WF-MTD: Evolutionary Decision Method for Moving Target Defense Based on Wright-Fisher Process

Jinglei Tan, Hui Jin, Hao Hu, Ruiqin Hu, Hongqi Zhang and Hengwei Zhang

***Abstract*—The limitations of the professional knowledge and cognitive capabilities of both attackers and defenders mean that moving target attack-defense conflicts are not completely rational, which makes it difficult to select optimal moving target defense strategies difficult for use in real-world attack-defense scenarios. Starting from the imperfect rationality of both attack-defense, we construct a Wright-Fisher process-based moving target defense strategy evolution model called WF-MTD. In our method, we introduce rationality parameters to describe the strategy learning capabilities of both the attacker and the defender. By solving for the evolutionarily stable equilibrium, we develop a method for selecting the optimal defense strategy for moving targets and describe the evolution trajectories of the attack-defense strategies. Our experimental results in our example of a typical network information system show that WF-MTD selects appropriate MTD strategies in different states along different attack paths, with good effectiveness and broad applicability. In addition, compared with no hopping strategy, fixed periodic route hopping strategy, and random periodic route hopping strategy, the route hopping strategy based on WF-MTD increase defense payoffs by 58.7%, 27.6%, and 24.6%, respectively.**

***Index Terms*— Moving target defense; Wright-Fisher process; evolutionary strategy; attack-defense conflict**

I. INTRODUCTION

The growing spread, complexity, and scale of network information systems have led to an increase in the diversity of attacks as well as the overall threat to network systems. Traditional static defense technologies are increasingly unable to adapt to attack-defense conflict scenarios [1, 2] in this environment. The moving target defense, a new active defense method, aims to increase the attack difficulty by changing the attack surface, i.e. the exposed interfaces to the network. The goal is not to build a perfect network security system but to make the network system dynamic, thereby increasing the uncertainty and unpredictability attackers face [3, 4].

Most of the existing MTD research focuses on the design and formulation of strategies while ignoring the selection of strategies. The research on the selection of MTD strategies lacks quantitative analysis and a decision-making framework [5]. Game theory is a decision theory that studies the direct interaction between decision-makers and incorporates non-cooperative relationships, strategic dependence, and conflicting objectives. These features are consistent with the needs of moving target attack-defense [6–8], which explains why game theory has attracted widespread research attention for moving target defense [9–13].

Most of the existing research on MTD strategy selection is based on the assumption of a completely rational game, which requires both attackers and defenders to act rationally with the optimal strategic decision as the goal. While considering the opponent's strategy, pursuing the maximization of payoffs and making the most favorable strategic decision during the game process can provide a basic theoretical framework for the selection of attack-defense strategies for moving targets, which has high theoretical research significance and value. However, the assumption of complete rationality is based on many difficult-to-achieve pre-conditions, such as perfect rational reasoning, recognition and judgment ability, memory and calculation ability, and analytical ability. It also assumes that everyone involved in the game knows how to maximize their own gains. If they do not meet any of these conditions, they are not considered completely rational. In the real world, neither attackers or defenders can meet the requirements of complete rationality; therefore, the application of the completely rational game model in the real world has limitations, which greatly reduces the effectiveness of the completely rational game model and method. For the moving target attack-defense conflict process in the real world, due to the differences in security knowledge, skill level, and experiences of the attackers and defenders, they have different cognitive abilities, and their rationality and evolutionary learning ability are also limited. Both attackers and defenders cannot fully grasp all the attack-defense strategies and payoffs, which will lead to distortion in the analysis and modeling of the moving target attack-defense behaviors, thus affecting the correctness and practicability of the MTD strategy selection method. To summarize, the assumption of complete rationality is too strict, which is not

Jinglei Tan and Hui Jin contributed equally to this work.

Jinglei Tan is with State Key Laboratory of Mathematical Engineering and Advanced Computing, Zhengzhou, 450001 China.

Hui Jin is with State Key Laboratory of Mathematical Engineering and Advanced Computing, Zhengzhou, 450001 China.

Hao Hu is with State Key Laboratory of Mathematical Engineering and Advanced Computing, Zhengzhou, 450001 China.

Ruiqin Hu is with State Key Laboratory of Mathematical Engineering and Advanced Computing, Zhengzhou, 450001 China.

Hongqi Zhang is with State Key Laboratory of Mathematical Engineering and Advanced Computing, Zhengzhou, 450001 China.

Hengwei Zhang (Corresponding author) is with State Key Laboratory of Mathematical Engineering and Advanced Computing, Zhengzhou, 450001 China.

conducive to the application of traditional game theory in the actual strategy selection of MTD.

Winterrose et al. [10] proposed a dynamic game model using incomplete information, modeling platform layer, moving target defense (MTD) migration, and zero-day vulnerability attack-defense conflict and compared the defensive effectiveness of random MTD and diverse MTD against adaptive attackers. The experimental results showed that the diversity-maximizing MTD strategy was more conducive to resisting short-term attacks. In 2019, Sengupta et al. [11] proposed a dynamic placement model for a network detection system using Markov game theory for the MTD strategy in cloud environments and deployed the model in a real-world cloud environment. In 2020 Sengupta et al. [12] proposed a moving target defense model using Bayesian Stackelberg Markov games and introduced a multiple agent reinforcement learning algorithm to solve the optimal MTD strategy. They proposed a Bayesian Strong Stenberg Q learning method with unknown transition probabilities and rewards. Finally, they proved that the BSS-Q learning strategy performed significantly better than existing benchmarks in web applications and cloud networks, the two different MTD scenarios.

Li et al. [13] proposed a spatiotemporal decision-making model for moving target defense using Markov Stackelberg game theory for adaptive and complex attackers and introduced a relative value iteration algorithm to determine the optimal MTD strategy. Their experiments showed that their method performed significantly better than the Bayesian Stackelberg game decision-making strategy and uniform random strategy. Zhang et al. [14] introduced random games with incomplete information into the MTD decision-making process. Based on the historical collection of attack-defense strategies and the strategy selection distribution, both the attacker and defender dynamically adjust the attack-defense payoffs and then use the Nash-Q learning algorithm to select optimal MTD strategies.

MTD game decision-making in different attack-defense scenarios has also garnered research attention. Lakshminarayana et al. [15] investigated coordinated cyber and physical attacks on power grids with an emphasis on minimizing the defensive cost. They determined the optimal link set of perturbance of the grid's transmission line reactance based on zero-sum game theory and proposed a robust hybrid strategy using a moving target defensive solution incorporating a reinforcement learning algorithm. He et al. [16] proposed a differential game-based IP address hopping model for the internet of vehicles, adaptively adjusting the IP hopping frequency of roadside units (RSUs) and maximizing its defensive rewards. Niu et al. [17] modeled the attack-defense interaction between the linear time-invariant (LTI) system controller and the attacker as a Stackelberg game for false data injection attacks on the LTI system, analyzed single- and multiple-stage optimal attacks, and obtained the optimal detection threshold for the controller by solving a convex optimization problem. Their experimental results showed that the method performed better than an attack detector designed with fixed parameters.

Most of the existing MTD game strategy selection methods based on complete rationality use the Markov game model to solve the Nash equilibrium strategy. In this paper, considering the repeatability and bounded rationality of attack-defense games, the bounded rational MTD game strategy selection method is used to solve the evolutionary stable equilibrium strategy based on the Wright-Fisher process. The attackers and defenders can adjust the update strategy according to the Wright-Fisher mechanism. While finding the optimal strategy to maximize payoffs, such adjustments can also ensure the robustness of the strategy, so that attack-defense game can gradually evolve to a stable state. Figure 1 shows the comparison between the completely rational MTD strategy selection method and bounded rational MTD game strategy selection method.

Although investigations into optimal strategy selection for moving target defenses using game theory have achieved measurable results, the existing studies have the following shortcomings:

(1) The completely rational game assumptions of the attackers and defenders are difficult to apply in realistic moving target defensive situations with actual attack-defense processes. The limitations of the assumptions reduce the value and practicality of the results.

(2) Existing studies into the selection of optimal strategies for moving target defenses using multiple stage games do not introduce learning mechanisms and lack descriptions of the dynamic learning process and learning effects of strategies.

To tackle these problems, we propose a moving target defense strategy evolution method using the Wright-Fisher process. First, we explain the moving target defense strategy from the dynamicity-redundancy-diversity perspective. Then, we use a stochastic process to describe the changes in the system state as the game stage progresses and an evolutionary game to describe the dynamic processes of the attacker and defender by observing opponents' behaviors to adjust their own strategies. We combine the stochastic process with an evolutionary game model to construct a Wright-Fisher process-based moving target defense strategy evolution model (WF-MTD) that extends the attack-defense game to multiple states and agents. We design the algorithm for solving the WF-MTD equilibrium strategy to select the optimal WF-MTD defense strategy.

The main contributions of our study are as follows:

1) We abstract the moving target defense strategy into the dynamicity-diversity-redundancy transformation of network vulnerability, called the DDR-MTD strategy. We then construct an evolutionary game model that scientifically describes the attacker's and defender's behaviors of a moving target. By using the Wright-Fisher process to characterize the evolutionary learning mechanism of moving target attack-defense strategies and by quantifying the degrees of rationality of the attacker and the defender of the moving target to distinguish different players, this model ensures good scalability and is applicable to diverse attacker behaviors.

2) We present the objective function of WF-MTD to find an optimal MTD strategy. The optimal MTD strategy is given by solving the evolutionarily stable equilibrium. Compared with the static Nash equilibrium, the strategy reveals the attack-defense strategies of the network system at different moments in time, depicts the evolutionary trajectories of different attack-defense strategies, improves the efficiency of dynamic analysis

in defense decision-making, and enhances the ability to predict situations.

3) When the proposed WF-MTD method is applied to the route hopping strategy selection in actual scenarios, it can effectively select the optimal hopping path, and the average delay is only 0.078 ms. Compared with different route hopping strategies, namely no hopping strategy, fixed periodic route hopping strategy, and random periodic route hopping strategy, the strategy based on the WF-MTD method increased the defense payoffs by 58.7%, 27.6%, and 24.6%, respectively.

