This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TNSM.2020.2995713, IEEE Transactions on Network and Service Management

Optimal Decision Making Approach for Cyber Security Defense Using Evolutionary Game

Hao Hu, Yuling Liu, Chen Chen, *Senior Member, IEEE*, Hongqi Zhang, and Yi Liu

***Abstract*—At present, there are many techniques for cyber security defense such as firewall, intrusion detection and cryptography. Despite decades of studies and experiences on this issue, there still exists a problem that we always pay great attention to technology while overlooking strategy. In the traditional warfare, the level of decision-making and the formulation of optimal strategies have a great effect on the warfare result. Similarly, the timeliness and quality of decision-making in cyber attack-defense also make great significance. Since the attackers and defenders are oppositional, the selection of optimal defense strategy with the maximum payoff is difficult. To solve this problem, the stochastic evolutionary game model is utilized to simulate the dynamic adversary of cyber attack-defense. We add the parameter λ to the Logit Quantal Response Dynamics (LQRD) equation to quantify the cognitive differences of real-world players. By calculating the evolutionary stable equilibrium, the best decision-making approach is proposed, which makes a balance between defense cost and benefit. Cases studies on ransomware indicate that the proposed approach can help the defender predict possible attack action, select the related optimal defense strategy over time, and gain the maximum defense payoff.**

***Index Terms*— cybersecurity; attack-defense; decision making; optimal defense strategy; evolutionary game; LQRD**

I. INTRODUCTION

With the complexity of large-scale cyber information system, security attacks become more and more diversified. According to the statistics [1], in the past 2016, the Alibaba cloud has suffered an average of about 300 million times of HTTPS SSL/CC attacks each day. These hackers tried to invade about 20 million Taobao user accounts. For different attackers, each attack-defense scenario may only occur once. However, for the defenders such as Alibaba cloud, they face many same attack scenarios every day. Considering the limited hardware resources of network devices, how to make a tradeoff between defense costs and benefits as well as to select the best action plan to maximize defense revenue, so that the defender can reach a balance between risk and investment has become one of the major challenges of cyber security [2]. As shown in Fig. 1, for the smart city, the more computer systems such as the public water system, power grid, vehicle traffic control systems, street lighting, sensor devices, and physical security systems, the more integration between the systems, the more vulnerable to cyberattacks [3][4][5][6]. To mitigate risks, developing an appropriate security strategy is of paramount importance.

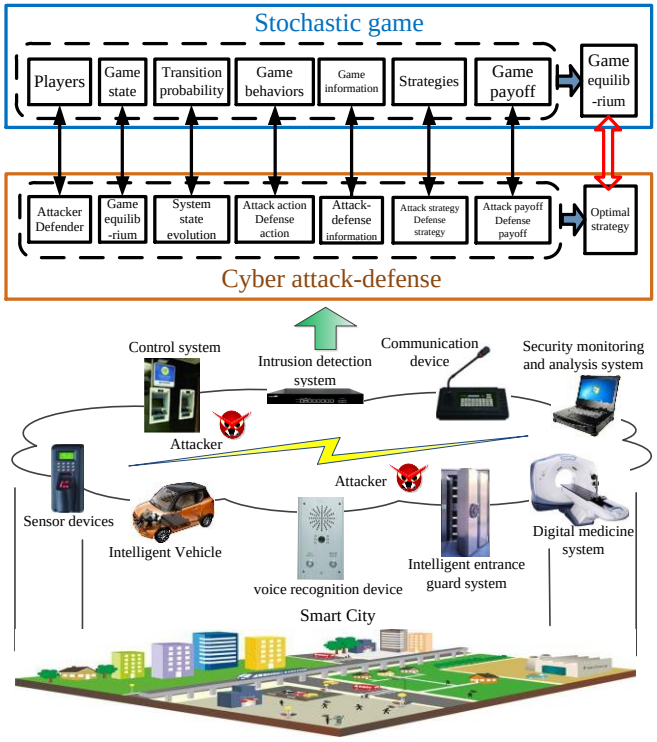


Fig. 1. Game theory for cyber security defense in smart city.

However, existing studies on decision making for cyber security pay more attention to technology while overlook the tactics strategies. Although there are many technologies and products for cyber security defense such as firewall, intrusion detection system (IDS) and data encryption system (DES), the confrontation of cyber attack-defense is essentially a confrontation between attack strategy and defense strategy. The application of defense mainly focuses on the scale of the number of defense methods, while ignoring the combination, coordination and linkage control of multiple defense methods in the space-time dimension. Especially for the emerging APT attack, an emerging comprehensive attack with strong concealment, long latent period and high confrontation. To defend against APT attack, it is necessary to use a variety of defense means in an appropriate combination so as to maximize defense effectiveness.

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

Yuling Liu (Corresponding author) is with the State Key Laboratory of Information Security, Chinese Academy of Sciences, Beijing 100190, China

Chen Chen is with the State Key Laboratory of Integrated Service Networks, Xidian University, Xian 710071, China

Hongqi Zhang is with the China National Digital Switching System Engineering & Technological Research Center, Zhengzhou 450001, China

Yi Liu is with the Zhengzhou Information Science and Technology Institute, Zhengzhou 450001, China.

The quality of decision-making and the implement of defense strategy still make a significant influence. When the gap of technical level between attack and defense is small, the quality of decision-making directly determines the situation of the cybersecurity. When the gap of technical level of attack-defense is large, the side with the lower technical level can strive for gaining best interests through scientific decision-making and even the turnaround the disadvantage situation. Therefore, research on defense decision-making is of great significance and has great value for enhancing adversarial capabilities as well as ensuring cybersecurity.

Game theory is a decision-making theory for studying the direct interaction among decision makers [7], whose goal is to maximize the earnings of players, and is suitable for analyzing the strategy selection issue when the behaviors of decision-makers interact directly. In general, it mainly includes player, state, behavior, information, strategy, payoff, and equilibrium elements. Game theory has the characteristics of objective opposition, non-cooperative and strategic interdependence, all of which are in line with the basic characteristics of cyber attack-defense [8]. Therefore, how to apply the game theory to model and analyze cyber attack-defense process has become a hot research issue in recent years [9][10][11]. However, there are still some challenges. To our best knowledge, existing game models for cyber attack-defense are mainly based on the hypothesis of complete rational players [12][15][16]. Complete rationality includes many preconditions that are difficult to achieve, such as perfect rational consciousness, perfect ability of analyzing and inferring, identifying and judging, memorizing and computing. If any of these conditions cannot be reached, it belongs to bounded rationality. The strict requirement of complete rationality is too harsh for the social attacker and defender. In fact, the real-world attackers and defenders have different abilities of cognizance, which is determined by their own interests such as safety knowledge, skill level, experience, and so on [13]. In a word, the selection of strategy affected by various uncertain factors leads to the bounded rational game. At present, this issue is still assumed as a great challenge.

The main contributions of this paper are as follows:

1) The population game for modeling attack-defense is promoted. The modeling from individual game to the population game is an improvement. In the population game, through population strategy learning and improvement, the proportion of each strategy in the final population is stable, and the greater the proportion, the higher the population acknowledges to the selected strategy.

2) **The stochastic evolutionary game model for describing the dynamic process of attack-defense is constructed**. Complex attack-defense scenarios are nonlinear, uncertain, dynamic, highly confrontational and time-varying evolutionary. The existing adversary modeling methods based on system identification or engineering experience cannot reflect the high-intensity confrontation of attack-defense. Especially in complex scenarios, both sides of attacker and defender will feed back and co-evolve, the proposed evolutionary game model try to depict the interaction of decision-making and behavior evolutions. We use Logit Quantal Response Dynamics (LQRD) equation to describe the evolutionary mechanism of attack-defense strategies. Our model is more flexible for social cyber players with different rationalities comparing with other game models. We distinguish different social players by measuring and quantifying their rational degrees. Moreover, we simulate the growth of the rationality. Our model tells decision-makers what actions and decisions they need to take, the maximum revenue, and what impact the strategy has on its adversary.

3) **The approach for selecting the optimal defense strategy over time is proposed**. The optimal defense strategy is given by solving the evolutionary stable equilibrium. Compared with the static Nash equilibrium, the dynamic selections at different evolutionary times and its earnings are depicted visually, and the evolution tracks with the best strategies for both sides of attacker and defender are exhibited at the same time. In addition, we give the convergence process from evolutionary equilibrium to Nash equilibrium, which improves the dynamic analysis and situation prediction of attack early warning and defense decision-making. Through strategy "learning and improvement", the defenders are guided to select the best strategies in the continuous confrontation process between the two sides of attack and defense so as to achieve strategic advantage.

4) **The formation of the optimal strategies of both sides of attacker and defender as well as the maximum defense revenue is analyzed.** In the classical complete rational game model, the best strategy is explained as the best reaction between adversaries, but no forming process of this strategy is given. In the emerging bounded rational game model of replicator dynamic, it only gives pure strategy for decision making, but not provides the practical mixed strategy like the combination of several defense techniques. From this aim, the emphasis of this paper is to analyze the dynamic evolution tracks of optimal mixed strategies. We first simulate the formation process of player strategy through repeated games. Our approach can predict the possible attack strategy in future and provide the corresponding best defense plan.

II. Related Works

The main existing techniques for defense decision-making are developed on finite-state machine, rule system, cybernetics, expert system, hierarchical task-network, case reasoning, biological evolution, influence network, game theory and so on. Most methods are essentially modeled based on expert domain knowledge, and are more applicable to the analysis and statistics of uncomplicated systems. Although their strategy results are explainable and gain low computational cost, while difficult to express human tacit knowledge. Therefore, they are weak in simulating the decision process of human brain and hard to depict the evolutions of complex attack-defense system.

The interests of attackers and defenders are in conflict, so it is suitable to use game theory to model the process of attack-defense confrontation. Depending on the rational degree of the players, existing researches can be divided into two categories: complete rational game and bounded rational game.

Complete rational game takes the hypothesis that the players have full cognition. Each player not only can select the best strategy to maximize itself payoff, but also can predict other players' strategy selections. Through maximizing the expected defense payoff, the optimal response of each player is

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TNSM.2020.2995713, IEEE Transactions on Network and Service Management

calculated by Nash equilibrium. Jiang *et al*. [14] regarded both attacker and defender as players in game and treated the attack-defense adversary as the zero-sum game. Jiang *et al*. regarded players have complete information and act simultaneously. In addition, he constructed the non-cooperative static game model based on defense graph to analyze attack intention and select optimal defense strategy. Aiming at the security issue of sensor networks, Li *et al*. [15] constructed a non-cooperative game model between attackers and trusted sensor nodes to balance costs and benefits. Due to the restrictions of Nash equilibrium solution, Li *et al*. [16] used the Pareto optimization to calculate the equilibrium. With the help of Bayes game model, Liu *et al*. [17] analyzed the impact of attack-defense strategy changes on the defense evaluation of worm attack-defense performance. However, this method is limited to pure strategy Nash equilibrium. Patcha *et al*. [18] analyzed the optimal mixed strategy of IDS intrusion response. Carroll e*t al.* [19] regarded the defender as the signal sender and the attacker as the signal receiver, and built attack-defense signal game model. The attacker identifies and adjusts the cognition of the defender according to the defense signal. Then the single-stage and multi-stage signal game models are developed respectively. By calculating the equilibrium of each stage, the best defense strategy is given to assist the decision for different types of defenders during different stages. Almost the same time, Lei *et al*. [20] constructed the multistage repeated game model of attack-defense, in which the attacker type was inferred depending on the priori attack strategy from the defender viewpoint, and the posterior inference was revised to improve the accuracy of the decision progressively. The above multi-stage models mainly analyze discontinuous behaviors in discrete time periods.

In summary, all the studies above are based on the complete rationality assumption of players. They first quantify the strategy payoffs according to their types, and then construct the payoff matrices to calculate the Nash equilibrium. However, they do not consider whether the players' complete rationality is in line with the reality of biology property of players. In fact, the environment and individual factors affect the attack-defense players, so their behaviors can hardly reach complete rationality. To a certain extent, they are bounded rational agents [21], and the strategy selection is the process of continuous learning and adjustment. Without the premise of bounded rationality, the modeling and analysis of attack-defense may be impractical. Therefore, studying the rules of attack-defense game under bounded rationality is an applied and promising research issue.

