A Reinforcement Protection Strategy against an Adversary in the IoT

Andrey Garnaev[1] and Wade Trappe[2]

***Abstract*— The significant scale of the Internet of Things (IoT), in conjunction with its heterogeneous nature since it involves many device types, could lead the IoT to be exposed many security threats and attacks from adversaries. Game theory is a promising tool that can be used to design anti-adversary protection strategies for different network security scenarios. Traditionally, such anti-adversary strategies aim to minimize the expected number of infected/corrupted nodes. A con of such a traditional approach is that, due to limited resources available for defense, the IoT controller might have to sacrifice the protection of some nodes in order to better protect others, thus minimizing the expected number of infected nodes. This could lead to drastic consequences for the protected network as a whole since, by merely changing its behavior, the adversary might succeed with a higher probability of corrupting a few of the nodes, and thereby corrupt the network's operation as a whole although the total number of infected nodes is reduced. In this paper, motivated by this observation, we suggest using the protection level as a payoff for the IoT controller. Under the protection level metric, we specify a minimal probability of being non-infected for each of the IoT nodes. Finally, we suggest an approach how to combine a strategy that maximizes the protection level strategy with a strategy that minimizes the expected number of infected/corrupted nodes so as to increase the joint efficiency. We provide examples to numerically illustrate the derived strategies.**

***Index Terms*— Detection probability, Equilibrium, Fairness**

I. Introduction

Given the rapid deployment of the IoT as a new technology paradigm, the need for securing the IoT has become essential, particularly as the IoT involves safety-critical processes and the online management of sensitive data. Therefore, improving IoT device security has emerged as a leading priority for both manufacturers and researchers. We refer to [1], [2] for recent comprehensive surveys on IoT attacks, their taxonomy, detection mechanisms, solutions to deal with threats and emerging challenges. Driven by the understanding that a minimum of two agents are engaged in a security problem (for instance, the IoT controller and the adversary), each with distinct objectives, game theory has been employed in research literature to study security since it provides a framework and foundation for formulating the type of solutions that should be used in such multi-agent environments. In [3], the readers can find a comprehensive survey on the application of game theory in IoT. As examples of employing different concepts of game theory to modeling network protection, we also refer to [4]–[10].

It is worth noting that traditionally anti-adversary strategies aim to minimize the expected number of damaged/infected/corrupted nodes for different network scenarios. A drawback of such an approach is that, because of resource limitations, the IoT controller might have to sacrifice the protection of some less protected nodes in order to reinforce others with the goal of minimizing the expected number of infected nodes. This might lead to a drastic consequence for the network as a whole since the adversary might adjust its behavior in response and have a higher probability of corrupting some of the nodes. Thus, the network as a whole becomes compromised even though the total number of infected nodes is reduced. In this paper, motivated by this observation, we suggest using the protection level as a payoff to the IoT controller. Under the protection level, we understand the minimal probability of being infected for each of the nodes. Finally, we suggest a way to combine both protection protocols (maximizing the protection level and minimizing the expected number of infected/corrupted nodes) to combine their pros and increase their joint efficiency and we numerically illustrate the derived strategies. To the best knowledge of the authors such a problem has not been studied in the literature.

II. A brief overview of the network model

In the paper, we consider an IoT system consisting of a set of IoT nodes (devices) located in a (protected) zone, connected to each other for communicating and sharing data, and facing a (malicious) adversary attack aimed to infect its nodes. In this paper, we abstract the network, avoiding the specification of any particular topology, and instead consider merely a set $\mathcal{N}$ of nodes located in the protected zone. This set consists of a finite number (say, n) of nodes, which are identified by their number, i.e., $\mathcal{N} = \{1, \dots, n\}$. The nodes are under attack by an adversary attempting to intrude on the protected zone in order to perform a damaging action (e.g. to infect the nodes). To execute an intrusion, the adversary possesses certain resources, such as a set of devices targeting the IoT network. The total adversary's resource is $\overline{y}$. To reinforce the security of the network, the IoT controller is equipped with various defense resources, such as the time allocated

[1]Andrey Garnaev is with WINLAB, Rutgers University, North Brunswick, NJ garnaev@yahoo.com

[2]Wade Trappe is with WINLAB, Rutgers University, North Brunswick, NJ trappe@winlab.rutgers.edu

for remote scanning and device attestation. The total IoT controller resource is $\overline{x}$. Let x_i be the reinforcement effort the IoT controller applies to protect node i, and y_i be the resource applied by the adversary to infect/corrupt node i. Thus, the set of feasible strategies for the IoT controller is $\mathcal{X} \triangleq \{\boldsymbol{x} = (x_1, \ldots, x_n) : x_i \geq 0, i \in \mathcal{N}, \sum_{i\in\mathcal{N}} x_i = \overline{x}\}$. Similarly, the set of feasible strategies for the adversary is $\mathcal{Y} \triangleq \{\boldsymbol{y} = (y_1, \ldots, y_n) : y_i \geq 0, i \in \mathcal{N}, \sum_{i\in\mathcal{N}} y_i = \overline{y}\}$.

