change behaviors in tracking environment to avoid automatic forensics analysis
Anti-debugging MTA	Change behavior in tracking environment to avoid automatic / manual investigation
Target vulnerability exploit MTA	Transform parameters and signatures to avoid automatic / manual investigation
Behavior change MTA	Wait for real user activity before execution

The above effective MTA methods win the asymmetric attack advantage for attackers, which make the traditional defense technology into a passive and adverse situation. The attacker knows his own attack target, attack time, and attack mode, while the defender is in an uncertain state. He can only use a lot of cost, time and resources to avoid any attack detection and intrusion activities that the attacker may launch. Therefore, there is asymmetry between the defender and attacker. For example, multimode MTA can effectively avoid the feature detection of IDS. On the one hand, multimode MTA uses multiple encryption keys to generate different instances of the same malware. Because the new instance has a new unknown static signature, the signature based anti-malware defense is invalid. On the other hand, multimode MTA payloads (code and data) are encrypted, bypassing the defense's deep static analysis. Multimode MTA complicates the attack detection process by changing the codes in memory.

4.2. MTD defense strategy

The defense and security committee of the White House gave the concept of moving target in the cyberspace security research progress report [11] in 2012, moving target is a technical mean that can reduce the attacker's advantage and increase the system flexibility through transformation in multiple dimensions organized in **Table 2**. In 2014, Department of Defense Intelligence Intelligence of U.S. defined MTD as a defense means to create, analyze, evaluate and deploy diversified and time-varying mechanisms and strategies, so as to increase the complexity and cost of attack implementation, limit and reduce the vulnerability exposure and the probability of being attacked, and improve the flexibility of the system [12].

Table 2. MTD strategy classification

Classification		Details
System layer MTD	Software-MTD	Transform applications, operating systems, data
	Hardware-MTD	Transform processor
Network layer MTD	MAC-MTD	Change MAC address
	IP-MTD	Change IP address
	Procotol-MTD	Change protocol
	Path-MTD	Change path
	OS-MTD	Change operating system
	Finger-MTD	Change fingerprint
	Port-MTD	Change port

4.3. Strategy payoff quantification

Ref. [9, 16, 18] has studied and summarized the classification of attack-defense strategies, the quantification of strategy cost / benefit, and the calculation of strategy payoff, but has not quantified the role of defense signals. This section is improved based on the existing researches.

Definition 2 [16] Defense Cost *DC* includes attack surface shifting cost, negative impact cost and attack identification cost. Attack surface transfers cost refers to the cost of changing system resources during attack surface transfer, which is related to the changed dimension of system attack surface; negative impact cost refers to the loss caused by changing network resources during attack surface transfer, which results in the network unable to work normally or the quality of service degradation. This cost is related to the transfer cycle of defense strategy, that is, the shorter the cycle, the larger negative impact cost. The cost of attack identification is the cost of detecting and identifying different types of attackers. The higher the ability level, the more difficult the attackers are to be detected and identified, the higher the cost.

$$DC(d_k,\theta_i,t^h)=ASSC(d_k)+NC(d_k,t^h)+AIC(d_k,\theta_i)$$

Definition 3 [16] Attack Cost *AC* refers to the cost that an attacker pays to discover and utilize the system resources on the attack surface, usually including the time cost, software and hardware resource cost, professional knowledge cost and risk cost of discovering and invading the system resources. Attackers with different ability levels pay different costs when using the same system resources. The higher the ability level is, the lower the cost is, which is expressed as $AC(a_j,\theta_i)$.

Definition 4 [16] System Damage *SD* refers to the damage to the system caused by the attacker's use of the system resources on the attack surface. It is usually described by the target resource criticality *C* (criticality), attack lethality *Al* (attack lethality) and security attribute damage *SAD* (security attribute damage). The system damage can be denoted as $SD(a_j)$. The longer the attacker controls the system resources on the attack surface, the greater the system loss.

Definition 5 [16] Defense Effectiveness *DE* refers to the impact on the attacker's payoff. After the defender implements the defense strategy and the attacker shifts the attack surface. If the attacker has obtained the control permission of a system resource, the attack surface transfer reduces the existing permission of the attacker. If the attacker has not obtained the control permission of a system resource, the attack surface transfer increases the difficulty of the attacker to obtain the control permission.