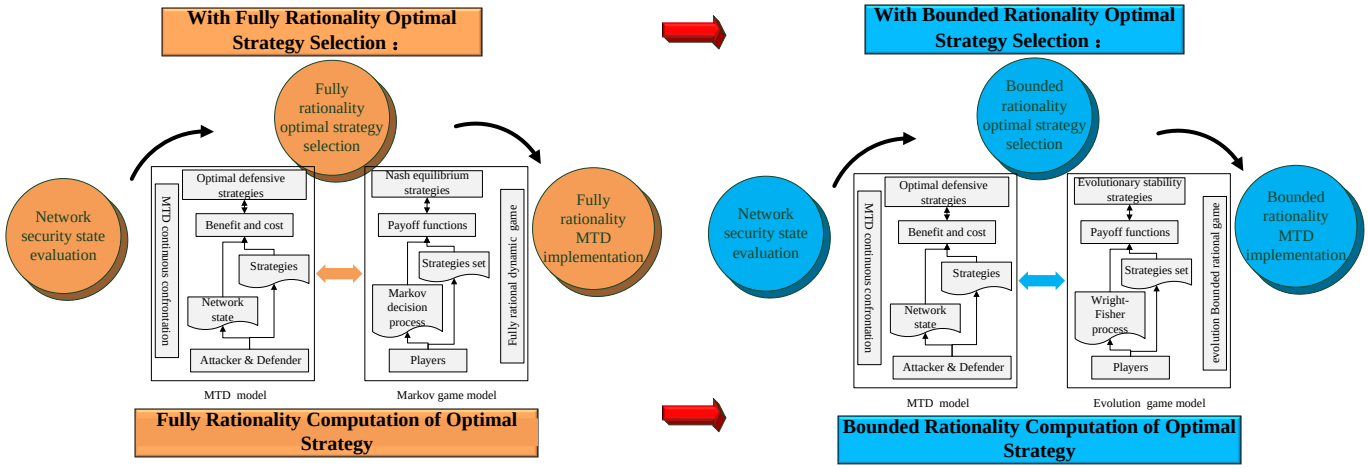

Figure 1. Comparison of selection methods of full rationality and bounded rationality MTD game strategies

II. MOVING TARGET DEFENSE AND STOCHASTIC EVOLUTIONARY GAME ANALYSIS

A. *MTD strategy analysis*

According to the core idea of moving target defense, the moving target defense strategy has three main characteristics: dynamicity, diversity, and redundancy:

Dynamicity: Dynamicity refers to dynamic network system configurations, including but not limited to IP hopping [18], host randomization [19], network topology reconfiguration [20], and VM migration [21]. Its core idea is to randomize the configuration of a network system to increase the uncertainty for the network attacker, making the information collected in the attacker's reconnaissance and identification process useless. As an example, a client-server communication process might change the IP address of the communication link in real time, making it difficult for malicious adversaries to obtain the real IP addresses of the legal client and the protected server cluster, and thus more difficult to launch effective attacks.

Diversity: Diversity refers to diversified network system configurations, including but not limited to multiple variants of servers, programming languages, operating systems, and hardware [22, 23], and thus providing alternative systems with the same function and structure. This greatly improves the resilience and fault tolerance of network systems in the face of attacks and forces the attackers to spend more time and effort to address new variants. For example, a clients might communicate with different computers in a protected server cluster composed of different operating systems such as Windows, Linux, and Unix. Doing so requires the attacker to scan all the variants to formulate the next attack plan effectively, greatly increasing the time and space cost for the attacker to detect targets.

Redundancy: Redundancy refers to redundant network system configurations, including but not limited to increasing duplications of server components: hardware, operating systems, software, services, and components [24, 25] to increase reliability and availability of the network system. In the case of client-server communication, the server cluster might consist of *n* components. Once a cluster is damaged by an attack, existing services can be migrated to another server cluster, greatly increasing service availability while ensuring security.

B. *Moving target attack-defense game analysis*

The moving target attack-defense process has characteristics of competition, bounded rationality, and multiple stage evolution.

1) Competition: In an environment where a moving target defense is deployed, the attacker's goals are to scan and discover the vulnerabilities of the target network system in the network attack-defense conflict and to use the detected vulnerabilities to launch as many attacks as possible to achieve the goal. The goal of the defender is to transfer vulnerabilities to avoid or reduce the attacker's use of the attack surface, thereby improving the security of the system.

2) Bounded rationality: It should be noted that bounded rationality is different from complete rationality. The traditional game theory-based model is based on the assumption of complete rationality of both attackers and defenders. However, complete rationality [26] requires meeting the following three conditions: a) each party must understand every factor that affects a decision made during the decision-making process; namely, the decision-makers have perfect identification and judgment ability and fully consider various factors in the decision-making process; b) each party must fully consider every possible outcome and compute its probability of occurrence when making a decision (i.e., the decision-makers have perfect memory and computation ability); c) each party has the ability to rank the preference for each result (decision-makers have perfect analytical reasoning ability). In the process of network attack-defense conflict, limitations in the different attackers' and defenders' understanding of the network security situation and the differences in their responses leads to different

prediction and decision-making mechanisms, including the defenders' attack experience, attack capabilities, and recognition of the attack target. The attacker's cognition of the target system's attack surface and attack surface transfer strategy also has limitations. This makes it difficult for the attacker and defender to make perfect determinations and to select strategies in the game. Therefore, both parties of the game constrained by resources and capabilities need to make rational decisions in the way of inductive reasoning according to the information they have gradually acquired through constant exploration and attempt, adjustment and optimization in the game process, i.e., to behave with bounded rationality.

3) Multiple stage evolution: In a network attack-defense conflict, the game between the attacker and defender in a hopping cycle or a change in network tasks causes the attack surface to shift and change, leading to a transformation of the target network system's state. The next interval in the offensive–defensive game behavior is to implement the evolution and optimal selections of attack-defense strategies using the current state of the network system. The attacker and defender obtain different payoffs after each action. Each player continuously improves its own security strategy by learning from the succeeding party and forming a new attack-defense situation. Driven by the players' continuous improvement of attack-defense strategies, both the attacker and defender dynamically adjust their own strategies based on feedback from actions by attackers and defenders [27].

This analysis makes clear that, in the environment where the network moving target defense is deployed, the network system has multiple states, and the various states of the network system in different hopping cycles transfer to each other according to the different network attack and defense strategies. Randomness when selecting defense strategies leads to randomness in the network system state. Therefore, a network attack-defense conflict is a conflict between multiple system states with the strategy payoff matrix different for each state. At the same time, within a given network system state, both the attacker and the defender only have bounded rationality. Both players dynamically adjust their own strategies by learning from feedback obtained from the network system for each attacker or defender behavior. The rates of attack-defense strategy learning and dynamic adjustment are limited.

C. Overview of Wright-Fisher process

In 2006, Imhof et al. [28] proposed the Wright-Fisher process for the first time and used it to describe the evolution of a limited biological population. It is a typical stochastic evolution mechanism. The process updates synchronously during a biological evolution iteration, modeling the evolution process of a finite game group as a random process, and taking bounded rationality and incomplete information as hypothetical premises. It has been used in applications such as network software [29] and infinite sensor networks [30]. In offensive–defensive conflicts, both the attackers and defenders are finite groups. Therefore, compared to the deterministic evolutionary game represented by the replication dynamic equation, a stochastic evolutionary game using the Wright-Fisher process effectively describes the finiteness of the strategic groups in an offensive–defensive conflict, analyzing the process of the game reaching equilibrium, which is typical of actual offensive–defensive conflict scenarios.

A stochastic evolution game based on the Wright-Fisher process is a dynamic stochastic game. Its stochastic nature is mainly reflected in the rule update process during the strategy evolution step. A Wright-Fisher process uses a synchronous update strategy in its evolution mechanism. At a certain stage k, all attacker and defender strategies are updated at the same time to generate alternative attacker and defender strategies for the stage k+1, and then choose the actual strategy in stage k+1 from them. The stage's algorithm ensures that the total number of attackers and defenders selecting strategies remains unchanged. The asynchronous update randomly selects a strategy from all attacker and defender strategies to update at a specific stage k. Typical asynchronous update processes include the Moran process, the Fermi process, and the vision update process [31]. Compared with an asynchronous update learning mechanism, the synchronous update method converges faster when learning and is more suitable for actual network defense decision-making scenarios in high-frequency offensive–defensive conflicts.

Therefore, our use of the Wright-Fisher process to describe the internal drive of both attackers and defenders to improve their behavioral strategies continuously is consistent with the dynamic evolution of actual network offensive–defensive conflicts, enhancing the accuracy and credibility of the game model when analyzing attacker and defender behaviors. In this paper, we construct an evolutionary game model using the Wright-Fisher process to describe and analyze the conflicts between attackers and defenders in a moving target defense. By combining the Wright-Fisher process and the evolutionary game model, we develop a multiple state and multiple agent evolutionary game model.

III. CONSTRUCTION OF MTD STRATEGY EVOLUTION MODEL BASED ON WRIGHT-FISHER PROCESS

The attacker and defender have bounded rationality, which is between complete rationality and irrationality. The main reason is that the professional knowledge, perception ability, and decision computation ability of the players are not perfect. The values and goals are not always consistent but rather conflict with each other very often. Moreover, the actual decision-making environment is complex and uncertain. In this section, based on the evolutionary game [32], under the condition of bounded rationality, an evolutionary MTD model based on the Wright-Fisher process was constructed, and then the method for solving evolutionary equilibrium was presented.

The incompleteness of game information is mainly reflected in the payoff information, that is, the participants know their own payoff, but do not fully understand the opponent's payoff. Because both sides of MTD attack and defense have dynamic and diverse uncertainties, the attack-defense game has incomplete information characteristics. The stochastic evolutionary game model constructed in this section contains incomplete information characteristics, which are reflected in the perception of the other party's incomplete payoff information by both offensive and defensive sides and the calculation of expected payoff, as well as the posterior correction of the probability of selecting different strategies in

the populations through game feedback, and then the correction of payoff calculation. See Section III-A for detailed analysis.