Let $P_i(x_i, y_i)$ be the probability of a successful infection/corruption of node i of the network, if the protection effort x_i and intrusion effort y_i are employed. In this paper, we assume that this probability is given by exponential law of effort put into the attack and protection [11] as follows:

$$P_i(x_i, y_i) = \gamma_i e^{-\mu_i x_i} \left(1 - e^{-\lambda_i y_i}\right), \tag{1}$$

where γ_i is the maximal probability for successful infection/corruption of node i, and λ_i and μ_i are node's parameters which reflect the average rates of successful infection/corruption and protection reinforcement, respectively, as feedback to applied resources. Then, probability that node i is not infected is:

$$Q_i(x_i, y_i) = 1 - P_i(x_i, y_i). \tag{2}$$

Since $P_i(x_i, y_i)$ is the probability of a successful infection of node i, the sum of over i reflects the expected number of infected nodes, if the IoT controller and adversary apply strategies $\boldsymbol{x}$ and $\boldsymbol{y}$, respectively, i.e.:

$$E(\boldsymbol{x}, \boldsymbol{y}) = \sum_{i\in\mathcal{N}} P_i(x_i, y_i). \tag{3}$$

Traditionally, the IoT controller aims to minimize the expected number of infected nodes, reflecting expected total harm to the network. Thus, $E(\boldsymbol{x}, \boldsymbol{y})$ is considered as a cost function to the IoT controller. Meanwhile, $E(\boldsymbol{x}, \boldsymbol{y})$ is considered as a payoff to the adversary, since it aims to maximize the expected number of infected nodes. In such a traditional formulation, the problem is a two player (adversary and IoT controller) zero-sum game. We denote this game by Γ^N, and note it is studied in the literature for different scenarios. Because of resource limitations, the defense of the IoT might involve sacrificing the protection of some nodes so as to reinforce others in order to minimize the expected number of infected nodes. As the security of a network rapidly decreases with just a few nodes being compromised, such an approach can mean that the network faces drastic security consequences because the adversary can adapt its strategy to focus on corrupting just some of the nodes.

Motivated by this observation, we propose using a different metric, the protection level, as a payoff to the IoT controller. We specify the protection level as the minimal probability of being non-infected for each of the nodes:

$$Q(\boldsymbol{x}, \boldsymbol{y}) = \min\{Q_i(x_i, y_i) : i \in \mathcal{N}\}. \tag{4}$$

Thus, the protection level $Q(\boldsymbol{x}, \boldsymbol{y})$ reflects the minimal level of resistance of each node to the fixed adversary attack $\boldsymbol{y}$, in the response to applied protection efforts $\boldsymbol{x}$.

The protection level $Q(\boldsymbol{x}, \boldsymbol{y})$ is a payoff to the IoT controller, while the expected number of infected nodes $E(\boldsymbol{x}, \boldsymbol{y})$ is considered as the payoff to the adversary. Each player (IoT controller and adversary) wants to maximize their payoffs. We look for (Nash) equilibrium [12], i.e., for such strategies $\boldsymbol{x}$ and $\boldsymbol{y}$ of the IoT controller and adversary, which are the best response to each other. In other words, they are a solution of the best response equations:

$$\boldsymbol{x} = \operatorname{argmax}\{Q(\tilde{\boldsymbol{x}}, \boldsymbol{y}) : \tilde{\boldsymbol{x}} \in \mathcal{X}\}, \tag{5}$$

$$\boldsymbol{y} = \operatorname{argmax}\{E(\boldsymbol{x}, \tilde{\boldsymbol{y}}) : \tilde{\boldsymbol{y}} \in \mathcal{Y}\}. \tag{6}$$

We denote this non-zero-sum game by Γ^D.

III. Equilibrium and its uniqueness

In the following theorem, using a constructive approach, we derive equilibrium strategies as functions of two parameter Lagrange multiplier of the NLP (6) and the maxmin value of (5).

Theorem 1: In game Γ^D the equilibrium strategies $(x_1^D, \ldots, x_n^D)$ and $(y_1^D, \ldots, y_n^D)$ of the IoT controller and adversary are given by (7) and (8) below where θ is Lagrange multiplier of nonlinear programming (NLP) problem (6) and ν is the value of the maxmin problem (5):

$$\begin{aligned} x_i^D &= x_i^D(\theta, \nu) \\ &= \begin{cases} 0, & \theta + \lambda_i(1-\nu) > \lambda_i\gamma_i \\ \dfrac{1}{\mu_i} \ln\left(\dfrac{\lambda_i\gamma_i}{(1-\nu)\lambda_i + \theta}\right), & \theta + \lambda_i(1-\nu) \leq \lambda_i\gamma_i, \end{cases} \end{aligned} \tag{7}$$

$$\begin{aligned} y_i^D &= y_i^D(\theta, \omega) \\ &= \begin{cases} 0, & \lambda_i\gamma_i \leq \theta \\ \dfrac{1}{\lambda_i} \ln\left(\dfrac{\lambda_i\gamma_i}{\theta}\right), & \theta < \lambda_i\gamma_i \leq \theta + \lambda_i(1-\nu), \\ \dfrac{1}{\lambda_i} \ln\left(1 + \dfrac{\lambda_i(1-\nu)}{\theta}\right), & \theta + \lambda_i(1-\nu) < \lambda_i\gamma_i. \end{cases} \end{aligned} \tag{8}$$

For the proof, please see in the appendix.

In the following theorem we prove the uniqueness of the equilibrium.