Definition 6 Signal Faking Cost *SDE* means the cost that the defender spends in order to conceal the real information of his defense ability. The defender releases the faked signal deception and induces the attacker in the attack-defense game. If the signal is the same as the real defense type, we assign SDE as zero. Through the gap between the real level and the faked level of the faking signal, we quantify the SDE by integer value in the interval [0, 100]. Then the calculation formula of attack payoff is

$$U_A(m_j,d_g,a_h,\theta_i)=\sum_{g,h}SDC(d_g,a_h)-AC_h$$

The defense payoff

$$U_D(m_j,d_g,a_h,\theta_i)=\sum_{g,h}SDC(d_g,a_h)-DC_g-SDE$$

The effect of defense strategies belonging to the same level is basically the same. If there are *g* defense strategies in total, it can be assumed that the defender selects the *g*-th strategy with equal probability $\beta=\frac{1}{g}$ to get the average defense payoff

$$U_D(\theta_i)=\sum_{g=1}^{k}\beta\cdot U_D(m_j,d_g,\mathrm{A}_h,\theta_i)$$

Definition 7 Attack Reward *AR* refers to the payoff obtained by the attacker using the system resources on the attack surface.

$$AR(\theta_i,a_j,d_k)=\varphi(t^h,t^*)(1+g_a(t^h,t^*))SD(a_j)+DC(d_k,\theta_i,t^h)-AC(a_j,\theta_i)$$

$$\varphi(x,y)=\begin{cases}1,x\ge y\\0,x<y\end{cases},\quad x\in R,y\in R$$

$g_a(t^h,t^*)$ indicates the impact of the attacker's control time over the system resources on the attack payoff and is a monotone increasing function. $t^h - t^*$ indicates the time when the attacker controls the system resources on the attack surface.

Definition 8 Defense Reward *DR* refers to the payoff gained by the defender by transferring the system resources of the attack surface.

$$DR(\theta_i,a_j,d_k) = g_d(t^h,t^*)SD(a_j) + AC(a_j,\theta_i) - DC(d_k,\theta_i,t^h)$$

$g_d(t^h,t^*)$ indicates the effect of the defender's control time over the system resources on the defense payoff and is a monotone increasing function。

There are two characteristics of defense effect: (1) the effectiveness of the same defense strategy to different attack strategies is different. (2) The effectiveness of the same defense strategy is different when different types of attackers implementing the same attack strategy, which is related to the attacker's ability level. For example, for the IP address hopping strategy, suppose that it takes 5 seconds for a low-level attacker to detect a new IP address, and only 1 second for a high-level attacker. If the IP address hopping period is 2 seconds, then a low-level attacker will never be able to detect a new IP address assuming that there are enough IP addresses.

5. Perfect Bayesian equilibrium

As shown in **Fig. 4**, the preconfigured strategy is implemented at the beginning of the confrontation to achieve the proactive confrontation. In the process of the confrontation, the strategy deployment is modified based on the confrontation signal to improve the pertinence of the confrontation. In one confrontation cycle, the implementation cost and benefit of the confrontation strategy are modified and updated by studying the change of network state to evaluate the effectiveness of the confrontation.

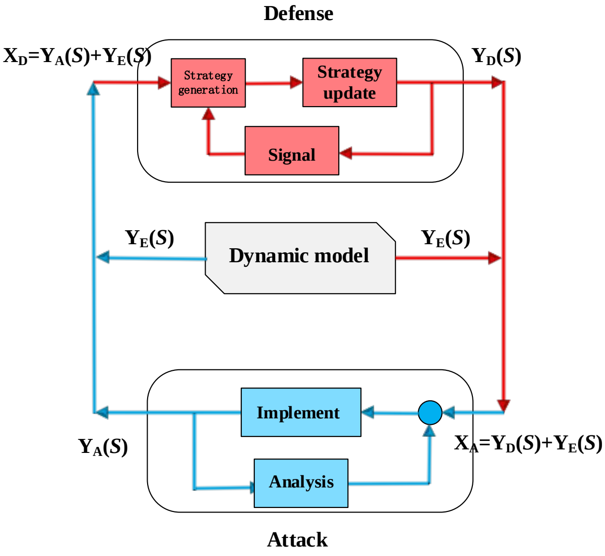

Fig. 4. The operation mechanism of game model

In the next confrontation cycle, the optimal strategy is selected according to the benefit of the new attack-defense strategy and the trend of the confrontation state, so as to ensure the continuous evolution of the confrontation. Therefore, through the attack-defense game model, we can predict confrontation process. The closed-loop feedback control ensures the continuous and effective implementation of attack and the convergence of defense effectiveness.

