Predictive Cyber Defense Remediation against Advanced Persistent Threat in Cyber-Physical Systems

Kamrul Hasan[1], Sachin Shetty[2],Tariqul Islam[3], Imtiaz Ahmed[4]
[1]Tennessee State University, Nashville, TN, USA
[2]Old Dominion University, Norfolk, VA, USA
[3]Syracuse University, Syracuse, NY, USA
[4]Howard University, Washington, D.C., USA
Email: {*mhasan1*@tnstate.edu, sshetty @odu.edu, mtislam@syr.edu, imtiaz.ahmed@howard.edu }

2022 International Conference on Computer Communications and Networks (ICCCN) | 978-1-6654-9726-8/22/$31.00 ©2022 IEEE | DOI: 10.1109/ICCCN54977.2022.9868886

***Abstract*—Advanced Persistent Threat (APT) has dramatically changed the landscape of cybersecurity. APT is carried out by stealthy, continuous, sophisticated, and well-funded attack processes for long-term malicious gain thwarting most current defense mechanisms. There is a need for a defense strategy that continuously combats APT over a long time-span in imperfect/incomplete information on attacker's actions. We propose the stochastic evolutionary game model to simulate the dynamic adversary to address this need in this work. We add the player's rationality parameter ς to the Logit Quantal Response Dynamics $(LQRD)$ model to quantify the cognitive differences of real-world players. We propose an optimal decision-making plan by calculating the stable evolutionary equilibrium that balances a trade-off between defense cost and benefit. Cases studies conducted on Energy Delivery Systems (EDS) indicate that the proposed method can help the defender predict possible attack action, select the related optimal cyber defense remediation over time, and gain the maximum defense payoff.**

***Index Terms*—APT, Cyber Defense, Cyber-Physical Systems (CPS), Energy Delivery Systems (EDS)**

I. Background

Advanced Persistent Threat (APT) has dramatically changed the landscape of cybersecurity. APT is carried out by stealthy, continuous, sophisticated, and well-funded attack processes for long-term malicious gain thwarting most current defense mechanisms. There is a need for a defense strategy that continuously combats APT over a long time-span in imperfect/incomplete information on attacker's actions. We propose the stochastic evolutionary game model to simulate the dynamic adversary to address this need in this work. We add the player's rationality parameter ς to the Logit Quantal Response Dynamics $(LQRD)$ model to quantify the cognitive differences of real-world players. We propose an optimal decision-making plan by calculating the stable evolutionary equilibrium that balances a trade-off between defense cost and benefit. Cases studies conducted on Energy Delivery Systems (EDS) indicate that the proposed method can help the defender predict possible attack action, select the related optimal cyber defense remediation over time, and gain the maximum defense payoff. In this work, we use EDS as an instance of CPS for case study.

Game theory is a decision-making theory for studying the direct interaction among decision-makers [1], 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. It mainly includes player, state, action, information, strategy, payoff, and equilibrium elements. Game theory has the characteristics of objective opposition, non-cooperative, and strategic interdependence and aligned with the essential attributes of cyber attack-defense [2]. Therefore, applying the game theory to model and analyze the cyber attack-defense process has become a hot research issue in recent years [3]. However, researchers have to address some challenges. To our best knowledge, existing game models for cyber attack-defense are mainly on the hypothesis of complete rational players [4] [5] [6]. Complete rationality includes several preconditions that are difficult to achieve, such as perfect rational consciousness, the perfectability of analyzing and inferring, identifying and judging, memorizing, and computing. If players can not reach any of these conditions, then it belongs to bounded rationality. The strict requirement of complete rationality is too harsh for the social attacker and defender. Real-world attackers and defenders have different cognizance abilities, which is determined by their interests, such as safety knowledge, skill level, experience, and so on [7]. In a word, the selection of strategy affected by various uncertain factors leads to the bounded rational game. At present, this issue assumes a significant challenge.

To sum up the above, APT calls for a framework that could characterize the continuous interplay of $advanced$ defense-attack on system resources with $imperfect/incomplete$ opponent's actions in a long time-span. This study involves (1) a model to accurately capture the continuously evolving process of the system status and how attackers influence it and defender's actions; and (2) dynamic defense/attack strategies that judiciously and continuously take steps to minimize/maximize the long-term system damage without knowing the opponent's behavior.

The main contributions of this paper are as follows:

(1). We construct the dynamic stochastic attack-defense

978-1-6654-9726-8/22/$31.00 ©2022 IEEE

game model for describing the evolutionary process of cyber attack-defense remediation.

(2). Then we analyze the improved strategy payoff calculation and formation of the optimal strategies.

(3) We also design the optimal cyber defense remediation selection approach for a multistep attack.

II. Related Works

A. Advanced Persistent Threat

The cybersecurity domain has been changed dramatically by a new class of threats, referred to as *Advanced Persistent Threat (APT)* by industry. The earliest well known APT case $Stuxnet$ [8] designed to modify industrial Programmable Logic Controllers and force them to diverge from the expected behaviors by exploiting a vast majority of security holes and tools. Another famous APT case is *Operation Aurora* [9], which targets at Google and dozens of other companies. The APT attacker can exploit the *zero-day* vulnerability in the Internet Explorer. [10] introduces the definition of APT and unique characteristics, making it different from traditional security issues. Sengupta et al. [11] propose a game-theoretic approach to model the *stealthy- takeover* property of APT and provide several guidelines for the system design based on the analytic results.

B. Game Theory in Cyber Security

Depending on the players' rational degree, existing research can be divided into two categories: complete rational game and bounded rational game.

A complete rationale game takes the hypothesis that the players have full cognition. Each player can select the best strategy to maximize its payoff and predict other players' strategy selections. The Nash equilibrium calculates the optimal response of each player through maximizing the expected defense payoff. Orojloo et al. [12] regarded both attacker and defender as players in the game and treated the attack-defense adversary as the zero-sum game. They considered players have complete information and act simultaneously. They constructed the non-cooperative static game model based on a defense graph to analyze attack intention and select the optimal defense strategy. Aiming at sensor networks' security issue, Li et al. [5] constructed a non-cooperative game model between attackers and trusted sensor nodes to balance costs and benefits. Due to the Nash equilibrium solution's restrictions, Li et al. [6] used the Pareto optimization to calculate the equilibrium. Do et al. [2] analyzed the impact of attack-defense strategy changes on the defense evaluation of worm attack-defense performance with the Bayes game model's help. However, this method is limited to pure strategy Nash equilibrium. Etesami et al. [13] analyzed the optimal mixed strategy of IDS intrusion response. Pawlick et al. [14] regarded the defender as the signal sender and the attacker as the signal receiver and built an 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. The identified optimal defense strategy by calculating each stage's equilibrium helps defenders decide during various stages. Almost the same time, Lei et al. [15] constructed the multi-stage repeated game model of attack-defense. The defender inferred attacker type depending on a priori attack strategy, and the posterior inference was revised to improve the accuracy of the decision progressively. The above multi-stage models mainly analyze erratic behaviors in discrete 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 players' biology property. 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 [16], and the strategy selection is the process of continuous learning and adjustment. Although Hu et al. [17] proposed an evolutionary game model of bounded rationality, this model is restricted from analyzing the payoff between two strategies only. Also, the stable equilibrium states, according to rationality changes, did not notify accordingly in the simulation result. So, without the premise of bounded rationality in a diversified strategy selection, the modeling and analysis of attack-defense may be impractical. Therefore, studying the attack-defense game rules under bounded rationality is an applied and promising research issue.