Theorem 2: In game Γ^D the equilibrium strategies $(x_1^D, \ldots, x_n^D)$ and $(y_1^D, \ldots, y_n^D)$ of the IoT controller and adversary are unique and given by (7) and (8), respectively, where $\theta = \Theta^D(\nu)$ and $\nu = \nu_*$ with ν_* is the unique root in $(\overline{\nu}, 1)$ (for $\overline{\nu}$ please see (12) below) of the following equation:

$$Y^D(\Theta^D(\nu_*), \nu_*) = \overline{y} \tag{9}$$

and $\Theta^D(\nu)$ for each fixed $\nu \in (\overline{\nu}, 1)$ is the unique root in $(0, \max_i \lambda_i\gamma_i)$ of the following equation:

$$X^D(\Theta^D(\nu), \nu) = \overline{x}, \tag{10}$$

where

$$X^D(\theta,\nu) \triangleq \sum_{i\in\mathcal{N}} x_i^D(\theta,\nu), Y^D(\theta,\nu) \triangleq \sum_{i\in\mathcal{N}} y_i^D(\theta,\nu), \quad (11)$$

and $\overline{\nu}$ *is the unique root in* $(0,1)$ *of the following equation:*

$$X^D(0,\overline{\nu}) = \overline{x}. \quad (12)$$

Moreover, for each fixed $\nu \in (0,1)$ the root $\Theta(\nu)$ can be found via the bisection method, and this $\Theta(\nu)$ is increasing on ν. The function $Y^D(\Theta(\nu),\nu)$ is decreasing on ν, and, so, ν_* also can be found via the bisection method.

Proof of Theorem 2 please see the appendix.

IV. How to combine both protection strategies

Maximizing the protection level strategy and minimizing the expected number of infected nodes strategy is beneficial, which raises the question: *How to combine both these strategies to incorporate all their benefits in one protocol?*

First note that game Γ^N is a boundary case of the game studied in [13], and its equilibrium is unique and can be found via a superposition of two bisection methods. Denote by $(\boldsymbol{x}^N, \boldsymbol{y}^N)$ its unique equilibrium.

Since the payoff metrics of the IoT controller in games Γ^D and Γ^N differ from each other, to combine their equilibrium $(\boldsymbol{x}^D, \boldsymbol{y}^D)$ and $(\boldsymbol{x}^N, \boldsymbol{y}^N)$ we have to reduce both metrics to one without changing their equilibration strategies. Note that $F^N(\boldsymbol{x},\boldsymbol{y}) \triangleq (n - E(\boldsymbol{x},\boldsymbol{y}))/C^N$, with C^N is a normalizing coefficient

$$C^N = \max\{Q(\boldsymbol{x}^N, \boldsymbol{y}^D), Q(\boldsymbol{x}^N, \boldsymbol{y}^N)\}, \quad (13)$$

can be interpreted as the normalized expected number of non-infected nodes. Moreover, the game Γ^N with payoff $F^N(\boldsymbol{x},\boldsymbol{y})$ to the IoT controller instead of $-E(\boldsymbol{x},\boldsymbol{y})$ are equivalent to each other since they have the same equilibrium $(\boldsymbol{x}^N, \boldsymbol{q}^N)$. Further, Γ^D with payoff $F^D(\boldsymbol{x},\boldsymbol{y}) \triangleq Q(\boldsymbol{x},\boldsymbol{y})/C^D$ with C^D is a normalizing coefficient

$$C^D = \max\{n - E(\boldsymbol{x}^D, \boldsymbol{y}^D), n - E(\boldsymbol{x}^D, \boldsymbol{y}^N)\} \quad (14)$$

to the IoT controller instead of $Q(\boldsymbol{x},\boldsymbol{y})$ are equivalent to each other in the sense that they have the same equilibrium $(\boldsymbol{x}^D, \boldsymbol{q}^D)$.

Moreover, due to normalization, the metrics for the IoT controller payoffs coincide, and, we can combine them into a payoff matrix (see below (15)). To do so, we assume that the IoT controller could apply one of two modes (mode N, minimizing the expected number of infected nodes and mode D, maximizing the protection level). The adversary also can apply one of two modes (mode N with the best response strategy minimizing the expected number of infected nodes and mode D with the best response to the strategy maximizing the protection level). The players choose their modes independently of each other, and each player does not know which mode is chosen by the other. Thus, in mode N and mode D, the adversary (IoT controller) applies strategies $\boldsymbol{y}^N$ and $\boldsymbol{y}^D$ ($\boldsymbol{x}^N$ and $\boldsymbol{x}^D$). Let p_D and $p_N = 1 - p_D$ be the frequency (probability) for the IoT controller to apply mode D and mode N, respectively. Let q_D and $q_N = 1 - q_D$ be the probability for the adversary to apply mode D and mode N, respectively. Thus, $\boldsymbol{p}^T = (p_D, p_N)$ and $\boldsymbol{q}^T = (q_D, q_N)$ are randomized (mixed) strategies for the IoT controller and the adversary, respectively. Thus, each player has the same set of feasible randomized (mixed) strategies denoted by $\mathcal{P}$. This allows us to combine both IoT controller and adversary modes in the following payoff bi-matrix, in which rows are the IoT controller's strategies and columns are the adversary' s strategies:

$$(A,B) = \begin{matrix} \\ \text{Mode } D: \\ \text{Mode } N: \end{matrix} \begin{matrix} \text{Mode } D: & \text{Mode } N: \\ \end{matrix} \begin{pmatrix} (a_{DD}, b_{DD}), & (a_{DN}, b_{DN}) \\ (a_{ND}, b_{ND}) & (a_{NN}, b_{NN}) \end{pmatrix} \quad (15)$$