Definition 9 Perfect Bayesian equilibrium is a combination of attack-defense strategy combination and a posteriori probability, which meets the following conditions.

① The defender N_d has an initial inference about the type of attacker N_a, and the inference value is the conclusion obtained after observing the attack strategy a_j, which is $p_{ij} = p(\theta_i \mid a_j)$.

$$\forall \theta, p(\theta_i \mid a_j) \geq 0, \sum_{\theta_i \in \Theta_a} p(\theta_i \mid a_j) = 1$$

$$d^*(a_j) = \arg\max_{d_k \in D} \sum_{\theta_i \in \Theta_a} p(\theta_i \mid a_j) f_d(\theta_i, a_j, d_k)$$

② Given inference $p_{ij} = p(\theta_i \mid a_j)$ and attack strategy a_j, strategy $d^*(a_j)$ selected by the defender should be optimal.

$$d^*(a_j) = \arg\max_{d_k \in D} \sum_{\theta_i \in \Theta_a} p(\theta_i \mid a_j) f_d(\theta_i, a_j, d_k)$$

③ Given the optimal strategy $d^*(a_j)$ of the defender, the strategy $a^*(\theta_i)$ selected by the attacker is optimal.

$$a^*(\theta_i) = \arg\max_{a_j \in A} f_a(\theta_i, a_j, d^*(a_j))$$

④ For any $a_j \in A$, if $\exists \theta_i \in \Theta_a$ making $d^*(a_j) = d(a_j)$, then in the information set of the corresponding a_j of defender, the next inference value of the player is obtained by Bayesian law

$$p(\theta_i \mid a_j) = \frac{p(a_j \mid \theta_i)}{\sum_{\theta_i \in \Theta_a} p(a_j \mid \theta_i)}$$

The algorithm of refining Bayesian equilibrium in dynamic game model of network dynamic defense is as follows.

Step1. The defender establishes a posterior probability inference on each information set $p(\theta_i \mid a_j)$.

Step2. The defender infers the best defense strategy $d^*(a_j)$.

Step3. Calculate perfect Nash equilibrium by $p(\theta_i \mid a_j)$. In the second step of the game, the defender observes the attack strategy a_j selected by the attacker in the first step. Under the assumption that inferring the type of attacker $p(\theta_i \mid a_j)$, the defender chooses to maximize his expected payoff $d^*(a_j)$. By $\arg\max_{d_k \in D} \sum_{\theta_i \in \Theta_a} p(\theta_i \mid a_j) f_d(\theta_i, a_j, d_k)$, the defender obtains the optimal strategy.

Step4. Infer the attack strategy that the attacker may adopt $a^*(\theta_i)$.

In the first step of the game, the attacker of type θ_i predicts that the best strategy action $d^*(a_j)$ of the defender. The attacker chooses $a^*(\theta_i)$ to maximize his expected payoff. By

$\arg\max_{a_j \in A} f_a(\theta_i, a_j, d^*(a_j))$, then the attacker's best strategy $a^*(\theta_i)$ can be obtained.

Step5. Calculate perfect Bayesian equilibrium

By using the subgame perfect Nash equilibrium $(a^*(\theta), d^*(a))$ derived from the former two steps ② and ③, the defender's inference $p^*_{ij} = p^*(\theta_i | a_j)$ to the attacker type satisfying the Bayesian rule is obtained. If $p(\theta_i | a_j)$ and $p^*(\theta_i | a_j)$ do not conflict, then $(a^*(\theta), d^*(a), p^*(\theta_i | a_j))$ is perfect Bayesian equilibrium.

6. Countermeasure: selecting the optimal defense strategy

The best defense strategy selecting algorithm based on signal game model is designed. It is based on the study of dynamic game model and its perfect Bayesian equilibrium solution in the network dynamic defense environment.