Considering that in actual adversarial situations, there are typically multiple attackers and defenders, this study has classified players into different groups. In the classic attack-defense game theory-based model based on the assumption of complete rationality, the game equilibrium has been interpreted as an optimal response of the attacker and defender, but it cannot provide the formation process of the game equilibrium. This study has focused on the evolution process of the attack and defense strategies by simulating the learning process and strategy adjustment of the attacker and defender. The proposed model can dynamically characterize the evolutionary trajectory of decision-making and can improve the fineness and accuracy of the results, and enhance the effectiveness of the defense decision.

Definition 1 The Wright-Fisher process-based moving target defense strategy evolution model WF-MTD can be represented as a quintuple (N, S, P, T, R) as follows:

1) $N = \{N_A, N_D\}$ represents the set of players in the attack-defense game. We consider the presence only of attackers and defenders, $|N| = 2$. N_A denotes attackers, and N_D denotes defenders.

2) $S = \{S_1, S_2, ..., S_k\}$ represents the set of states in the attack-defense game, where each state is the network attack surface at a specific moment. A transition between network system states is a transfer or reduction of the attack surface.

3) $P = \{AS, DS\}$ represents the set of the strategies in the attack-defense game. We express the attacker 's optional strategy set as $AS = \{AS_1, AS_2 \cdots AS_m\}, m \in N^+$ and $m \geq 2$ and the defender's optional strategy set as $DS = \{DS_1, DS_2 \cdots DS_n\}$, $n \in N^+$ and $n \geq 2$.

4) $$T = Pr\{S_{t+1} = S_j^{'} \mid S_t = S_j, AS = AS_i, DS = DS_j\}$$

represents the state transition probability in the attack-defense game. The next state S_{t+1} in the attack-defense game depends only on the current state S_t and the attack-defense strategies AS_i and DS_j, with no dependency on earlier states and strategies. The value of the state transition probability generally depends on the network environment and attack-defense process, such as the network configuration, node operating system environment and attack-defense strategies[33]. During the attack and defense process of MTD, the state transition probability is equal to the transition probability of the attack and defense strategy. See Section III-A and Formula (6) for details.

5) $R = \{R_A, R_D\}$ represents the set of payoff functions in the attack-defense game and are jointly determined by the strategies of all participants. $R_A(S, AS_i, DS_j)$ represents the attack payoff of the attacker adopting strategy AS_i and the defender adopting strategy DS_j under state S. $R_D(S, AS_i, DS_j)$ represents the defense payoff of the attacker adopting strategy AS_i and the defender adopting strategy DS_j under state S. The specific calculation method is shown in Formula (1) and (2).

Since the reward functions of both the attacker and the defender consider the cost and reward of launching an attack or implementing a defense. According to the work of Lei et al. [39][42], the rewards of the attacker and defender can be characterized by the performance consumption and outcomes of changing the MES and the MAS.

$$R_A(S, AS_i, DS_j) = ASR[f_{\Delta MAS}(DDR, DIR) + f_{\Delta MES}(DDR, DIR)] + (1 - ASR) f_{\Delta MES}(DIR) \quad (1)$$

$$R_D(S, AS_i, DS_j) = ASR[f_{\Delta MES}(DDR, PC) + f_{\Delta MAS}(DDR, PC)] + (1 - ASR)[f_{\Delta MES}(PC) + f_{\Delta MAS}(PC)] \quad (2)$$

where ASR represents the attack success rate of the attacker when the attacker and defender implement the corresponding strategies; ΔMES is the change in the MES of a target system; ΔMAS is the change in the MAS of the target system; DDR is short for direct defense reward, which indicates changes in the loss of the target system resources caused by the attack after the successful defense; DIR represents the indirect defense reward, which denotes the change in the attack cost after the successful defense, and it is determined by the attacker's capability and prior knowledge; PC denotes the performance cost caused by network hopping.

A. MTD strategy evolution model based on the Wright-Fisher process

In a moving target attack-defense conflict, the decision-makers of the attacker A and the defender D have multiple strategies to choose from. At different stages of the game, the probability that the strategy is adopted by the decision makers differs and changes constantly under the action of the learning mechanism over time. Thus, the attack-defense strategy selections form a dynamic process. The use of different strategies for an offensive–defensive conflict generates the corresponding attack-defense payoffs. The payoff values are expressed by the following specific payoff matrix, where a_{ij} and b_{ij} represent the payoffs when the attacker and the defender take AS_i and DS_j, respectively

$$\begin{Bmatrix} a_{11}, b_{11} & a_{12}, b_{12} & a_{13}, b_{13} & \cdots & a_{1n}, b_{1n} \\ a_{21}, b_{21} & a_{22}, b_{22} & a_{23}, b_{23} & \cdots & a_{2n}, b_{2n} \\ a_{31}, b_{31} & a_{32}, b_{32} & a_{33}, b_{33} & \cdots & a_{3n}, b_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{m1}, b_{m1} & a_{m2}, b_{m2} & a_{m3}, b_{m3} & \cdots & a_{mn}, b_{mn} \end{Bmatrix}$$

In this case, the expected payoff of the attacker's strategy AS_i is

$$f_{AS_i} = \frac{M_1}{M} a_{i1} + \frac{M_2}{M} a_{i2} + \cdots + \frac{M_n}{M} a_{in} = \frac{1}{M} \sum_{j=1}^{n} M_j a_{ij} \quad (3)$$

The expected payoff of the defender's strategy DS_j is

$$f_{DS_j} = \frac{N_1}{N} b_{1j} + \frac{N_2}{N} b_{2j} + \cdots + \frac{N_m}{N} b_{mj} = \frac{1}{N} \sum_{i=1}^{m} N_i b_{ij} \quad (4)$$

In these equations, $i \in \{1, 2, ..., m\}$, $j \in \{1, 2, ..., n\}$, N_i represents the number of attackers who choose the strategy

AS_i, and M_j represents the number of defenders who choose the strategy DS_j. Also, $\sum_{i=1}^{m} N_i = N, \sum_{j=1}^{n} M_j = M$. At the initial stage of the game, the attack and defense sides evaluate the probability of choosing different strategies in their own populations according to prior experience, set the number of attackers M_j and the number of defenders N_i, and then make a posteriori correction according to the feedback information.

To reflect the bounded rationality of attackers and defenders, rationality factors ω_A, ω_D are introduced to describe the rationality of attackers and defenders, respectively, in the network attack-defense process. As the attack-defense game progresses, both the attackers and defenders gain better understanding of the unknown game payoff information, and their calculation of the game payoffs is more accurate. In this process, both the attackers and the defenders always choose the strategy that maximizes the game payoffs as the optimal strategy.

Therefore, the actual expected payoffs of different attack-defense strategies are

$$\begin{cases} F_{AS_i} = 1-\omega_A + \omega_A f_{AS_i} \\ F_{DS_j} = 1-\omega_D + \omega_D f_{DS_j} \end{cases} \tag{5}$$

where F_{AS_i} and F_{DS_j} are linear functions of payoffs f_{AS_i} and f_{DS_j}, respectively. ω_A, $\omega_D \in [0,1]$ control the strength of the strategy selection of the attackers and defenders, respectively. When $\omega_A = \omega_D = 1$, the attackers and defenders completely grasp the gain information of the strategy, and the game is a completely rational game. An actual attack-defense game is a process of continuous trial-and-error learning, and it is impossible to fully grasp the payoff information. As ω_A and ω_D gradually approach 0, the degree of rationality diminishes. The specific parameter updating process could be divided into three grades: (1) when $\omega_A = \omega_D \in (0, 0.5]$, the rationality of players is low; (2) when $\omega_A = \omega_D = 0.5$, players are moderately rational; (3) when $\omega_A = \omega_D \in (0.5, 1)$, the players have a high level of rationality.

Under different states S_t of the attack and defense process, the probability of attackers and defenders choosing strategies will be updated and changed. We use the Wright Fisher process to describe it. The Wright-Fisher process is a dynamic evolution process that updates the strategy using synchronous update. This process performs the N-fold Bernoulli experiment in the offspring set, individuals with different strategies obey the binomial distribution [34]. The attack and defense sides use the follow-up game feedback to update and adjust the probability π_{AS_i}, π_{DS_j} and change rate $E(\Delta p_i)$, $E(\Delta q_j)$ of different strategies.