with $a_{DD} \triangleq F^D(\boldsymbol{x}^D, \boldsymbol{y}^D)$, $a_{DN} \triangleq F^D(\boldsymbol{x}^D, \boldsymbol{y}^N)$, $a_{ND} \triangleq F^N(\boldsymbol{x}^N, \boldsymbol{y}^D)$, $a_{NN} \triangleq F^N(\boldsymbol{x}^N, \boldsymbol{y}^N)$, $b_{DD} \triangleq E(\boldsymbol{x}^D, \boldsymbol{y}^D)$, $b_{DN} \triangleq E(\boldsymbol{x}^D, \boldsymbol{y}^N)$, $b_{ND} \triangleq E(\boldsymbol{x}^N, \boldsymbol{y}^D)$ and $b_{NN} \triangleq E(\boldsymbol{x}^N, \boldsymbol{y}^N)$. Note that, by (6), we have that

$$b_{NN} > b_{ND} \text{ and } b_{DD} > b_{DN}. \quad (16)$$

The IoT aims to maximize the fairness of allocating the expected playoffs $(a_{DD}q_D + a_{DN}q_N)p_D$ and $(a_{ND}q_D + a_{DN}q_N)p_N$ with respect to the applied modes. Here, as a utility we consider α-fairness criteria (please see for example [14], [15]). Then, the payoff to the IoT if the IoT controller and adversary apply strategies $\boldsymbol{p}$ and $\boldsymbol{q}$, respectively, is

$$\begin{aligned} V_C(\boldsymbol{p},\boldsymbol{q}) = {} & \Psi_\alpha((a_{DD}q_D + a_{DN}q_N)p_D) \\ & + \Psi_\alpha((a_{ND}q_D + a_{DN}q_N)p_N) \end{aligned} \quad (17)$$

with $\Psi_\alpha(t)$ is α-fairness utility given as follows:

$$\Psi_\alpha(t) = \begin{cases} t^{1-\alpha}/(1-\alpha), & \alpha \neq 1, \\ \ln(t), & \alpha = 1. \end{cases} \quad (18)$$

The adversary payoff is given as follows:

$$V_A(\boldsymbol{p},\boldsymbol{q}) = \boldsymbol{p}^T B \boldsymbol{q}. \quad (19)$$

We look for Nash equilibrium, i.e., for such strategies $\boldsymbol{p}$ and $\boldsymbol{q}$ which are a solution of the best response equations:

$$\boldsymbol{p} = \text{argmax}\{V_C(\tilde{\boldsymbol{p}}, \boldsymbol{q}) : \tilde{\boldsymbol{p}} \in \mathcal{P}\}, \quad (20)$$

$$\boldsymbol{q} = \text{argmax}\{V_A(\boldsymbol{p}, \tilde{\boldsymbol{q}}) : \tilde{\boldsymbol{q}} \in \mathcal{P}\}. \quad (21)$$

Denote this game by Γ_α. For $\alpha = 0$, the game Γ_0 is a classical bi-matrix game [12]. To solve the game for $\alpha > 0$ first let us introduce an auxiliary function $\Phi_\alpha(t)$ and its monotonous properties in the following lemma:

Lemma 1: Let

$$\Phi_\alpha(t) \triangleq \frac{1}{1 + \dfrac{(a_{ND}t + a_{NN}(1-t))^{(1-\alpha)/\alpha}}{(a_{DD}q_D + a_{DN}(1-t))^{(1-\alpha)/\alpha}}}. \quad (22)$$

Then $\Phi_1(t) = 1/2$ *for* $\alpha = 1$, *meanwhile for* $\alpha \neq 1$*:*

(a) if $a_{DN}a_{ND} > a_{DD}a_{NN}$ *then* $\Phi_\alpha(t)$ *is decreasing in* $[0,1]$ *if* $\alpha \in (0,1)$*, and it is increasing in* $[0,1]$ *if* $\alpha > 1$;

(b) if $a_{DN}a_{ND} < a_{DD}a_{NN}$ *then* $\Phi_\alpha(t)$ *is increasing in* $[0,1]$ *if* $\alpha \in (0,1)$*, and it is decreasing in* $[0,1]$ *if* $\alpha > 1$.

Proof please see in the appendix.

THEOREM 3: *In game* Γ_α *equilibrium* $(\boldsymbol{p},\boldsymbol{q}) = ((p_D,p_N),(q_D,q_N))$ *exists. For* $\alpha \neq 1$ *it is an unique except for the case (b) below. Moreover* $p_D = \Phi_\alpha(q_D)$ *and*

(a) if $\Phi_\alpha(0) > \Phi_\alpha(1)$ *then*

$$q_D = \begin{cases} 0, & \Phi_\alpha(0) \le p_*, \\ \Phi_\alpha^{-1}(p_*), & \Phi_\alpha(1) < p_* < \Phi_\alpha(0), \\ 1, & p_* \le \Phi_\alpha(1) \end{cases} \tag{23}$$

with

$$p_* \triangleq \frac{b_{NN} - b_{ND}}{b_{NN} + b_{DD} - b_{ND} - b_{DN}}, \tag{24}$$

(b) if $\Phi_\alpha(0) < \Phi_\alpha(1)$ *then*

$$q_D = \begin{cases} 0, & \Phi_\alpha(1) < p_*, \\ \in \{0, 1, \Phi_\alpha^{-1}(p_*)\}, & \Phi_\alpha(0) < p_* < \Phi_\alpha(1), \\ 1, & p_* < \Phi_\alpha(0). \end{cases} \tag{25}$$

If $\alpha = 1$, then the IoT equilibrium strategy is fifty-fifty, i.e., $p_D = 1/2$ and it is indifferent to the network's parameter. The adversary's strategy is $q_D = 0$ if $1/2 < p_*$, $q_D = 1$ if $p_* < 1/2$ and is any probability otherwise.