III. The Systems Model of Our Optimal Cyber Defense Remediation

The architecture of our optimal defense decision-making approach illustrates in Fig. 1. The input includes evidence such as vulnerability database, Nessus scanned logs, Attack Graph (AG), Intrusion Detection System's (IDS) real-time alert, security configuration, network topology, and MITRE ATT&CK [18], and the output is the optimal defense strategy. The decision-making process involves five steps: (1) Determine the targets and critical attack paths to strategy selection targets. (2) Extract candidate attack-defense strategies from the input security data according to the enhanced AG of attack evidence and abnormal evidence. (3) Model the attack-defense process as the stochastic evolutionary game based on the $LQRD$ model. (4) Evaluate the strategy payoff based on cost-benefit analysis. (5) Generate optimal defense strategy.

Besides, as a typical cyber adversary scenario consists of multiple players, we also extract the set of candidate defense strategies by analyzing the network environment information, including the vulnerability repairs, firewall access rules, security configuration, etc. 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 tactics and techniques in the MITRE ATT&CK framework [18].

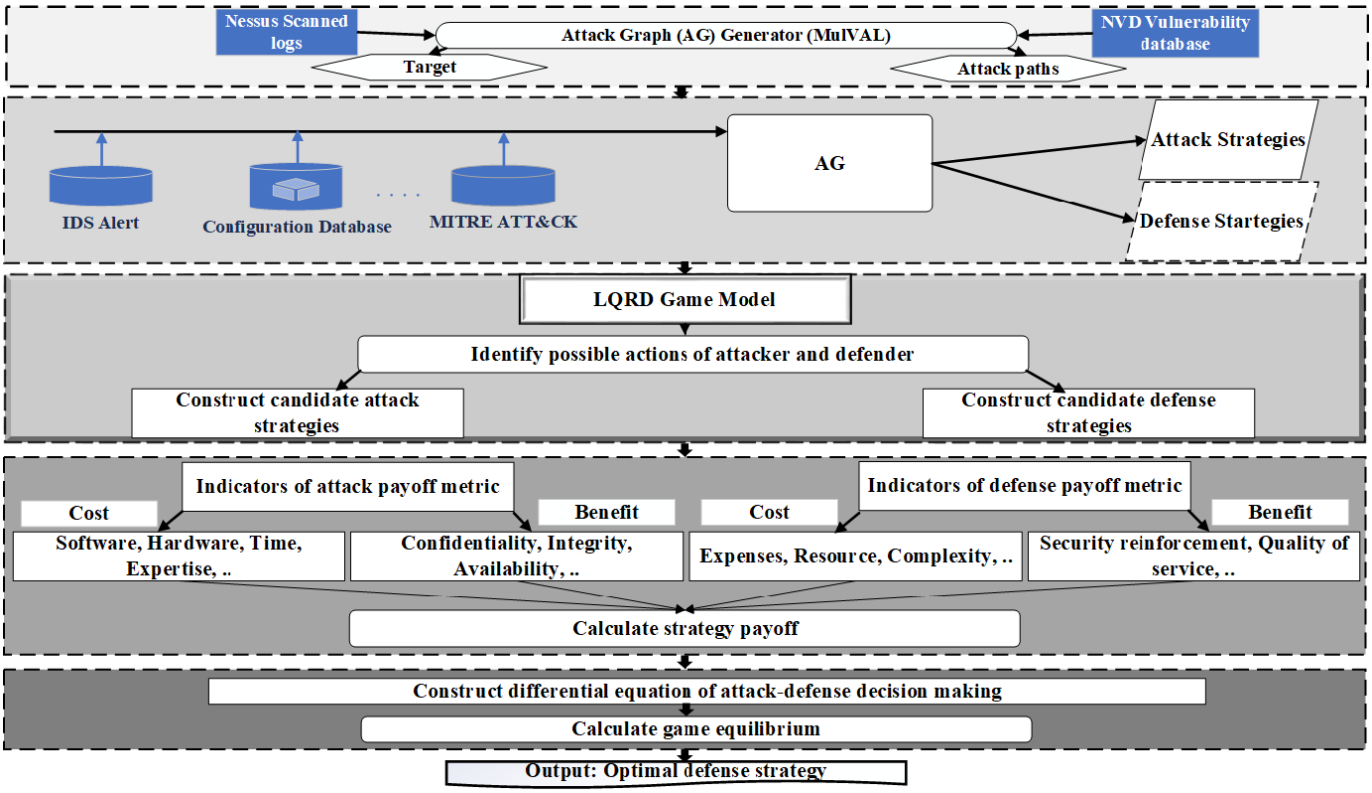


Fig. 1: The architecture of our cyber defense remediation method

A. *Game Modeling of Attack-Defense Based on* $LQRD$:

The evolutionary game model includes four essential elements: player sets, candidate strategy set, belief set, and payoff set.

Definition 1. A four-tuple can denote the model of the Attack-defense Stochastic Evolutionary Game (ASEGM).

(1) $\mathfrak{N} = (\mathfrak{N}_{\mathfrak{A}}, \mathfrak{N}_{\mathfrak{D}})$ is the population set of attack-defense players, where $\mathfrak{N}_{\mathfrak{A}}$ and $\mathfrak{N}_{\mathfrak{D}}$ are the populations of attackers and defenders, respectively.

(2) $\mathfrak{S} = (\mathfrak{S}_{\mathfrak{A}}, \mathfrak{S}_{\mathfrak{D}})$ is the set of candidate attack-defense strategies, in which $\mathfrak{S}_{\mathfrak{A}} = \{\mathfrak{A}_1, \mathfrak{A}_2, \ldots.., \mathfrak{A}_n\}$ is the set of the candidate strategies for attackers, $\mathfrak{S}_{\mathfrak{D}} = \{\mathfrak{D}_1, \mathfrak{D}_2, \ldots.., \mathfrak{D}_m\}$ is the set of candidate strategies for defenders. n and m are the numbers of attack and defense strategies, respectively. where $m, n \epsilon N^+$ and $n, m \geq 2$

(3) $\Theta = (\mathfrak{P}, \mathfrak{Q})$ is the belief set of the attack-defense game, where $p_i \epsilon \mathfrak{P}$ represents the probability that the attacker selects candidate strategy $\mathfrak{A}_i$, $q_i \epsilon \mathfrak{Q}$ represents the probability that defender chooses candidate strategy $\mathfrak{D}_j$, 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) $\mathfrak{U} = (\mathfrak{U}_{\mathfrak{A}}, \mathfrak{U}_{\mathfrak{D}})$ is the payoff function set. $\mathfrak{U}_{\mathfrak{A}}$ and $\mathfrak{U}_{\mathfrak{D}}$ represent the payoff functions of attack and defense, respectively.

B. *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 immediate benefit is the level of security reinforcement. Security measures only consider the direct benefits, and we further add the indirect benefits of the defender through the counterattack. For example, the electronic evidence of port scanning time, port number, source IP address, and destination IP address can use to reconstruct the attack chain. Through which the defender can earn indirect benefits through investigating 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 the implementation of security devices, and the economic value of affecting the regular operation of service (a.k.a.negative impact of control measures.).