The bounded rational attack-defense games take the thought of biological population evolution. Biologists found that using game theory can better explain many biological problems such as the evolution of biological population. Combining biological knowledge with game theory, they proposed evolutionary game theory [38], which reduces the limitation on the complete rationality of players. They think that individual players first select strategies from the population randomly and then study the best strategy through repeated games. By developing the learning and imitating mechanism, they can improve their selections constantly, and evolve to a stable state gradually. Compared with the Nash equilibrium under complete rational, which only gives the final solution, but does not explain the formation of the solution, we consider that the equilibrium is usually not achieved in one-step. If the attack-defense is a continuous process, the evolutionary game can better explain the dynamic process of equilibrium formation, which has been widely used in transportation, sociology and bioinformatics. Recently, some scholars have tried to use evolutionary game models to describe the population behaviors of attackers and defenders for cyber security. James *et al*. [22] analyzed the cooperative behavior cascades in human social networks, he exploit a seminal set of laboratory experiments that originally showed that voluntary costly punishment can help sustain cooperation in social network, which broadens the viewpoint of strategy evolution with social attacker and defender. Tosh *et al*. [23] proposed the framework of evolutionary game model for sharing cyber information. Zhu *et al*. [24] used system dynamics to explore the evolution characteristic of attack-defense for cyber security governance. However, this study only divides the security strategies into two categories: attack/no attack and defense/no defense, and the detailed defense for specific attack is not given. Ruan *et al*. [25] and Abass *et al*. [26] used the replicator dynamics equation to calculate the evolutionary stable strategy for against denial of service attacks (DoS) and advanced persistent threat attacks (APT) respectively. Hayel *et al*. [27] established an evolutionary Poisson game framework to capture the stochastic, dynamic, and heterogeneous interactions of agents in a holistic fashion, and designed mechanisms to control their behaviors to achieve a system-wide objective. Considering that the randomness of attack-defense can lead to the state transition of the network system, in order to analyze the equilibrium of the system under different states, [28] analyzed the optimal defense under different states through combining Markov chain and intelligent learning.

From the existing researches, although the application of evolutionary games in the field of cybersecurity has achieved remarkable progress, but there still exist some challeges. In terms of describing evolutionary behavior, existing studies depend on the replicator dynamic equation, which requires that players always select the strategy from high-payoff candidate strategies in strategy learning. This is the deterministic evolution essentially. Affected by incomplete information, cognitive difference, local shortsightedness and other factors of biological players, the rationality of different players are quite different. We believe that strategy selection is a social behavior having the characteristics of inertia and randomness. For example, when attackers implement APT attacks, the techniques are similar with different attackers. Most attackers have inertia in their selections. However, in some large-scale and high-intensity adversarial network, combining with experience, some attackers may try to change the strategy to obtain better attack reward. Nevertheless, the existing model cannot depict the inertia and variation of network population behavior. Therefore, the scientificality of modeling of cyber attack-defense is still assumed as a challenge.

This paper starts from the bounded rationality of both sides of attacker and defender, and builds an attack-defense evolutionary game model based on LQRD. As a stochastic game model, it is applicable to the repeated game for describing the inertia and mutation of player population. We improve the

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TNSM.2020.2995713, IEEE Transactions on Network and Service Management

LQRD to consider the cognitive differences and strategy preferences. Furthermore, we design the approach of optimal defense strategy selection. Moreover, the evolution tracks of the attack-defense strategies are depicted, which improve the defense accuracy and attack early warning. Case studies on ransomware verify the scientificity and effectiveness of our approach.

III. Stochastic Evolutionary Game Model for Cyber Attack-Defense

Different from the general replicator dynamics game model, this section builds the attack-defense game model based on improved LQRD stochastic evolutionary game. Stochastic evolutionary game is a game analysis method for multi-player and multi-state decision-making, which conforms to cyber attack-defense. It can effectively deal with the impact of randomness on the attack-defense process. By adding the parameter λ, we quantify the cognitive differences of diverse players. Through which, we improve the previous approach by depicting the inertia, randomness, dynamics and diversity of real-world biological players. In this section, we first demonstrate our motivation and then construct the ASEGM attack-defense game model using LQRD. Finally, we give the metric of strategy payoff based on cost and benefit analysis.

A. Motivation

From the perspective of decision-making, we can abstract the security adversary as a stochastic game. If we treat the time of the whole attack-defense process containing a series of time slices as shown in Fig. 2, each time slice corresponds to one security state of the cyber system, and then the attack-defense actions can be treated as a series of discrete events occurring at discrete times. In this way, we can process the cyber attack-defense process discretely. In each time slice, the player detects the current network state. Taking the time slice t_1 as an example, the player evaluates the present optimal action according to the system state and the adversary's action. The game ends when the network transfers to the security state. During this process, when one side changes its strategy, the game system moves to an unbalanced state. Then, the security state of time slice t_1 transfers to the next state of time slice t_2 under the actions of both attacker and defender. From the dimension of time, the attacker and defender makes continuous decision and dynamic adversary over time. The game equilibrium strategy is treated as the controller of the track depicting state evolution. By predicting the optimal strategy at each game moment, we can improve the accuracy and timeliness of security decision-making.

Let's explain our motivation with an example as follows.

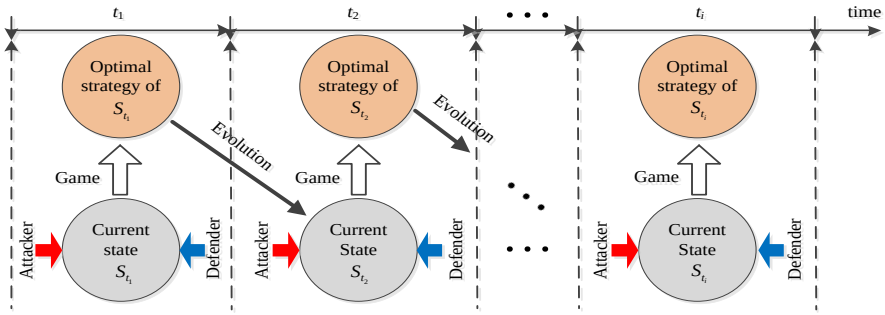


Fig. 2. The process of optimal strategy selection for attack-defense game varying with time.

1) For the bounded rational attack-defense modeling, existing researches generally adopt the replicator dynamic [30]. The core idea is that the population with lower rationality will gradually adopt the strategy with higher average payoff, which essentially reflects the deterministic selecting behavior [22]. In other words, the mechanism of selection determines that players always select the high-payoff strategies during evolution. For instance, we assume a population containing three game players, in the first evolution round, the obtained payoffs are respectively 3, 8, 10. Then we can calculate the average payoff 7. In the second round, due to that the player with the payoff 3 earns less than the average level. He prefers to select the strategy with payoff 10 and obtains the corresponding payoff level. Then the three payoffs are 10, 8, 10 and the average payoff is 9.3. In the third round, since the player with payoff 8 earns less than the average level, he adopts the strategy of payoff 10. In this way, all the three players obtain the payoff 10 after three rounds of game. Population realizes the evolution from low payoff to high payoff. In this process, it requires that the game players can always identify and learn the high-payoff strategy accurately. However, due to the influence of the attack-defense players' own factors (skill level, safety knowledge, priori experience, etc.), the cognitive abilities of different players are usually different. Not all individuals can correctly calculate the expected strategy payoff. These individuals have certain probabilities of changing planned strategies. We call this as stochastic disturbance. Moreover, for some multi-variant replicator dynamic equations, there are no polymorphic equilibrium solutions, which reduces the operability.

2) The LQRD [31] considers that the player belief is keeping rational, while the limitation of learning capability leads to the gap of achieving an ideal Nash equilibrium. We denote the payoff as $U = V + \varepsilon$, where V denotes the payoff generated by deterministic factors, while ε denotes the payoff generated by uncertain disturbance. The players make decisions for gaining maximum payoff U. Compared with the general replicator dynamics [30], we further consider the individual preferences and cognitive differences. On this basis, we introduce the parameter λ to quantify the degree of player rationality so as to reflect the diversities of population behaviors. Meanwhile, with the increase of λ by strategy learning and improving, players can obtain indirect decision information by observing their own experience or other players' decisions in similar

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TNSM.2020.2995713, IEEE Transactions on Network and Service Management

environments. Meanwhile, the players can also obtain direct decision-making information of the population's strategy distribution from the observed game history. Through repeated games, the cognitions of players are enhanced. We exhibit the best strategy selection varying with time, which improves the interpretation and prediction for strategy formation.

B. Architecture of Our Optimal Defense Decision Making Approach

The architecture of our optimal defense decision making approach is illustrated in Fig. 3. The input includes evidences such as vulnerability database, IDS real-time alert, security configuration, network topology and other security information, and the output is the optimal defense strategy. The decision making process contains four steps: (1) Determine the targets and principles of strategy selection. (2) Extract candidate attack-defense strategies from the input security data according to the feature analysis of attack evidence and abnormal evidence. (3) Model the process of attack-defense as the stochastic evolutionary game based on LQRD model. (4) Evaluate the strategy payoff based on cost-benefit analysis. (5) Generate optimal defense strategy.

Two additional notes:

(1) We consider that actual cyber adversary scenario usually consists of multiple players. To simplify the analysis, we divided players into different populations according to their types, such as the moderate type attacker, aggressive type attacker and dangerous type attacker.

(2) In Step (3), we extract the set of candidate defense strategies by analyzing the network environment information, including the vulnerability repairs, firewall access rules, security configuration and so on. We further collect alert data of firewall, IDS and virus detection system and host audit log. By analyzing the attack behavior information, we can extract the set of candidate attack strategies by referring to the network behaviors database of MIT laboratory [35].

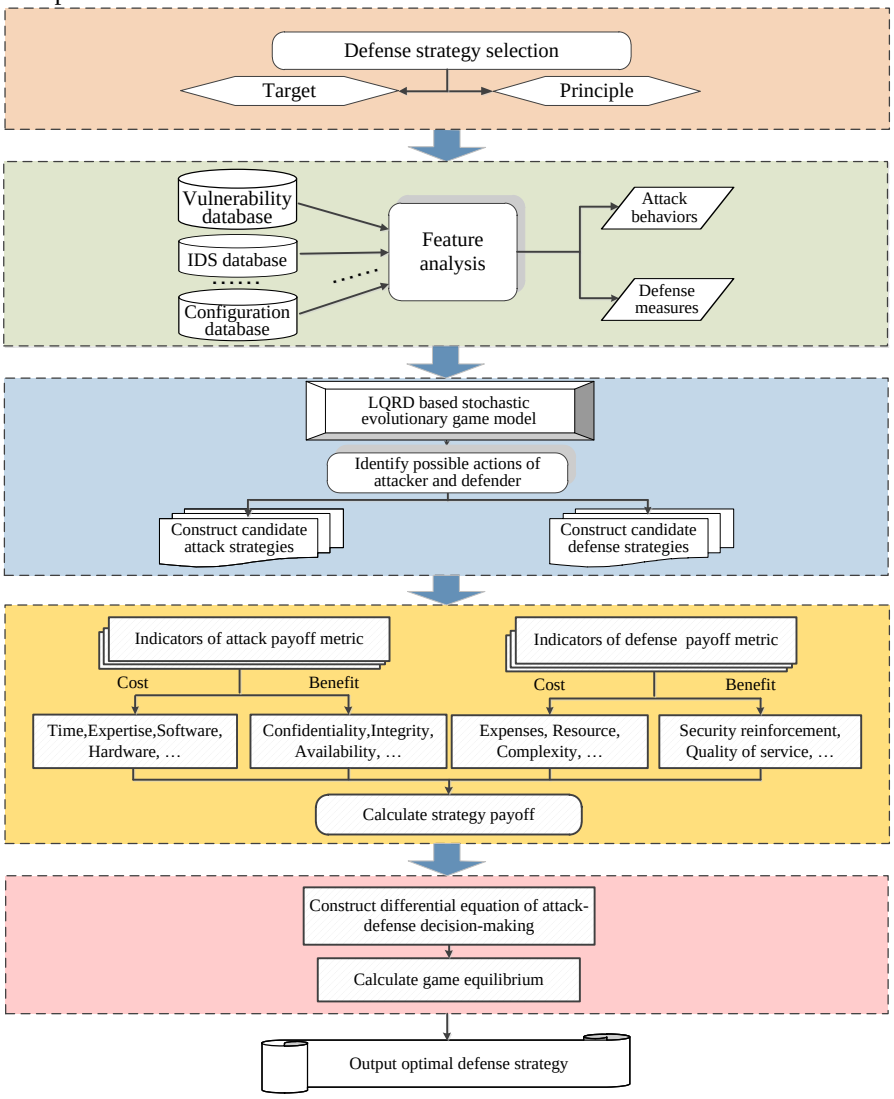


Fig. 3. The architecture of our decision making approach.

C. Game Modeling of Attack-Defense Based on LQRD

Based on the analysis in section III-A, we give the definition of evolutionary game model as follows. It includes 4 basic elements: player sets, candidate strategy set, belief set and payoff set.

