sensors

MDPI

Article

Stackelberg Dynamic Game-Based Resource Allocation in Threat Defense for Internet of Things

Bingjie Liu, Haitao Xu * and Xianwei Zhou

School of Computer and Communication Engineering, University of Science and Technology Beijing, Beijing 100083, China; b20160290@xs.ustb.edu.cn (B.L.); xwzhouli@sina.com (X.Z.)

* Correspondence: alex_xuht@hotmail.com; Tel.: +86-10-61647796

Received: 9 November 2018; Accepted: 19 November 2018; Published: 21 November 2018

Abstract: With the rapid development of the Internet of Things, there are a series of security problems faced by the IoT devices. As the IoT devices are generally devices with limited resources, how to effectively allocate the restricted resources facing the security problems is the key issue at present. In this paper, we study the resource allocation problem in threat defense for the resource-constrained IoT system, and propose a Stackelberg dynamic game model to get the optimal allocated resources for both the defender and attackers. The proposed Stackelberg dynamic game model is composed by one defender and many attackers. Given the objective functions of the defender and attackers, we analyze both the open-loop Nash equilibrium and feedback Nash equilibrium for the defender and attackers. Then both the defender and attackers can control their available resources based on the Nash equilibrium solutions of the dynamic game. Numerical simulation results show that correctness and effeteness of the proposed model.

Keywords: resource allocation; threat defense; Internet of Things; Stackelberg dynamic game; Nash equilibrium

1. Introduction

The Internet of Things (IoT) [1] refers to a huge network of various information-sensing devices combined with the Internet. These sensing devices include infrared sensors, Radio Frequency Identification Devices (RFID) [2], laser scanners, global positioning systems (GPS), and other devices. In recent years, with the development of computer intelligence technology, communication technology and perceptual recognition technology, the IoT has been widely used in smart home, smart medical, smart grid, Intelligent Transportation System (ITS) and other fields, and brought great convenience to people's lives [3–5].

Generally, the IoT system is composed of a large number of nodes that are often exposed to public situations, lack effective protection, and are easily attacked [6–8]. Therefore, the security threats faced by IoT devices are more serious than those of the traditional network. In addition, the IoT environment is complex and IoT devices with limited resources are more vulnerable to cyber-attack [9]. Faced with limited resources, how to effectively allocate resources [10] to defend against threats in the IoT system has become a serious problem that desperately needs to be solved.

Lots of work has been done on resource allocation problems in threat defense for IoT systems. In order to build up the overall security of the IoT, studies [11,12] propose an overall security architecture for the IoT from different perspectives. In order to promote multiple resource-sharing and heterogeneous resource-demanding allocations, Intrusion Detection Systems (IDS) architecture and resource allocation are recommended [13]. Zhang et al. [14] evaluated the four levels of a security index system of the Internet of Things through fuzzy analytic hierarchy process, and concluded that the key indicators for improving the security of the Internet of Things are privacy protection, WSN

anti-attack capability, intelligent node security, and information application security. Leusse et al. [15] analyzed the security threats of the IoT, and proposed a self-organized community security structure.

Since most IoT devices are micro-embedded devices, their hardware and software resources are very limited, and only a small number of computing tasks can be performed. There are not enough resources to implement defense against the attacks. For most of the time, the devices of the IoT system are unprotected and the resources of each device are limited. The authors of [16] developed a distributed algorithm to detect anomalous activities in the information flow in a wireless sensor network-based IoT system. Considering the limited computational and communication resources, Eschenauer et al. [17] proposed a key management scheme for the wireless sensor networks, which relies on probabilistic key sharing among the nodes of a random graph to make a trade-off between sensor memory cost and connectivity. The research into attack defense strategy plays a crucial role in defending against malicious attacks and protecting the security of the Internet of Things. Tague et al. [18] mapped the data traffic of the captured wireless mobile network and calculated the minimum cost of node capture attacks using a key protocol to enhance the security of data privacy. The literature [19–21] mainly focused on the research of node capture attacks on RFID and WSN in the IoT perception layer. Liu et al. [22] proposed a dynamic defense framework for IoT security. The literature [23] studies the optimal invasive differential game theory, which effectively reduces the intrusion frequency of intruders. Zhang et al. [24] proposed a lightweight defense algorithm for IoT network environment attacks.

Game theory is a particularly effective mathematical tool to study problems in diverse networks, such as power control in wireless networks [25,26], channel allocation in cognitive radio networks [27,28], congestion control in telecommunications networks [29], marketing and economics, and security problems [30,31]. In this paper, we propose Stackelberg dynamic game-based resource allocation model in threat defense for a resource-constrained IoT system. We will try to find optimal solutions for both the defender and attackers for resource allocation problems. In summary, the key contributions of this paper are as follows:

- Firstly, we research a cyber-security IoT system, which consists of one defender and N attackers. The defender tries to find the optimal allocated resources for threat defense, and the attackers try to use their resources to attack the IoT system.
- Secondly, a Stackelberg dynamic game model is proposed to formulate the resource allocation problem in threat defense for the Internet of Things. The Stackelberg game is a one-leader-many-followers game, where the defender is the leader and the attackers are the followers.
- For the dynamic game, we use the risk level as the system state. The objectives for the defender and attackers are to optimize the cost during the threat defense to find the optimal allocated resources for both the defender and attackers.
- Finally, the open-loop control solutions and the feedback control solutions for both the defender and attackers are given based on Bellman dynamic programming.

The remainder of the paper is organized as follows. Section 2 introduces the system model and problem formulation. Section 3 provides the Nash equilibrium solutions for the proposed game model. Numerical simulations are given in Section 4. Finally, we conclude the work in Section 5.

2. System Model and Problem Formulation

We consider a cyber-security IoT system that is composed of one defender and N attackers. The defender can control its resources, such as energy resources, computing resources, and bandwidth resources, to resist intrusion from all sorts of attackers. The attackers will use all available resources to successfully break the defense and break into the IoT system. Based on these, we will try to formulate a dynamic model for both the defender and the attackers, to find their optimal strategies for resources allocation of the cyber-security IoT system in the process of defense and attack. In our proposed

model, the relationship between the defender and the attackers is considered a one-leader-N-followers Stackelberg dynamic game, where the defender is the leader and the attackers are the followers. At the beginning of the Stackelberg dynamic game, the defender will choose a resource strategy to protect the system. After observing the defender's strategy, the attackers will choose their optimal resource strategies for intrusion based on the observed defense strategy. Then the defender will allocate its resources to defense the invasion based on the attackers' strategies. The strategies for both the defender and the attackers are dynamic and time-dependent.