For proof of Theorem 3 please see in the appendix.

V. Numerical illustrations and discussions

Let us consider an example of the network to showcase how the equilibrium strategies derived in Theorem 1 and Theorem 2 are influenced by the adversary's and IoT controller's resources. This particular example involves a network consisting of $N = 4$ nodes, and the maximal detection probabilities $\gamma = (0.9, 0.4, 0.7, 0.6)$ and network's parameters $\lambda = (0.09, 0.8, 0.07, 0.6])$ and $\mu = (0.1, 0.3, 0.2, 0.6)$, the adversary's total resource $\overline{y}$ varies from 1 to 10 and protection resource $\overline{x}$ is 1 or 10. Fig. 1(a) and Fig. 1(b) illustrate that an increase in the adversary's resource or a decrease in the IoT controller resources leads to an increase in the expected number of infected nodes and a decrease in protection level. The protection level is higher in Γ^D than in Γ^N, meanwhile, the expected number of infected nodes is less in Γ^N than in Γ^D. Fig 1(c) and Fig. 1(d) illustrate the dependence of players' strategies on the adversary's resource for the IoT controller resource $\overline{x} = 10$.

Fig 1(e) illustrates switching strategies (Theorem 3) of the players between mode D and mode N for two boundary cases of players' resources $(\overline{x},\overline{y}) = (1,1)$ and $(\overline{x},\overline{y}) = (10,10)$. It illustrates that, for any network parameters and fairness coefficient, the IoT controller always applies a randomized strategy, i.e., $0 < p < 1$, beginning from fifty-fifty strategy $p = 1/2$ in contrast to the adversary's strategy q that might be non-random (specifically, $q = 0$). Theorem 3 gives us a continuum

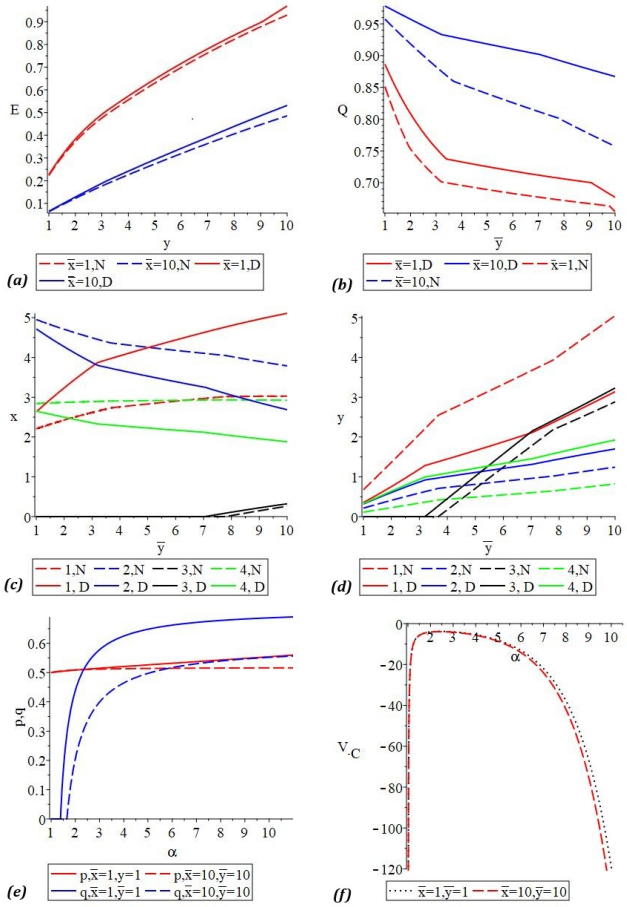

Fig. 1. (a) Expected number of infected nodes, (b) protection level, (c) IoT controller strategies in Γ^N and Γ^D, (d) adversary's strategies in Γ^N and Γ^D, (e) strategies of switching between modes in Γ_α and (f) α-fairness utility for IoT controller.

of fair strategies (a strategy per a fairness coefficient) of applying modes D and N. Fig 1(f), reflecting α-fairness utility for the IoT controller, allows us to find the fairness coefficient corresponding to the most fair solution. Such a solution is given by α where the IoT controller payoff achieves its maximum. In the considered example, it is $\alpha = 2.24$ for $(\overline{x},\overline{y}) = (1,1)$ with equilibrium $(p,q) = (0.511, 0.490)$ and $\alpha = 2.41$ with equilibrium $(p,q) = (0.517, 0.321)$ for $(\overline{x},\overline{y}) = (10,10)$.

VI. Conclusions

Motivated by an observation that traditional anti-adversary strategies, which aim to minimize the expected number of infected/corrupted nodes, might sacrifice the protection of some nodes to reinforce other nodes so as to minimize the expected number of infected nodes, this paper aims to deal with the drawbacks of such an approach. We have proposed using the protection level as a payoff to the IoT controller. Under the protection level, we understand the minimal probability of being non-infected for each of the nodes. The anti-adversary strategy has been designed in the framework of game theory. Finally, we have developed an approach that combines both pro-

tection protocols (maximizing the protection level and minimizing the expected number of infected/corrupted nodes) to combine their advantages and increases their joint efficiency. A goal of our future work is to generalize the suggested one-step network protection model to a multi-step Bayesian learning network protection model.

Appendix

Proof of Theorem 1: First, in Proposition 1 and Proposition 2 below, we derive adversary's best response strategy and IoT controller's best response strategy, respectively.