Definition 1. The model of Attack-defense Stochastic Evolutionary Game (ASEGM) can be denoted by a four-tuple.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TNSM.2020.2995713, IEEE Transactions on Network and Service Management

(1) $N=(N_A,N_D)$ is the population set of the attack-defense players, in which N_D and N_A are the populations of defenders and attackers respectively.

(2) $S=(S_A,S_D)$ is the set of candidate attack-defense strategies, in which $S_A=\{A_1,A_2,\cdots,A_n\}$ is the set of the candidate strategies for attackers, $S_D=\{D_1,D_2,\cdots,D_m\}$ is the set of candidate strategies for defenders. n and m are the number of attack and defense strategies respectively, where $m,n\in N^+$ and $m,n\geq 2$.

(3) $\theta=(P,Q)$ is the belief set of attack-defense game, where $p_i\in P$ represents the probability that attacker selects candidate strategy A_i, $q_j\in Q$ represents the probability that defender selects candidate strategy D_j, where $1\leq i\leq n$, $1\leq j\leq m$, $\sum_{i=1}^{n}p_i=1$, $\sum_{j=1}^{m}q_j=1$.

(4) $U=(U_A,U_D)$ is the payoff function set. U_A and U_D represent the payoff functions of attack and defense respectively.

Modeling the game parameters like the attack/defense players, actions, strategies, payoffs, network states and so on are significant. Its basic idea is to integrate network topologies, assets, vulnerabilities, attack means, protection mechanisms, alert logs and other network operation and maintenance information in depth. Then, describe the multi-source heterogeneous network security elements uniformly. Afterwards, through deep fusion and association analysis, the knowledge can be displayed using visual technologies such as attack graph, attack tree, privilege graph and knowledge graph. Finally, we can extract attack-defense actions and network states information. The final effect is shown in Fig.4.

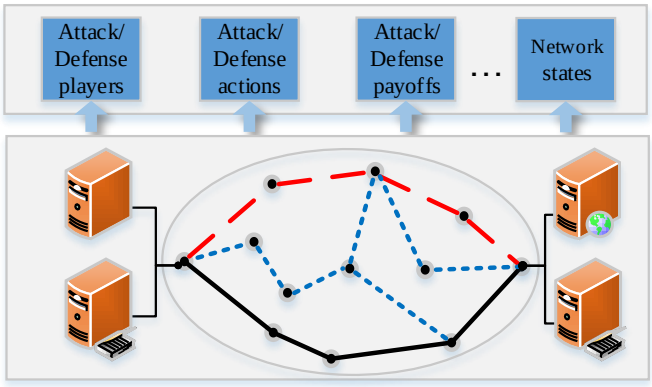


Fig.4. The effect of extracting the game parameters from the network.

To model the above process, we need to comprehensively use the theory and methods including ontology, finite state machine, graph theory and big data analysis. The related modeling languages contain OWL, SQWRL, LAMBDA, JIGSAW, etc. With the help of these theory and language tools, the unified expression of attack-defense elements and their relationship knowledge bases can be developed. Related techniques can be referred to literature [40, 41, 42]. Meanwhile, some tools like SIEM system [39] and MulVAL toolkit [37] also can be accessed to achieve the same effects automatically.

D. Game Payoff Quantification of Attack-Defense Strategy

Considering the condition 4) of Definition 1, the payoff quantification of the attack-defense strategy is the basis of defense strategy selection. Therefore, its accuracy directly affects the selecting results. We summarized the types of different attack-defense strategies and proposed the payoff metric based on cost-benefit analysis.

Definition 2. Attack Benefit (AB) is the earned network resources through a series of attack actions or the level of network damage, which reflects the capability of controlling the targeted network system.

Definition 3. Attack Cost (AC) is the cost or effort that an attacker pays to obtain network resources or cause losses to the network system.

Definition 4. Defense Benefit (DB) includes direct benefit and indirect benefit. The direct benefit is the level of security reinforcement through security measures. [14][16] only consider the direct benefits, and we further add the indirect benefits of defender through counterattack. For example, the electronic evidences of port scanning time, port number, source IP address and destination IP address can be used to reconstruct the attack chain. Through which, the defender can earn indirect benefits through investigating the criminal responsibility.

Definition 5. Defense Cost (DC) is the cost or effort that defenders take against the possible attacks, including the human and time cost of implement of security devices, and the economic cost of affecting the normal operation of service.

Definition 6. Attack-defense payoff matrices M are as follows. In which, a_{ij} and d_{ij} represent the attack and defense payoff of selecting strategy combination (A_i,D_j) respectively, $a_{ij}=AB-AC$, $d_{ij}=DB-DC$.

$$M=\begin{bmatrix} a_{11},d_{11} & a_{12},d_{12} & \cdots & a_{1m},d_{1m} \\ a_{21},d_{21} & a_{22},d_{22} & \cdots & a_{2m},d_{2m} \\ \cdots & \cdots & \ddots & \cdots \\ a_{n1},d_{n1} & a_{n2},d_{n2} & \cdots & a_{nm},d_{nm} \end{bmatrix} \quad (1)$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TNSM.2020.2995713, IEEE Transactions on Network and Service Management

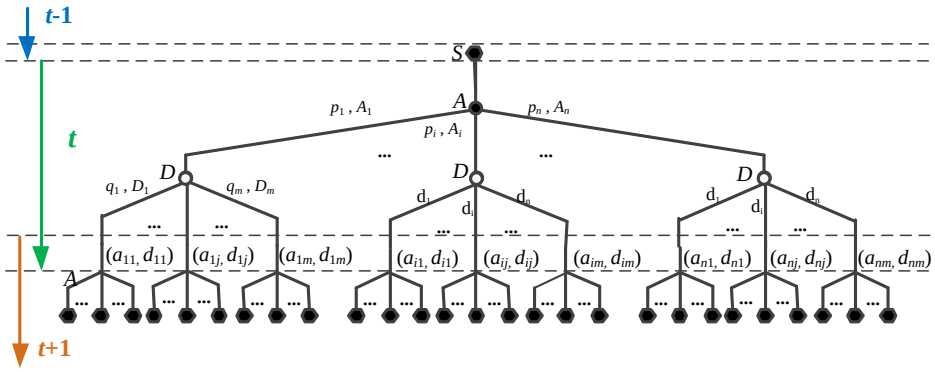


Fig. 5. The game tree of attack-defense.

The attack-defense game tree shown in Fig. 5 depict the attack-defense payoff with different combinations of candidate strategies visually. When the defender detects attack events, the game analysis begins at time *t*-1. The details of the process are as follows:

(1) Both sides of attacker and defender detect the current security state of network system in the period of *t*,

(2) Players select their own optimal actions according to the candidate strategies and their payoffs in the period of *t*.

(3) Players earn actual rewards through implementing strategies.

(4) Some players change their strategies through learning and modifying in the period of *t*+1.

(5) Repeat step (1-4) until the game system transfers to a stable state, that is, no player can make higher earnings by changing itself strategy alone.

IV. Optimal Defense Decision Making Approach

A. Construction of Evolution Equations for Attack-Defense Decision Making

Evolutionary stable strategy (ESS) is an optimal decision of the game system in the process of long-time strategy evolution. We obtain the best strategy, which is balanced and stable and can against the invasion of other strategies. The definition of evolutionary stable strategy of cyber attack-defense is as follows.

Definition 7. Suppose the attacker population selects the candidate strategy set $S_A=(A_1,A_2,\cdots,A_n)$ with the probability distribution $P=(p_1,p_2,\cdots,p_n)$, and the defender population selects the candidate strategy set $Q=(q_1,q_2,\cdots,q_m)$ with the probability distribution $S_D=(D_1,D_2,\cdots,D_m)$, it means that individual in the attacker and defense population randomly selects and implements their pure strategies with the probability distribution *P* and *Q* in actual adversary. We call $\sigma^*=(P,\ Q)$ as the stable strategy of attack-defense if the following conditions hold. $U(\sigma^*,\sigma^*)$ denotes the payoff when attacker and defender both select σ^*. For any $\sigma\neq\sigma^*$, $U(\sigma^*,\sigma)$ is the payoff when either side changes its strategy. Then the following conditions hold.

1) (stability) $U(\sigma^*,\sigma^*)\geq U(\sigma,\sigma^*)$

2) (balance) $U(\sigma^*,\sigma^*)=U(\sigma,\sigma^*)\Rightarrow U(\sigma^*,\sigma)>U(\sigma,\sigma)$

Condition 1) guarantees that both attacker and defender cannot earn more if either side changes strategy. In a strategy containing a large number of σ^* and a small number of σ, it is necessary to meet that *A* is the best response to itself, otherwise other strategies may invade and develop. Condition 2) guarantees that if there is another optimal strategy σ, σ^* is required to react better than σ, which ensures that σ cannot develop even if the strategy mutates to σ.

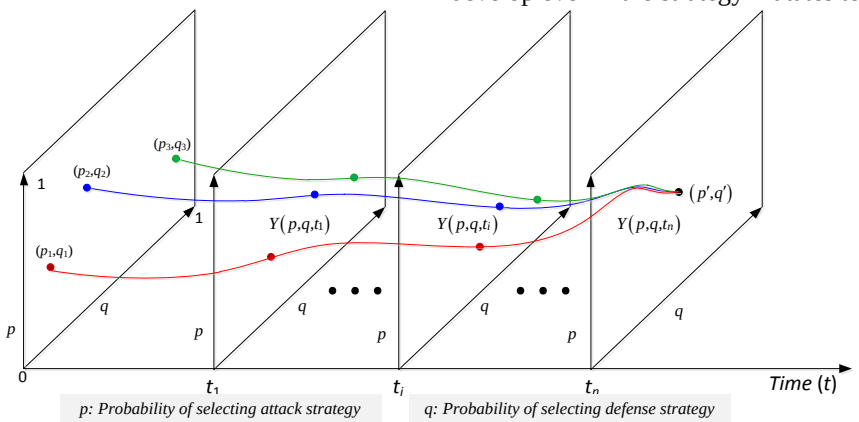


Fig. 6. The evolution tracks of strategy selection of the game system over time.

To explain the Definition 5, Fig. 6 shows a possible dynamic track of the strategy evolution over time. $Y(p,q,t_1)$ is the probability function of selecting strategy. *t* is the evolution times. *p* and *q* denote the probabilities of selecting attack and

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TNSM.2020.2995713, IEEE Transactions on Network and Service Management

defense strategy respectively. The red, blue and green curves depict the evolution tracks of the game system with different initial states $(p_1,q_1),(p_2,q_2),(p_3,q_3)$. We can derive that the system can evolve to the same stable state (p',q') through multiple repeated games. The middle strategy at time t_0, t_1, t_i, t_n can be obtained through intercepting the time surfaces of game system. The optimal strategy we want to receive meets the Fig. 6, that is, no matter what action the attack-defense players take at the initial moment, through strategy learning, imitation and improvement, we can get the best strategy ultimately. The strategy is stable and antijamming, which can defend the invade of other strategies. The key point is how to give the construction of $Y(p,q,t_1)$.

Definition 7 gives the condition of whether a strategy is an evolutionary stable strategy, but does not characterize the track of players' selection on this strategy. As described in section III-A, the attack-defense players search the best strategy along with being disturbed by stochastic error. This section describes the strategy evolution track by modifying the LQRD equation to indicate the randomness of selection. The LQRD uses Fisher-Tippett (an independent-identical-distribution) to depict the degree of noise influence on different players [33]. That is to say, the player selects the strategy with the exponential probability distribution, which is in line with the law of evolution of most things in the real world. Herein, we first give the deduction of the proposed LQRD equation in combination with Definition 8 - Definition 10.

Definition 8. The differential equation [32] of the probability of selecting this strategy is

$$\frac{dq_j}{dt}=\left(\sum_{k=1}^{m}q_k c_{kj}-\sum_{y=1}^{m}q_j c_{jy}\right) \tag{2}$$

q_j is the probability of selecting strategy D_j, and c_{kj} is the conditional transition probability of the defender's selection from strategy D_k to strategy D_j, which describes the updating rules of strategy selections. $\frac{dq_j}{dt}$ denotes the change rate of the population proportion of selecting strategy D_j in the entire defender population varying with time, which also reflects the defender's learning and adjusting speed of selecting strategy D_j through repeated games.

The core of attack-defense evolutionary game is to study the dynamic change speed of proportion of individual selecting strategy in the total population. That is, we need to calculate the selecting probabilities of different strategies. For conditional transition probability, we use LQRD equation to describe the rules of strategy updating, and add an extra rationality parameter λ to quantify the cognitive capabilities of different game players. The improved LQRD equation is defined as follows.