Algorithm 1. Optimal MTD defense strategy selection algorithm

Input: Security elements of network attack-defense game

Output: Best defense strategy

```
Initial signal game model $ND^2MSGM = (N, \Theta, A, D, P, P', U)$;
Initialize attacker type space $\Theta_a = \{\theta_i, 1 \le i \le n\}$;
Initialize attack policy space $A = \{a_j, 1 \le j \le g\}$;
Initialize defense strategy space $D = \{d_k, 1 \le k \le h\}$;
Calculate attack payoff $U_a = f_a(\theta_i, a_j, d_k), 1 \le i \le n, 1 \le j \le g, 1 \le k \le h$;
Calculate defense payoff $U_d = f_d(\theta_i, a_j, d_k), 1 \le i \le n, 1 \le j \le g, 1 \le k \le h$;
Define perfect Bayesian equilibrium solution function $EQ()$
{
    Attacker release signal $a_j, 1 \le j \le g$;
    The defender builds a posteriori inference to the attacker $p(\theta_i | a_j)$;
    The defender selects the best defense strategy $d^*(a_j)$;
    Optimal attack strategy $a^*(\theta_i)$;
    A posteriori inference for the type of attacker satisfying Bayes rule $p^*(\theta_i | a_j)$;
    if $p^*(\theta_i | a_j)$ and $p(\theta_i | a_j)$ do not conflict
    then Obtain perfect Bayesian equilibrium solution $(a^*(\theta), d^*(a), p^*(\theta_i | a_j))$;
    return Optimal defense strategy $d^*(a_j)$;
}
```

```
for（t=1, t<=s, t++）/* s is the total stage number of dynamic game
  {
    if (t=1)
      EQ ();
    else
    {
      Calculate $p(\theta_i | a_j(T_t), h_a(T_t))$;
      Update $p(\theta_i | a_j)$;
      EQ ();
    }
  }
end for
```

If the type space of the attacker is *m*, then the number of information sets is *m*, and the time complexity of post inference is $O(m)$. The defender selects the optimal defense strategy and the attacker predicts the optimal attack strategy as the subgame. The average time complexity is $O((\max(g,h))^3)$. Hence, the time complexity of posterior probability inference for solving the attacker type satisfying the Bayesian law is $O(m)$. Therefore, the time complexity of the algorithm is $O(2(\max(g,h))^3 + 2m)$. The space storage of the algorithm is to save the payoff value and equilibrium solution of the attack-defense strategy. Hence the spatial complexity is $O(\max(g,h)m)$.

7. Experiments and analysis

The simulation experiment is carried out by deploying the network information system topology as shown in the **Fig. 5** to verify the proposed game model and strategy selection approach.The topology consists of the web server in the DMZ zone and the database server and the file server in the trusted zone. The rule of firewall is that the user in this Internet can only access the web server, while cannot access the servers in the trusted zone directly.

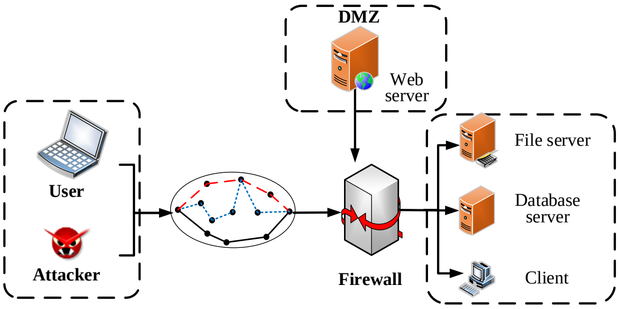

Fig. 5. The experiment network topology

Assuming that the attacker starts from the Internet and has root privilege on the host, the target is to obtain the important information of the database server. The vulnerability of each host in the network is shown in the **Table 3** below.

Table 3. Vulnerability Information

Host	OS	Vulnerability	CVE
Web server	Linux	Apache Chunked Enc.	CVE-2017-11176
Database server	Linux	Local buffer overflow	CVE-2018-18344

An attacker can only gain access privileges to access the web server. Using web server as a springboard, the attacker can gain access to the database through a series of atomic attacks in further. The atomic attack information is shown in **Table 4**.