The selection of attack-defense strategies is that the two sides repeatedly select different strategies from the attack-defense strategy space many times, then continuously adjust the strategy based on the payoffs generated, and finally obtain the optimal strategy that obtains the greatest payoffs. The strategy selection principle of the Wright-Fisher process is to optimize according to the proportion of the payoffs of a specific strategy in the payoffs of the overall strategy space. Then the probability of the attack decision maker to select the strategy AS_i after each game is π_{AS_i}, and the probability of the defense decision maker creating the strategy DS_j after each game is π_{DS_j}:

$$\begin{cases} \pi_{AS_i} = \dfrac{N_i F_{AS_i}}{\sum_{i=1}^{m} N_i F_{AS_i}} \\ \pi_{DS_j} = \dfrac{M_j F_{DS_j}}{\sum_{j=1}^{n} M_j F_{DS_j}} \end{cases} \tag{6}$$

According to the features of the Wright-Fisher process, the individuals in the Wright-Fisher process perform a synchronous update and obey the binomial distribution. Let $Y_A(t)$ be the number of attackers who use the attacking strategy AS_i after the t-th attack-defense game, where $Y_A(t) = N_i$; let $Y_D(t)$ be the number of defenders who use the defending strategy DS_j after the t-th attack-defense game, where $Y_D(t) = M_j$. $N_i^{'}$ and $M_j^{'}$ are the number of attackers and defenders who use the strategies AS_i and DS_j after the t+1-th game, respectively. Thus, the probabilities of using strategies AS_i and DS_j for the t+1-th game are

$$\begin{cases} P\{Y_A(t+1) = (N_1^{'}, N_2^{'}, \ldots, N_m^{'}) \mid Y_A(t) = (N_1, N_2, \ldots N_m)\} \\ = \dfrac{N!}{N_1^{'}! N_2^{'}! \ldots N_m^{'}!} \prod_{i=1}^{m} \left(\dfrac{N_i F_{AS_i}}{\sum_{i=1}^{m} N_i F_{AS_i}} \right)^{N_i^{'}} \\ P\{Y_D(t+1) = (M_1^{'}, M_2^{'}, \ldots, M_n^{'}) \mid Y_D(t) = (M_1, M_1, \ldots, M_n)\} \\ = \dfrac{M!}{M_1^{'}! M_2^{'}! \ldots M_n^{'}!} \prod_{j=1}^{n} \left(\dfrac{M_j F_{DS_j}}{\sum_{j=1}^{n} M_j F_{DS_j}} \right)^{M_j^{'}} \end{cases} \tag{7}$$

Let $p = \{p_1, p_2, \cdots, p_m\}$ represent the proportion of the individuals who select specific attack strategies in the attack group, where $p_i = \dfrac{N_i}{N}, i = 1, 2, \cdots, m$. Let $q = \{q_1, q_2, \cdots, q_n\}$ denote the proportion of the individuals who select specific defensive strategies in the defense group, where $q_j = \dfrac{M_j}{M}, j = 1, 2, \cdots, n$. Let $E(\Delta p_i)$ and $E(\Delta q_j)$ denote the change in frequency for the attacker and the defender when selecting a strategy. Δt is the step size of the update time. The rates of change with respect to time $\dfrac{dp_i}{dt}$ and $\dfrac{dq_j}{dt}$ of p_i and

q_j , respectively, of the Wright-Fisher process can be approximated by the Langevin equation [35], to study the dynamic evolution process of attacking and defending strategies, and when $N \to \infty$ and $M \to \infty$, $\frac{dp_i}{dt} = \frac{E(\Delta p_i)}{\Delta t}$ and $\frac{dq_j}{dt} = \frac{E(\Delta q_j)}{\Delta t}$. Therefore, we can calculate $E(\Delta p_i)$ and $E(\Delta q_j)$ as follows:

$$\begin{aligned} E(\Delta p_i) &= \frac{E(\Delta N_i)}{N} \\ &= \frac{\sum_{N_i^{'}=0}^{N} (N_i^{'} - N_i) P(Y_A(t+1) = N_i^{'} \mid Y_A(t) = N_i)}{N} \\ &= \frac{N_i F_{AS_i}}{\sum_{i=1}^{m} N_i F_{AS_i}} - p_i \end{aligned} \tag{8}$$

and

$$\begin{aligned} E(\Delta q_j) &= \frac{E(\Delta M_j)}{M} \\ &= \frac{\sum_{M_j^{'}=0}^{M} (M_j^{'} - M_j) P(Y_D(t+1) = M_j^{'} \mid Y_D(t) = M_j)}{M} \\ &= \frac{M_j F_{DS_j}}{\sum_{j=1}^{n} M_i F_{DS_j}} - q_j \end{aligned} \tag{9}$$

To summarize, we can obtain the replication dynamic evolution equations in the Wright-Fisher process for the attack-defense strategies. By solving the equations, we determine the equilibrium state of the network offensive–defensive evolution, enabling the analysis and prediction for the selection of security defense strategies.

$$\begin{cases} \dfrac{dp_i}{dt} = \dfrac{1}{\triangle t}\left(\dfrac{p_i F_{AS_i}}{\sum_{i=1}^{m} p_i F_{AS_i}} - p_i\right), i = 1, 2, \cdots, m \\ \dfrac{dq_j}{dt} = \dfrac{1}{\triangle t}\left(\dfrac{q_j F_{DS_j}}{\sum_{j=1}^{n} q_j F_{DS_j}} - q_j\right), j = 1, 2, \cdots, n \end{cases} \tag{10}$$

B. Algorithm for selecting the optimal MTD strategy

In our model, we use the following algorithm to select the optimal MTD strategy.

Algorithm 1 Optimal MTD strategy selection algorithm in WF-MTD

Input Network environment information NetInf, safety protection equipment configuration information SafetyInf, and intrusion alarm information AlertInf

Output Optimal MTD strategy q

BEGIN

1) **Initialize** WF-MTD = (N, S, P, T, R)

/*Initialize the Wright-Fisher process-based MTD strategy evolution model*/

{

1-1) **Construct** $DS = \{DS_1, DS_2 \cdots DS_n\}$, $n \in N^+$ and $n \geq 2$

/*Analyze the configuration information of safety protection equipment SafetyInf, collect the defensive strategy, and initialize the defensive strategy space DS */

1-2) **Construct** $AS = \{AS_1, AS_2 \cdots AS_m\}, m \in N^+$ and $m \geq 2$

/*Collect real-time alert data AlertInf, analyze the characteristics of the attack behavior, and initialize the attacker's strategy space AS */

1-3) **Construct** $p = \{p_i\}$, $0 \leq p_i \leq 1$, $\sum_{i=1}^{m} p_i = 1$

/*Initialize the attack belief set p. The attacker chooses the attack strategy AS_i with the probability $p_i \in p$ */

1-4) **Construct** $q = \{q_j\}$, $0 \leq q_j \leq 1$, $\sum_{j=1}^{n} q_j = 1$

/*Initialize the defense belief set q. The defender chooses the defense strategy DS_j with the probability $q_j \in q$ */

}

2) **For** $(i = 1;\ i \leq m;\ i++)$

For $(j = 1;\ j \leq n;\ j++)$

{

Calculate $R = \max_{AS_i} \min_{DS_j}(R_A, R_D)$ by formula (1) (2)

/*Calculate the attack-defense payoffs of the different strategy combinations AS_i and DS_j under the state of network*/

}

3) **Assign** ω_A, ω_D, $0 \leq \omega_A, \omega_D < 1$

/*Select data based on past strategies of the game, and set the rationality parameters ω_A and ω_D of the attacker and the defender, respectively*/

4) **For** $(i = 1;\ i \leq m;\ i++)$

For $(j = 1;\ j \leq n;\ j++)$

{

Calculate $f_{AS_i} = \sum_{j=1}^{n} q_j a_{ij}$

Calculate $F_{AS_i} = 1 - \omega_A + \omega_A f_{AS_i}$

Construct $\dfrac{dp_i}{dt} = \dfrac{1}{\triangle t}\left(\dfrac{p_i F_{AS_i}}{\sum_{i=1}^{m} p_i F_{AS_i}} - p_i\right), i = 1, 2, \cdots, m$

}

/*Construct the Wright-Fisher evolution equation of the attacker's strategy*/

5) **For** $(j = 1;\ j \leq n;\ j++)$

For $(i = 1;\ i \leq m;\ i++)$

{

Calculate $f_{DS_j} = \sum_{i=1}^{m} p_i b_{ij}$

Calculate $F_{DS_j} = 1 - \omega_D + \omega_D f_{DS_j}$

Construct $\frac{dq_j}{dt} = \frac{1}{\Delta t}\left(\frac{q_j F_{DS_j}}{\sum_{j=1}^{n} q_j F_{DS_j}} - q_j\right), j = 1, 2, \cdots, n$

}

/*Construct the Wright-Fisher process-based evolution equation of the moving target defense strategy*/

6) Calculate $Y = \begin{bmatrix} dp_i / dt \\ dq_j / dt \end{bmatrix} = 0$

/*Calculate the evolutionarily stable equilibrium solution*/

7) Output $q = \{q_1^*, q_2^*, \cdots, q_n^*\}$

/*Output the optimal MTD strategy set*/

END

In Algorithm 1, step 1 initializes the WF-MTD parameters. Step 2 calculates the payoff matrix of the attack-defense game. Step 3 combines historical data to set rationality parameters for both the attacker and the defender Steps 4 and Step 5 construct the evolution equation of the moving target attack-defense strategies using the Wright-Fisher process. Step 6 calculates the evolution equilibrium solution through the simultaneous equations, and step 7 outputs the optimal MTD strategy set.

IV. APPLICATION EXAMPLES AND ANALYSIS

In this section, we present an attack intrusion and moving target defense in a network information network system as an example, verify the model and algorithm proposed above, and compare and analyze the evolution processes and results of the attack–defense strategies along different attack paths. Based on this, we then summarize the general trends of the evolution of the strategies. Finally, the proposed method is applied to the SDN route hopping strategy selection scenario to test the effectiveness of the proposed WF-MTD method in practical applications.

A. Experimental apparatus

We chose a medical information network system [36] as our test case, as such systems provide great convenience for patient visits and diagnosis but store sensitive data such as doctor–patient information and often interface with various types of medical equipment. As such, they are common attack targets. Cyber attacks such as ransomware [37] pose a serious threat to the medical information network system, from data breaches and network paralysis. In severe cases, the loss of network functionality endangers the lives of patients. A typical medical information network system architecture is shown in Figure 1.

There are two attack paths that the attacker can take according to Figure 2.