In order to formulate the Stackelberg dynamic game, there should be a system state for both the defender and the attackers. As the risk level of the IoT system concerns both the defender and the attackers, we use it as the system state of the Stackelberg dynamic game. Generally speaking, the risk level is affected by the input variables, which are the defense strategy and the attack strategies. In our proposed model, we assume that the risk level is not only related to the current input variables, but also is affected by the current risk level with a degradation coefficient. Assuming $u_0(t)$ and $u_i(t)$ are the control variables of the defender and the attackers for the resource allocation, respectively, let $x(t)$ denote the risk level of the IoT system, which can be given by the following differential equation [32]:

$$\frac{dx(t)}{dt} = \alpha u_0(t) + \sum_{i \in N} \beta_i u_i(t) + \varepsilon x(t), \tag{1}$$

where α is a negative weighted factor and β_i is a positive weighted factor that denote the strategies' relative importance on the risk level. Through an effective defense strategy, the system will be more robust as time elapses. Then the risk level of the system will decrease with a random degradation coefficient, which is denoted by ε. Based on Equation (1), we find that, in our proposed IoT system, the risk level will decrease with the defender's action, and increase with the attackers' actions. Meanwhile, the longer the system continues, the lower the risk level, which is denoted by the random degradation coefficient ε.

Given the system state, we can now discuss the objective functions for the defender and the attackers. As the leader of the Stackelberg dynamic game, we will formulate the objective function for the defender first. The IoT system is usually composed of devices with limited resources, such as low power supply and limited battery capacity, so the defense cost is the main concern for the IoT system protection. For the defender, it aims to minimize the cost for protecting the IoT system and resisting the attackers with limited resources. Its objective function can be given as follows:

$$U_0 = \mu_0 \sum_{i=1}^{N} u_i^2(t) + \nu_0 u_0^2(t) + \rho_0(\widetilde{x} - x(t)), \tag{2}$$

where μ_0, ν_0, and ρ_0 are positive weighted factors that denote the relative importance of the components. In our paper, we assume that the weighted factors add up to 1, which means the weighted factors are decimals larger than 0 and less than 1. The physical meanings of the weighted factors are the importance of the components in the cost function. The objective function of the defender has three components. The first part is $\sum_{i=1}^{N} u_i^2(t)$, which means the observing cost. The defender should observe the attackers' strategies to generate its own strategy for system defense. The observing cost is a direct ratio to the attackers' resource strategies; when the attackers spend more resources, the defender should pay more attention for observation. $u_0^2(t)$ is the second part of the objective function, and means the defending cost of the IoT system. Generally, the defending cost is directly related to the resources allocated for defending. The third part of the objective function of the defender is given by $(\widetilde{x} - x(t))$, which is the disparity between the maximum permissible risk level and the real-time risk level. The purpose of the defender is to reduce the risk level to ensure data security, even with a risk level of zero. Let $\widetilde{x}$ denote the maximum permissible risk level; we should try to make the risk level no higher than the threshold. According to Equation (2), we find that the total instantaneous cost of the defender is a

function of the allocated resources $u_0(t)$, $u_i(t)$, and the risk level $x(t)$. The defender wants to find the optimal allocated resources $u_0(t)$ that minimize its cost function over time interval $[0, T]$:

$$\min_{u_0(t)} L_0(t) = \min_{u_0(t)} \left\{ \int_0^T \left\{ \mu_0 \sum_{i=1}^{N} u_i^2(t) + \nu_0 u_0^2(t) + \rho_0(\widetilde{x} - x(t)) \right\} e^{-rt} dt \right\}, \tag{3}$$

subject to Equation (1). Here, r is the discount rate.

The attackers also want to invade the entire IoT system at a low cost. Therefore the aims of the attackers are to minimize the cost of breaking into the IoT system with limited resources. Its objective function can be given as follows:

$$U_i = \mu_i u_i^2(t) + \nu_i u_0(t)(u - u_i(t)) + \rho_i(x(t) - \widetilde{x}), \tag{4}$$

where μ_i, ν_i, and ρ_i are positive weighted factors that denote the relative importance of the components. The objective function of the defender has three components. The first part is $u_i^2(t)$, the resource cost to the attacker. Generally, the attacker should allocate enough energy resources to attack the IoT system, and we use the linear quadratic form to denote the attacking resource cost, the energy or power cost during attack. The second part is the cost for observing and weakening the IoT system. We use $u - u_i(t)$ to denote the resources available for observing, except for the allocated attacking resources, where u denotes the maximum resources that can be allocated. $u_0(t)$ denotes the difficulty of weakening the IoT system; the more resources allocated for protecting by the defender, the harder it is to weaken the IoT system. We combine the above components to denote the system observing and weakening cost. The third part is the cost caused by the risk level, which is denoted by $(x(t) - \widetilde{x})$. Each attacker wants to increase the risk level of the IoT system higher than the maximum permissible risk level $\widetilde{x}$. Based on the objective function, attacker i wants to find the optimal allocated resources $u_i(t)$ that minimize its cost function over time interval $[0, T]$:

$$\min_{u_i(t)} L_i(t) = \min_{u_i(t)} \left\{ \int_0^T \left\{ \mu_i u_i^2(t) + \nu_i u_0(t)(u - u_i(t)) + \rho_i(x(t) - \widetilde{x}) \right\} e^{-rt} dt \right\}, \tag{5}$$

subject to Equation (1). Here, r is the discount rate.

3. Game Analysis

In this section, we will discuss both the open-loop Nash equilibrium solutions and the feedback Nash equilibrium solutions to the proposed game model established in the previous section, and analyze the optimal strategies for the attackers and defender in the IoT system. The open-loop Nash equilibrium solutions will be given first, followed by the feedback Nash equilibriums. Both solutions are given based on Bellman's dynamic programming principle.

3.1. Open-Loop Nash Equilibrium Solutions

During the Stackelberg relations in the threat defense, the IoT system will first consume certain resources for implementing a defense strategy, then the attackers will attack the system based on the initial defending resource. After observing the attackers' strategies, the resources of the IoT system will be recalibrated to cope with all kinds of risk. Because both the IoT system and attackers' resources are limited, it is important to effectively allocate resources during the defense and attack based on the proposed Stackelberg dynamic game. If the defenders and attackers choose to commit their strategies from the outset, their information structure can be seen as an open-loop pattern, which means the optimal strategies for the defender and attackers are functions of the initial risk level $x(0)$ and the time instant t. In this section, the open-loop Nash equilibrium will be given to the game Equations (3) and (5) to obtain the optimal strategies in a finite time horizon $[0, T]$.

3.1.1. Open-Loop Solutions for the Attackers

In order to minimize the cost function, each attacker needs to choose their optimal resource strategies based on the observed defense strategy. Before getting the optimal allocated resources, we first give some definitions for understanding the proposed game model.