Table 4. The atomic attack information

No.	Atomic attack	Type	*AL*
a_1	Ftp.rhost attack on Web Sever	User	6
a_2	Wu-Ftpd Sockprintf ()	User	7
a_3	Ftp.rhost attack on Database Sever	User	7
a_4	Apache chunk overflow	Root	8
a_5	Local buffer overflow	Root	9

To illustrate the effectiveness of the model, we assume that the attacker has two types: strong attacker and weak attacker. After the attack graph is generated using Mulval tools, we can extract the candidate MTA strategies using the edges (vulnerability exploitation) of attack graph. The attack strategies adopted by attackers with different attack ability levels are also different. The following **Table 5** shows the strategies adopted by different types of attackers.

Table 5. Candidate MTA strategy

Attacker type	Signal type	Strategy
Strong	High-level threat	A_1 { a_3 $a_{4,}$ a_5}
		A_2 :{ $a_{1,}$ $a_{4,}$ a_5}
		A_3 :{ $a_{1,}$ $a_{5,}$ a_4}
Weak	Low-level threat	A_4 :{ $a_{1,}$ $a_{2,}$ a_4}
		A_5 :{ $a_{1,}$ a_5}
		A_6 :{ $a_{3,}$ a_5}

The defense strategy selected by the defender is often a set of various defense actions. Based on the defense cost, benefit and expert knowledge, we can get the set of candidate defense strategies in **Table 6**. In order to illustrate the our design motivation concisely, we select two defense strategies for example analysis.

Table 6. Candidate MTD strategy

Atomic action of defense	Strategy	
	D_1	D_2
Patch Ftp .rhost on Web Sever		√
Patch Apache		√
Patch Ftp.rhost on Database Sever	√	
Close rsh on Database Sever	√	
Patch Ssh on Ftp Sever		√
Close Ssh on Ftp Sever	√	

The payoff value of the different candidate attack-defense strategies can be calculated by Section4, and the game tree is illustrated in **Fig. 6** as below.

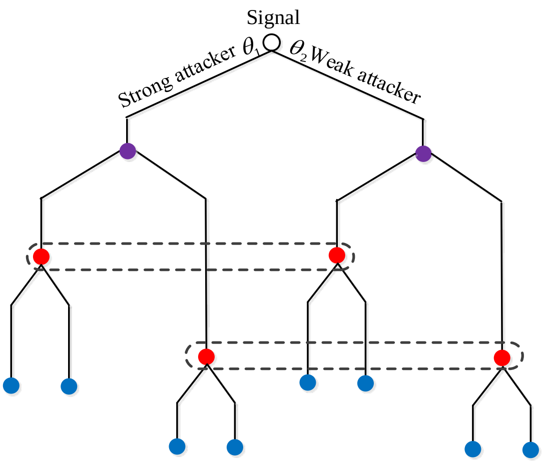

Fig. 6. Signal attack-defense game tree

(1) The first stage of attack-defense

The posterior probability constructed on different information sets can be inferred as $p_1^* = 0.42, q_1^* = 0.52$. The equilibrium solutions are as follows.

- If $p_1 > p_1^*, q_1 > q_1^*$, we can get the perfect Bayesian equilibrium $EQ_1 =$ $[(m_H, m_H) \to (A_1, A_1),\ p_1 = 0.63,\ q_1 = 0.55]$. The best strategy is to implement mixed strategy {D_1, D_2} with probability {q_1=0.55, 1-q_1=0.45}.
- If $p_1 > p_1^*, q_1 < q_1^*$, we can get the perfect Bayesian equilibrium $EQ_2 = [(m_H, m_H)$ $\to (A_1, A_2),\ p_1 = 0.63,\ q_1 = 0.55]$. The best strategy is to implement mixed strategy {D_1, D_2} with probability {q_1=0.55, 1-q_1=0.45}.
- If $p_1 < p_1^*, q_1 > q_1^*$, we can get the perfect Bayesian equilibrium $EQ_3 =$ $[(m_L, m_L) \to (A_2, A_2),\ p_1 = 1,\ q_1 = 0]$. The best strategy is to implement pure strategy {D_1, D_2} with probability {q_1=0, 1-q_1=1}.

- If $p_1 < p_1^*, q_1 < q_1^*$, we can get the perfect Bayesian equilibrium $EQ_4 =$ $[(m_L, m_L) \to (A_2, A_2),\ p_1 = 0.63,\ q_1 = 0.55]$. The best strategy is to implement mixed strategy {*D*1, *D*2} with probability {*q*1=0.55, 1-*q*1=0.45}.