Definition 6: Attack-defense payoff matrices M are as follows. In which, $\mathfrak{a}_{ij}$ and $\mathfrak{d}_{ij}$ represent the attack and defense payoff of selecting strategy combination $(\mathfrak{A}_i, \mathfrak{D}_j)$ respectively, $\mathfrak{a}_{ij} = AB - AC$, $\mathfrak{d}_{ij} = DB - DC$. The payoff matrices M is

as below:

$$M = \begin{pmatrix} \mathfrak{a}_{11}, \mathfrak{d}_{11} & \mathfrak{a}_{12}, \mathfrak{d}_{12} & \cdots & \mathfrak{a}_{1m}, \mathfrak{d}_{1m} \\ \mathfrak{a}_{21}, \mathfrak{d}_{21} & \mathfrak{a}_{22}, \mathfrak{d}_{22} & \cdots & \mathfrak{a}_{2m}, \mathfrak{d}_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ \mathfrak{a}_{n1}, \mathfrak{d}_{n1} & \mathfrak{a}_{n2}, \mathfrak{d}_{n2} & \cdots & \mathfrak{a}_{nm}, \mathfrak{d}_{nm} \end{pmatrix} \quad (1)$$

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

Evolutionary stable strategy (ESS) is an optimal decision of the game system in long-time strategy evolution. The definition of the permanent evolutionary strategy of cyber attack-defense is as follows:

Definition 7: Suppose the attacker population selects the candidate strategy set $\mathfrak{S}_{\mathfrak{A}} = (\mathfrak{A}_1, \mathfrak{A}_2, .., \mathfrak{A}_n)$ with the probability distribution $\mathfrak{P} = (p_1, p_2, .., p_n)$, and the defender population selects the candidate strategy set $\mathfrak{S}_{\mathfrak{D}} = (\mathfrak{D}_1, \mathfrak{D}_2, .., \mathfrak{D}_m)$ with the probability distribution $\mathfrak{Q} = (q_1, q_2, .., q_m)$. It means that individuals in the attacker and defense population randomly selects and implements their pure strategies with the probability distribution $\mathfrak{P}$ and $\mathfrak{Q}$ in the actual adversary. We call $\sigma^* = (\mathfrak{P}, \mathfrak{Q})$ as the stable strategy of attack-defense if the following conditions hold.

(1). (stability) $\mathfrak{U}(\sigma^*\sigma^*) \geq \mathfrak{U}(\sigma, \sigma^*)$

(2). (balance) $\mathfrak{U}(\sigma^*, \sigma^*) = \mathfrak{U}(\sigma, \sigma^*) \Rightarrow \mathfrak{U}(\sigma^*, \sigma) \geq \mathfrak{U}(\sigma, \sigma)$

Where, $\mathfrak{U}(\sigma^*, \sigma^*)$ denotes the payoff when attacker and defender both select σ^*. For any $\sigma \neq \sigma^*$, $\mathfrak{U}(\sigma^*, \sigma)$ is the payoff when either side changes its strategy.

Condition (1) guarantees that both attacker and the defender cannot earn more if either side changes strategy. In a policy containing a large number of σ^* and a small number of σ, it is necessary to meet that $\mathfrak{S}$ 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 approach mutates to σ.

Definition 7 provides the condition of whether a strategy is an evolutionarily stable strategy, but does not characterize the track of players' selection on this strategy and the attack-defense players search for the best approach and are disturbed by stochastic error. This Sub-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 [19]. That is to say, and 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 combined with Eqns. 2-4. So, the differential equation of the probability of selecting this strategy is [20]:

$$\frac{dp_i}{dt} = \sum_{k=1}^{n} p_k c_{ki} - \sum_{y=1}^{n} p_i c_{iy} \quad (2)$$

Where, p_i is the probability of selecting strategy $\mathfrak{A}_i$, $\frac{dp_i}{dt}$ is the probability that selects strategy $\mathfrak{A}_i$ varying with time. c_{ki} is the conditional transition probability of the attackers selection from strategy $\mathfrak{A}_k$ to strategy $\mathfrak{A}_i$, which describes the updating rules of strategy selections.

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

$$c_{ki} = \frac{exp(\varsigma \mathfrak{U}_{\mathfrak{A}_i})}{\sum_{k=1}^{n} exp(\varsigma \mathfrak{U}_{\mathfrak{A}_i})} \quad (3)$$

Set the rational parameters $\varsigma(\varsigma \geq 0)$ based on the historical rational degrees of players. The bigger ς is, the higher the degree of rationality is. The payoff is $\mathfrak{U} = \mathfrak{V} + \epsilon$, where $\mathfrak{V}$ is the payoff of observable factors, ϵ is the payoff of uncertain factors. The deduction of c_{ki} in Eqn. 3 can be referred to [19]. Take the formula in Eqn. 3 into Eqn. 2 and get the $LQRD$ equation as following:

$$\frac{dp_i}{dt} = \frac{exp(\varsigma \mathfrak{U}_{\mathfrak{A}_i})}{\sum_{k=1}^{n} exp(\varsigma \mathfrak{U}_{\mathfrak{A}_i})} - p_i \quad (4)$$

The Eqn. 4 shows that the change rate of the population proportion of player selecting strategy $\mathfrak{A}_i$ is proportional to the difference between the ratio of individual expected payoff to the total gain and the balance of unique numbers of choosing this strategy to the whole numbers. It also shows that in the attacker population composed of bounded rational players, the number change rate of players selecting a specific candidate strategy varies with the proportion of this strategy payoff to the total gain.

To construct the $LQRD$ equations of attack-defense, from condition (3) of **Definition 1**, we denote the strategy of probability vectors $\mathfrak{P}$ and $\mathfrak{Q}$ is the mixed probability of selecting $\mathfrak{S}_{\mathfrak{A}}$ and $\mathfrak{S}_{\mathfrak{D}}$ respectively. The evolution equations are as follows:

(1). Evolution equation of attack strategy over time The expected payoff $\mathfrak{U}_{\mathfrak{A}_i}$ of an attacker selecting candidate strategy $\mathfrak{A}_i$ is as follows, $i = 1, 2, ..., n$

$$\begin{aligned} \mathfrak{U}_{\mathfrak{A}_1} &= q_1\mathfrak{a}_{12} + q_2\mathfrak{a}_{12} + ... + q_m\mathfrak{a}_{1m} \\ \mathfrak{U}_{\mathfrak{A}_2} &= q_1\mathfrak{a}_{21} + q_2\mathfrak{a}_{22} + ... + q_m\mathfrak{a}_{2m} \\ \mathfrak{U}_{\mathfrak{A}_i} &= q_1\mathfrak{a}_{i1} + q_2\mathfrak{a}_{i2} + ... + q_m\mathfrak{a}_{im} = \sum_{j=1}^{m} q_j\mathfrak{a}_{ij} \\ &.... \\ \mathfrak{U}_{\mathfrak{A}_n} &= q_1\mathfrak{a}_{n1} + q_2\mathfrak{a}_{n2} + ... + q_m\mathfrak{a}_{nm} \end{aligned}$$

The changing rate of the proportion of individuals selecting strategy $\mathfrak{A}_i$ in the attacker population overtime is $\frac{dp_i}{dt}$. It reflects the learning and improving selecting strategy $\mathfrak{A}_i$ for bounded rational attacker through repeated games. The $LQRD$ differential equation of change rate from Eqn. 4 is:

$$\frac{dp_i}{dt} = \frac{exp(\varsigma \sum_{j=1}^{m} q_j \mathfrak{a}_{ij})}{\sum_{k=1}^{n} exp(\varsigma \sum_{j=1}^{m} q_j \mathfrak{a}_{kj})} - p_i \tag{5}$$

(2). Evolution equation of defense strategy over time.