Definition 9. Improved transition probability of LQRD is

$$c_{kj}=\frac{\exp(\lambda U_{D_j})}{\sum_{k=1}^{m}\exp(\lambda U_{D_k})} \tag{3}$$

Set the rational parameters $\lambda(\lambda\geq 0)$ based on the historical rational degrees of players. The bigger λ is, the higher the degree of rationality is. As described in section III-A, the payoff is $U=V+\varepsilon$, where V is the payoff of observable factors, ε is the payoff of uncertain factors. The deduction of c_{kj} in Definition 9 can be referred to [33]. Take the formula in Definition 9 into Definition 8 and get the LQRD equation as following definition 10.

Definition 10. Evolution equation of LQRD is

$$\frac{dq_j}{dt}=\frac{\exp(\lambda U_{D_j})}{\sum_{k=1}^{m}\exp(\lambda U_{D_k})}-q_j \tag{4}$$

The formula shows that the change rate of the population proportion of player selecting strategy D_j is proportional to the difference between the proportion of individual expected payoff to the total payoff and the proportion of individual numbers of selecting this strategy to the total numbers.

Definition 10 shows that in the defender population composed of bounded rational players, the number change rate of players selecting certain candidate strategy varies with the proportion of this strategy payoff to the total payoff.

In order to construct the LQRD equations of attack-defense, from condition 3) of the Definition 1, we denote the strategy of probability vectors $P=(p_1,p_2,\cdots,p_n)$ and $Q=(q_1,q_2,\cdots,q_m)$ as the mixed probability of selecting S_A and S_D respectively. The evolution equations are as follows.

1) Evolution equation of defense strategy over time

The expected payoff U_{D_j} of defender selecting candidate strategy D_j is as follows, $j=1,2,\cdots,m$

$$\begin{cases}U_{D_1}=p_1d_{11}+p_2d_{21}+\cdots+p_nd_{n1}\\ U_{D_2}=p_1d_{12}+p_2d_{22}+\cdots+p_nd_{n2}\\ U_{D_j}=p_1d_{1j}+p_2d_{2j}+\cdots+p_nd_{nj}=\sum_{i=1}^{n}p_id_{ij}\\ \cdots\\ U_{D_m}=p_1d_{1m}+p_2d_{2m}+\cdots+p_nd_{nm}\end{cases} \tag{5}$$

The changing rate of proportion of individuals selecting strategy D_j in the defender population over time is $\frac{dq_j}{dt}$. It reflects the learning and improving process of selecting strategy D_j for bounded rational defender through repeated games. The LQRD differential equation of change rate $\frac{dq_j}{dt}$ is

$$\frac{dq_j}{dt}=\frac{\exp(\lambda\sum_{i=1}^{n}p_id_{ij})}{\sum_{k=1}^{m}\exp(\lambda\sum_{i=1}^{n}p_id_{ik})}-q_j \tag{6}$$

2) Evolution equation of attack strategy over time

The expected payoff U_{A_i} of selecting candidate strategy A_i for attacker is as follows, $i=1,2,\cdots,n$.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TNSM.2020.2995713, IEEE Transactions on Network and Service Management

$$\begin{cases} U_{A_1} = q_1 a_{12} + q_2 a_{12} + \cdots + q_m a_{1m} \\ U_{A_2} = q_1 a_{22} + q_2 a_{22} + \cdots + q_m a_{2m} \\ U_{A_i} = q_1 a_{i1} + q_2 a_{i2} + \cdots + q_m a_{im} = \sum_{j=1}^{m} q_j a_{ij} \\ \cdots \\ U_{A_n} = q_1 a_{n1} + q_2 a_{n2} + \cdots + q_m a_{nm} \end{cases} \quad (7)$$

Similarly, the changing rate of proportion of selecting strategy A_i in the attacker population over time is $\frac{dp_i}{dt}$. It shows the learning and improving process of attacker selecting strategy A_i through repeated games. The LQRD differential equation $\frac{dq_j}{dt}$ is

$$\frac{dp_i}{dt} = \frac{\exp(\lambda \sum_{j=1}^{m} q_j a_{ij})}{\sum_{k=1}^{n} \exp(\lambda \sum_{j=1}^{m} q_j a_{ij})} - p_i \quad (8)$$

The realistic significance of the above evolution equation is as below. Taking the defense strategy D_j as an example, if the number proportion of individual selecting the pure strategy D_j is smaller than the payoff proportion of individual obtaining from D_j, then the growth rate of defender number selecting D_j is larger than 0. Otherwise, the growth rate is less than 0. If the number proportion is exactly equal to the payoff proportion, then the growth rate of number of defender selecting strategy D_j is 0. Set $F(p_i) = \frac{dp_i}{dt}$, $G(q_j) = \frac{dq_j}{dt}$, and then combine the above equations to calculate $Y(p_i, q_j) = \begin{bmatrix} F(p_i) \\ G(q_j) \end{bmatrix} = 0$, we can derive the stable equilibrium of attack-defense adversary.

B. The Algorithm for Optimal Defense Strategy Selection

Based on the attack-defense evolutionary game model proposed in section III, the algorithm for selecting the optimal defense strategy is as follows. The decision making process of our approach is shown in Fig. 7, which consists of five steps. It receives two types of input data: attack evidence and abnormal evidence. All these pieces of evidence come from real-time IDS, VDS and log information. Through decision making, the optimal security strategy with maximized earnings is calculated to against the possible attacks.

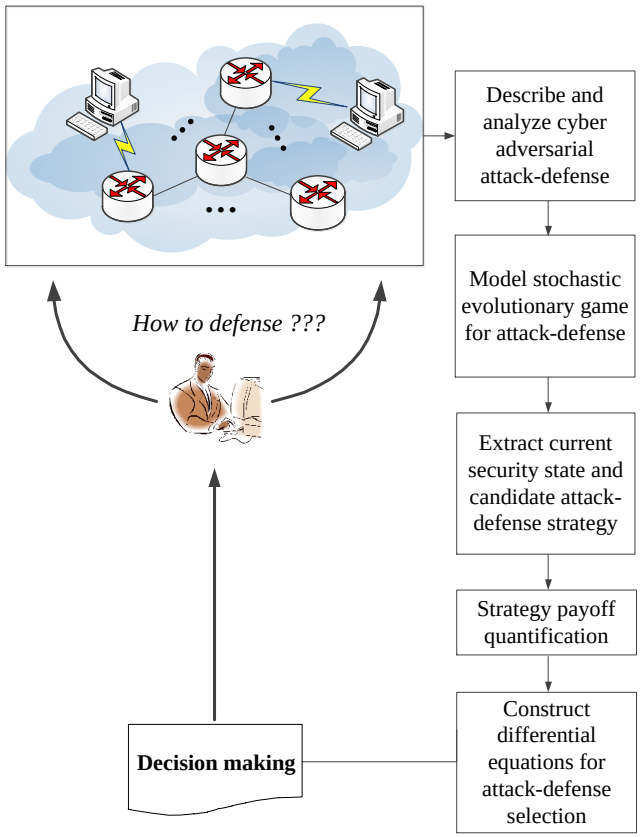


Fig. 7. The process of our decision making approach.

In the following Algorithm 1, the Step 1) corresponds to first three steps of Fig. 7, where we initialize the relevant parameters of the ASEGM model. The Step 2) corresponds to the fourth step of Fig. 7, where we calculate the payoff matrices of candidate strategies. Step 3)-Step 5) correspond to the fifth step of Fig. 7, we assign value to the rationality parameters of the attack-defense players in combination with their historical selections in Step 3). The Step 4) and Step 5) is implemented according to section IV-A, where the equations are constructed for describing strategy selection. Step 6)-Step 7) correspond to the decision making step in Fig. 7. In Step 6), we calculate the evolutionary equilibrium. Step 7) output the optimal strategy set.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TNSM.2020.2995713, IEEE Transactions on Network and Service Management

Algorithm 1 Optimal Defense Strategy Selection Algorithm

Input Network environment information NetInf, Security device configuration information SafetyInf and intrusion alert information AlertInf

Output Optimal defense strategy Q

BEGIN

1) **Initialize** $\mathrm{AEGM}=(N,S,\theta,U)$

/* Initialize attack-defense evolutionary game model */

{

1-1) **Construct** $S_D=\{D_j\}$, $1\le j\le m$

/*Analyze security device configuration information SafetyInf, use [35] to extract the candidate defense strategy set S_D */

1-2) **Construct** $S_A=\{A_i\}$, $1\le i\le n$

/*Collect real-time alert data AlertInf, use [35] to analyze the features of attack behaviors and extract the candidate attack strategy set S_A */

1-3) **Construct** $P=\{p_i\}$, $0\le p_i\le 1$, $\sum_{i=1}^{n}p_i=1$

/* Initialize the attack belief set P in which the attacker selects the attack strategy A_i with the probability $p_i\in P$ */

1-4) **Construct** $Q=\{q_j\}$, $0\le q_j\le 1$, $\sum_{j=1}^{m}q_j=1$

/* Initialize the defense belief set Q in which the defender selects the defense strategy D_j with the probability $q_j\in Q$ */

}

2) **For** $(i=1;\ i\le n;\ i++)$

For $(j=1;\ j\le m;\ j++)$

{

Calculate $\begin{cases}a_{ij}=AB(A_i,D_j)-AC(A_i,D_j)\\ d_{ij}=DB(A_i,D_j)-DC(A_i,D_j)\end{cases}$

/* Calculate the attack-defense payoffs of different strategy combinations A_i and D_j in turn */

}

3) **Assign** λ_1,λ_2, $0\le\lambda_1,\lambda_2$

/* Set the parameter value of the rational degree of λ_1 and λ_2 according to the historical selection information*/

4) **For** $(i=1;\ i\le n;\ i++)$

{

Construct $F(p_i)=\dfrac{\exp(\lambda_1\sum_{j=1}^{m}q_ja_{ij})}{\sum_{k=1}^{n}\exp(\lambda_1\sum_{j=1}^{m}q_ja_{ij})}-p_i$

}

/* Construct the LQRD equation for selecting attack strategy A_i */

5) **For** $(j=1;\ j\le m;\ j++)$

{

Construct $G(q_j)=\dfrac{\exp(\lambda_2\sum_{i=1}^{n}p_id_{ij})}{\sum_{k=1}^{m}\exp(\lambda_2\sum_{i=1}^{n}p_id_{ik})}-q_j$

}

/* Construct the LQRD equation for selecting defense strategy D_j */

6) **Calculate** $Y=\begin{bmatrix}F(p_i)\\ G(q_j)\end{bmatrix}=0$

/* Calculate the evolutionary stable equilibrium point */

7) **Output** $Q=\{q_1,q_2,\cdots,q_m\}$

/* Output the optimal strategy */

END

The time cost of Algorithm 1 mainly generates from Step 2) and Step 6). The computational complexity of Step 2) is $O(mn)$. The complexity of equations calculation of Step 6) is $O\left((m+n)^3\right)$. Therefore, the order of the time complexity of Algorithm 1 is $O(n^3)$. The storage cost is generated by $n\times m$ payoff matrices in step 2) and the transition vectors of equilibrium calculation in step 6), and the highest storage complexity is $O(nm)$. Hence, the order of storage complexity is $O(n^2)$.

V. Experiments and Analyses

Ransomware, a class of self-propagating malware that uses encryption to hold the victims' data ransom, has emerged in recent years as one of the most dangerous cyber threats, with widespread damage: e.g., zero-day WannaCry has caused world-wide catastrophe, from knocking U.K. National Health Service hospitals offline to shutting down a Honda Motor Company in Japan [36]. In a few days, this ransomware virus targeting Microsoft windows systems infected more than 230,000 computers in 150 countries. Once activated, the virus demanded ransom payments in order to unlock the infected system. In this section, we take the invasion and proactive defense against ransomware in the real-world hospital network system as an example. We analyze the adversarial attack-defense process against ransomware, verify the proposed approach for optimal defense strategy selection. The results of the two scenarios with different strategy payoffs are compared and analyzed. In addition, we summarize the general evolution rules of the best defense strategy in the targeted network system. Finally, we compare our methods with the existing research comprehensively.