Therefore, when (p_1^*, q_1^*) and (p_1, q_1) are not in conflict, the perfect Bayesian equilibrium of first stage is the mixed equilibrium *EQ*2. It can be seen that the optimal defense strategy is to select the low defense level and send the high defense signal in the first stage initialy. It shows that the defender uses the signal to show the defense ability beyond the actual situation. The defender cheats and induces the attacker to reduce the possible losses and plays an active defense effect. When the attacker observes signal m_H, he mistakenly thinks that the probability of the defender being defense type (θ_H, θ_L) is $(1 - p_1, p_1) = (0.37, 0.63)$. It shows that the deception signal has an effect on the attacker, which leads to an increase in the probability that the attacker infers the defender as a low-level defense type. Hence, the attacker increases the wrong probability of inferring the defender type.

(2) The second stage of attack-defense

In the first stage, the attacker gets the posterior probability $(0.37, 0.63)$ of defender type. The posterior probability can be used to infer of defender type in the first stage of game. At the same time, with the result of the feedback of game process in the first stage, the attacker enhances the measuring ability of fake defense signals. The signal decreases in this stage and assigns as $\delta_2 = 0.5$. The game results of the second stage of attack-defense are as follows.

The posterior probability constructed on different information sets is $p_2^* = 0.59, q_2^* = 0.43$, then we can derive the optimal strategies at the second stage as follows.

- If $p_2 > p_2^*, q_2 > q_2^*$, we can get the perfect Bayesian equilibrium $EQ_5 =$ $[(m_H, m_H) \to (A_1, A_1),\ p_2 = 0.92,\ q_2 = 0.38]$. The best strategy is to implement mixed strategy {*D*1, *D*2} with probability {*q*2=0.38, 1 - *q*2=0.62}.
- If $p_2 > p_2^*, q_2 < q_2^*$, we can get the perfect Bayesian equilibrium $EQ_6 = [(m_H, m_H)$ $\to (A_1, A_2),\ p_2 = 0.92,\ q_2 = 0.38]$. The best strategy is to implement mixed strategy {*D*1, *D*2} with probability {*q*2=0.38, 1 - *q*2=0.62}.
- If $p_2 < p_2^*, q_2 > q_2^*$, we can get the perfect Bayesian equilibrium $EQ_7 =$ $[(m_L, m_L) \to (A_2, A_2),\ p_2 = 1,\ q_2 = 0]$. The best strategy is to implement pure strategy {*D*1, *D*2} with probability {*q*2=0, 1 - *q*2=1}.
- If $p_2 < p_2^*, q_2 < q_2^*$, we can get the perfect Bayesian equilibrium $EQ_8 =$ $[(m_L, m_L) \to (A_2, A_2),\ p_2 = 0.92,\ q_2 = 0.38]$. The best strategy is to implement mixed strategy {*D*1, *D*2} with probability {*q*2=0.38, 1 - *q*2=0.62}.

When (p_2^*, q_2^*) and (p_2, q_2) do not conflict with each other, the perfect Bayesian equilibrium of the second stage is *EQ*5. After the attacker observes the signal m_H, the posterior probability of defender type (θ_H, θ_L) is $(1 - p_2, p_2) = (0.08, 0.92)$. It indicates that the deception and camouflage of the defense signal decrease, and the attacker increases the probability that the defender is low ability type.

From the analysis of the above experiments, we can obtain the following conclusions.

(1) Effective defense signal can improve the proactive defense ability and help the defender to maximize the defense effect. From the first two game stages, the optimal defense strategy is to select low-level defense and release high defense signal. Using the defense signal, the

defender can deceive the attacker and induce the attacker. The attacker makes a wrong judgment on the type of real defender and the actual defense strategy. This will bring a lot of defense payoffs to defenders. When the real defense ability is low, the defender can obtain a stronger defense deterrent effect with a smaller defense input.