The expected payoff $\mathfrak{U}_{\mathfrak{D}_j}$ of an attacker selecting candidate strategy $\mathfrak{D}_j$ is as follows, $j = 1, 2, ..., m$

$$\begin{aligned}
\mathfrak{U}_{\mathfrak{D}_1} &= p_1\mathfrak{d}_{11} + p_2\mathfrak{d}_{21} + ... + p_n\mathfrak{d}_{n1} \\
\mathfrak{U}_{\mathfrak{D}_2} &= p_1\mathfrak{d}_{12} + p_2\mathfrak{d}_{22} + ... + p_n\mathfrak{d}_{n2} \\
\mathfrak{U}_{\mathfrak{D}_j} &= p_1\mathfrak{d}_{1j} + p_2\mathfrak{d}_{2j} + ... + p_n\mathfrak{d}_{nj} = \sum_{i=1}^{n} p_i\mathfrak{d}_{ij} \\
&.... \\
\mathfrak{U}_{\mathfrak{D}_m} &= p_1\mathfrak{d}_{1m} + p_2\mathfrak{d}_{2m} + ... + p_n\mathfrak{d}_{nm}
\end{aligned}$$

The changing rate of the proportion of individuals selecting strategy $\mathfrak{D}_j$ in the defender population overtime is $\frac{dq_j}{dt}$. It reflects the learning and improving selecting strategy $\mathfrak{D}_j$ for bounded rational defender through repeated games. So, the $LQRD$ differential equation of change rate is:

$$\frac{dq_j}{dt} = \frac{exp(\varsigma \sum_{i=1}^{n} p_i \mathfrak{d}_{ij})}{\sum_{k=1}^{m} exp(\varsigma \sum_{i=1}^{n} p_i \mathfrak{d}_{ik})} - q_j \tag{6}$$

The practical significance of the above evolution equation is: 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. The growth rate of the defender number choosing D_j is larger than zero. Otherwise, the growth rate is less than zero. If the number proportion is exactly equal to the payoff proportion, then the growth rate of the number of defender selecting strategy D_j is zero. Set $F(p_i) = \frac{dp_i}{dt}$, $G(q_j) = \frac{dq_j}{dt}$, and then combine the above equations to equate below condition:

$$Y(p_i, q_j) = \begin{pmatrix} F(p_i) \\ G(q_j) \end{pmatrix} = 0 \tag{7}$$

This will give us the stable equilibrium of attack-defense adversary.

IV. Implementation, Result, and Analysis:

The Industroyer malware has unleashed a major escalation in cyber-attacks on Industrial Control Systems (ICS) by combining a multi-stage APT attack with in-depth domain knowledge. Industroyer (also referred to as Crash-override) is a malware framework considered to have been used in the cyber-attack on Ukraine's power grid on December 17, 2016. The attack cut a fifth of Kyiv, the capital, off power for one hour and considered a large-scale test [21]. The Kyiv incident was the second cyber-attack on Ukraine's power grid in less than a year. The first attack occurred on December 23, 2015. Industroyer is the first-ever known malware specifically designed to attack electrical grids. Simultaneously, it is the fourth malware publicly revealed to target industrial control systems, after Stuxnet, Havex, and BlackEnergy.

In this section, we take the invasion and proactive defense against Industroyer in the real-world Energy Delivery System (EDS) network as an example. We analyze the adversarial attack-defense process against Industroyer, verify the proposed approach for optimal defense strategy selection. The results of the two scenarios with different strategy payoffs are compared and analyzed. Besides, 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. *EDS Network Implementation:*

We implemented an EDS network that is shown in Fig. 2 from [22]. The entire test-bed is connected to a network switch and a router, and the zoning is implemented using VLAN and firewall rules. There are five subnets created by an external and internal firewall. The IT Workstations (WSs) were located at the IT subnet. A Web Server (WebS) is located at the DMZ subnet and is directly accessible from the Internet through an external firewall. Supervisory Control and Data Acquisition (SCADA) servers (L3/L2), Remote Transmit Unit (RTUs) (L1) are in different subnets under larger Operational Technology (OT) subnet that holds critical communication. The SCADA1 servers and SCADA2 servers are only accessible from the WebS of the DMZ. The WebS is accessible from user WS and other hosts from level 4 or 5. The user subnet contains the user's WS. The firewalls allow all outbound traffic from users subnet. The test-bed also includes Intrusion Detection System (IDS) running both IT and OT specific rules and a commercial OT Asset Discovery and Management (ADM). They are both connected to the span port of the switch to inspect the entire ICS traffic. For the Industroyer attack simulation, we injected vulnerabilities on the test-bed machines. The user workstations contained the vulnerability CVE-2009-1918 in Internet Explorer (IE). If a user accesses malicious content using the vulnerable IE browser, the device may be compromised. The WebS contained the vulnerability CVE-2006-3747 in the Apache HTTP service, resulting in a remote attacker executing arbitrary code on the machine. The SCADA1 and SCADA2 server had the vulnerability CVE-2018-5313, allowing privilege escalation up to the administrator level. The SCADA1 server controls 10 RTUs of substation 1, whereas the SCADA2 server controls 7 RTUs of substation 2. We assume that if an attacker acquires control over the SCADAs, the RTUs can be acquired as well.

As a defender, the network center's administrator is responsible for the security of the EDS's whole intranet. The attacker comes from the external network and attacks the

intranet through the Internet. The purpose is to erase system-crucial registry keys and overwrite all ICS configuration files to make the system unbootable and recovery from the attack harder. Industroyer attacks can be divided mainly 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 only communicate with the IT network's Work Station (WS) and mail server but cannot access the Operation Technology (OT) network. The security protection devices are composed of the firewall, Intrusion Protection System (IPS), virus detection system (VDS), and patch management system. We used the Nessus scanning tool to scan the EDS network. Table I shows the results of the principal vulnerabilities.

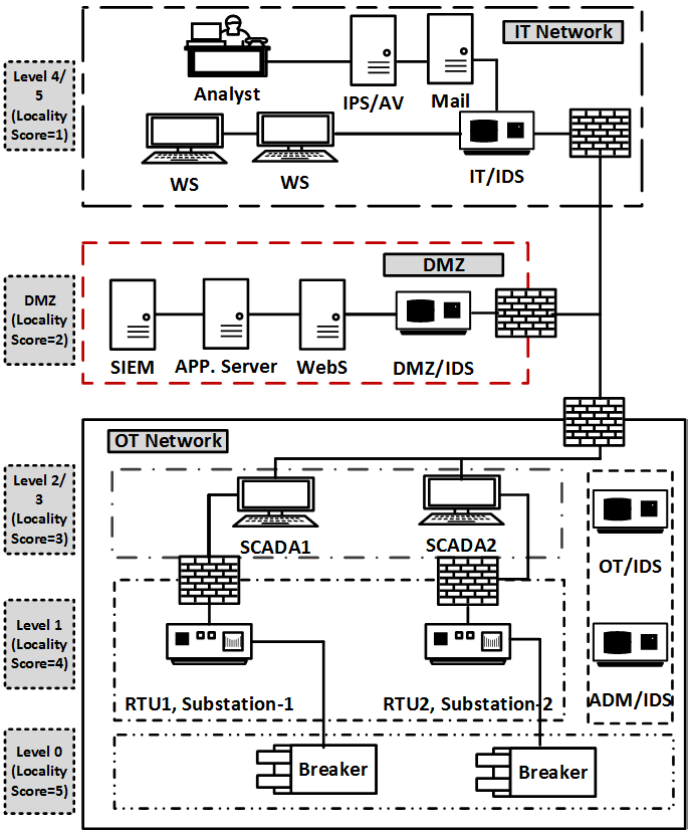


Fig. 2: Logical view of EDS Test-bed

B. Candidate Strategy Extraction and Payoff Calculation:

In this experiment, based on the network topology and vulnerabilities, the logical Attack Graph (AG) is created using the open-source tool $MulVAL$ as illustrated in Fig. 3 [22]. The $MulVAL$ is a reasoning toolkit for automatically identifying vulnerabilities in IT and OT networks [23]. The different shapes represent the network state, and the edge represents the atomic attack action. By referring to the attack-defense behavior database of MITRE for Industrial Control Systems (ICS) [24], 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.