A. Experiment Setup

Hospital information system is special, such as medical records, data and patient information in it, which is urgent, private and sensitive. After encrypted by ransomware, the patients will try their best to decrypt data as quickly as possible. Otherwise, it will directly affect the normal medical treatment, and even affect the life safety of patients. The architecture of typical digital hospital network system is shown in Fig. 8.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TNSM.2020.2995713, IEEE Transactions on Network and Service Management

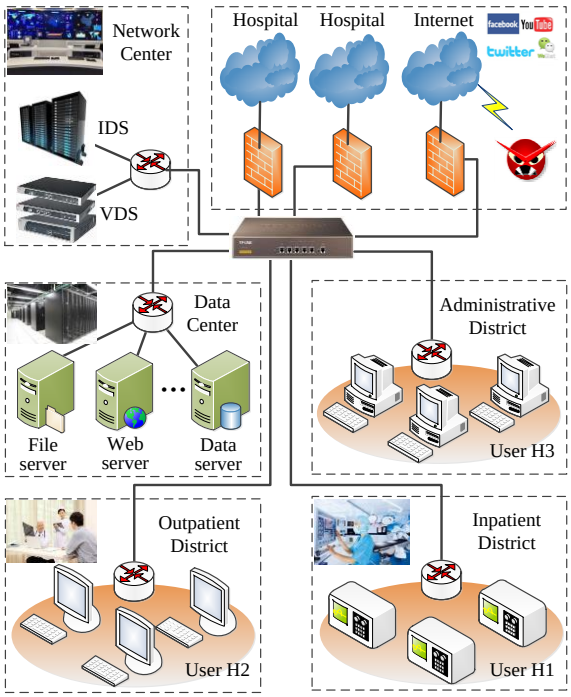


Fig. 8. Architecture of digital hospital network system.

As a defender, the administrator of network center is responsible for the security of the whole intranet of the hospital. The attacker comes from the external network and attacks the intranet through the Internet. The purpose is to encrypt the documents of the data center to obtain the payment of ransom. Ransomware attacks can be divided into two steps, the first is to break through the boundary, and the second is to penetrate the intranet horizontally. Due to the firewall rules, external attackers can initially only communicate with the inpatient, outpatient, administrative district, but cannot access the data center. The security protection devices are composed of the firewall, intrusion prevention system (IPS), virus detection system (VDS), and patch management system. We used the Nessus [34] scanning tool to scan the target network and obtain the results of main vulnerabilities as shown in Table I.

TABLE I
NETWORK CONFIGURATION AND VULNERABILITY INFORMATION

Name	Configuration	CVE#	Vulnerability description
Web server	Windows server 2016	CVE 2017-8464	Allowing local users or remote attackers to execute arbitrary code via a crafted
Data server	Linux	CVE 2017-7494	Allowing a malicious client to upload a shared library to a writable share, and then cause the server to load and execute it.
User H1	Windows 7	CVE 2017-0143	Windows SMB remote code execution vulnerability
User H2	Windows 7	CVE 2017-0144	Allowing remote attackers to execute arbitrary code via crafted packets
User H3	Windows 7	CVE 2017-0199	Microsoft office/WordPad remote code execution vulnerability
		CVE 2017-0145	Allowing remote attackers to execute arbitrary code via crafted packets

B. *Candidate Strategy Extraction and Payoff Calculation*

In this experiment, based on the network topology and vulnerabilities, the logical AG is created using the open source tool MulVAL as illustrated in Fig. 9. The MulVAL is a reasoning toolkit for automatically identifying vulnerabilities in enterprise networks [37]. The ellipses represent the network state node $s_i =< host, privilege, x >$, where x = 0 or 1 respectively indicates that the node file is not encrypted or has been encrypted, and the edge represents the atomic attack action. By referring to attack-defense behavior database of MIT Lincoln Laboratory [35], we extracted the atomic attack and defense actions that can be launched in the network system. All the possible atomic actions are shown in Table II.

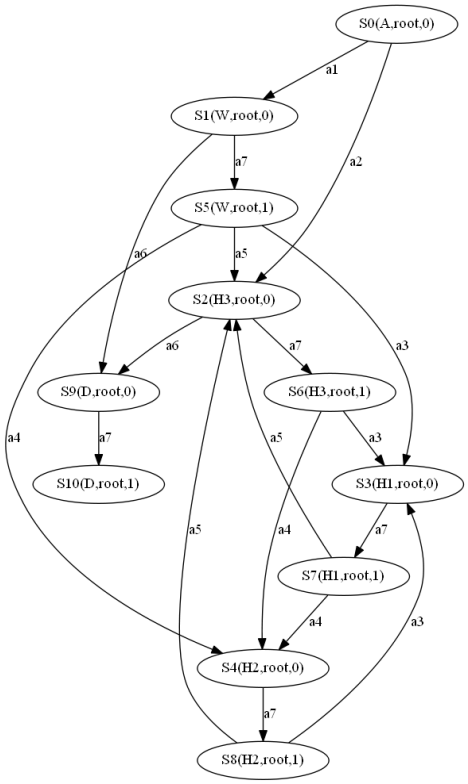


Fig. 9. The attack graph of the experiment.

TABLE II
DESCRIPTION OF CYBER ATTACK AND DEFENSE ACTIONS

No.	Atomic attack action	No.	Atomic defense action
a_1	Scan port	D_1	Close unused port
a_2	Obtain root privilege	D_2	Restart device
a_3	Buffer overflow	D_3	Offline network
a_4	Execute arbitrary code	D_4	Install patches
a_5	Infect office macro virus	D_5	Forbid office macro scripts
a_6	SQL injection	D_6	Modify account password
a_7	Encrypt database	D_7	Backup data

According to Step 1-1) and Step 1-2) of Algorithm 1, we initialize the parameters of the game model. By collecting the alert information generated by firewall, IPS, VDS and host security audit log, the attack behavior information is analyzed. We find that the attacker first conducted port scanning action a_1 through port 135 of Web server. Furthermore, the attacker collected open service information to prepare for subsequent

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TNSM.2020.2995713, IEEE Transactions on Network and Service Management

attacks. Since port scanning is a concealed mean of attacking, which is the passive attack essentially. We denote it as A_1 = Scan port.

Based on further detections and analyses of alert information, we find that some adventurous attackers execute atomic attack a_7 shown in Table II. The unauthenticated attackers exploit the vulnerability CVE 2017-0199 of oracle server H3 to implement arbitrary code. In detail, the attacker inserted the SQL command into the query string of requested page, and then cheated the servers to execute malicious SQL commands, leaked out user privacy information, stole the cookie data and finally encrypted the patient medical file in the Intranet data center. We denote this candidate strategy as A_2 = Encrypt database , which is an active attack actually.

In this experiment, to simplify the analysis, we only consider two candidate defense strategies D_1 = Close unused ports and D_2 = Backup data in Table II. The security administrator performed atomic defense action d_1 by adjusting firewall configuration to close seldom or unused service ports. To mitigate the security risks of port scanning attack, once a SQL injection attack is detected, the atomic defense action d_8 is launched to pre-backup the important data to against the possible data encryption attack.

From definition 6, the payoff matrix of attack-defense is as follows.

$$M = \begin{bmatrix} a_{11}, d_{11} & a_{12}, d_{12} \\ a_{21}, d_{21} & a_{22}, d_{22} \end{bmatrix} \tag{9}$$

C. Constructions of Evolution Equations for Decision Making

Firstly, we set the attackers and defenders with equal degrees of rationality. Furthermore, we assign the proportion of number of players selecting A_1 in the attacker population as p and that of selecting strategy A_2 as $1-p$, $0 \le p \le 1$. Secondly, according to Step 1-4), we assign the proportion of defender population selecting strategy D_1 as q and assign the proportion of defender population selecting D_2 as $1-q$, $0 \le q \le 1$.

In addition, we construct the LQRD equation of attack-defense strategy as follows respectively.

1) Evolution equation of attack strategy

The expected payoff of attacker selecting strategy " A_1 = port scan attack" is $U_{A_1} = a_{11}q + a_{12}(1-q)$, the expected payoff of "Encrypt database" is $U_{A_2} = a_{21}q + a_{22}(1-q)$. Then, according to the Step 3) of Algorithm 1, we can obtain the evolution equation of strategy A_1 as follows.

$$\frac{dp}{dt} = \frac{\exp\left(\lambda(a_{11}q + a_{12}(1-q))\right)}{\exp\left(\lambda(a_{11}q + a_{12}(1-q))\right) + \exp\left(\lambda(a_{21}q + a_{22}(1-q))\right)} - p$$

2) Evolution equation of defense strategy

The expected payoff of defender selecting strategy "close unused ports" is $U_{D_1} = d_{11}p + d_{21}(1-p)$, the expected payoff of selecting " D_2 = Encrypt database" is $U_{D_2} = d_{12}p + d_{22}(1-p)$. Then, from the Step 4) of Algorithm 1, we can derive that the evolution equation of strategy as below.

$$\frac{dq}{dt} = \frac{\exp\left(\lambda(d_{11}p + d_{21}(1-p))\right)}{\exp\left(\lambda(d_{11}p + d_{21}(1-p))\right) + \exp\left(\lambda(d_{12}p + d_{22}(1-p))\right)} - q$$

Then, according to Step 6) of Algorithm 1, assign

$$\begin{cases} \dfrac{\exp\left(\lambda(a_{11}q + a_{12}(1-q))\right)}{\exp\left(\lambda(a_{11}q + a_{12}(1-q))\right) + \exp\left(\lambda(a_{21}q + a_{22}(1-q))\right)} - p = 0 \\ \dfrac{\exp\left(\lambda(d_{11}p + d_{21}(1-p))\right)}{\exp\left(\lambda(d_{11}p + d_{21}(1-p))\right) + \exp\left(\lambda(d_{12}p + d_{22}(1-p))\right)} - q = 0 \end{cases}$$

The solution of the above equation is the evolutionary stable equilibrium of attack-defense decision-making, and defender's optimal defense strategy is selecting strategy $\{D_1, D_2\}$ with mixed probability $\{q, 1-q\}$.

D. Case Studies and Analyses

We take two numerical experiments: case study 1 (without considering counterattack payoff) and case study 2 (considering counterattack payoff). The comprehensive comparison and analysis is given finally.

1) Case study 1

According to Step 2) of Algorithm 1, we combine the Definition 2 - Definition 5 and security behaviors database [35], and then obtain the game payoff of attack-defense as organized in Table III. Because the attack-defense competition is the resource-consuming activity, the payoff value is less than 0.

TABLE III
GAME PAYOFF OF CASE STUDY 1

Candidate attack strategy	Candidate defense strategy	
	D_1=Close unused ports	D_2=Backup data
A_1=Scan port	(-40, -60)	(-10, -5)
A_2=Encrypt database	(-20, -10)	(-15, -30)

From Section IV-A, we can derive the dynamic equations of LQRD optimal response for the evolution of attack-defense strategies as follows.

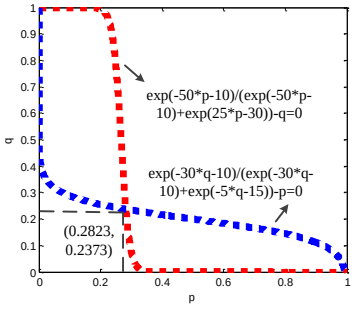


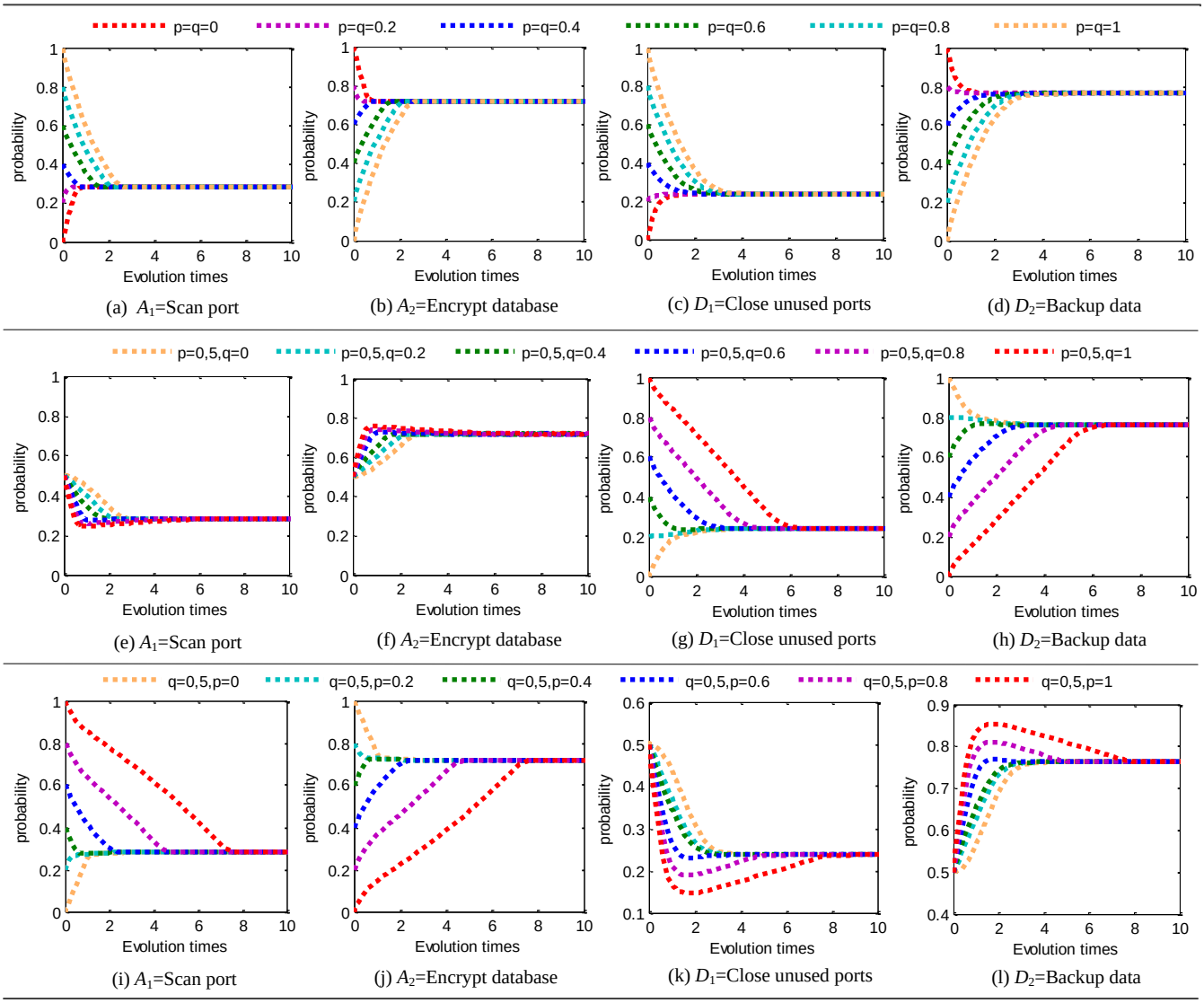
Fig. 10. Equilibrium point of case study 1 when $\lambda = 1$, $q = 0.5$, $p = 0.5$.