Definition 1. *For attacker i, the resource strategy $u_i^*(t)$ is optimal if the following inequality holds for all feasible control $u_i(t) \neq u_i^*(t)$:*

$$L_i(u_i^*(t), x^*(t), t) \leq L_i(u_i(t), x(t), t). \tag{6}$$

Definition 2. *A set of controls $\{u_i^*(t)\}$ constitutes an open-loop Stackelberg equilibrium to the problem in Equation (5), and $x^*(t)$ is the corresponding state trajectory, if there exists a costate function $\Lambda_i(t)$ such that the following relations are satisfied,*

$$u_i^*(t) = \underset{u_i(t)}{\operatorname{argmin}}\{U_i + \Lambda_i(t)\dot{x}(t)\}, \tag{7}$$

$$\dot{\Lambda}_i(t) = -\frac{\partial[U_i + \Lambda_i(t)\dot{x}(t)]}{\partial x(t)}, \tag{8}$$

where $\Lambda_i(t)$ in Equation (8) is an adjoint equation to describe the dynamics of a costate variable. The costate function $\Lambda_i(t)$ is a function associated with the state variable $x(t)$. Generally, Equation (7) can be considered a Hamiltonian system $H_i(t)$ of the proposed game model, and $H_i(t) = U_i + \Lambda_i(t)\dot{x}(t)$.

Based on the definitions given above, we can solve the attacker's optimal resource strategy problem based on the Bellman's dynamic programming principle.

Lemma 1. The optimal resource strategy to attacker i is

$$u_i^*(t) = \frac{v_i u_0(t) - \beta_i \Lambda_i(t)}{2\mu_i}, \tag{9}$$

where $\Lambda_i(t)$ is given by the following:

$$\Lambda_i(t) = \frac{e^{\varepsilon(t-T)} - \rho_i}{\varepsilon}. \tag{10}$$

Proof. See Appendix A.

Equation (9) shows that the optimal resource strategy of the attacker will be affected by $u_0(t)$ and costate functions $\Lambda_i(t)$. We can see that the optimal resource strategy of attacker i is in positive proportion to the resource strategy $u_0(t)$ of the defender. The attackers will choose their optimal resource strategies for intrusion based on the initial defense strategy.

3.1.2. Open-Loop Solutions for the Defender

The defender will allocate its resources to defend the attackers based on the attackers' strategies. In this subsection, we will give the open-loop Nash equilibrium to the defender.

Definition 3. *For the defender, the resource strategy $u_0^*(t)$ is optimal if the following inequality holds for all feasible control $u_0(t) \neq u_0^*(t)$:*

$$L_0(u_0^*(t), x^*(t), t) \leq L_0(u_0(t), x(t), t). \tag{11}$$

Definition 4. *A set of controls* $\{u_0^*(t)\}$ *constitutes an open-loop Stackelberg equilibrium to the problem in Equation (3), and* $x^*(t)$ *is the corresponding state trajectory, if there exist costate functions* $\lambda_0(t)$ *and* $\lambda_i(t)$ *such that the following relations are satisfied:*

$$u_0^*(t) = \underset{u_0(t)}{\operatorname{argmin}} H_0(t), \tag{12}$$

$$\dot{\lambda}_0(t) = -\frac{\partial H_0(t)}{\partial x(t)}, \tag{13}$$

$$\dot{\lambda}_i(t) = -\frac{\partial H_0(t)}{\partial \Lambda_i(t)}, \tag{14}$$

where the Hamiltonian system $H_0(t)$ *of the defender can be expressed as follows:*

$$\begin{aligned} H_0(t) &= U_0 + \lambda_0(t)\dot{x}(t) + \sum_{i=1}^{N} \lambda_i(t)\Lambda_i(t) \\ &= \left[\mu_0 \sum_{i=1}^{N} u_i^2(t) + \nu_0 u_0^2(t) + \rho_0(\widetilde{x} - x(t))\right] \\ &+\lambda_0(t)\left[\alpha u_0(t) + \sum_{i\in N} \beta_i u_i(t) + \varepsilon x(t)\right] \\ &+\sum_{i=1}^{N} \lambda_i(t)(-\rho_i - \varepsilon\Lambda_i(t)) \end{aligned}. \tag{15}$$

Calculate the partial derivative for $\lambda_0(t)$ and $\lambda_i(t)$ in the Hamiltonian system $H_0(t)$, we can obtain,

$$\dot{\lambda}_0(t) = \rho_0 - \varepsilon\lambda_0(t), \tag{16}$$

$$\dot{\lambda}_i(t) = \varepsilon\lambda_i(t). \tag{17}$$

Solving Equations (16) and (17), we have,

$$\lambda_0(t) = \frac{\rho_0 - e^{\varepsilon(t-T)}}{\varepsilon}, \tag{18}$$

$$\lambda_i(t) = \frac{e^{\varepsilon(T-t)}}{\varepsilon}. \tag{19}$$

Calculating the partial derivative for $u_0(t)$ in Equation (15), we obtain,

$$u_0^*(t) = -\frac{\alpha\lambda_0(t)}{2v_0}. \tag{20}$$

Based on the above analysis, the optimal solutions for both the attackers and defender are obtained, we get the corresponding state trajectory $x^*(t)$ using Equations (9) and (20) as follows:

$$x^*(t) = \frac{1}{\varepsilon}\left[e^{\varepsilon(T-t)} - \alpha u_0^*(t) - \sum_{i=1}^{N} \beta_i u_i^*(t)\right] = \frac{1}{\varepsilon}\left[e^{\varepsilon(T-t)} + \frac{\alpha^2\lambda_0(t)}{2v_0} - \sum_{i=1}^{N} \beta_i \frac{v_i u_0(t) - \beta_i\Lambda_i(t)}{2\mu_i}\right]. \tag{21}$$

3.1.3. Open-Loop Control Algorithm

In this subsection, we discuss the implementation open-loop control algorithm for the proposed game analysis. Algorithm 1 is the open-loop control algorithm for the attackers and defenders. The whole algorithm cycling can be divided into two parts. One is the "open-loop control of attackers" part, which is used to calculate the optimal strategies of resource allocation during the attacks. The other is the "open-loop control of defender" part, to make a decision on the resource level for threat

defense. The time complexity of the algorithm will be $O(n)$, because the algorithm should be solved for all the attackers and the defender, and should be solved in a finite time horizon $[0, T]$. The space complexity of the presented solution is $O(n)$, because the function for the open-loop solution should be invoked at each time. The progress can be described as follows.

Algorithm 1. Open-loop control algorithm for the attackers and defender.