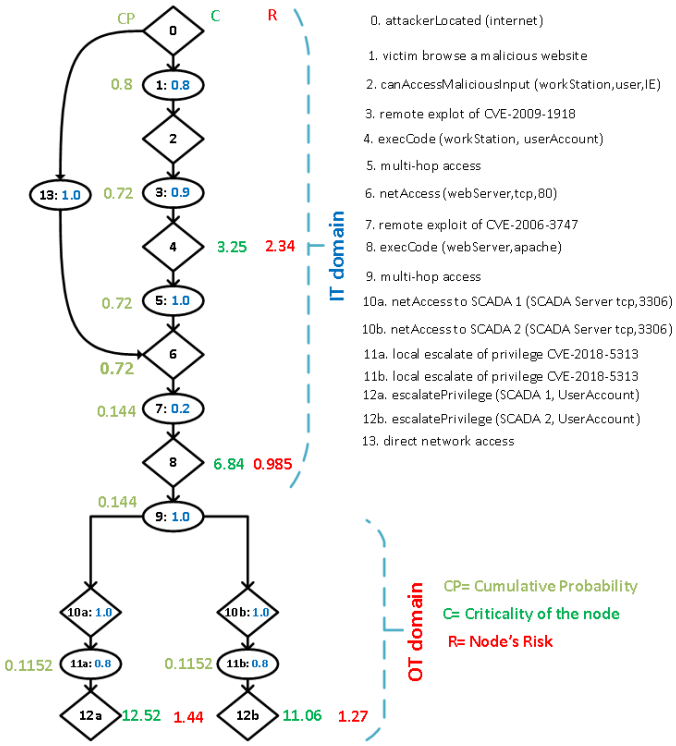


Fig. 3: The AG of test-bed based EDS [22]

We find that the attacker first conducted port scanning action $\mathfrak{A}_1$ through port 25 of the mail server at the IT domain. Furthermore, the attacker collected open service information to prepare for subsequent attacks. Since port scanning is a concealed means of attacking, which is the passive attack virtually, we denote it as $\mathfrak{A}_1$ =*Scan Port*. Based on further detections and analyses of alert information, we find that some adventurous attackers may execute atomic attacks, $\mathfrak{A}_4$, and $\mathfrak{A}_5$ shown in Table II along the most critical path from the alert node to a goal SCADA 1/SCADA 2 [22]. The unauthenticated attackers exploit the vulnerability *CVE-2006-3747* of Webs at DMZ to allows remote attackers to cause a denial of service (application crash) and possibly execute arbitrary code via crafted URLs that are not adequately handled using certain rewrite rules. We denote this candidate strategy as $\mathfrak{A}_4$=*Denial Of Service (DoS)*, which is an active attack. After the WebS is compromised as the next stage of an APT, the attacker starts exploiting *CVE-2018-5313* of SCADA 1/SCADA 2. We denote this candidate strategy as $\mathfrak{A}_5$ = *Execute Arbitrary Code*, which is also an active attack. So, in this experiment three candidate defense strategies $\mathfrak{D}_1$ = *Close Unused Ports*, $\mathfrak{D}_4$ =*Block Unwanted IP Address*, and $\mathfrak{D}_5$ = *Install Patches* are mapped from Table II as an extraction for that critical APT chain.

TABLE I: Network Configuration and Vulnerability Information

Nodes	$Configuration$	CVE	$Description$
WS	Microsoft Internet Explorer (IE)	CVE-2009-1918	Allows remote attackers to execute arbitrary code via a crafted HTML document
WebS	Apache Web Server	CVE-2006-3747	allows remote attackers to cause a denial of service (application crash) and possibly execute arbitrary code via crafted URLs
SCADA 1	SCADA Master server	CVE-2018-5313	An attacker can leverage this vulnerability to execute arbitrary code under the context of Administrator
SCADA 2	SCADA Master server	CVE-2018-5313	An attacker can leverage this vulnerability to execute arbitrary code under the context of Administrator

TABLE II: Cyber Attack and Defense Actions

No.	Attack Action	No.	Defense Option
$\mathfrak{A}_1$	Scan Port	$\mathfrak{D}_1$	Close Unused Port
$\mathfrak{A}_2$	Obtain Root Privilege	$\mathfrak{D}_2$	Restart Device
$\mathfrak{A}_3$	Buffer Overflow	$\mathfrak{D}_3$	Offline Network
$\mathfrak{A}_4$	Denial of Service	$\mathfrak{D}_4$	Block unwanted IPs
$\mathfrak{A}_5$	Execute Arbitrary Code	$\mathfrak{D}_5$	Install Patches

From Eqn. 1, the payoff matrix of attack-defense is as follows: $M = \begin{pmatrix} \mathfrak{a}_{11}, \mathfrak{d}_{11} & \mathfrak{a}_{12}, \mathfrak{d}_{12} & \mathfrak{a}_{13}, \mathfrak{d}_{13} \\ \mathfrak{a}_{21}, \mathfrak{d}_{21} & \mathfrak{a}_{22}, \mathfrak{d}_{22} & \mathfrak{a}_{23}, \mathfrak{d}_{23} \\ \mathfrak{a}_{31}, \mathfrak{d}_{31} & \mathfrak{a}_{32}, \mathfrak{d}_{32} & \mathfrak{a}_{33}, \mathfrak{d}_{33} \end{pmatrix}$

C. Evolution of Equations for Decision Making:

Firstly, we set the attackers and defenders with equal degrees of rationality. Furthermore, we assign the proportion of the number of players selecting strategy $\mathfrak{A}_1$, $\mathfrak{A}_4$, and $\mathfrak{A}_5$, in the attacker population as p_1,p_2, and p_3 respectively. Secondly, we assign the proposition of defender population selecting strategy $\mathfrak{D}_1$, $\mathfrak{D}_4$, and $\mathfrak{D}_5$, as q_1, q_2, and q_3 respectively. Besides, we construct the $LQRD$ equation of attack-defense strategy as follows, respectively.