$$\begin{cases} \dfrac{\exp\left(\lambda(-30q - 10)\right)}{\exp\left(\lambda(-30q - 10)\right) + \exp\left(\lambda(-5q - 15)\right)} - p = 0 \\ \dfrac{\exp\left(\lambda(-50p - 10)\right)}{\exp\left(\lambda(-50p - 10)\right) + \exp\left(\lambda(25p - 30)\right)} - q = 0 \end{cases}$$

In general, the degree of player rationality in the real world is medium and we set $\lambda = 1$, set the initial state of the game system as $p = 0.5$, $q = 0.5$ according to Step 5 of Algorithm 1. That is, the attacker randomly selects a strategy from candidate A_1 and A_2 with equal probability 0.5 at initial time. Similarly, the defender randomly selects a strategy from candidate D_1 and D_2 with equal probability. With the simulation tool Matlab2017,

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TNSM.2020.2995713, IEEE Transactions on Network and Service Management

the equilibrium point can be calculated in Fig. 10, in which the color of evolution curve of attack strategy A_1 is blue and that of defense strategy D_1 is red. The cross point $p = 0.2823,\ q = 0.2373$ is the stable equilibrium solution. From Fig. 10, the attacker is more probable to select $\{A_1, A_2\}$ with mixed probability $\{0.2823, 0.7177\}$. Meanwhile, the optimal defense strategy for the defender is to randomly implement $\{D_1, D_2\}$ with mixed probability $\{0.2373, 0.7627\}$. The above results show that the attacker is more likely to select the aggressive strategy A_2 = Encrypt database with probability 0.7177. Since the attack of Encrypt database is more harmful, in order to avoid the serious attack influence, the corresponding optimal defense strategy is to select D_2=backup data with the probability 0.7627.



(a) A_1=Scan port (b) A_2=Encrypt database (c) D_1=Close unused ports (d) D_2=Backup data

(e) A_1=Scan port (f) A_2=Encrypt database (g) D_1=Close unused ports (h) D_2=Backup data

(i) A_1=Scan port (j) A_2=Encrypt database (k) D_1=Close unused ports (l) D_2=Backup data

Fig. 11. The strategy evolution tracks with different initial selections of p and q.

Secondly, in order to analyze the influence of system initial state on strategy selections, we simulate the evolution tracks of strategy selections with different initial p and q in Fig. 11. The abscissa t represents the number of evolutions in decision-making. The ordinate *probability* represents the probability of selecting strategy. With Fig. 11, we can predict the defender's best strategy selection at different game moments.

Fig. 11(a)-Fig. 11(d) respectively show the evolution tracks of A_1, A_2, D_1, D_2 when the initial states of attacker and defender are the same with $p = q = 0,\ 0,2,\ 0,4,\ 0.6,\ 0.8,\ 1$. From Fig. 11(c-d), we assume that the attacker and defender initially select the strategy A_1 and D_1 with probability $p = q = 1$ when t = 0. Then from the yellow curve of Fig, 8(c), the probability of selecting strategy D_1 is falling over time and stabilize to *probability*=0.2373 when $t = 3.2690$. Accordingly, the probability of selecting strategy D_2 is rising and stabilize to *probability*=0.7627 from the yellow curve of Fig. 11(d). Herein, the optimal defense strategy is selecting $\{D_1, D_2\}$ with mixed probability $\{0.2373, 0.7627\}$. This selection is stable and best when against different candidate attack strategies.

We further fix a parameter in $\{q, p\}$ and obtain the evolution tracks as shown in Fig. 11(e)-Fig. 11(l). Among which, Fig. 11(e-h) depict the evolution tracks with the same initial p but

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TNSM.2020.2995713, IEEE Transactions on Network and Service Management

different initial q. Taking Fig. 11(g) and Fig. 11(h) as examples, we assume that the defender selects the strategy D_1 with a higher probability q=1 initially, namely, the larger the gap between the defender's initial selection and the optimal selection q=0.2373, the more evolution times needed to achieve the best strategy, as can be observed that the inflection point $t = 6.3338$ of the red curve appears later than that $t = 3.0376$ of the yellow curve. In contrast to Nash equilibrium game model [16], our approach can better explain the strategy evolution rules in adversarial attack-defense and have stronger performance of attack prediction.

Fig. 11(i-l) depict the evolution tracks with the same initial q while different initial p. The greater the probability of selecting strategy A_1 at the initial time, the later the curve inflection point appears. Indicating that more number of repeated games is required for decision-making and longer time takes. This is due to that the attacker selects A_1 with a high probability at the initial time. The false signal deceived the defender. It caused the defender mistakenly assume that the attacker will select the moderate attack strategy A_1 while overlook the real attack purpose encrypt the database. Therefore, rational defenders need to implement many evolution times so as to discover the attacker's real purpose and obtain the best defense strategy. For example, when q=0.5, p=1, the probability of selecting strategy D_1 denoted by the red curve in Fig. 11(k) first reduces to $p = 0.1462$ at $t = 1.777$ and then rebounds and finally stabilizes to $p = 0.2373$ at $t = 9.356$. The reason is that the proportion of defender population selecting strategy D_1 at the initial time decreases too much. With the decrease of the D_1 payoff to the total payoff, the number of individuals selecting D_1 rises gradually to ensure that the proportion of population selecting D_1 to the total populations is equal to the proportion of payoff selecting D_1 to the total payoffs.

As can be seen from each column in Fig. 11, the optimal strategy for both defender and attacker are the same regardless of their initial p and q selections. It is only related to the candidate strategy set and the strategy payoff. Moreover, the initial state can only affect the stabilization time of the game system.

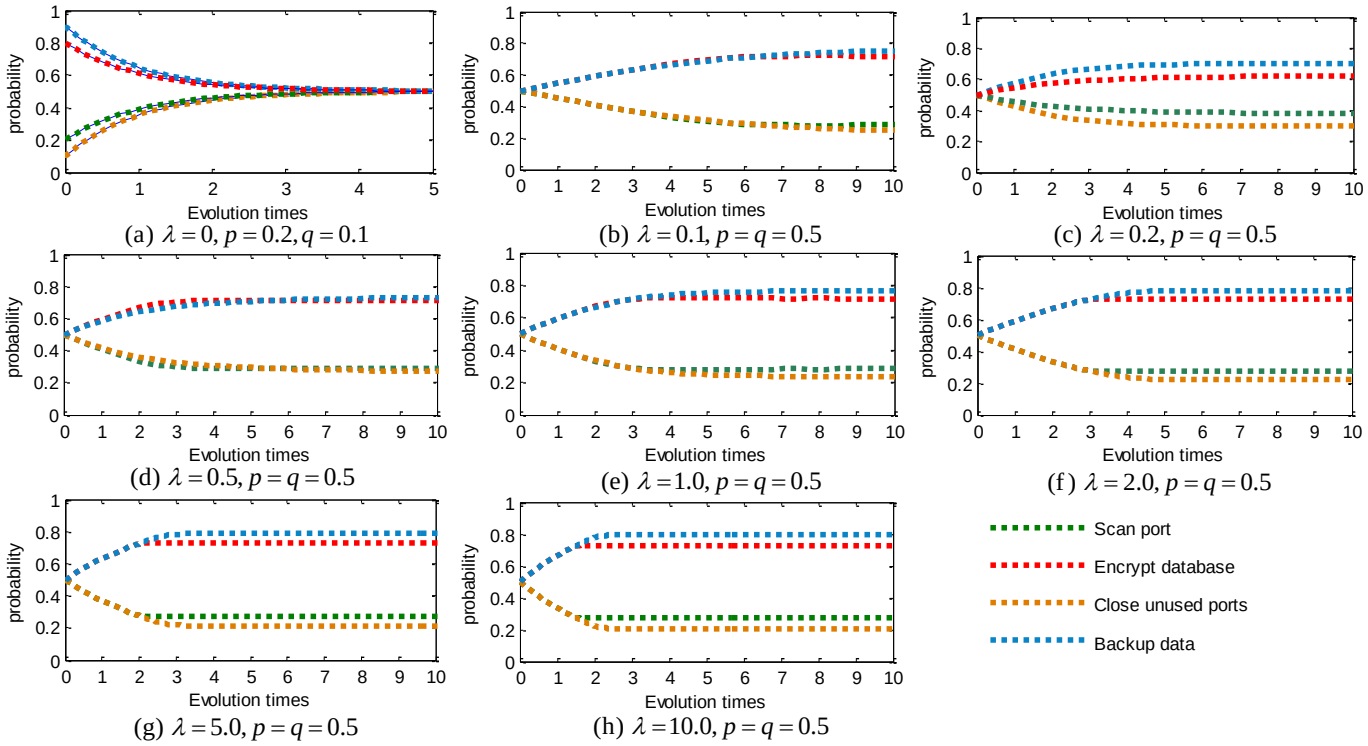


Fig. 12. The strategy evolution tracks with different rationality λ.

Finally, to analyze the influence of degrees of players' rationality on strategy evolution, some simulations shown in Fig. 12 and discussions are as follows.

1) According to Step 5) of algorithm 1, we assume that the players are irrational and set $\lambda = 0$, assign initial $p = 0, 2, q = 0.1$, then obtain the strategy evolution tracks in Fig. 12(a). Herein, the final result is to select the different candidate strategy with the same probability 0.5. It means that players cannot distinguish the advantages and disadvantages of different candidate strategies since they have no cognitive abilities. Meanwhile, from the LQRD equation of attack-defense in section V-C, there is only one solution $p = 0.5, q = 0.5$ when $\lambda = 0$. Above results both show that when the game players are irrational, regardless of their initial selections, they cannot distinguish the merits and demerits of each strategy since they do not have any learning and cognitive capabilities. The candidate strategies are still selected by game players randomly.

2) Suppose that the player rational degree $\lambda > 0$, we simulate the strategy evolutions in Fig. 12(b-h). With time going by, all

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TNSM.2020.2995713, IEEE Transactions on Network and Service Management

the players can finally obtain the correct strategy through several times of repeated games. The main difference is that when the players have a high degree of rationality, they can find the optimal strategy more quickly. For example, when $\lambda = 1$, the game system can reach the stable state through about 7 times of game evolutions (shown in Fig 9(e)), while when $\lambda = 10$, they can be stable only through 2 times of game evolutions (shown in Fig 9(h)). Above results demonstrate that when the defenders have a high degree of rationality (have rich knowledge, skilled techniques, etc.), their cognition, learning and adjustment abilities are strong, which help the defenders identify the optimal strategy more quickly.