(2) The effect of defense signal will decline rapidly in the dynamic process of multi-stage attack-defense. Therefore, the proposed mechanism of defense signal must be used together with other defense approaches. In the process of multi-stage dynamic attack-defense game, with the development of attack-defense game, the attacker can enhance the ability to distinguish the fake defense signal by analyzing the former game results. Therefore, in the game, the defender should recognize the limitation of deception of defense signal. The defense signal should be taken as an emergency defense mean to gain time for adjusting defense system and improving defense effect. Defense signal can enhance defense efficiency by cooperating with other defense means.

8. Conclusions

This paper is aimed at the problem of selecting the optimal strategy for MTD, we analyze the characteristics of attack-defense under the environment of MTD. The dynamic attack-defense game model for network dynamic defense is constructed by using the signal game. On this basis, the perfect Bayesian equilibrium is calculated. We analyze the changing mechanism of defense signal effect and put forward the signal attenuation factor to represent the defense signal in different stages of the game. In addition, we analyze the characteristics of the attack-defense game model for MTD environment and put forward the algorithm of selecting the optimal MTD strategy from the candidate strategies set. The general rules of MTD using equilibrium solution of the game are summarized. This paper has a certain practical significance for the selection of the optimal MTD strategy, and provides a useful direction for improving the payoff of MTD. The future work is to study the compatibility and flexibility of honeypot, deception defense, trusted computing and other security means to improve the overall proactive defense ability.

Acknowledgment

This work was supported by the National Key Research and Development Program of China (2019QY1301, 2019QY1303, and 2018YFB0803602), the National Natural Science Foundation of China (61902427, 61702508). In particular, we thank Prof. Yuling Liu from Institute of Information Engineering Chinese Academy of Sciences for his beneficial help in improving the paper.

References

[1] Alshamrani A, Myneni S, Chowdhary Ankur, et al., “A survey on advanced persistent threats: techniques, solutions, challenges, and research opportunities,” *IEEE Communications Surveys and Tutorials*, 21(2), 1851-1877, 2019. Article (CrossRef Link)

[2] Chen C, Hu J, Qiu T, et al., “CVCG: Cooperative V2V-aided transmission scheme based on coalitional game for popular content distribution in vehicular ad-hoc networks,” *IEEE Transactions on Mobile Computing*, 18(12), 2811-2828, 2019. Article (CrossRef Link)

[3] Chen C, Liu L, Qiu T, et al., “ASGR: an artificial spider-web-based geographic routing in heterogeneous vehicular networks,” *IEEE Transactions on Intelligent Transportation Systems*, 20(5), 1604-1620, 2019. Article (CrossRef Link)

[4] Sengupta S, Chowdhary A, Sabur A, et al., "A survey of moving target defenses for network security," *arXiv: Cryptography and Security*, 2019. Article (CrossRef Link)

[5] Song F, Zhou Y T, Wang Y, et al., "Smart collaborative distribution for privacy enhancement in moving target defense," *Information Sciences*, 479, 593-606, 2019. Article (CrossRef Link)

[6] Zhuang R, Deloach S A, Ou X, "Towards a theory of moving target defense," in *Proc. of the First ACM Workshop on Moving Target Defense*, pp. 31-40, 2014. Article (CrossRef Link)

[7] Cho J H, Sharma D P, Alavizadeh H, et al., "Toward proactive, adaptive defense: A survey on moving target defense," *IEEE Communications Surveys & Tutorials*, 22(1), 709-745, 2020. Article (CrossRef Link)

[8] Pan K, Palensky P, Esfahani P M, "From static to dynamic anomaly detection with application to power system cyber security," *IEEE Transactions on Power Systems*, 35(2), 1584-1596, 2020.Article (CrossRef Link)

[9] Hu H, Liu Y, Zhang H, et al., "Optimal network defense strategy selection based on incomplete information evolutionary game," *IEEE Access*, 6, 29806-29821, 2018. Article (CrossRef Link)

[10] Hu H, Liu Y, Chen C, et al., "Optimal decision making approach for cyber security defense using evolutionary game," *IEEE Transactions on Network and Service Management*, 17(3), 1683-1700, 2020.Article (CrossRef Link)

[11] Zimba A, Chen H, Wang Z, "Bayesian network based weighted APT attack paths modeling in cloud computing," *Future Generation Computer Systems*, 96, 525-537, 2019. Article (CrossRef Link)

[12] Jha S, Sheyner O, and Wing J, "Two formal analyses of attack graphs," in *Proc. of the 15th IEEE Workshop on Computer Security Foundations*, 49-63, 2002. Article (CrossRef Link)

[13] Poolsappasit N, Dewri R, and Ray I, "Dynamic security risk management using Bayesian attack graphs," *IEEE Transaction on Dependable and Secure Computing*, 9(1), 61-74, 2012. Article (CrossRef Link)

[14] Chen X J, Shi J Q, Xu F, "Algorithm of optimal security hardening measures against insider threat," *Chinese Journal of Computer Research and Development*, 51(7), 1565-1577, 2014.