The expected payoff of attacker selecting strategy $\mathfrak{A}_1$ = *port scan attack* is $\mathfrak{U}_{\mathfrak{A}_1} = \mathfrak{a}_{11}q_1 + \mathfrak{a}_{12}q_2 + \mathfrak{a}_{13}q_3$, the expected gain of *denial of service* is $\mathfrak{U}_{\mathfrak{A}_4} = \mathfrak{a}_{21}q_1 + \mathfrak{a}_{22}q_2 + \mathfrak{a}_{23}q_3$, and the expected payoff of attacker selecting strategy $\mathfrak{A}_5$ = *Execute Arbitrary Code* is $\mathfrak{U}_{\mathfrak{A}_5} = \mathfrak{a}_{31}q_1 + \mathfrak{a}_{32}q_2 + \mathfrak{a}_{33}q_3$. Then, we can obtain the evolution equation of strategy $\mathfrak{A}_1$, $\mathfrak{A}_4$, and $\mathfrak{A}_5$ from Eqn. 5:

The expected payoff of defender selecting strategy $\mathfrak{D}_1$ = *close unused port* is $\mathfrak{U}_{\mathfrak{D}_1} = \mathfrak{d}_{11}p_1 + \mathfrak{d}_{21}p_2 + \mathfrak{d}_{31}p_3$, the expected gain of *denial of service* is $\mathfrak{U}_{\mathfrak{D}_4} = \mathfrak{d}_{12}p_1 + \mathfrak{d}_{22}p_2 + \mathfrak{d}_{32}p_3$, and the expected payoff of defender selecting strategy $\mathfrak{D}_5$ = *Execute Arbitrary Code* is $\mathfrak{U}_{\mathfrak{D}_5} = \mathfrak{d}_{13}p_1 + \mathfrak{d}_{23}p_2 + \mathfrak{d}_{33}p_3$. Then, we can obtain the evolution equation of strategy $\mathfrak{D}_1$,$\mathfrak{D}_4$, and $\mathfrak{D}_5$ from Eqn. 6:

Then, according to Eqn. 7, equalize all equations to zero. The solution of those equations is the stable evolutionary equilibrium of attack-defense decision-making, and defender's optimal defense strategy is selecting strategy $\{\mathfrak{D}_1, \mathfrak{D}_4, \mathfrak{D}_5\}$ with mixed probability $\{q_1, q_2, q_3\}$.

D. Result and Analysis:

We consider two numerical experiments: *Scenario 1* (without considering counterattack payoff) and *Scenario 2* (considering counterattack payoff). In this work, we only consider Scenario 1.

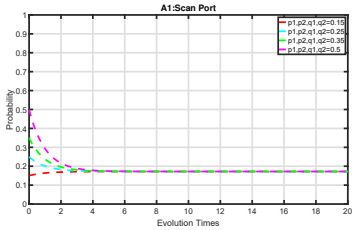
Scenario 1: We combine the *Definition 2 - Definition 5* and security behaviors database and then obtain the game payoff of attack-defense as organized in Table III.

TABLE III: Game Pay-off of Scenario 1

Candidate Attack Strategy	Candidate Defense Strategy		
	$\mathfrak{D}_1$	$\mathfrak{D}_4$	$\mathfrak{D}_5$
$\mathfrak{A}_1$	(0.16,0.06)	(0.16,-0.15)	(0.16,-0.3)
$\mathfrak{A}_4$	(0.24,-0.2)	(0.24,0.39)	(0.24,-0.3)
$\mathfrak{A}_5$	(0.4,-0.2)	(0.4,-0.15)	(0.4,0.7)

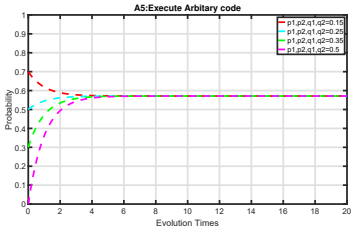
In general, the degree of player rationality in the real world is medium, and here we set $\varsigma = 5.0$, and set the initial state of the game system as $p_1 = p_2 = p_3 = q_1 = q_2 = q_3 = 0.33$. That is, the attacker randomly selects a strategy from candidate $\mathfrak{A}_1$,$\mathfrak{A}_4$, and $\mathfrak{A}_5$ with equal probability 0.33 at the initial time. Similarly, the defender randomly selects a action from candidate $\mathfrak{D}_1$, $\mathfrak{D}_4$, and $\mathfrak{D}_5$ with equal probability. With the simulation tool $Matlab$ 2021, the stable equilibrium point is calculated by function $fsolve()$ for $\varsigma = 5.0$. The calculated stable equilibrium point is $\{p_1, p_2, p_3\} = \{0.172, 0.257, 0.571\}$ and $\{q_1, q_2, q_3\} = \{0.087, 0.179, 0.734\}$. In this context, the attacker is more likely to select $\{\mathfrak{A}_1, \mathfrak{A}_4, \mathfrak{A}_5\}$ with mixed probability of $\{0.172, 0.257, 0.571\}$. Meanwhile, the optimal defense strategy for the defender is to randomly implement $\{\mathfrak{D}_1, \mathfrak{D}_4, \mathfrak{D}_5\}$ with mixed probability $\{0.087, 0.179, 0.734\}$. The results show that the attacker is more likely to select the aggressive strategy $\mathfrak{A}_5$ = *Execute Arbitrary Code* with probability 0.571. Since the attack of the *Execute Arbitrary Code* is more harmful, to avoid the severe attack influence, the corresponding optimal defense strategy is to select $\mathfrak{D}_5$ = *Install Patch* with a probability of 0.571.

Secondly, to analyze the influence of the system's initial state on strategy selections, we simulate the evolution tracks of strategy selections with different first $p_1, p_2, p_3, q_1, q_2, q_3$ in Fig. 4a-4c and in Fig. 5a-5c. The abscissa t represents the number of evolutions in decision-making. The ordinate $probability$ represents the probability of selecting a strategy.



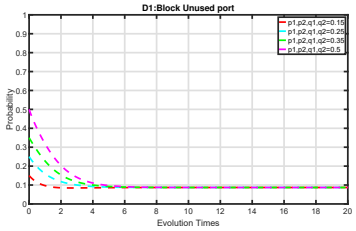
(a) A_1 =Scan Port

(b) A_4 =Denial of Service

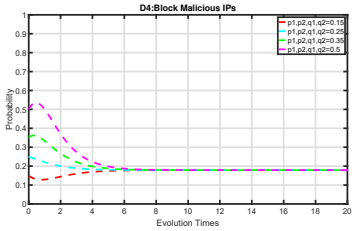


(c) A_5 =Execute Arbitrary code

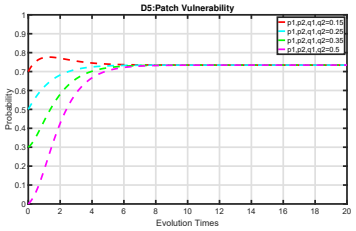
Fig. 4: Strategy Evolution of an Attacker



(a) D_1 =Close Unused ports



(b) D_4 =Block malicious IPs



(c) D_5 =Patch Vulnerability

Fig. 5: Strategy Evolution of a Defender

Fig. 4a-4c and Fig. 5a-5c can predict the defender's best strategy selection at different game moments.

Fig. 4a-4c and Fig. 5a-5c respectively show the evolution tracks of $\{\mathfrak{A}_1, \mathfrak{A}_4, \mathfrak{A}_5, \mathfrak{D}_1, \mathfrak{D}_4, \mathfrak{D}_5\}$, when the initial states of attacker and defender are the same with $p_1, p_2, q_1, q_2 = \{0.15, 0.25, 0.35, 0.5\}$. From Fig. 4a-4c and Fig. 5a-5c, we assume that the attacker and defender initially select the strategy $\{\mathfrak{A}_1, \mathfrak{A}_4\}$ and $\{\mathfrak{D}_1, \mathfrak{D}_4\}$ with probability $p_1 = p_2 = q_1 = q_2 = 0.5$, when t = 0. Then from the magenta curve of Fig. 5a, the likelihood of selecting strategy $\mathfrak{D}_1$ is falling over time and stabilize to $probability = 0.179$, when $t = 10$. Also, the possibility of choosing a strategy $\mathfrak{D}_4$ is falling and stabilize to $probability = 0.017$ from the magenta curve of Fig. 5b. Herein, the optimal defense strategy is selecting $\mathfrak{D}_1, \mathfrak{D}_4, \mathfrak{D}_5$ with mixed $probability = \{0.087, 0.179, 0.734\}$. This selection is stable and best when against different candidate attack strategies.