Start algorithm
Step 1. Set up the parameter for the attackers and defender;
Step 2. The defender controls its initial strategy for resource allocation for threat defense;
Step 3. Start the open-loop control of the attackers and the defender;
Step 4. Start to calculate the open-loop control solutions for the attackers,
Step 4.1. Set up the objective function for the attackers;
Step 4.2. Calculate the solutions for the attackers.
Step 5. Get the open-loop solutions of the attackers for the defender;
Step 6. Start to calculate the open-loop control solutions for the defender;
Step 6.1. Set up the objective function for the defender;
Step 6.2. Calculate the solutions for the defender.
Algorithm End

3.2. Feedback Nash Equilibrium Solutions

To eliminate information nonuniqueness in the derivation of Nash equilibria, we can obtain the optimal solutions for the proposed game mode to satisfy the feedback Nash equilibrium property. In the feedback situation, the information structures of the defender and attackers follow a closed-loop perfect state pattern, and the optimal strategies for the defender and attackers become functions of the initial risk level $x(t)$, the current risk level $x(t)$ at time instant t, and the current time t. In this subsection, the feedback Nash equilibrium solutions to the proposed Stackelberg dynamic game are discussed based on the dynamic optimization programming technique developed by Bellman [33]. In the following subsections, we first discuss the optimal resource strategies for each attacker in a finite time horizon $[0, T]$. Then, the optimal strategy of the defender is obtained based on the attackers' solutions.

3.2.1. Feedback Solutions for the Attackers

In this section, we first discuss the optimal resource strategies for the attackers, the feedback Nash equilibrium solutions to the game Equations (1) and (5) will be discussed.

Definition 5. *A set of control $\{u_i^*(t)\}$ constitutes a feedback solution to Equations (1) and (5); if there exists a continuously differentiable function $V^i(t,x)$, and $V^i(t,x)$ satisfies the following differential equation:*

$$-V_t^i(t,x) = \min_{u_i(t)}\left\{ \left[\mu_i u_i^2(t) + \nu_i u_0(t)(u - u_i(t)) + \rho_i(x(t) - \widetilde{x})\right]e^{-rt} + V_x^i\left[\alpha u_0(t) + \sum_{i\in N}\beta_i u_i(t) + \varepsilon x(t)\right]\right\}. \quad (22)$$

Calculating the partial derivative for $u_i(t)$ in Equation (22), we can then obtain

$$u_i^*(t) = \frac{v_i u_0(t) - \beta_i V_x(t,x)e^{rt}}{2\mu_i}. \quad (23)$$

Lemma 2. *The value function $V^i(t,x)$ admits a solution that satisfies,*

$$V^i(t,x) = [A_i(t)x + B_i(t)]e^{-rt}, \quad (24)$$

where $A_i(t)$ *is given by*

$$A_i(t) = \frac{e^{(r-\varepsilon)(T-t)} + \rho_i}{r-\varepsilon}, \tag{25}$$

and $B_i(t)$ *are satisfied,*

$$\begin{aligned} \dot{B}_i(t) = rB_i(t) - A_i(t)\left[\alpha u_0(t) + \sum_{i\in N}\beta_i u_i(t)\right] \\ -\mu_i u_i^2(t) - \nu_i u_0(t)\big(u - u_i(t)\big) + \rho_i\widetilde{x} \end{aligned} \tag{26}$$

Proof. See Appendix B.

3.2.2. Feedback Solutions for the Defender

In this subsection, the feedback Nash equilibrium solution for the defender will be discussed.

Definition 6. *A set of control* $\{u_0^*(t)\}$ *constitutes an feedback solution to Equations (1) and (3), if there exists a continuously differentiable function* $V^0(t,x)$*, and* $V^0(t,x)$ *satisfies the following differential equation:*

$$-V_t^0(t,x) = \min_{u_0(t)}\left\{\left[\mu_0\sum_{i=1}^{N}u_i^2(t) + \nu_0 u_0^2(t) + \rho_0(\widetilde{x} - x(t))\right]e^{-rt} + V_x^0\left[\alpha u_0(t) + \sum_{i\in N}\beta_i u_i(t) + \varepsilon x(t)\right]\right\}. \tag{27}$$

As the game leader, the defender should consider the resource strategies of the attackers before making a decision on the resource strategies. Calculating the partial derivative for $u_0(t)$ in (27), we obtain

$$u_0^*(t) = -\frac{\alpha V_x^0(t)e^{rt}}{2v_0}. \tag{28}$$

Lemma 3. *The value function* $V^0(t,x)$ *admits a solution that satisfies,*

$$V^0(t,x) = [A_0(t)x + B_0(t)]e^{-rt}, \tag{29}$$

where $A_0(t)$ *and* $B_0(t)$ *are given by*

$$\begin{cases} \dot{A}_0(t) = (r-\varepsilon)A_0(t) + \rho_0 \\ \dot{B}_0(t) = rB_0(t) - A_0(t)\left[\alpha u_0(t) + \sum_{i\in N}\beta_i u_i(t)\right] \\ -\mu_0\sum_{i=1}^{N}u_i^2(t) - \nu_0 u_0^2(t) - \rho_0\widetilde{x} \end{cases}. \tag{30}$$

Proof. By taking the derivative of $V^0(t,x)$ with respect to t and x, we obtain,

$$V_t^0(t,x) = \left[-rA_0(t) + \dot{A}_0(t)\right]x + \left[-rB_0(t) + \dot{B}_0(t)\right], \tag{31}$$

$$V_x^0(t,x) = A_0(t)e^{-rt}. \tag{32}$$

Solving Equations (27), (31), and (32), $A_0(t)$ and $B_0(t)$ are satisfied:

$$\begin{cases} \dot{A}_0(t) = (r-\varepsilon)A_0(t) + \rho_0 \\ \dot{B}_0(t) = rB_0(t) - A_0(t)\left[\alpha u_0(t) + \sum_{i\in N}\beta_i u_i(t)\right] - \mu_0\sum_{i=1}^{N}u_i^2(t) - \nu_0 u_0^2(t) - \rho_0\widetilde{x} \end{cases}. \tag{33}$$

Solving the above equation, we can obtain the expression of $A_0(t)$ as follows:

$$A_0(t) = \frac{e^{(r-\varepsilon)(T-t)} - \rho_0}{r - \varepsilon}. \tag{34}$$

Substituting Equation (34) into Equation (28), we can derive the optimal resource strategy for the defender as follows:

$$u_0^*(t) = -\frac{\alpha A_0(t)}{2v_0}. \tag{35}$$

Solving Equation (1), we can get the optimal state:

$$x^*(t) = \frac{1}{\varepsilon}\left[e^{\varepsilon(T-t)} - \alpha u_0^*(t) - \sum_{i=1}^{N} \beta_i u_i^*(t)\right] = \frac{1}{\varepsilon}\left[e^{\varepsilon(T-t)} + \frac{\alpha^2 A_0(t)}{2v_0} - \sum_{i=1}^{N} \beta_i \frac{v_i u_0(t) - \beta_i A_i(t)}{2\mu_i}\right]. \tag{36}$$

3.2.3. Feedback Control Algorithm

In this subsection, we will discuss the implementation feedback control algorithm for the proposed game analysis, which is given in Algorithm 2. Similarly, the whole algorithm cycling can be divided into the attackers' part, and the defender's part. The time complexity of the feedback control algorithm will be $O(n)$, and the space complexity is $O(n)$. The progress can be described as follows.