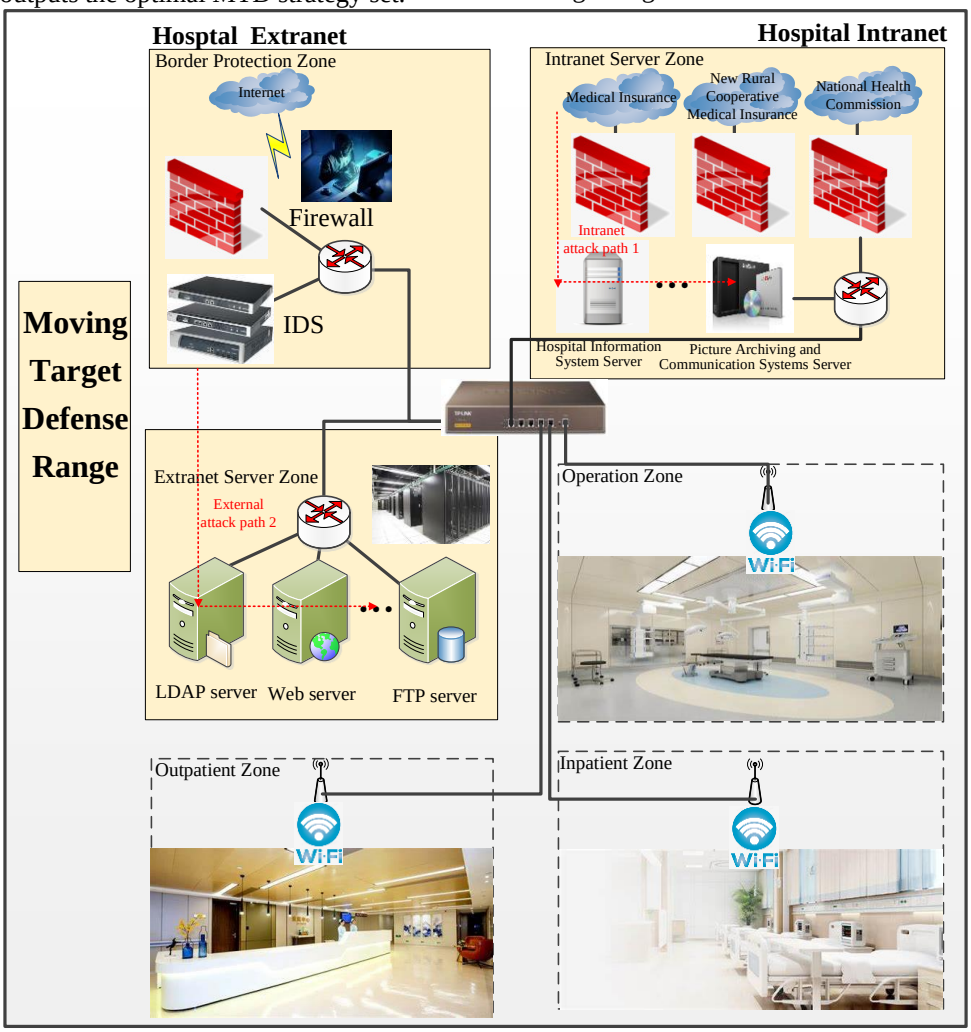

Figure 2. Architecture of typical medical information network system.

Intranet attack path 1: Inpatient Department/Administrative Department User Host → Key System Server;

External network attack path 2: LDAP Server → Web server.

The medical information network system security administration team is the network defender responsible for ensuring secure operation of the medical information network system. Table 1 presents the configuration and vulnerability information of a typical medical information network system, with the security protection measures composed of passive defense strategies such as a firewall, an intrusion detection system (IDS), and moving target defense strategies. Among them, the passive defense strategy mainly provides border security, and the moving target defense strategy provides global protection. For the external network of a hospital, according to the network access policy preset by the firewall, external hosts only have user-level access to the file server. The attacker's purpose is to steal outpatient medical data stored on the file server accessible from the external network.

TABLE 1. NETWORK CONFIGURATION AND VULNERABILITY INFORMATION

Configuration	CVE #	Vulnerability description	Vulnerability level	Vulnerability type
LDAP server	CVE-2015-5330	Mishandles string length	Medium	Information leakage
Web server	CVE-2014-0098	Allows remote attackers to cause a denial of service	Medium	Input verification
FTP server	CVE-2019-12815	Allows for remote code execution and information disclosure without authentication	Critical	Access control error
HIS server	CVE-2020-1938	Reads or includes arbitrary files in all webapp directories on Tomcat.	Critical	Input verification error
PACS server	CVE-2012-6694	Allows remote attackers to control workstation	Critical	Trust management

From the hospital intranet, the worm WannaCry is a suitable example [38]. The attacker first employs common user-facing applications to trick medical staff into downloading them. It then uses a virtualization sandbox to evade defense and run DLL32 and finally invades the intranet server area. Once inside, it can destroy key system servers such as the inpatient medical information system, hospital information system (HIS) server, image archives, and picture archiving and communication system (PACS) servers.

B. Numerical experiments and analysis

Using the hospital external network attack as an example, we carried out the following numerical experiments using the two attack paths to explore the evolution trends of the MTD attack and defense strategies in different network states. Based on the definition of the attack-defense strategies and methods of quantifying payoffs described previously [39], according to Algorithm 1, we first use the Nmap tool to scan the medical information network system. We constructed a set of attack-defense strategies in different states of the two attack paths in the medical information network system using MITRE ATT&CK™ and moving target defense strategies, and calculated the payoffs of different attack-defense strategies in the conflict. The attack-defense strategy sets and payoffs in different states in attack path 1 are shown in Tables 2 and 3. The attack-defense strategy sets and payoffs in different states in attack path 2 are shown in Tables 4 and 5. We then conducted numerical experiments from three dimensions (i.e., network states), initial strategy selection probabilities, and rationality parameters.

For the internal network attack path 1, the set of the states of the experimental network system were $S=\{S_1,S_2\}$, where S_1 denotes the normal state, and S_2 denotes the server user permission obtained by using the vulnerability of the user host.

TABLE 2. DEFENSE STRATEGY SET IN DIFFERENT STATES WITHIN ATTACK PATH 1

S_1
$AS=\{Overflow\ attack, Data\ destroy, Non\}$
$DS=\{Patch\ upgrade, ASD_1+time, Non\}$
S_2
$AS=\{Overflow\ attack, Data\ destroy, Privilege\ gaining\}$
$DS=\{Patch\ upgrade, ASD_1+ASD_3, ASD_1\}$

The payoffs obtained from the strategies used by both the attacker and the defender in each state are shown in the following table:

TABLE 3. ATTACK-DEFENSE PAYOFFS IN THE DIFFERENT STATES WITHIN ATTACK PATH 1

S_1	
Offensive payoff	Defensive payoff
$\begin{bmatrix}15&10&10\\15&20&20\\0&0&0\end{bmatrix}$	$\begin{bmatrix}-15&-10&-10\\-15&-20&-20\\0&0&0\end{bmatrix}$
S_2	
Offensive payoff	Defensive payoff
$\begin{bmatrix}20&72&37\\10&50&30\\20&40&20\end{bmatrix}$	$\begin{bmatrix}-20&-72&-37\\-10&-50&-30\\-20&-40&-20\end{bmatrix}$

For the external network attack path 2, the set of the states of the experimental network system were $S=\{S_1,S_2\}$, where S_1 denotes the server user permission obtained by exploiting a vulnerability of the LDAP server, and S_2 denotes the server root permission obtained by exploiting a vulnerability of the web server.

TABLE 4. DEFENSE STRATEGY SET IN DIFFERENT STATES WITHIN ATTACK PATH 2

S_1
$AS=\{Semi\text{-}blind\ \ scan, Follow\ scan, Overflow\ attack\}$
$DS=\{ASD_1, ASD_1+time, Patch\ upgrade\}$
S_2
$AS=\{Semi\text{-}blind\ \ scan, Injection\ attack, Non\}$
$DS=\{ASD_3, ASD_3+time, Patch\ upgrade\}$

The payoffs obtained by the strategies used by both the attacker and the defender in each state are shown in the following tables.

TABLE 5. ATTACK-DEFENSE PAYOFFS IN DIFFERENT STATES WITHIN ATTACK PATH 2

S_1	
Offensive payoff	Defensive payoff
$\begin{bmatrix} 20 & 40 & 20 \\ 30 & 50 & 10 \\ 37 & 72 & 20 \end{bmatrix}$	$\begin{bmatrix} -20 & -72 & -20 \\ -30 & -50 & -10 \\ -37 & -40 & -20 \end{bmatrix}$
S_2	
Offensive payoff	Defensive payoff
$\begin{bmatrix} 20 & 33 & 15 \\ 15 & 30 & 10 \\ 0 & 0 & 0 \end{bmatrix}$	$\begin{bmatrix} -20 & -33 & -15 \\ -15 & -30 & -10 \\ 4 & 11 & 0 \end{bmatrix}$

(1) Evolution of attack-defense strategies under different network conditions

Assuming the rationality parameters w_A=w_D=0.5 and setting the probability of the initial strategy selection of both attacker and defender to be 1/3, we first calculated the expected payoffs of both the attacker's and defender's strategies and then determined the expected payoffs of both the attacker and the defender. We then obtained the evolution trend of attack-defense strategies for different network states by calculating the evolution equation of the Wright-Fisher process, as shown in Figure 3, where the horizontal axis represents the number of plays of the offensive–defensive game, and the vertical axis represents the probability for selecting an offensive or defensive strategy.

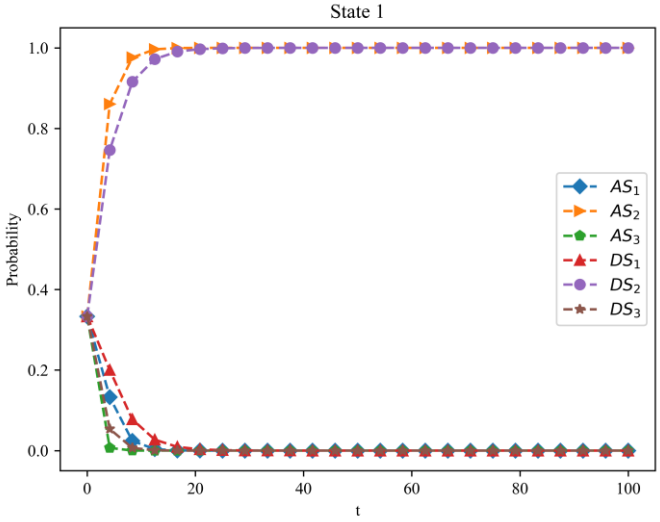

Figure 3. Evolution trajectories of attack-defense strategies in state 1 within attack path 1

First, the attack path in state 1, the attacker of the moving target chose the optimal attack strategy $AS_{2|S_1}$, destroying data, after about 49 games, while the defender of the moving target chose to implement the best defense strategy $DS_{2|S_1}=ASD_1+time$ after about 73 games. The optimal strategy selections of both sides maintained a continuous and stable evolutionary state. This stable state was only broken when the network environment changed. In the offensive–defensive conflict, the attacker prioritized the overflow attack for greatest effectiveness, while the defender adopted attack surface hopping to resist the overflow attack.