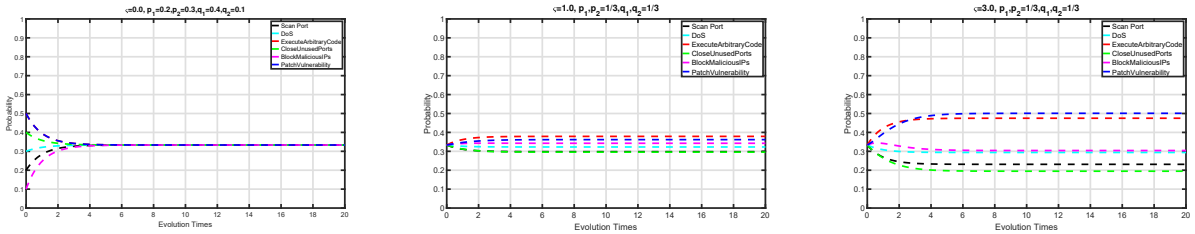
Moreover, as we assume that the defender selects the strategy $\{\mathfrak{D}_1, \mathfrak{D}_4, \mathfrak{D}_5\}$ with a probability $\{q_1, q_2, q_3\}$=$\{0.5, 0.5, 0.0\}$ initially, namely, the larger the gap between the defender's initial selection and the optimal selection $\{q_1 = 0.087, q_2 = 0.179, q_3 = 0.734\}$, the more evolution times needed to achieve the best strategy. In contrast to Nash equilibrium game model [6], our approach can better explain the strategy evolution rules in adversarial attack-defense and have stronger performance of attack prediction.

Again, the higher the probability of selecting a strategy from $\{\mathfrak{A}_1, \mathfrak{A}_4\}$ at the initial time, the later the curve inflection point appears. They are indicating that more number of repeated games is required for decision-making and longer time takes. The condition is due to that the attacker selects $\mathfrak{A}_1$ or $\mathfrak{A}_4$ with a very high probability at the initial time. The false signal deceived the defender. It caused the defender mistakenly to assume that the attacker will select the moderate attack strategy about $\mathfrak{A}_1$ and $\mathfrak{A}_4$ while overlooking the ultimate attack purpose $\mathfrak{A}_5$ = *Execute Arbitrary Code*. Therefore, rational defenders need to implement many evolution times to discover the attacker's real purpose and obtain the best defense strategy. For example, when $\{p_1 = 0.5, q_2 = 0.5\}$, the probability of selecting the strategy $\mathfrak{D}_4$ denoted by the magenta curve in Fig. 5b first increases to $q_2 = 0.54$ at $t = 0.466$ and then rebounds and finally stabilizes to $q_2 = 0.178$ at $t = 11$. The reason is that the proportion of the defender population selecting strategy $\mathfrak{D}_4$ at the initial time increases to high. With the increase of the $\mathfrak{D}_4$ payoff to the total payoff, the number of individuals selecting $\mathfrak{D}_4$ decreases gradually to ensure that the proportion of population selecting $\mathfrak{D}_4$ to the total population is equal to the proportion of payoff selecting $\mathfrak{D}_4$ to the total payoffs.

As can be seen from each column in Fig. 4a-4c and in Fig. 5a-5c, the optimal strategy for both defender and attacker are the same regardless of their initial p_1, p_2, p_3 and q_1, q_2, q_3 selections. It is only related to the candidate strategy set, player, and the strategy pays off. Moreover, the initial state can only affect the stabilization time of the game system.

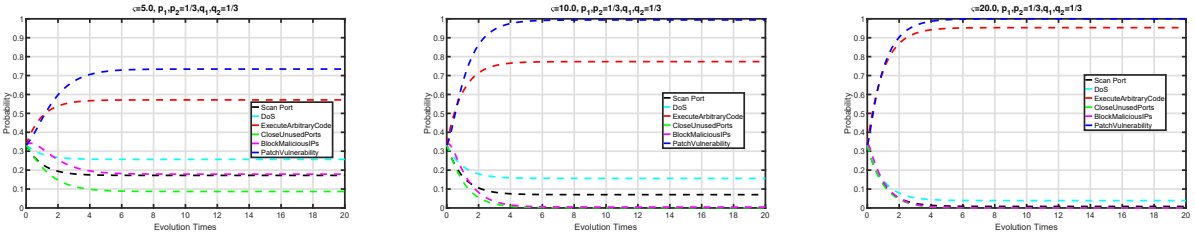
Finally, to analyze the influence of degrees of players' rationality on strategy evolution, some simulations show in Fig. Fig. 6a-6c and in Fig. 7a-7c and discussions are as follows:

1) When, we assume that the players are irrational and set $\varsigma = 0$, assign initial $p_1 = 0.2, p_2 = 0.3, q_1 = 0.4, q_2 = 0.1$, then obtain the strategy evolution tracks in Fig. 6a. Herein, the final result is to select the different candidate



(a) $\varsigma = 0, p_1, p_2 = 0.2, q_1, q_2 = 0.1$ (b) $\varsigma = 1.0, p_1, p_2 = 0.33, q_1, q_2 = 0.33$ (c) $\varsigma = 3.0, p_1, p_2 = 0.33, q_1, q_2 = 0.33$

Fig. 6: The strategy evolution tracks with different rationality ς.



(a) $\varsigma = 5.0, p_1, p_2 = 0.33, q_1, q_2 = 0.33$ (b) $\varsigma = 10.0, p_1, p_2 = 0.33, q_1, q_2 = 0.33$ (c) $\varsigma = 15.0, p_1, p_2 = 0.33, q_1, q_2 = 0.33$

Fig. 7: The strategy evolution tracks with different rationality ς.

strategy with the same probability of 0.33. It means that players cannot distinguish the advantages and disadvantages of varying candidate strategies since they have no cognitive abilities. Meanwhile, from the $LQRD$ Eqn. 7 of attack-defense, there is only one solution $\{p_1, p_2, p_3, q_1, q_2, q_3\} = \{0.33, 0.33, 0.33, 0.33, 0.33, 0.33\}$ when $\varsigma = 0$. The results show that when the game players are irrational, regardless of their initial selections, they cannot distinguish each strategy's merits and demerits since they do not have any learning and cognitive capabilities. The candidate strategies are still selected by game players randomly.

2) Suppose that the rational player degree $\varsigma > 0$, we simulate the strategy evolution in Fig. 6a-6c and in Fig. 7a-7c. As time goes by, all 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 $\varsigma = 5$, the game system can reach the stable state through about 6 times of game evolution (shows in Fig. 7a), while when $\varsigma = 10$, they can be stable only through 4 times of game evolution (shows in Fig. 7a). The 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 helps the defenders identify the optimal strategy more quickly.