Algorithm 2. Feedback control algorithm for the attackers and defender.

Start algorithm
Step 1. Set up the parameter for the attackers and defender;
Step 2. The defender control its initial strategy for resource allocation for threat defense;
Step 3. Start the feedback control of the attackers and the defender;
Step 4. Start to calculate the feedback control solutions for the attackers,
Step 4.1. Set up the objective function for the attackers;
Step 4.2. Calculate the solutions for the attackers.
Step 5. Get the feedback solutions of the attackers for the defender;
Step 6. Start to calculate the feedback control solutions for the defender;
Step 6.1. Set up the objective function for the defender;
Step 6.2. Calculate the solutions for the defender.
Algorithm End

4. Numerical Simulations

In this section, we will use MATLAB software to simulate the proposed Stackelberg dynamic game model. We will analyze the resource strategies of attackers and defender, in the form of open-loop and feedback. The simulation parameters are shown in Table 1. To simplify the simulations, we assume all the attackers are uniform with the same simulation parameters.

Table 1. Parameter settings.

Parameter	α	β_i	ε	μ_0	v_0	ρ_0	μ_i	v_i	ρ_i
Value	−0.85	0.6	−0.5	0.3	0.5	0.2	0.1	0.5	0.4

4.1. Numerical Simulations of Open-Loop Nash Equilibrium Solutions

We first simulate the open-loop Nash equilibrium solutions of the model to get the optimal defense resource strategies of the defender and attackers.

Figure 1 describes the relationship between the optimal strategies $u_i^*(t)$ and $u_0^*(t)$ over time $t(t \in [0, 10])$. As shown in Figure 1, the optimal resource strategy of both the attacker and defenders monotonically decrease with time t. In order to protect the security of the system, the defender adopts

a strategy to consume its own resources when the attacker attacks. The attacker adopts a strategy to attack the system and consumes its own resources. As the time changes, the optimal strategies for the defender and attackers tend to convergence.

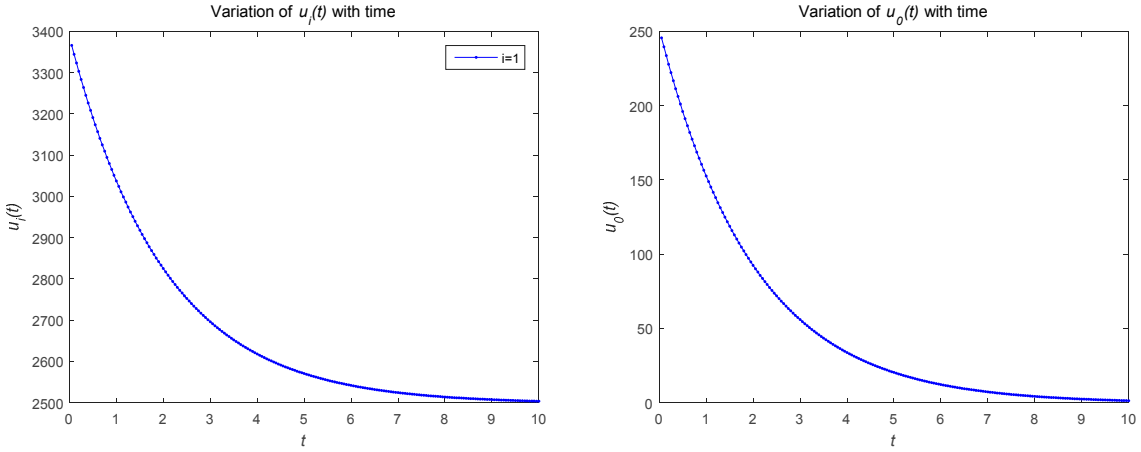


Figure 1. (**a**) Optimal strategy of the attackers; (**b**) optimal strategy of the defender.

Figure 2 describes the changes in the attacker's optimal resource strategy when $u_0(t)$ takes different values. We find that the smaller $u_0(t)$, the smaller the optimal resource strategy $u_i(t)$. This is because the attacker will choose their optimal resource strategies for attacks based on the observed defense strategy. Figure 3 describes the relationship between the risk level of the system and time t. The risk level at the initial moment is the highest, and with the effective defense of the defender, the risk level shows a decreasing trend. As shown in Figure 3b, the risk level is a decreasing function with respect to time t, which is proportional to the number of attackers. The number of attackers are set to 1, 5, and 20, respectively. Figure 4 shows the risk level variation of the system, when the number of the devices in the IoT system becomes a large number, to analyze the scalability of the proposed model. We can prove that the proposed model can be used for IoTs with a large number of devices based on Figure 4.

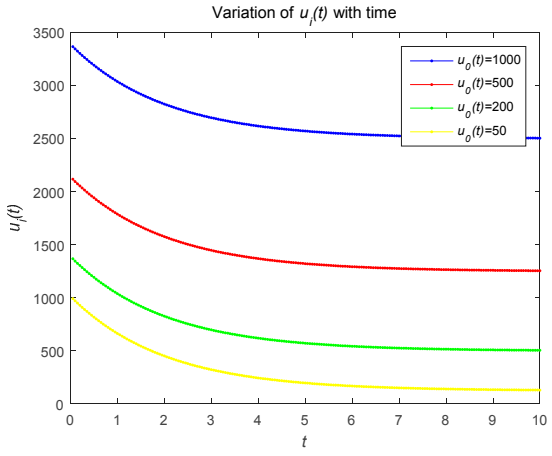


Figure 2. Optimal strategy of the attacker with different $u_0(t)$ over time.