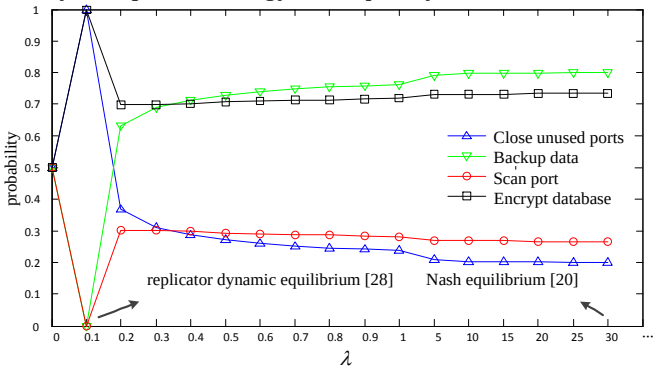


Fig. 13. The impact of different rationality λ on the strategy selections when $p = q = 0.5$.

In general, both sides of attacker and defender gain increased decision-making experience through adversarial attack-defense. Hence, the rational degree λ increases during the game process. To reveal this scenario, Fig. 13 illustrate the results under different λ, where the abscissa λ represents the rational degree and the ordinate represents the probability of strategy selection. When $\lambda = 0$, players have no rationality, so they choose candidate strategies randomly. When $\lambda = 0.1$, the rational degree of the players is very low as the replicator dynamics [16]. From Fig. 13, the probability of defender selecting strategy D_1 and D_2 rapidly increases to 1 and decreases to 0 respectively, which reflects the sensitivity of the decision-making system. The corresponding equilibrium solution is $p = 0, q = 1$. The result corresponds to the replicator dynamic equilibrium [28]. Since the rational degree of replicator dynamic game is very low, its equilibrium solution is pure strategy. When $\lambda > 0.1$, the player rational degree increases, and both sides of attacker and defender constantly approach to complete rational Nash equilibrium as λ increases. When $\lambda = 30$, the solution $p = 0.2665, q = 0.2013$ of LQRD in this paper is very close to the Nash equilibrium solution $p = 0.2667, q = 0.2000$ [15]. It indicates that the player rationality is very close to complete rationality over time and the difference with the Nash equilibrium decreases gradually through obtaining experience in the game process. It is foreseeable that when $\lambda \to \infty$, the proposed LQRD equilibrium will approach to Nash equilibrium. Compared with the complete rational Nash equilibrium [15] and the bounded rational replicator dynamic equilibrium [28], our approach can depict the diversity of rationality of attacker and defender players and reflect the real strategy selection rules.

2) Case study 2

Based on case study 1, this section further explores the impact of payoff changes (consider / do not consider counterattack) on the selection of attack-defense strategies. Compared with case study 1 only considering the directly security rewards, in case study 2, we further consider the indirect rewards through counterattacks and legal penalty, economic and time rewards (business recovery time, data recovery time, etc.). Table IV gives the strategy payoff of case study 2, which depicts the difference from Table III.

TABLE IV
GAME PAYOFFS OF CASE STUDY 2

Candidate attack strategy	Candidate defense strategy	
	D_1=Close unused ports	D_2=Encrypt database
A_1=Scan port	(-50, -55)	(-20, 0)
A_2=Encrypt database	(-30, -5)	(-25, -25)

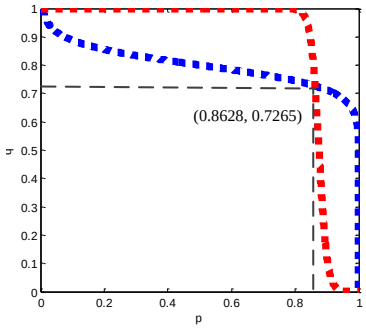


Fig. 14. Equilibrium point of case study 2 when $\lambda = 1$, $q = 0.5$, $p = 0.5$.

Similarly, we first consider the medium rationality $\lambda = 1$, and set the initial selection as $p = 0.5$, $q = 0.5$. The calculation process is the same with that of case study 1. We obtain the evolution equilibrium point $p = 0.8628, q = 0.7265$ in Fig. 14. Comparing the equilibrium points of case study 1 (see in Fig. 10) with that of case study 2 (see in Fig. 14), we can derive that:

1) For the attacker in case study 2, the probability of adopting strategy A_1 increased from 0.2823 to 0.8628 and that of strategy A_2 reduced from 0.7177 to 0.1372 comparing with case study 1, indicating that the attacker is deterred effectively once the defender made counterattack. Meanwhile, more attackers are likely to change their strategy from selecting Encrypt database to Scan port. Because Encrypt database is a more serious and aggressive attack, while Scan port is a relative covert and moderate attack. Taking the moderate strategy can effectively reduce the risk of countermeasures for the attacker.

2) For the defender of case study 2, the probability of selecting strategy close unused ports improves from 0.2373 to 0.7265, and that of strategy backup database reduces from 0.7627 to 0.2735. Since we calculate the indirect revenue of counterattack, the attacker was deterred to adopt a more moderate action Scan port instead of Encrypt database. Accordingly, the defender changed to adopt the simpler defense action Close unused ports instead of Backup database, which not only can reduce the defense cost but also can mitigate the risk of the total network system.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TNSM.2020.2995713, IEEE Transactions on Network and Service Management

Secondly, we set different initial p and q and simulate the evolution tracks of both attacker and defender in Fig. 15. The abscissa denotes the number of repeated games and the ordinate denotes the probability of strategy selection. Similarly, we can derive that the stable equilibrium of the game is not determined by game system's initial state but is impacted by the candidate strategy set and payoff value. Different initial states can only affect the stabilization time, namely, the moment of inflection point of curve appears, but cannot determine the final trend of the game. Moreover, the rational players can always find the optimal strategy through strategy learning and improving in the repeated game process. Moreover, the proposed optimal defense strategy has stronger foreseeability and robustness when facing different candidate attack strategies.

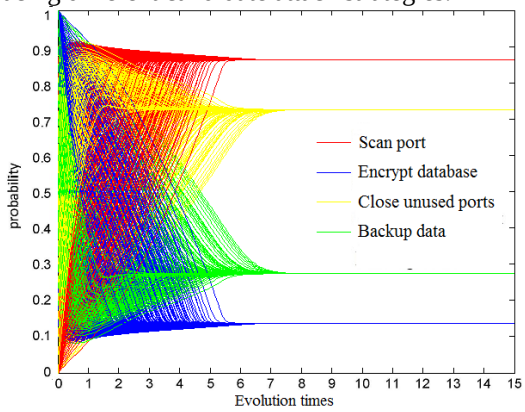


Fig. 15. Evolution tracks of attack-defense strategies in case study 2.

To sum up, through the above two case studies, we can conclude that:

1) **We established an evolutionary game model based on the bounded rationality of both sides of attackers and defenders.** We considered that the process of adversarial attack-defense reach a stable state progressively through repeated games. Fig. 11 and Fig. 15 depict the tracks of the strategy evolution over time, which can be used to predict the best strategy selections at different game moments. It presents the formation of the best defense strategy, and enhances the performance of attack early warning and the corresponding countermeasures for cyberdefense. Our approach helps the security manager win the time warfare of cyber attack-defense effectively.

2) **We simulate the defense evolution process against WannaCry attack.** Since the radical attackers often select the strategy Encrypt database. Due to that these medical data are private, sensitive and significant, which is related to the treatment and health of patients, so most hospitals usually choose to pay the ransom once infected. Through our analyses on case study 1, if we cannot strengthen the defense and counterattack, the best strategy for the defender is to take the action of Backup data in advance. Although the cost of this strategy is high, from the game process of the whole attack-defense, it can ensure the defender maximize the defense revenue.

3) **We quantify the rationality of different attackers and defenders by introducing a flexible parameter λ.** Through which, we can depict the diversities of evolution behaviors. Fig. 13 shows the generality and interpretability of the proposed approach in explaining the attack-defense adversarial process. Our model is more flexible for decision-making of social players with different rationalities. In fact, different attackers and defenders have different abilities of cognizance, which is influenced by their own interests such as safety knowledge, skill level, experience, external environment and self emotion, etc. Therefore, the strategy selections affected by various uncertain factors lead to the bounded rational with many rounds repeated game. At present, the research on attack-defense game with player's bounded rationality is still in its infancy as the most used fixed rationality game model. For this aim, we design the rationality factors to quantify the rational degrees of different game players. Fig.12 and Fig.13 illustrated the strategy evolution tracks with different rationality and its impact on strategy selections. The flexibility and interpretability of the proposed approach is explained in Fig. 13. Ours is more generalized as it not only can be converted to the low rational replicator dynamic equilibrium [24, 26, 27, 28], but also can be converted to the complete rational Nash equilibrium [15, 16] (Two models that have been widely used recently). Moreover, we explain the dynamic approach process from low rational replicator dynamic equilibrium to complete rational Nash equilibrium.

4) **We find that evolution stable strategy is only related to the candidate strategy and payoff value but not to the initial selections of attacker and defender.** In case study 2, by adjusting the payoff of candidate strategy, we can change the results of strategy selection so as to turnaround the defense situation. For example, by increasing the payoff reward of attack penalty for defender and strengths on defense and counterattack, it is beneficial to guide attackers and defenders to adopt more moderate candidate strategies, avoid escalation of conflicts, and promote cyber security governance.

E. Comparisons and Discussions

The comparisons among ours and others are organized in Table V, some discussions are as follows:

TABLE V
PERFORMANCES COMPARISON AMONG OURS AND OTHERS

Ref.	Rationality	Game type	Game structure	Strategy type	Equilibrium solution	Generality
[12]	Complete	Static	*n*	Mixed	Nash equilibrium	Medium
[15]	Complete	Static	*n*	Mixed	Nash equilibrium	Medium
[16]	Complete	Static	*n*	Mixed	Nash equilibrium	Medium
[20]	Complete	Dynamic	*n*	Pure	Bayesian equilibrium	Medium
[24]	Bounded	Dynamic	2	Mixed	Replicator dynamics evolutionary equilibrium	Medium
[26]	Bounded	Dynamic	*n*	Mixed	Replicator dynamics evolutionary equilibrium	High
[27]	Bounded	Dynamic	*n*	Mixed	Replicator dynamics evolutionary equilibrium	High
[28]	Bounded	Dynamic	*n*	Pure	Replicator dynamics evolutionary equilibrium	Medium
Ours	Bounded	Dynamic	*n*	Mixed	LQRD equilibrium	High

1) **Player rationality**. [15][16][12][20] assume that attackers and defenders are totally rational. For instance, Nash equilibrium [15][16] requires that all attackers and defenders can predict adversary's optimal strategy correctly at the same time. However, the cognitive capabilities of different players are quite different, so the hypothesis of complete rationality is deviated from the reality. In contrast, [24][26][27][28] and ours regard that the players are bounded rationality. We analyze the strategy learning and improving mechanism of game players. Therefore, ours significantly improve the scientificity of modeling of cyber attack-defense.

2) **Game type and structure**. [15][16] combine the payoff matrix and the Nash equilibrium to calculate the optimal strategies of both sides of attacker and defender, which are static games essentially. [20] considers that the order of action of attack-defense has the priority, and the latter actor can modify his belief according to the strategy of the former actor. Based on this, it divides the game process into several stages according to the order of strategy implement. [24][26][27][28] and ours are designed based on evolutionary game theory, the best strategy is obtained through simulating attack-defense process, which depict the dynamics of strategy evolution and present the strategy selections at different game moments. For game structures, only two simple candidate strategies are abstracted in [24], namely, attack and no attack, defense and no defense. On the contrary, this paper considers the extended game structure including n kinds of candidate attack strategies and m kinds of candidate defense strategies.

3) **Strategy type and equilibrium calculation**. As the most used bounded rationality game model, the replicator dynamics equilibrium is limited to pure strategy [28], which is a special case of mixed strategy actually. The others and ours consider the more general mixed strategy. In [15], the optimal strategy against node replication attacks in sensor networks is given by calculating Nash equilibrium. Nash equilibrium [12, 15, 16] is explained as the optimal reaction between both two sides of attacker and defender, but no forming process of Nash equilibrium is given. The emphasis of this paper is to analyze the evolutionary process of attack-defense strategies. We revealed the dynamic process and hidden law from low rationality to complete rationality. [16] makes use of the Pareto algorithm to optimize the calculation process so as to reduce the complexity. [24][26][27][28] describe the strategy updating rules using the replicator dynamic mechanism of biological evolution, which are still limited to pure strategy selection. Moreover, the depicted player rationality is very low, which deviates from the characteristics of fast learning and improvement of cyber warfare. They are essentially deterministic evolutionary behaviors without considering the stochastic disturbance in the real network environment. In this paper, we use LQRD equations to describe diverse evolutionary behaviors. Meanwhile, we use the flexible parameter λ to quantify the rational degree to reflect the randomness and inertia of population social behaviors of realistic attackers and defenders. With the improvement of rationality, we present the formation of best strategy and simulate the approach process and from bounded rational replicator dynamic equilibrium to complete rational Nash equilibrium.