In general, both sides of attackers and defenders gain increased decision-making experience through adversarial attack-defense. Hence, a rational degree of ς increases during the game process. Fig. 8 illustrates the results under different ς, where the abscissa ς represents the reasonable degree, and the ordinate represents the probability of strategy selection. When $\varsigma = 0$, players have no rationality, so they choose candidate strategies randomly. When $\varsigma = 0.1$, the reasonable degree of the players is very low as the replicator dynamics [6]. From Fig. 8, the probability of defender selecting strategy $\mathfrak{D}_1$ and $\mathfrak{D}_4$ rapidly decreases to 0 and $\mathfrak{D}_5$ increases to 1, respectively, which reflects the sensitivity of the decision-making system. The corresponding equilibrium solution is $\{p_1 = 0.33, p_2 = 0.33, p_3 = 0.34\}$ and $\{q_1 = 0.33, q_2 = 0.33, q_3 = 0.34\}$. The result corresponds to the replicator dynamic equilibrium [6]. Since the rational degree of dynamic replicator game is very low, its equilibrium solution is pure strategy. When $\varsigma > 0.1$, the player rational degree increases, and both sides of the attacker and defender always approach to complete balanced Nash equilibrium as ς increases. When $\varsigma > 15$, the solution $\{p_1 = 0.0003, p_2 = 0.27, p_3 = 0.97, q_1 = 0.02, q_2 = 0.1, q_3 = 0.88\}$ of $LQRD$ in this paper is very close to the Nash equilibrium solution. 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 ς towards infinity, then the proposed $LQRD$ equilibrium will approach Nash equilibrium. Compared with the complete rational Nash equilibrium [6] and the bounded rational replicator dynamic equilibrium, our approach can depict the diversity of rationality of attacker and defender players and reflect the real strategy selection rules.

V. Conclusion and Future Works:

This paper studies the strategy selection with a maximum payoff in the EDS attack-defense dispute based on the evolutionary bounded rationality game model. Advanced Persistent Threat (APT) becomes more diverse with the complexity and large-scale network information systems, leading the cyber

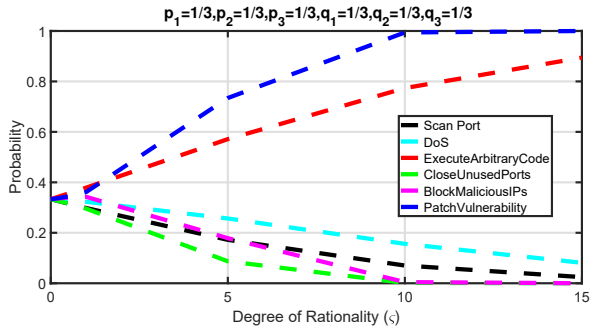


Fig. 8: The impact of rationality (ς) on the strategy selections

attack-defense situation to 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 a useful tool to model the adversarial cyber attack-defense. At present, game modeling of attack-defense with bounded rationality is still in its infancy. 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 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 build 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 CrashOverride to depict the best strategy formation. By analyzing the stable evolutionary equilibrium, we can obtain the optimal defense strategy at different game moments. Our approach is more generalized comparing with replicator dynamics and the Nash equilibrium model. Two case studies on Crash Override both show that the proposed method is effective and practical. The performances of attack prediction and defense decision-making are improved significantly for winning cyber attack-defense warfare. In the future, we will quantify the players' rationality from network logs, host logs, and communication protocols. We will then apply the machine learning and Artificial Intelligence (AI) technique to achieve the automatic analysis of attack-defense strategies to implement faster strategy implementation.

REFERENCES

[1] K. Hasan, S. Shetty, J. Sokolowski, and D. K. Tosh, "Security game for cyber physical systems," in *Proceedings of the Communications and Networking Symposium*, 2018, pp. 1–12.

[2] C. T. Do, N. H. Tran, C. Hong, C. A. Kamhoua, K. A. Kwiat, E. Blasch, S. Ren, N. Pissinou, and S. S. Iyengar, "Game theory for cyber security and privacy," *ACM Computing Surveys (CSUR)*, vol. 50, no. 2, pp. 1–37, 2017.

[3] K. Merrick, M. Hardhienata, K. Shafi, and J. Hu, "A survey of game theoretic approaches to modelling decision-making in information warfare scenarios," *Future Internet*, vol. 8, no. 3, p. 34, 2016.

[4] J.-l. Tan, C. Lei, H.-q. Zhang, and Y.-q. Cheng, "Optimal strategy selection approach to moving target defense based on markov robust game," *Computers & Security*, vol. 8, no. 5, pp. 63–76, 2019.

[5] Y. Li, D. E. Quevedo, S. Dey, and L. Shi, "A game-theoretic approach to fake-acknowledgment attack on cyber-physical systems," *IEEE Transactions on Signal and Information Processing over Networks*, vol. 3, no. 1, pp. 1–11, 2016.

[6] X. Li, C. Zhou, Y.-C. Tian, and Y. Qin, "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.

[7] W. M. Czarnecki, G. Gidel, B. Tracey, K. Tuyls, S. Omidshafiei, D. Balduzzi, and M. Jaderberg, "Real world games look like spinning tops," *arXiv preprint arXiv:2004.09468*, 2020.

[8] A. Cherepanov and R. Lipovsky, "Industroyer: Biggest threat to industrial control systems since stuxnet," *WeLiveSecurity, ESET*, vol. 12, 2017.

[9] W. Niu, X. Zhan, K. Li, G. Yang, and R. Chen, "Modeling attack process of advanced persistent threat," in *International Conference on Security, Privacy and Anonymity in Computation, Communication and Storage*. Springer, 2016, pp. 383–391.

[10] E. Cole, *Advanced persistent threat: understanding the danger and how to protect your organization*. Newnes, 2012.

[11] S. Sengupta, A. Chowdhary, A. Sabur, A. Alshamrani, D. Huang, and S. Kambhampati, "A survey of moving target defenses for network security," *IEEE Communications Surveys & Tutorials*, 2020.

[12] H. Orojloo and M. A. Azgomi, "A game-theoretic approach to model and quantify the security of cyber-physical systems," *Computers in Industry*, vol. 88, pp. 44–57, 2017.

[13] S. R. Etesami and T. Başar, "Dynamic games in cyber-physical security: An overview," *Dynamic Games and Applications*, vol. 9, no. 4, pp. 884–913, 2019.

[14] J. Pawlick, E. Colbert, and Q. Zhu, "A game-theoretic taxonomy and survey of defensive deception for cybersecurity and privacy," *ACM Computing Surveys (CSUR)*, vol. 52, no. 4, pp. 1–28, 2019.

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

[16] X. Wang, J. Quan, and W. Liu, "Study on evolutionary games and cooperation mechanism within the framework of bounded rationality," *Systems Engineering Theory & Practice*, vol. 31, no. 1, pp. 82–93, 2011.

[17] H. Hu, Y. Liu, C. Chen, H. Zhang, and Y. Liu, "Optimal decision making approach for cyber security defense using evolutionary game," *IEEE Transactions on Network and Service Management*, 2020.

[18] O. Alexander, M. Belisle, and J. Steele, "Mitre att&ck® for industrial control systems: Design and philosophy," 2020.

[19] S. P. Anderson, J. K. Goeree, and C. A. Holt, "The logit equilibrium: A perspective on intuitive behavioral anomalies," *Southern Economic Journal*, pp. 21–47, 2002.

[20] 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.

[21] J. Slowik, "Anatomy of an attack: Detecting and defeating crashoverride," *VB2018, October*, 2018.

[22] K. Hasan, S. Shetty, S. Ullah, A. Hassanzadeh, and E. Hadar, "Towards optimal cyber defense remediation in energy delivery systems," in *2019 IEEE Global Communications Conference (GLOBECOM)*. IEEE, 2019, pp. 1–7.

[23] X. Ou, S. Govindavajhala, and A. W. Appel, "Mulval: A logic-based network security analyzer." in *USENIX security symposium*, vol. 8. Baltimore, MD, 2005, pp. 113–128.

[24] Z. Zheng and A. Reddy, "Towards improving data validity of cyber-physical systems through path redundancy," in *Proceedings of the 3rd ACM Workshop on Cyber-Physical System Security*. ACM, 2017, pp. 91–102.