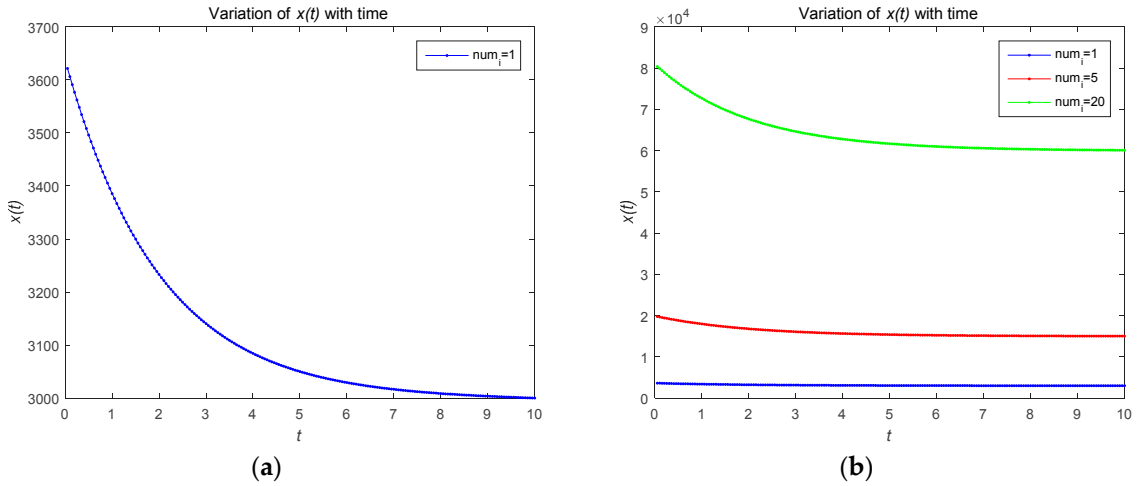


Figure 3. (**a**) Risk level variation for a system with one attacker; (**b**) risk level variation for a system with different numbers of attackers.

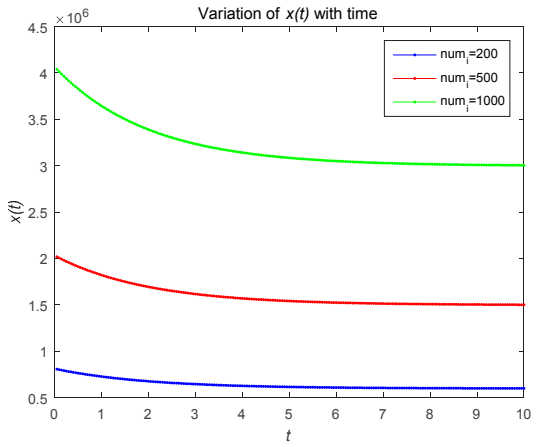
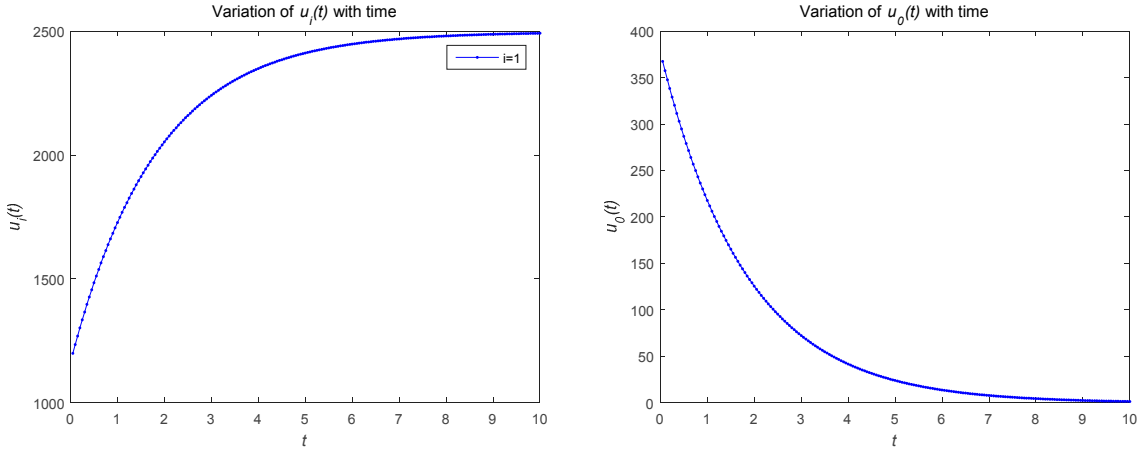


Figure 4. Risk level with a large number of attackers under open-loop control.

4.2. Numerical Simulations of Feedback Nash Equilibrium Solutions

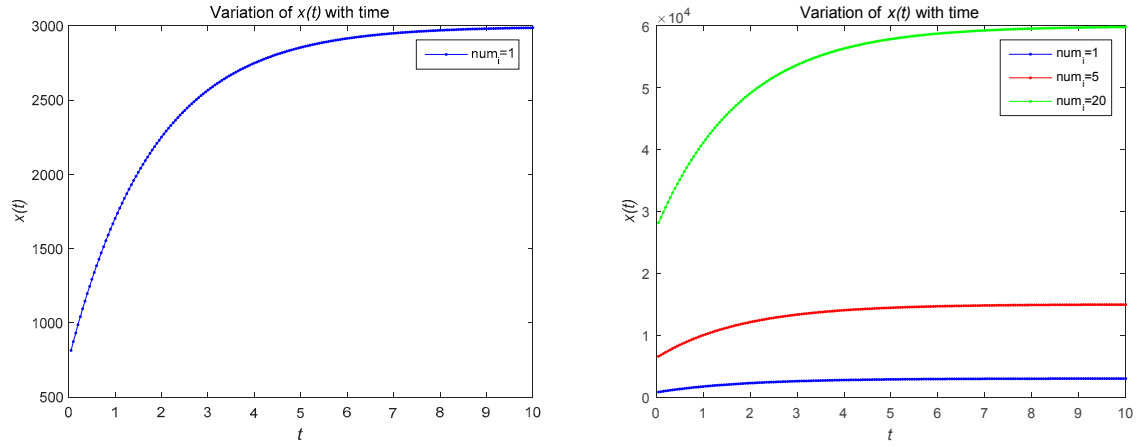
This subsection mainly simulates the feedback Nash equilibrium solution of the model. Figure 5 describes the relationship between the optimal strategy $u_i^*(t)$ and $u_0^*(t)$ with time $t(t \in [0,10])$. As shown in Figure 5a, the attacker's optimal resource strategy is an increasing function with respect to time t. The attackers control their own resource strategies. The aim of the attack is to increase the risk level, so they allocate more resources for attacks under the feedback control situation. As shown in Figure 5b, the defender's optimal resource strategy is a decreasing function with respect to time t. The defender controls its own resource strategy to minimize risk, but, because of limited resources, may not have enough for defense as the time changes.

Figure 6 describes the relationship between the risk level of the system and time t. As shown in Figure 6, the risk level is proportional to the number of attackers. In the feedback Nash equilibrium solution, the attacker uses more attacks to increase the risk. Figure 7 shows the risk level variation of the system, when the number of devices in the IoT system becomes large. Figure 8 gives the time complexity of the proposed Stackelberg dynamic game. As shown in Figure 6, the time complexity of both the open-loop and feedback control algorithm will be $O(n)$.