[15] Lye K W and Wing J M, "Game strategies in network security," *International Journal of Information Security*, 4(1), 71-86, 2005.Article (CrossRef Link)

[16] Jiang W, Fang B, Tian Z H, et al., "Evaluating network security and optimal active defense based on attack-defense game model," *Chinese Journal of Computers*, 32(4), 817-827, 2009. Article (CrossRef Link)

[17] Jiang W, Fang B, Tian Z H, et al., "Research on defense strategies selection based on attack-defense stochastic game model," *Chinese Journal of Computer Research and Development*, 47(10), 1714-1723, 2010. Article (CrossRef Link)

[18] Tan J, Lei C, Zhang H, et al., "Optimal strategy selection approach to moving target defense based on Markov robust game," *Computers & Security*, 8(5), 63-76, 2019. Article (CrossRef Link)

[19] Colbaugh R and Glass K, "Predictability-oriented defense against adaptive adversaries," in *Proc. of the IEEE International Conference on Systems, Man, and Cybernetics*, 2721-2727, 2012. Article (CrossRef Link)

[20] Manadhata P K, "Game theoretic approaches to attack surface shifting," *Moving Target Defense II*, pp. 1-13, 2012.Article (CrossRef Link)

[21] Zhu Q and Basar T, "Game-theoretic approach to feedback-driven multi-stage moving target defense," *Decision and Game Theory for Security*, 246-263, 2013. Article (CrossRef Link)

[22] Carter K M, Riordan J F, and Okhravi H, "A game theoretic approach to strategy determination for dynamic platform defenses," in *Proc. of the First ACM Workshop on Moving Target Defense*, 21-30, 2014. Article (CrossRef Link)

[23] Winterrose1 M L, Carter1 K M, Wagner N, et al., "Adaptive attacker strategy development against moving target cyber defenses," *Advances in Cyber Security Analytics and Decision Systems*, pp 1-14, 2020. Article (CrossRef Link)

[24] Prakash A and Wellman M P, "Empirical game-theoretic analysis for moving target defense," in *Proc. of the Second ACM Workshop on Moving Target Defense*, 57-65, 2015. Article (CrossRef Link)

[25] Jones S, Outkin A, Gearhart J, et al., *Evaluating moving target defense with PLADD*, Sandia National Laboratories, United States, 2015.Article (CrossRef Link)

[26] Vadlamudi S G, Sengupta S, Taguinod M, et al., "Moving target defense for web applications using Bayesian stackelberg games," in *Proc. of the 15th International Conference on Autonomous Agents and Multiagent Systems*, 1377-1378, 2016. Article (CrossRef Link)

[27] Maleki H, Valizadeh S, Koch W, et al., "Markov modeling of moving target defense games," in *Proc. of the 2016 ACM Workshop on Moving Target Defense*, 81-92, 2016. Article (CrossRef Link)

Hao Hu received his M.S. degree and Ph.D. degree in the Zhengzhou Information Science and Technology Institute in 2015 and 2018, respectively. He is currently a lecturer with the Zhengzhou Information Science and Technology Institute. His research interests include system security and cryptography.

Jing Liu is currently studying the B.S. degree in the Department of Computer Science and Information Engineering at Zhengzhou Information Science and Technology Institute. Her research interest is network security.

Jinglei Tan received his B.S. degree in the Ningxia University in 2017. He is currently studying the Ph.D. degree in the Department of Computer Science and Information Engineering at Zhengzhou Information Science and Technology Institute. His research interests include game theory and proactive defense.

Jiang Liu received his Ph.D. degree in the Department of Computer Science and Information Engineering at Zhengzhou Information Science and Technology Institute. His research interests include moving target defense and proactive defense.