Similarly, as shown in Figure 4, in state 2 of the attack path 1, the attacker of the moving target selected the optimal attack strategy $AS_{1|S_2}=Overflow\ attack$ after about 64 games, while the defender of the moving target selected the optimal defense strategy $DS_{2|S_2}=ASD_1+ASD_3$ after about 38 games. Although the defensive strategies ASD_1+ASD_3 and ASD_1 both involved MTD hopping, the defender selected the ASD_1+ASD_3 strategy to implement its defense. This occurred because the strategies ASD_1+ASD_3 simultaneously selected the IP address and the protocol for coordinated hopping. The defense payoff of the mixed attack strategies in state 2 was much higher than that of single IP address hopping as implemented in strategy ASD_1. Therefore, to remove the effect of the network state on the defensive payoff, the payoff of the multiple element coordinated hopping strategy had better defense than single-element hopping. However, it required the coordinated implementation of different transition elements, rather than a simple superposition of multiple transition elements.

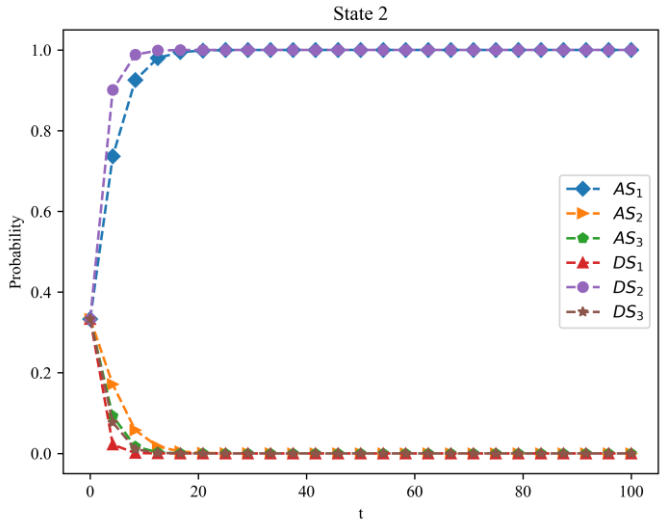

Figure 4. Evolution trajectories of attack-defense strategies in state 2 within attack path 1

The experimental results associated with attack path 2 are presented in Appendix.

(2) Evolution of the optimal MTD strategy with different initial MTD strategy probabilities

Assuming the rationality parameters w_A=w_D=0.5, we tested the evolution of attack-defense strategies with different initial MTD strategy probabilities using the following scenarios. At the beginning, the attacker of the moving target randomly selected an attack strategy with an equal probability of one-third,

This article has been accepted for publication in IEEE Transactions on Dependable and Secure Computing. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TDSC.2022.3232537

and the defensive strategy selection changed. We then observed the evolution trajectory of the optimal defense strategy of the moving target.

For different moving target defense strategies on path 1, the initial probabilities of the moving target defense strategies correspond to the following three situations: 1) DS_1=0.8, DS_2=0.1, DS_3=0.1; 2) DS_1=0.1, DS_2=0.8, DS_3=0.1; and 3) DS_1=0.1, DS_2=0.1, DS_3=0.8, with the attacker adopting a random attack strategy, AS_1=AS_2=AS_3=1/3. Through experiments, we determined the evolution trajectories of the defense strategy for state 1 on attack path 1 in these three cases, as shown in Figure 5.

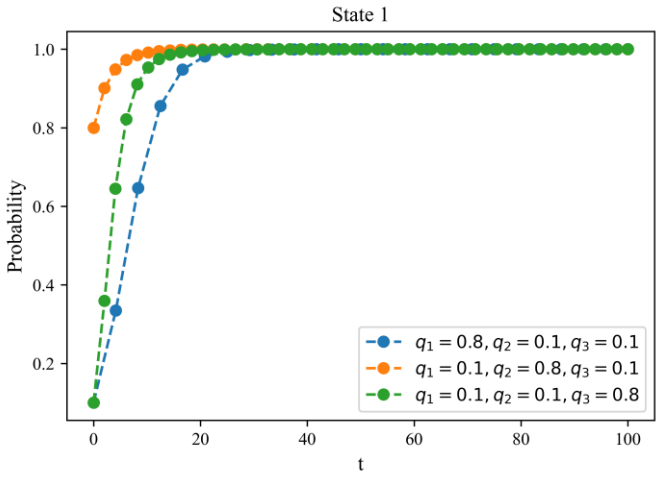

Figure 5. (a) For a specific attack strategy, evolution trajectories of defense strategy in state 1 within attack path 1

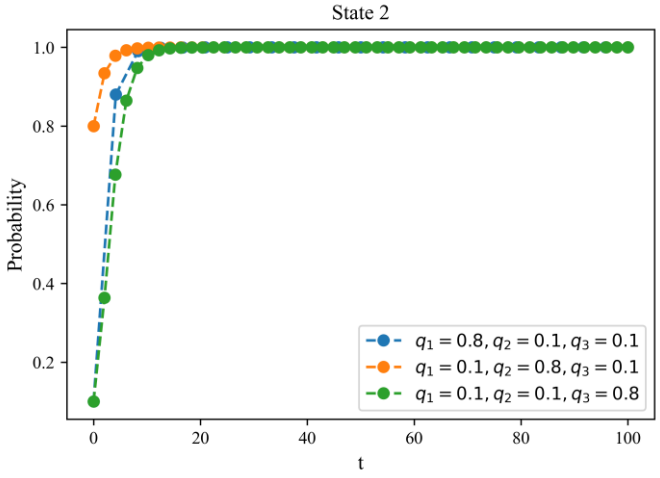

Figure 5. (b) For a specific attack strategy, evolution trajectories of defense strategy in state 2 within attack path 1.

At the beginning, the attacker randomly selected an attack strategy with an equal probability of 1/3, and the defender implemented the defense strategy in the three strategies above. Although at the initial moment the defender choses the strategies with different probabilities, after multiple games with the attacker, it continued to learn, adjust, and optimize the strategy. It ultimately selected the corresponding optimal MTD strategy, and maintained a continuous and stable evolution state. The different strategy selection probabilities of the defender at the initial moment only affected the time for the optimal defense strategy to reach a stable state and did not affect the selection of the optimal defense strategy.

The experimental results associated with attack path 2 are presented in Appendix.

(3) Evolution of optimal MTD strategy under different rationality parameters

Using state 1 of attack path 1 as an example, the initial strategy selection probability of both the attacker and the defender was set at one-third. Assuming that the attacker and the defender adopted the same degree of rationality, we explored the effect of changes in the degree of rationality on the evolution of the optimal moving target defense strategy, as shown in Figure 6.

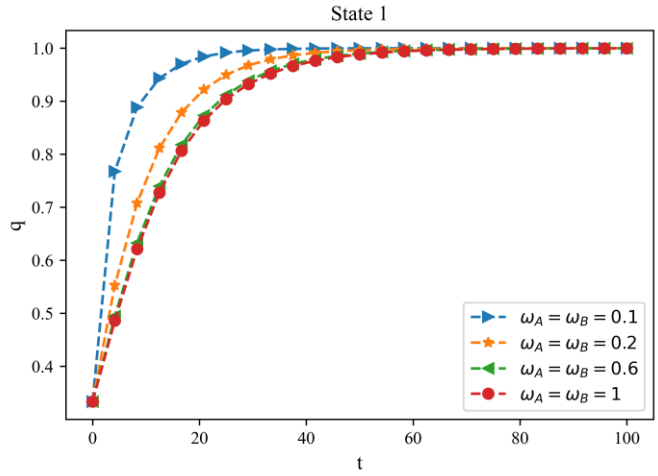

Figure 6. Evolution trajectories of optimal MTD strategies under different degrees of rationality.

Figure 6 shows that as the degree of rationality decreased, the convergence rate of the strategy learning mechanism of our method increased. When the degree of rationality increased to near complete rationality, our method was basically consistent with the convergence rate of a fully rational game. When the rationality was low and close to incomplete rationality, our method achieved better convergence. It can be seen that our method was still able to converge faster when the rationality was limited. The synchronous update learning mechanism of the Wright-Fisher process-base strategy not only guided the defender to make optimal decisions, but also had a faster learning rate, overcoming the disadvantages of the slow convergence rate of traditional bounded rational games [34].

At the same time, to reflect the actual offensive–defensive scenarios, we assumed that the attacker had bounded rationality, and we explored the convergence speeds of the optimal strategy selections in our method and in the fully rational game method. As shown in Figure 7, the convergence speed of the optimal MTD strategy in our method was significantly better than in the fully rational game method. Thus, our method was more effective and practical.

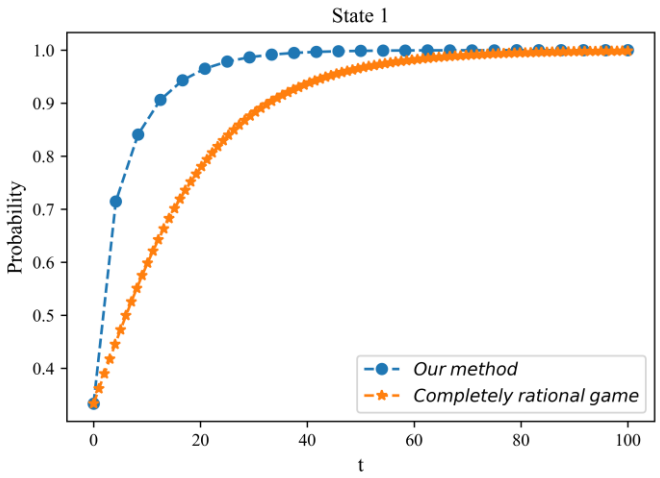

Figure 7. Evolution trajectories of optimal MTD strategies of our method and fully rational game method when facing attacker with bounded rationality.