(**a**) (**b**)

Figure 5. (**a**) Optimal strategy of the attacker; (**b**) optimal strategy of the defender.



(**a**) (**b**)

Figure 6. (**a**) Risk level variation for a system with one attacker; (**b**) risk level variation for a system with different numbers of attackers.

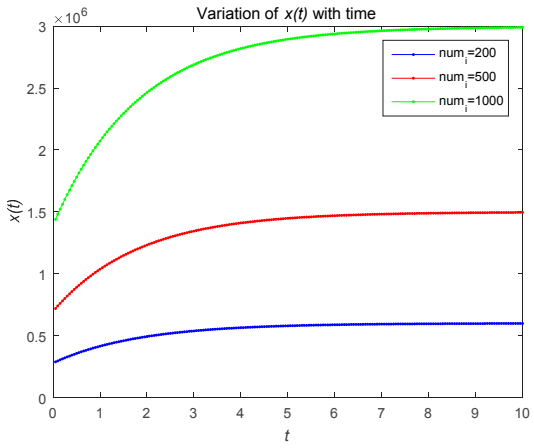


Figure 7. Risk level with a large number of attackers under feedback control.

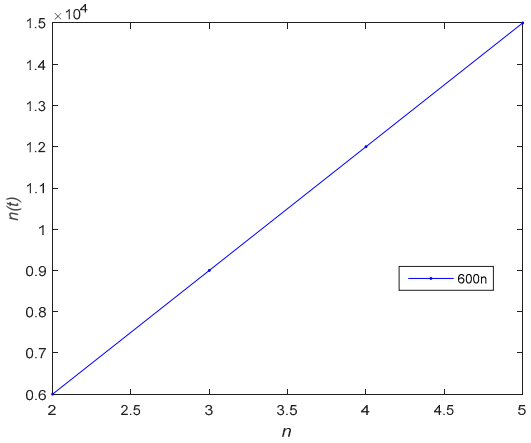


Figure 8. The time complexity of the proposed game model.

5. Conclusions

This paper proposes a Stackelberg dynamic game-based resource allocation model in the cyber-security IoT system that is composed by one defender and N attackers. We formulate a dynamic model for both the defender and the attackers to find their optimal strategies for resource allocation in the process of defense and attack. By solving the open-loop Nash equilibrium solution and the feedback Nash equilibrium solution, we find that the optimal resource solution for the defender is the open-loop Nash equilibrium solution, and under the open-loop situation, the defender can effectively reduce the risk level of the system. However, attackers can obtain more profit under the feedback situation.

Author Contributions: H.X. conceived the main idea and the dynamic game theory model; all authors contributed to data analysis, simulations, and the writing of this paper.

Funding: This work was supported by the Natural Science Foundation of China, Nos. 61501026, 61873026, U1603116, and the Foundation of Science and Technology on Information Assurance Laboratory, No. KJ-17-101.

Acknowledgments: The authors would like to thank the editor and the anonymous reviewers for their valuable comments and suggestions that improved the quality of this paper.

Conflicts of Interest: The authors declare no conflicts of interest. The founding sponsors had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

Appendix A

Proof of Lemma 1.

To solve the game model, we use the maximum principle proposed by Pontryagin et al. [34]. The Hamiltonian system of attacker i is given by

$$\begin{aligned} H_i(t) &= U_i + \Lambda_i(t)\dot{x}(t) \\ &= \left[\mu_i u_i^2(t) + \nu_i u_0(t)\left(u - u_i(t)\right) + \rho_i(x(t) - \widetilde{x})\right] \\ &+\Lambda_i(t)\left[\alpha u_0(t) + \sum_{i\in N}\beta_i u_i(t) + \varepsilon x(t)\right] \end{aligned} \tag{37}$$

Calculating the partial derivative for $u_i(t)$ and $x(t)$ in Equation (37), we have

$$\frac{\partial H_i(t)}{\partial u_i(t)} = 2\mu_i u_i(t) - v_i u_0(t) + \beta_i \Lambda_i(t), \tag{38}$$

$$\frac{\partial H_i(t)}{\partial x(t)} = \rho_i + \varepsilon\Lambda_i(t). \tag{39}$$

Then we have the optimal solution for resource allocation as follows:

$$u_i^*(t) = \frac{v_i u_0(t) - \beta_i \Lambda_i(t)}{2\mu_i}, \tag{40}$$

and the costate function $\Lambda_i(t)$ can be given by the following differential equation,

$$\dot{\Lambda}_i(t) = -\rho_i - \varepsilon \Lambda_i(t). \tag{41}$$

Solving Equation (41), we can obtain the expression of the costate function $\Lambda_i(t)$:

$$\Lambda_i(t) = \frac{e^{\varepsilon(t-T)} - \rho_i}{\varepsilon}. \tag{42}$$

Hence, Lemma 1 follows.

Appendix B

Proof of Lemma 2.

By taking the derivative of $V^i(t,x)$ with respect to t and x, we obtain,

$$V_t^i(t,x) = \left[-rA_i(t) + \dot{A}_i(t)\right]x + \left[-rB_i(t) + \dot{B}_i(t)\right], \tag{43}$$

$$V_x^i(t,x) = A_i(t)e^{-rt}. \tag{44}$$

Substituting Equations (43–44) into Equation (27), $A_i(t)$ and $B_i(t)$ are satisfied:

$$\begin{cases} \dot{A}_i(t) = (r-\varepsilon)A_i(t) - \rho_i \\ \dot{B}_i(t) = rB_i(t) - A_i(t)\left[\alpha u_0(t) + \sum\limits_{i\in N} \beta_i u_i(t)\right] \\ -\mu_i u_i^2(t) - \nu_i u_0(t)\left(u - u_i(t)\right) + \rho_i \widetilde{x} \end{cases}, \tag{45}$$

Then, we obtain

$$A_i(t) = \frac{e^{(r-\varepsilon)(T-t)} + \rho_i}{r-\varepsilon}. \tag{46}$$

Using Equation (44), we can derive the optimal resource strategies:

$$u_i^*(t) = \frac{v_i u_0(t) - \beta_i A_i(t)}{2\mu_i}, \tag{47}$$

where $A_i(t)$ is given by Equation (46).

Hence, Lemma 2. follows.