CONCLUSION

Along with the complexity and large-scale of network information systems, security attacks become more diverse, leading the cyber attack-defense situation change dynamically. How to comprehensively analyze defense costs and benefits, maximize defense revenue, predict the possible attack strategy, select the optimal defense strategy from the candidate strategies and measure the strategy revenue is still assumed as a big challenge. Game theory is an effective tool to model the adversarial cyber attack-defense. At present, the issue on game modeling of attack-defense with bounded rationality is still in its infancy, and there are many limitations such as player rationality quantification, game structure, strategy type, and equilibrium calculation. To a certain extent, it affects the scientificity and effectiveness of application of game theory for cybersecurity. For this purpose, we construct a novel evolutionary game model to describe attack-defense using LQRD and expand the strategy set and type of existing game structure. We construct the differential equations of strategy evolution varying with time for attackers and defenders with customized rational degrees. The strategy evolution tracks are simulated in the real-world attack scenario of WannaCry to depict the formation of best strategy. By analyzing the evolutionary stable equilibrium, we can obtain the optimal defense strategy at different game moments. Our approach is more generalized comparing with replicator dynamics and Nash equilibrium model. Two case studies on WannaCry both show that the proposed approach is effective and practical. The performances of attack prediction and defense decision-making are improved significantly for winning the time warfare of cyber attack-defense.

ACKNOWLEDGMENTS

This work was supported by the National Key Research and Development Program of China (2019QY1301, 2019QY1303, 2018YFB0803602), the National Natural Science Foundation of China (61902427, 61702508, 61802404). We thank the anonymous reviewers for their valuable comments.

REFERENCES

[1] H. Duan, "Forwarding-Loop attacks in content delivery Networks," *in Proc. Network & Distributed System Security Symposium*, 2016.

[2] W. Douglas and S. Richard, "How to measure anything in cybersecurity risk?" *John Wiley & Sons*, 2016, pp. 213-227.

[3] C. Chen, J. Hu, T. Qiu, M. Atiquzzaman, Zhiyuan Ren, "CVCG: Cooperative V2V-aided transmission scheme based on coalitional game for popular content distribution in vehicular ad-hoc networks," *IEEE Transactions on Mobile Computing*, vol. 18, no. 12, pp. 2811-2828, 2019.

[4] C. Chen, L. Liu, T. Qiu, K. Yang, F. K. Gong, H. B. Song, "ASGR: an artificial spider-web-based geographic routing in heterogeneous vehicular networks," *IEEE Transactions on Intelligent Transportation Systems*, vol. 20, no. 5, pp. 1604-1620, 2019.

[5] C. Chen, L. Liu, T. Qiu, D. O. Wu, Z. Ren, "Delay-aware grid-based geographic routing in urban VANETs: a backbone approach," *IEEE/ACM Transactions on Networking*, vol. 27, no. 6, pp. 2324-2337, 2019.

[6] A. Gharaibeh, M. A. Salahuddin, S. J. Hussini, A. Khreishah, *et al*., "Smart cities: A survey on data management, security, and enabling technologies," *IEEE Communications Surveys & Tutorials*, vol. 19, no. 4, pp. 2456-2501, 2017.

[7] R. B. Myerson, "Game theory," *Harvard University Press*, 2013.
[8] C. T. Do, N. H. Tran, C. Hong, *et al*., "Game theory for cyber security and privacy," *ACM Computing Surveys*, vol. 50, no. 2, Article. 30, 2017.
[9] L., Kong-wei and J. M. Wing, "Game strategies in network security," *International Journal of Information Security*, vol. 4, no. 1-2, pp. 71-86, 2005.
[10] X. Liang and Y. Xiao, "Game theory for network security," *IEEE Communications Surveys & Tutorials*, vol. 15, no.1, pp. 472-486, 2012.
[11] S. Roy, C. Ellis, S. Shiva, D. Dasgupta, *et al*., "A survey of game theory as applied to network security," *in Proc. the IEEE 43rd Hawaii International Conference on System Sciences*, 2010.
[12] J. Tan, C. Lei, H. Q. Zhang, *et al*., "Optimal strategy selection approach to moving target defense based on Markov robust game," *Computers & Security*, vol. 8, no. 5, pp. 63-76, 2019.
[13] M. D. James, "Game theory for political scientists," *Princeton University Press*, no. 30, 519.83, 1994.
[14] W. Jiang, B. X. Fang, Z. H. Tian, *et al*., "Evaluating network security and optimal active defense based on attack-defense game model," *Journal of Computers*, vol. 32, no. 4, pp. 817-827, 2009.
[15] Y. Li, D. E. Quevedo, S. Dey, *et al*., "A game-theoretic approach to fake-acknowledgment attack on cyber-physical systems," *IEEE Transactions on Signal & Information Processing Over Networks*, vol. 3, no.1, pp.1-11, 2017.
[16] X. Li, C. Zhou, Y. C. Tian, *et al*., "A dynamic decision-making approach for intrusion response in industrial control systems," *IEEE Transactions on Industrial Informatics*, vol. 15, no.5, pp. 2544-2554, 2018.
[17] Y. Liu, D. Feng, Y. Lian, *et al*., "Optimal defense strategies for DDoS defender using bayesian game model," *in Proc. IEEE Conference on Information Security Practice and Experience*, pp. 44-59, 2013.
[18] A. Patcha and J. M. Park, "A game theoretic formulation for intrusion detection in mobile Ad Hoc networks," *Internet Journal of Network Security*, vol. 2, no.2, pp. 131-137, 2006.
[19] T. Carroll, E. Thomas, G. Daniel, "A game theoretic investigation of deception in network security," *Security and Communication Networks* vol. 4, no. 10, pp. 1162-1172, 2011.
[20] C. Lei, H. Q. Zhang, L. M. Wang, et al., "Incomplete information Markov game theoretic approach to strategy generation for moving target defense," *Computer Communications*, vol. 116, no. 184-199, 2018.
[21] X. J. Wang, J. Quan, W. B. Liu, "Study on evolutionary games and cooperation mechanism within the framework of bounded rationality," *Systems Engineering-Theory & Practice*, vol. 31, no. s1, pp. 82-93, 2011.
[22] H. James and A. C. Nicholas, "Cooperative behavior cascades in human social networks," *in Proc. of the National Academy of Sciences*, no. 107.12, pp. 5334-5338, 2010.
[23] D. Tosh, S. Sengupta, C . Kamhoua, *et al*., "An evolutionary game-theoretic framework for cyber-threat information sharing," *in Proc. IEEE International Conference on Communications*, pp. 7341-7346, 2015.
[24] J. M. Zhu, B. Song, Q. Huang, "Evolution game model of offense-defense for network security based on system dynamics," *Journal on Communications*, vol. 35, no. 1, pp. 54-61, 2014.
[25] R. Na, L Gao, H. Zhu, *et al.*, "Toward Optimal DoS-resistant authentication in crowd sensing networks via evolutionary game," *in Proc. IEEE, International Conference on Distributed Computing Systems*, pp. 364-373, 2016.
[26] A. A. Abass, X. Liang, N. Mandayam, *et al.*, "Evolutionary game theoretic analysis of advanced persistent threats against cloud storage," *IEEE Access*, vol. 5, pp. 8482-8491, 2017.
[27] Y. Hayel and Q. Zhu, "Epidemic. Protection over heterogeneous networks using evolutionary poisson games," *IEEE Transactions on Information Forensics and Security*, vol. 12, no. 8, pp. 1786-1800, 2017.
[28] H. Hu, Y. Liu, H. Zhang, *et al.*, "Optimal network defense strategy selection based on incomplete information evolutionary game," *IEEE Access*, vol. 6, pp. 29806-29821, 2018.
[29] Y. Zhang and J. Liu. "Optimal Decision-Making Approach for Cyber Security Defense Using Game Theory and Intelligent Learning." Security and Communication Networks, Article no. 3038586, 2019.
[30] P. D. Taylor and L. B. Jonkerm, "Evolutionarily stable strategies and game dynamics," *Mathematical Biosciences*, vol. 40, no.2, pp. 145-156, 1978.
[31] C. Alosferrer and N. Netzer, "The logit-response dynamics," *Games and Economic Behavior*, vol. 68, no. 2, pp. 413-427, 2010.
[32] T. G. Kurtz, "Solutions of ordinary differential equations as limits of pure jump Markov processes," *Journal of Applied Probability*, vol. 7, no. 1, pp. 49-58, 1970.
[33] S. P. Anderson, J. K. Goeree, C. A. Holt, "The logit equilibrium: a perspective on intuitive behavioral anomalies," *Southern Economic Journal*, vol. 69, no. 1, pp. 21-47, 2002.
[34] N. Tenable, "Nessus vulnerability scanner" https://www.tenable.com/downloads/nessus
[35] L. Gordon, M. Loeb, W. Lucyshyn, *et al*., "CSI/FBI computer crime and security survey," *Computer Security Journal*, vol. 12, no. 3, pp. 11-34, 2005.
[36] Q. Chen and A. B. Robert, "Automated behavioral analysis of malware: A case study of wannacry ransomware," *in Proc. 16th IEEE International Conference on Machine Learning and Applications*, 2017.
[37] X. Ou, S. Govindavajhala, A. W. Appel, "MulVAL: A logic-based network security analyzer," *in Proc. 14th Usenix Security Symposium*, pp. 113-127. 2005.
[38] G. Herbert, "Game theory evolving," *Boston: Priceton University Press*, 2015.
[39] J. Steve, "Learning to love SIEM," *Network Security*, vol. 4, pp. 18-19, 2011.
[40] W C Reco, "OWL web ontology language overview," *February*, vol. 63, no. 45, pp. 990–996, 2004.
[41] H. Hu, Y Liu, H Zhang, *et al*., "Security metric methods for network multistep attacks using AMC and big data correlation analysis," *Security and Communication Networks*, no. 5787102, pp.1-14, 2018.
[42] O. Sheyner, J. Haines, S. Jha, R. L ippmann, and J. Wing, "Automated generation and analysis of attack graphs," *in Proc. the IEEE Symposium on Security and Privacy*, 2002.

Hao Hu received his Ph.D. degree in the Zhengzhou Information Science and Technology Institute in 2018. He has been a lecturer with State Key Laboratory of Mathematical Engineering and Advanced Computing since 2018. His research interests include attack-defense modeling and proactive defense.

Yuling Liu received his M.S. degree and Ph.D. degree in the Trusted Computing and Information Assurance Laboratory, Institute of Software, Chinese Academy of Sciences in 2013. His main research interests include network and system security assessment and big data security. Since 2013, he has he has been an Associate Researcher with State Key Laboratory of Information Security, Chinese Academy of Sciences.

Chen Chen received the B.Eng., M.Sc., and Ph.D. degrees in electrical engineering and computer science (EECS) from Xidian University, Xi'an, China, in 2000, 2006, and 2008, respectively. He was a Visiting Professor with the Department of EECS, University of Tennessee, and with the Department of CS with the University of California. He is currently an Associate Professor with the Department of EECS, Xidian University. He is also the Director of the Engineering Technology Transfer Center, Xi'an, in edge computing, and the Associate Director of the Intelligent Transportation Research Laboratory, Xidian University. He has authored/coauthored two books and over 80 scientific articles in international journals and conference proceedings. He has contributed to the development of five copyrighted software systems and invented over 50 patents. He is also a Senior Member of China Computer Federation (CCF) and member of

the ACM and the Chinese Institute of Electronics. He serves as a General Chair, a PC Chair, a Workshop Chair, a Financial Chair, or TPC Member of a number of conferences.

Hongqi Zhang received the Ph.D. degree in Zhengzhou Information Science and Technology in 1998. His main research interests include network security and classication protection, etc. Since 2002, he has been a Professor and a Ph.D. Supervisor with the China National Digital Switching System Engineering & Technological Research Center. He is a member of the advisory committee of cyber security teching in higher education, ministry of eduction. P.R. China. He is also the Editor of the *Chinese Journal of Network and Information Security*. Mr. Zhang received prizes, such as the second prize of the National teaching prize in 2009, the National Network Security Outstanding Teacher Award in 2017 and the first prize of the National Science and Technology Progress Award in 2018. He is a senior member of China Computer Federation.

Yi Liu received the B.S. degree in information engineering and M.S. degree in computer science and technology from Zhengzhou Information Science and Technology Institute, China. She is currently pursuing the Ph.D. degree in computer science and technology with Zhengzhou Information Science and Technology Institute. She is also a visiting scholar in State Key Laboratory of Mathematical Engineering and Advanced Computing. Her research interests include information security, software defined network, service function chain.