Proposition 1: Adversary's strategy $\boldsymbol{y}$ is the best response to IoT controller's strategy $\boldsymbol{x}$ if and only if:

$$y_i = \begin{cases} \ln\left(\lambda_i\gamma_i e^{-\mu_i x_i}/\theta\right)/\lambda_i, & \lambda_i\gamma_i e^{-\mu_i x_i} > \theta, \\ 0, & \lambda_i\gamma_i e^{-\mu_i x_i} \le \theta, \end{cases} \tag{26}$$

where θ is Lagrange multiplier.

Proof: Since (6) is a concave optimization problem to derive the adversary's best response we introduce its Lagrangian $\mathcal{L}_{\boldsymbol{x},\theta}(\boldsymbol{y}) \triangleq E(\boldsymbol{x},\boldsymbol{y}) + \theta\left(\overline{y} - \sum_{i\in\mathcal{N}} y_i\right)$, where θ is Lagrange multiplier. Then, adversary's strategy $\boldsymbol{y} \in \mathcal{Y}$ is the best response to IoT controller's strategy $\boldsymbol{x}$ if and only if the following condition holds:

$$\frac{\partial\mathcal{L}_{\boldsymbol{x},\theta}(\boldsymbol{y})}{\partial y_i} = \lambda_i\gamma_i e^{-\mu_i x_i - \lambda_i y_i} - \theta \begin{cases} = 0, & y_i > 0, \\ \le 0, & y_i = 0, \end{cases} \tag{27}$$

and Proposition 1 follows. ∎

Proposition 2: The IoT controller's strategy $\boldsymbol{x}$ is the best response to an adversary's strategy $\boldsymbol{y}$ if and only for a positive ν following relations hold:
(A) if $y_i = 0$ then $x_i = 0$;
(B) if $y_i > 0$ and $x_i > 0$ then $Q_i(x_i, y_i) = \nu$;
(C) if $y_i > 0$ and $x_i = 0$ then $Q_i(0, y_i) \ge \nu$.

Proof: First note for $Q_i(x_i, y_i)$ we have that:
(P-I) $Q_i(x_i, 0) = 1$ *for any* $x_i \in [0, 1]$;
(P-II) *For each fixed $y_i > 0$ function $Q_i(x_i, y_i)$ is strictly increasing on x_i, and* $0 < Q_i(x_i, y_i) < 1$, $x_i \in [0, 1]$.

(a) Let $y_i = 0$ and assume that $x_i > 0$. Since $\boldsymbol{y} \in \mathcal{Y}$ there is a $j \ne i$ such that $y_j > 0$. Then $Q_j(x_j, y_j) < 1$. Let $\boldsymbol{x}^\epsilon = (x_1^\epsilon, \ldots, x_n^\epsilon)$ be such that

$$x_k^\epsilon = \begin{cases} x_i - \epsilon, & k = i, \\ x_j + \epsilon, & k = j, \\ x_k, & k \notin \{i, j\}. \end{cases} \tag{28}$$

Since $x_i > 0$ we have that $\boldsymbol{x}^\epsilon \in \mathcal{X}$ for enough small positive ϵ, and, by (P-I) and (P-II), we have that:

$$Q_k(x_k^\epsilon, y_k) = \begin{cases} Q_i(x_i - \epsilon, y_i) = 1, & k = i, \\ Q_j(x_j + \epsilon, y_j) > Q_j(x_j, y_j), & k = j, \\ Q_k(x_k, y_k), & k \notin \{i, j\}. \end{cases} \tag{29}$$

Thus, $\boldsymbol{x}$ cannot be the best response to $\boldsymbol{y}$. This contradiction implies that if $y_i = 0$ then $x_i = 0$, and (A) follows.

(b) Let $y_i y_j > 0$ and $x_i x_j > 0$. First let us prove than

$$Q_i(x_i, y_i) = Q_j(x_j, y_j). \tag{30}$$

Assume that $Q_i(x_i, y_i) \ne Q_i(x_j, y_j)$. Without loss of generality we can assume that

$$Q_i(x_i, y_i) < Q_j(x_j, y_j). \tag{31}$$

Then $\boldsymbol{x}^\epsilon \in \mathcal{X}$ for $\boldsymbol{x}^\epsilon$ given by (28) and enough small positive ϵ. Thus, by (P-II), we have that:

$$Q_k(x_k^\epsilon, y_k) \begin{cases} > Q_i(x_i, y_i), & k = i, \\ < Q_j(x_j, y_j), & k = j, \\ = Q_k(x_k, y_k), & k \notin \{i, j\}. \end{cases} \tag{32}$$

This and (31) imply that $\boldsymbol{x}$ cannot be the best response to $\boldsymbol{y}$. This contradiction implies (30), i.e., $Q_i(x_i, y_i) = \nu$ for i such that $x_i > 0$ and $y_i > 0$, and (B) follows.