C. Example application of route hopping strategy selection based on WF-MTD

To verify the practical application performance of the proposed method, we used Mininet to build an SDN network for typical communication services. The topology is shown in Figure 8. There are three communication paths, and the bandwidth of each path was configured as 10.0 Mbps, and route hopping strategy implemented through the Ryu controller. Attackers paralyze the target network based on link flooding attacks. We used the remaining bandwidth of the network communication data link to construct an attack-defense payoff matrix. The collection of the remaining bandwidth of the network communication data link can be automatically triggered based on attack events or pre-sets, or it can be manually triggered by the network administrator. The collection of forwarded data volume can be automatically triggered by pre-sets, or manually triggered by the network administrator. Finally, we applied the proposed method WF-MTD to route hopping strategy selection, and solved the optimal route hopping strategy in link flooding attack scenarios based on the WF-MTD model, thereby verifying the effectiveness of the practical application of the method proposed in this paper.

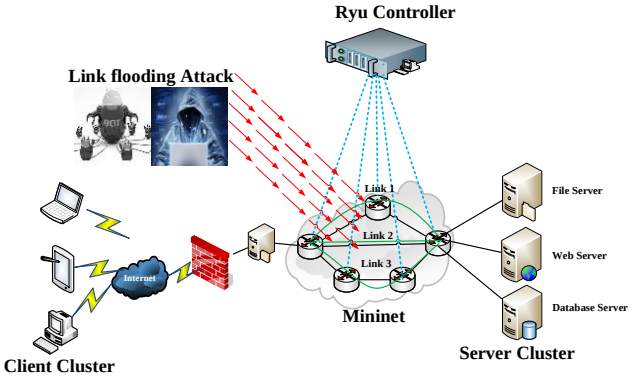

Figure 8. SDN network topology diagram

We set the normal communication bandwidth of path 1 to 1.0 M/s, the normal communication bandwidth of path 2 to 1.5 M/s, the normal communication bandwidth of path 3 to 2.0 M/s, and the link flooding attack traffic to 5.0 M/s. The attack-defense zero-sum payoff matrix is constructed based on the remaining bandwidth of the link at the time of the attack. When the attacker attacks path 1, the remaining bandwidth of the path 1 link is $10.0 - 1.0 - 5.0 = 4.0$, so the theoretical payoff of the attack is 4.0. At this time, other paths communicate normally, the remaining bandwidth of the path 2 link is $10.0 - 1.5 = 8.5$, the theoretical payoff of the attack is 8.5, the remaining bandwidth of the path 3 link is $10.0 - 2.0 = 8.0$, and the theoretical payoff of the attack is 8.0. Similarly, the theoretical payoff when other paths are attacked can be obtained, as shown in Table 6.

TABLE 6 ATTACK-DEFENSE THEORETICAL PAYOFFS

Link flooding attacker	Route hopping defender		
	Path 1	Path 2	Path 3
Path 1	(4.0, -4.0)	(8.5, -8.5)	(8.0, -8.0)
Path 2	(9.0, -9.0)	(3.5, -3.5)	(8.0, -8.0)
Path 3	(9.0, -9.0)	(8.5, -8.5)	(3.0, -3.0)

Based on the above attack-defense theoretical payoffs, we applied WF-MTD to select the optimal route hopping strategy. We set the rational parameter w_A=w_D=0.5 to obtain the attack-defense evolution strategies of different communication paths as shown in Figure 9. As can be seen from the figure, the attacker will launch a link flooding attack on path 3 with a probability of 0.7. Therefore, we should focus on traffic detection on path 3. When there is no clear attack path, the defender prefers path 1 for communication.

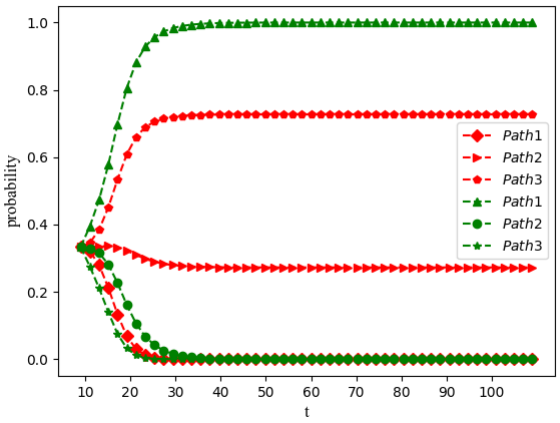

Figure 9. Attack-defense evolution strategy of different communication paths

Furthermore, the method in this paper is compared with the no hopping route strategy and fixed periodic route hopping strategy [40]. Based on the availability of the proposed method for delay quantification, the client and server communicate continuously for 10 min, the delay is collected every second, and the average delay is calculated every 60 s. The delays of the no route hopping strategy, the 5-s fixed periodic route hopping strategy, and the method in this paper are compared, as shown in Figure 10, where the abscissa is the time slot in minutes, and the ordinate is the average delay in milliseconds. Thus, the experimental results show that in the 10-min communication process the average delay of the WF-MTD route hopping strategy is 0.078 ms, and there is no packet loss. Compared with the 5-s fixed cycle route hopping strategy the delay increases by only 20%. For low-latency communication scenarios such as real-time two-way communication, the normal communication

service quality can be guaranteed if the delay is within the range of 100–600 ms.

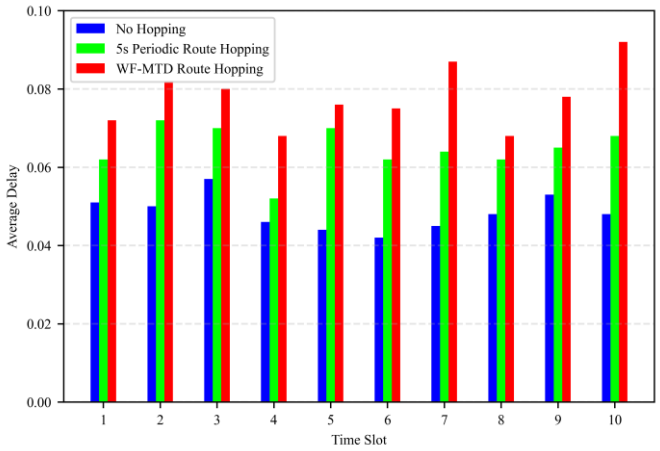

Figure 10. Comparison of communication delay performance between proposed method and other methods

To summarize, the use of route hopping strategy to resist link flooding attacks will affect the system network communication delay, but it can effectively ensure system security. We can trigger the WF-MTD route hopping strategy on demand or based on the attack event response, thereby reducing the system network communication delay and further improving the availability of the network system.

To verify the defense effectiveness of the method proposed in this paper, the proposed method is compared with the fixed periodic route hopping strategy [40], random periodic route hopping strategy [41], and no hopping strategy. Assuming that the link flooding attacker adopts the optimal attack strategy to attack link 3, we repeated 20 groups of experiments, collected the actual defense gain, and calculated the average value, as shown in Figure 11. The average actual benefit of the method in this paper is 9.10, that of the fixed periodic route hopping strategy is 6.59, that of the random periodic route hopping strategy is 6.86, and that of the no hopping strategy is 3.76. Compared with the no hopping strategy, the method proposed in this paper increased the defense payoff by 58.7%. Compared with the fixed periodic route hopping strategy, the increase was 27.6%, and compared with the random periodic route hopping strategy, the increase was 24.6%.

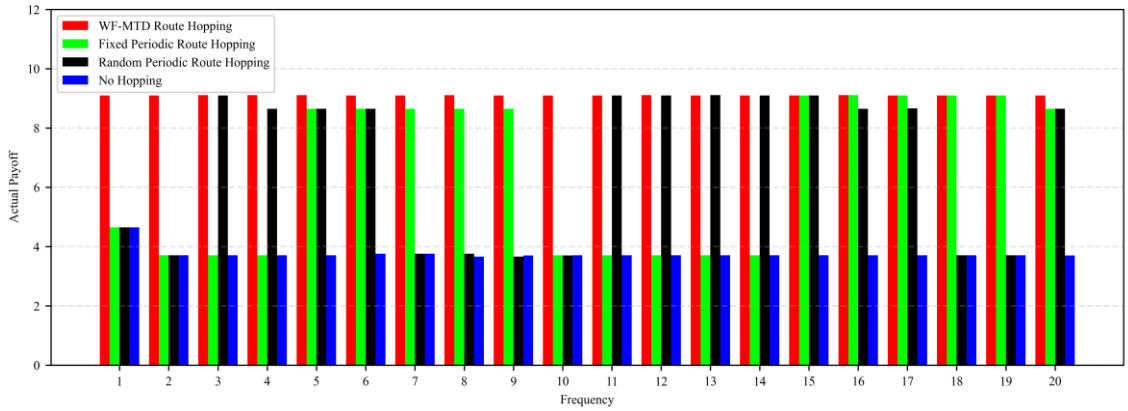

Figure 11. Comparison of actual gains of method proposed in this paper and other methods

V. CONCLUSION

The increasing complexity and scale of network information systems has led to an increase in the diversity of security attacks, making dynamic changes in network attack-defense conflicts. Comprehensively analyzing defense costs and gains, maximizing defense payoff, predicting possible attack strategies, selecting an optimal defense strategy from candidate strategies, and measuring the gains from a strategy remain huge challenges. Game theory is an effective method for studying the decision-making problems of moving target defense. At present, game research into the bounded rationality of attackers and defenders is still in early development. Many limitations in the rationality quantification, game structure, strategy types, and equilibrium solutions for attackers and defenders all affect the science and effectiveness of moving target defense game decision analysis models and methods.

From the perspective of bounded rationality of both attack-defense, we construct a moving target defense evolution strategy model based on the Wright-Fisher process. By quantifying the rationality of both attack-defense, we depict the evolution trajectories of attack-defense strategies and show the dynamic convergence process of these strategies. Our numerical experiments and application examples show that our model and method are versatile, effective, and practical, and greatly improve the performance of attack prediction and defensive decision-making.

ACKNOWLEDGMENTS

We thank all the reviewers for their valuable comments. This work was supported by the National Key Research and Development Program of China (Grant No. 2016YFF0204003) and National Natural Science Foundation of China (Grant No.61902427).

DATA AVAILABILITY

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

CONFLICTS OF INTEREST

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

REFERENCES

[1] Taylor P J, Dargahi T, Dehghantanha A, et al., "A systematic literature review of blockchain cyber security," *Digital Communications and Networks*, vol. 6, no. 2, pp. 147-156, 2020.

[2] Sengan S, Subramaniyaswamy V, Nair S K, et al., "Enhancing cyber–physical systems with hybrid smart city cyber security architecture for secure public data-smart network," *Future generation computer systems*, vol. 112, pp. 724-737, 2020.

[3] Cho J H, Sharma D P, Alavizadeh H, et al., "Toward proactive, adaptive defense: A survey on moving target defense," *IEEE Communications Surveys & Tutorials*, vol. 22, no. 1, pp. 709-745, 2020.

[4] Zhuang R, DeLoach S A, Ou X, "Towards a theory of moving target defense," in *Proceedings of the First ACM Workshop on Moving Target Defense*, 2014, pp. 31-40.

[5] Islam M M, Duan Q, Al-Shaer E, "Specification-driven Moving Target Defense Synthesis," in *Proceedings of the 6th ACM Workshop on Moving Target Defense*, 2019, pp. 13-24.

[6] Ghourab E M, Azab M, Mansour A, "Spatiotemporal diversification by moving-target defense through benign employment of false-data injection for dynamic, secure cognitive radio network," *Journal of Network and Computer Applications*, vol. 138, pp. 1-14, 2019.

[7] Lei C, Zhang H Q, Tan J L, et al., "Moving target defense techniques: A survey," *Security and Communication Networks*, 2018, doi: 10.1155/2018/3759626.

[8] Sharma D P, Enoch S Y, Cho J H, et al., "Dynamic Security Metrics for Software-Defined Network-based Moving Target Defense," *Journal of Network and Computer Applications*, vol. 170, 2020, doi:10.1016/j.jnca.2020.102805.

[9] Feng X, Zheng Z, Cansever D, et al., "A signaling game model for moving target defense" in *IEEE INFOCOM 2017 - IEEE Conference on Computer Communications*, pp. 1-9, 2017, doi:10.1109/INFOCOM.2017.8057200.

[10] Winterrose M L, Carter K M, Wagner N, et al., "Adaptive attacker strategy development against moving target cyber defenses," *Advances in Cyber Security Analytics and Decision Systems*, Springer, Cham, pp. 1-14, 2020.

[11] Sengupta S, Chowdhary A, Huang D, et al., "General sum Markov games for strategic detection of advanced persistent threats using moving target defense in cloud networks," in *International Conference on Decision and Game Theory for Security*, vol. 11836, pp. 492-512, 2019.

[12] Sengupta S, Kambhampati S, "Multi-agent reinforcement learning in bayesian stackelberg markov games for adaptive moving target defense," *arXiv preprint*, 2020, doi:10.48550/arXiv.2007.10457.

[13] Li H, Shen W, Zheng Z, "Spatial-Temporal Moving Target Defense: A Markov Stackelberg Game Model," *arXiv preprint*, 2020, doi:10.48550/arXiv.2002.10390.

[14] Zhang H, Zheng K, Wang X, et al., "Strategy Selection for Moving Target Defense in Incomplete Information Game," *Computers, Materials & Continua*, vol. 62, no. 2, pp. 763.786, 2020.

[15] Lakshminarayana S, Belmega E V, Poor H V, "Moving-Target Defense Against Cyber-Physical Attacks in Power Grids via Game Theory," *IEEE Transactions on Smart Grid*, vol. 12, no. 6, pp. 5244-5257, 2021.

[16] He Y, Zhang M, Yang X, et al., "The Intelligent Attack-defense Mechanism of Internet of Vehicles Based on the Differential Game-IP Hopping," *IEEE Access*, vol. 8, pp. 115217-115227, 2020.

[17] Niu L, Clark A, "A Framework for Joint Attack Detection and Control Under False Data Injection," in *International Conference on Decision and Game Theory for Security*, vol. 11836, pp. 352-363, 2019.

[18] Chang S Y, Park Y, Babu B B A, "Fast IP hopping randomization to secure hop-by-hop access in SDN," *IEEE Transactions on Network and Service Management*, vol. 16, no. 1, pp. 308-320, 2018.

[19] Al-Shaer E, Duan Q, Jafarian J H, "Random host mutation for moving target defense," in *International Conference on Security and Privacy in Communication Systems*. Springer, Berlin, Heidelberg, vol. 106, pp. 310-327, 2012.

[20] Gu Z, Zhang J, Ji Y, et al., "Network topology reconfiguration for FSO-based fronthaul/backhaul in 5G+ wireless networks," *IEEE Access*, vol. 6, pp. 69426-69437, 2018.

[21] Karthikeyan K, Sunder R, Shankar K, et al., "Energy consumption analysis of Virtual Machine migration in cloud using hybrid swarm optimization (ABC–BA)," *The Journal of Supercomputing*, vol. 76, no. 5, pp. 3374-3390, 2020.

[22] Wu J, *Analysis on Diversity, Randomness, and Dynamicity*, Cyberspace Mimic Defense, Springer, Cham, pp. 159-205, 2020.

[23] X. Yang, J. Hu, Y. Ji, L. Ge and X. Zeng, "Design of a Metasurface Antenna with Pattern Diversity," *IEEE Antennas and Wireless Propagation Letters*, vol. 19, no. 12, pp. 2467-2471, 2020, doi:10.1109/LAWP.2020.3035656.

[24] Hu H, Wu J, Wang Z, et al., "Mimic defense: a designed-in cybersecurity defense framework," *IET Information Security*, vol. 12, no. 3, pp. 226-237, 2017.

[25] Wang Y, Wu J, Guo Y, et al., "Scientific workflow execution system based on mimic defense in the cloud environment," *Frontiers of Information Technology & Electronic Engineering*, vol. 19, no. 12, pp. 1522-1536, 2018.

[26] Ye W, Fan S, "Evolutionary snowdrift game with rational selection based on radical evaluation," *Applied Mathematics and Computation*, vol. 294, pp. 310-317, 2017.

[27] Song B, Zhu J M, "Evolution entropy risk assessment of ERP information security based on the business process," *Journal on Communications*, vol. 33, no. Z1, pp. 210-215, 2012.

[28] Imhof L A, Nowak M A, "Evolutionary game dynamics in a Wright-Fisher process," *Journal of mathematical biology*, vol. 52, no. 5, pp. 667-681, 2006.

[29] Yin G S, Wang Y J, Dong Y X, et al., "Wright-Fisher muti-strategy trust evolution model of Internetware," *Expert Systems with Applications*, vol. 40, no. 18, pp. 7367-7380, 2013.

[30] Wang D, Cao Q Y, Xu H Y, "Stochastic evolutionary trust strategy of WSNs based on Wright-Fisher process," Computer Application and Software, vol. 34, no. 1, pp. 110-116, 2017.

[31] Wang X, Gu C, Zhao J, et al., "A Review of Stochastic Evolution Dynamics and Its Cooperative Mechanism," *Journal of Systems Science and Mathematical Sciences*, vol. 39, no. 10, pp. 1533-1552, 2019.

[32] Yang Y, Che B, Zeng Y, et al., "MAIAD: a multistage asymmetric information attack and defense model based on evolutionary game theory," *Symmetry*, vol. 11, no. 2, 2019.

[33] Wang S, Zhang Z, Kadobayashi Y, "Exploring attack graph for cost-benefit security hardening: A probabilistic approach," *Computers & security*, vol. 32, pp. 158-169, 2013.

[34] Harper M, "Inherent randomness of evolving populations," *Physical Review E*, vol. 89, no. 3, 2014, doi:10.1103/PhysRevE.89.032709.

[35] Ohtsuki H, Pacheco J M, Nowak M A, "Evolutionary graph theory: Breaking the symmetry between interaction and replacement," *Journal of Theoretical Biology*, vol. 246, no. 4, pp. 681-694, 2007.

[36] H. Hu, Y. Liu, C. Chen, H. Zhang and Y. Liu, "Optimal Decision Making Approach for Cyber Security Defense Using Evolutionary Game," *IEEE Transactions on Network and Service Management*, vol. 17, no. 3, pp. 1683-1700, 2020.

[37] Ramesh G, Menen A, "Automated dynamic approach for detecting ransomware using finite-state machine," *Decision Support Systems*, vol. 138, 2020, doi:10.1016/j.dss.2020.113400.

[38] T. Bossert, "It's Official: North Korea Is Behind WannaCry," *The Wall Street Journal*, 18 December 2017. [Online]. Available: https://www.wsj.com/articles/its-official-north-korea-is-behind-wannacry-1513642537.

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

[40] A. Aydeger, N. Saputro, K. Akkaya and M. Rahman, "Mitigating Crossfire Attacks Using SDN-Based Moving Target Defense," in *2016 IEEE 41st Conference on Local Computer Networks (LCN)*, pp. 627-630, 2016, doi:10.1109/LCN.2016.108.

[41] Zhang H, Lei C, Chang D, et al., "Network moving target defense technique based on collaborative mutation," *computers & security*, vol. 70, pp. 51-71, 2017.

[42] Cheng Lei, Duo-he Ma, Hong-qi Zhang, Li-ming Wang, "Moving Target Network Defense Effectiveness Evaluation Based on Change-Point Detection", *Mathematical Problems in Engineering*, vol. 2016, Article ID 6391502, 11 pages, 2016.