(c) Let $y_i y_j > 0$ and $x_i = 0$ and $x_j > 0$. Assume that

$$Q_i(0, y_i) < \nu = Q_j(x_j, y_j). \tag{33}$$

Let $\boldsymbol{x}^\epsilon = (x_1^\epsilon, \ldots, x_n^\epsilon)$ be such that

$$x_k^\epsilon = \begin{cases} \epsilon, & k = i, \\ x_j - \epsilon, & k = j, \\ x_k, & k \notin \{i, j\}. \end{cases} \tag{34}$$

Since $x_j > 0$ we have that $\boldsymbol{x}^\epsilon \in \mathcal{X}$ for enough small positive ϵ, and, by (P-I) and (P-II), we have that

$$Q_k(x_k^\epsilon, y_k) \begin{cases} > Q_i(0, y_i), & k = i, \\ < Q_j(x_j, y_j), & k = j, \\ = Q_k(x_k, y_k), & k \notin \{i, j\}. \end{cases} \tag{35}$$

This and (33) imply that $\boldsymbol{x}$ cannot be the best response to $\boldsymbol{y}$. This contradiction implies that (33) cannot hold and (C) follows. ∎

Now we prove Theorem 1. Let us separately consider two cases: (I) $y_i = 0$ and (II) $y_i > 0$.

(I) Let $y_i = 0$. Then, by Proposition 2, $x_i = 0$. Substituting such $x_i = 0$ and $y_i = 0$ into Proposition 1 implies

$$\lambda_i\gamma_i \le \theta, \tag{36}$$

and the first raw of (8) follows.

(II) Let $y_i > 0$. Then two cases arise to consider separately: (II-A) $x_i = 0$ and (II-B) $x_i > 0$.

(II-A) Let $x_i = 0$. Substituting such $x_i = 0$ and $y_i > 0$ into Proposition 1 and Proposition 2 imply $\lambda_i\gamma_i e^{-\lambda_i y_i} = \theta$ and $\gamma_i\left(1 - e^{-\lambda_i y_i}\right) \le 1 - \nu$. Thus, $y_i = \frac{1}{\lambda_i}\ln\left(\frac{\lambda_i\gamma_i}{\theta}\right)$ and $\theta < \lambda_i\gamma_i \le \theta + \lambda_i(1 - \nu)$. This implies the second raw of (8), and jointly with (I) the first raw of (7) also follows.

(II-B) Let $y_i > 0$ and $x_i > 0$. Substituting such $x_i > 0$ and $y_i > 0$ into Proposition 1 and Proposition 2 imply $\lambda_i\gamma_i e^{-\lambda_i y_i - \mu_i x_i} = \theta$ and $\gamma_i e^{-\mu_i x_i}\left(1 - e^{-\lambda_i y_i}\right) = 1 - \nu$. Thus, $x_i = \frac{1}{\mu_i}\ln\left(\frac{\lambda_i\gamma_i}{(1-\nu)\lambda_i + \theta}\right)$ and $y_i = \frac{1}{\lambda_i}\ln\left(1 + \frac{\lambda_i(1-\nu)}{\theta}\right)$ and $\theta + \lambda_i(1-\nu) < \lambda_i\gamma_i$. This implies the second raw of (7), and the third raw of (8) follows ∎

.

Proof of Theorem 2: By Theorem 1, equilibrium strategies have to be given by (7) and (8), where θ and ν are given by the following conditions:

$$X^D(\theta,\nu)=\overline{x} \text{ and } Y^D(\theta,\nu)=\overline{y} \tag{37}$$

with $X^D(\theta,\nu)$ and $Y^D(\theta,\nu)$ given by (11). By (7) and (11), $X^D(\theta,\nu)$ is decreasing on θ and increasing on ν. Thus, for each fixed $\nu\in(0,1)$ there is a unique $\Theta(\nu)$ given by (10). This $\Theta(\nu)$ can be found via the bisection method, and function $\Theta(\nu)$ is increasing on ν. This, (8) and (11) imply that function $Y^D(\Theta(\nu),\nu)$ is decreasing on ν, and, so, ν_* uniquely given by (9) via the bisection method.

Moreover, for each fixed $\nu\in(0,1)$ $\Theta(\nu)$ can be found via the bisection method, and this $\Theta(\nu)$ is increasing on ν. Function $Y^D(\Theta(\nu),\nu)$ is decreasing on ν, and, so, ν_* also can be found via the bisection method.

Proof of Lemma 1: Let $\phi(t)=(a_{ND}t+a_{NN}(1-t))/(a_{DD}q_D+a_{DN}(1-t))$. Then $d\phi(t)/dt=(a_{ND}a_{ND}-a_{DD}a_{NN})/(a_{DD}q_D+a_{DN}(1-t))^2$. This and (22) imply the result. ■

Proof of Theorem 3: First note that since $V_C(\boldsymbol{p},\boldsymbol{q})$ is concave on $\boldsymbol{p}$ and $V_A(\boldsymbol{p},\boldsymbol{q})$ is linear on $\boldsymbol{q}$, the game Γ_α has at least one equilibrium. We find these equilibrium via solving the best response equations. To find the best response $\boldsymbol{p}$ to $\boldsymbol{q}$ we have to maximize $V_C(\boldsymbol{p},\boldsymbol{q})$ with $\boldsymbol{p}=(p_D,p_N)$ such that $p_D+p_N=1$, $p_N\geq 0$ and $p_D\geq 0$. This problem is a concave optimization problem and to solve it we have to introduce its Lagrangian $\mathcal{L}_{\boldsymbol{p},\omega}(\boldsymbol{q})\triangleq V_\alpha(\boldsymbol{p},\boldsymbol{q})+\omega\,(1-p_D-p_N)$ with ω is its Lagrange multiplier.

Then, IoT controller strategy $\boldsymbol{p}=(p_D,p_N)$ is the best response to adversary strategy $\boldsymbol{q}=(q_D,q_N)$ if and only if the following relations hold for $\tau\in\{D,N\}$:

$$\frac{\partial\mathcal{L}_{\boldsymbol{p},\omega}(\boldsymbol{q})}{\partial p_\tau}=\frac{(a_{\tau D}q_D+a_{\tau N}q_N)^{1-\alpha}}{p_\tau^\alpha}-\omega\begin{cases}=0, & p_\tau>0,\\ \leq 0, & p_\tau=0,\end{cases} \tag{38}$$

Since $\boldsymbol{p}=(p_D,p_N)\in\mathcal{P}$, by (38), $p_D>0$ and $p_N>0$. Thus, first row of (38) has to hold for both $\tau=D$ and $\tau=N$. Since $p_D+p_N=1$, such two equations given by first rows of (38) with $\tau\in\{D,N\}$ imply

$$p_\tau=\frac{(a_{\tau D}q_D+a_{\tau N}q_N)^{(1-\alpha)/\alpha}}{\sum_{\kappa\in\{D,N\}}(a_{\kappa D}q_D+a_{\kappa N}q_N)^{(1-\alpha)/\alpha}},\tau\in\{D,N\}. \tag{39}$$

Substituting $q_N=1-q_D$ into (39) implies

$$p_D=\Phi_\alpha(q_D) \tag{40}$$

with $\Phi_\alpha(\cdot)$ given by (22). Since $V_A(\boldsymbol{p},\boldsymbol{q})$ is linear on $\boldsymbol{q}$, by (15) and (19), we have that adversary strategy $\boldsymbol{q}=(q_D,q_N)$ is the best response to IoT controller strategy $\boldsymbol{p}=(p_D,p_N)$ if and only if

$$q_D\begin{cases}=0, & b_{DD}p_D+b_{ND}p_N<b_{DN}p_D+b_{NN}p_N,\\ \in[0,1], & b_{DD}p_D+b_{ND}p_N=b_{DN}p_D+b_{NN}p_N,\\ =1, & b_{DD}p_D+b_{ND}p_N>b_{DN}p_D+b_{NN}p_N.\end{cases} \tag{41}$$

Substituting $p_N=1-p_D$ into (41) implies:

$$q_D\begin{cases}=0, & p_D<p_*,\\ \in[0,1], & p_D=p_*,\\ =1, & p_D>p_*\end{cases} \tag{42}$$

with p_* given by (24). By (16), $0<p_*<1$. Substituting (42) into (40) and solving the obtained fixed point equation on x_D, by Lemma 1, implies the result. ■

References

[1] X. Liang and Y. Kim, "A survey on security attacks and solutions in the iot network," in *Proc. IEEE 11th Annual Computing and Communication Workshop and Conference (CCWC)*, pp. 0853–0859, 2021.

[2] T. Sasi, A. H. Lashkari, R. Lu, P. Xiong, and S. Iqbal, "A comprehensive survey on iot attacks: Taxonomy, detection mechanisms and challenges," *Journal of Information and Intelligence*, vol. 2, no. 6, pp. 455–513, 2024.

[3] C. Chi, Y. Wang, X. Tong, , M. Siddula, and Z. Cai, "Game theory in internet of things: A survey," *IEEE Internet of Things Journal*, vol. 9, no. 14, pp. 12125–12146, 2022.

[4] N. Namvar, W. Saad, N. Bahadori, and B. Kelleys, "Jamming in the internet of things: A game-theoretic perspective," in *Proc. IEEE Global Communications Conference (GLOBECOM)*, 2016.

[5] Q.D. La, T.Q.S. Quek, J. Lee, S. Jin, and H. Zhu, "Deceptive attack and defense game in honeypot-enabled networks for the internet of things," *IEEE Internet of Things Journal*, vol. 3, pp. 1025–1035, 2016.

[6] G. Margelis, R. Piechocki, T. Tryfonas, and P. Thomas, "Smart attacks on the integrity of the internet of things: Avoiding detection by employing game theory," in *Proc. IEEE Global Communications Conference (GLOBECOM)*, 2016.

[7] S.B.H Shah, L. Wang, P. Reddy, and A. Carie, "Non-cooperative game to balance energy and security in resource constrained iot networks," in *Proc. IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, pp. 502–507, 2020.

[8] A. Garnaev and W. Trappe, "A bandwidth monitoring strategy under uncertainty of the adversary's activity," *IEEE Transactions on Information Forensics and Security*, vol. 11, pp. 837–849, 2016.

[9] X. Zhang, "Access control mechanism based on game theory in the internet of things environment," in *Proc. IEEE 8th International Conference on Computer and Communications (ICCC)*, pp. 1–6, 2022.

[10] S. Surekha and M.Z.U.. Rahman, "Spectrum sensing and allocation strategy for iot devices using continuous-time markov chain-based game theory model," *IEEE Sensors Letters*, vol. 6, no. 4, pp. 1–4, 2022.

[11] B.O. Koopman, *Search and screening: general principles with historical applications*. Pergamon Press, 1980.

[12] D. Fudenberg and J. Tirole, *Game theory*. Boston, MA: MIT Press, 1991.

[13] V.J. Baston and A.Y. Garnaev, "A search game with a protector," *Naval Research Logistics*, vol. 47, pp. 85–96, 2000.

[14] E. Altman, K. Avrachenkov, and A. Garnaev, "Fair resource allocation in wireless networks in the presence of a jammer," *Performance Evaluation*, vol. 67, pp. 338–349, 2010.

[15] A. Garnaev and W. Trappe, "Bargaining over the fair trade-off between secrecy and throughput in OFDM communications," *IEEE Transactions on Information Forensics and Security*, vol. 12, pp. 242–251, 2017.