References

1. Keoh, S.L.; Kumar, S.S.; Tschofenig, H. Securing the Internet of Things: A Standardization Perspective. *IEEE Internet Things J.* **2014**, *1*, 265–275. [CrossRef]
2. Tsiropoulou, E.E.; Baras, J.S.; Papavassiliou, S.; Qu, G. On the Mitigation of Interference Imposed by Intruders in Passive RFID Networks. In Proceedings of the International Conference on Decision and Game Theory for Security, New York, NY, USA, 2–4 November 2016.
3. Alaba, F.A.; Othman, M.; Hashem, I.A.T.; Alotaibi, F. Internet of things Security: A Survey. *J. Netw. Comput. Appl.* **2017**, *88*, 10–28. [CrossRef]
4. Atzori, L.; Iera, A.; Morabito, G. The Internet of Things: A survey. *Comput. Netw.* **2010**, *54*, 2787–2805. [CrossRef]

5. Miorandi, D.; Sicari, S.; De Pellegrini, F.; Chlamtac, I. Internet of things. *Ad Hoc Netw.* **2012**, *10*, 1497–1516. [CrossRef]
6. An, X.; Su, J.; Lü, X.; Lin, F. Hypergraph clustering model-based association analysis of DDOS attacks in fog computing intrusion detection system. *EURASIP J. Wirel. Commun. Netw.* **2018**, *2018*, 249. [CrossRef]
7. Li, S.; Xu, L.D.; Zhao, S. *The Internet of Things: A Survey*; Kluwer Academic Publishers: Dordrecht, The Netherlands, 2015; pp. 243–259.
8. Oleshchuk, V. Internet of things and privacy preserving technologies. In Proceedings of the 2009 1st International Conference on Wireless Communication, Vehicular Technology, Information Theory and Aerospace & Electronic Systems Technology, Aalborg, Denmark, 17–20 May 2009; pp. 336–340.
9. An, X.; Zhou, X.; Xing, L.; Lin, F.; Yang, L. Sample Selected Extreme Learning Machine Based Intrusion Detection in Fog Computing and MEC. *Wirel. Commun. Mob. Comput.* **2018**, *2018*, 1–10. [CrossRef]
10. Lin, F.; Lü, X.; You, I.; Zhou, X. A novel utility based resource management scheme in vehicular social edge computing. *IEEE Access* **2018**. [CrossRef]
11. Mavropoulos, O.; Mouratidis, H.; Fish, A.; Panaousis, E.; Kalloniatis, C. A conceptual model to support security analysis in the internet of things. *Comput. Sci. Inf. Syst.* **2017**, *14*, 557–578. [CrossRef]
12. Yang, J.C.; Fang, B.X. Security model and key technologies for the Internet of things. *J. China Univ. Posts Telecommun.* **2011**, *18*, 109–112. [CrossRef]
13. Lin, F.; Zhou, Y.; An, X.; You, I.; Choo, K.R. Fair Resource Allocation in an Intrusion-Detection System for Edge Computing: Ensuring the Security of Internet of Things Devices. *IEEE Consum. Electron. Mag.* **2018**, *7*, 45–50. [CrossRef]
14. Zhang, B.; Zou, Z.; Liu, M. Evaluation on security system of internet of things based on Fuzzy-AHP method. In Proceedings of the 2011 International Conference on E-Business and E-Government (ICEE), Shanghai, China, 6–8 May 2011; pp. 1–5.
15. Leusse, P.D.; Periorellis, P.; Dimitrakos, T.; Nair, S.K. Self Managed Security Cell, a Security Model for the Internet of Things and Services. In Proceedings of the 2009 First International Conference on Advances in Future Internet, Athens, Greece, 18–23 June 2009; pp. 47–52.
16. Acharya, R.; Asha, K. Data integrity and intrusion detection in Wireless Sensor Networks. In Proceedings of the 2008 16th IEEE International Conference on Networks, New Delhi, India, 12–14 December 2008; pp. 1–5.
17. Yuan, X. *Key Management Schemes for Distributed Sensor Networks*; University of Western Ontario: London, ON, Canada, 2008.
18. Tague, P.; Slater, D.; Rogers, J.; Poovendran, R. Vulnerability of Network Traffic under Node Capture Attacks Using Circuit Theoretic Analysis. In Proceedings of the IEEE INFOCOM 2008—The 27th Conference on Computer Communications, Phoenix, Arizona, 13–18 April 2008; pp. 161–165.
19. Choi, Y.S.; Shin, S.H. *A Study on Sensor Node Capture Defense Protocol for Ubiquitous Sensor Network*; InTech: Vienna, Austria, 2007.
20. Ho, J.W. *Distributed Detection of Node Capture Attacks in Wireless Sensor Networks*; InTech: Vienna, Austria, 2010; pp. 661–666.
21. Tague, P.; Poovendran, R. Modeling adaptive node capture attacks in multi-hop wireless networks. *Ad Hoc Netw.* **2007**, *5*, 801–814. [CrossRef]
22. Liu, C.; Zhang, Y.; Zhang, H. A Novel Approach to IoT Security Based on Immunology. In Proceedings of the 2013 Ninth International Conference on Computational Intelligence and Security, E'Mei Mountain, China, 14–15 December 2013; pp. 771–775.
23. An, X.; Lin, F.; Xu, S.; Miao, L.; Gong, C. A Novel Differential Game Model-Based Intrusion Response Strategy in Fog Computing. *Secur. Commun. Netw.* **2018**. [CrossRef]
24. Zhang, C.; Green, R. Communication security in internet of thing: preventive measure and avoid DDoS attack over IoT network. In Proceedings of the 18th Symposium on Communications & Networking, San Jose, CA, USA, 11–13 December 2015; pp. 8–15.
25. Tsiropoulou, E.E.; Vamvakas, P.; Papavassiliou, S. Joint utility-based uplink power and rate allocation in wireless networks: A non-cooperative game theoretic framework. *Phys. Commun.* **2013**, *9*, 299–307. [CrossRef]
26. Kastrinogiannis, T.; Tsiropoulou, E.; Papavassiliou, S. Utility-based uplink power control in CDMA wireless networks with real-time services. In *International Conference on Ad-Hoc*; Springer: New York, NY, USA, 2008; pp. 307–320.

27. Bloem, M.; Alpcan, T. A Stackelberg Game for Power Control and Channel Allocation. In Proceedings of the International Conference on Performance Evaluation Methodolgies & Tools, Nantes, France, 22–27 October 2007; pp. 1–9.
28. Daoud, A.A.; Alpcan, T.; Agarwal, S.; Alanyali, M. A stackelberg game for pricing uplink power in wide-band cognitive radio networks. In Proceedings of the 47th IEEE Conference on Decision and Control, Cancun, Mexico, 9–11 December 2008; pp. 1422–1427.
29. Basar, T.; Srikant, R. Revenue-maximizing pricing and capacity expansion in a many-users regim. In Proceedings of the Joint Conference of the IEEE Computer & Communications Societies, New York, NY, USA, 23–27 June 2002; pp. 294–301.
30. Zhu, Q.; Basar, T. Game-Theoretic Methods for Robustness, Security, and Resilience of Cyberphysical Control Systems: Games-in-Games Principle for Optimal Cross-Layer Resilient Control Systems. *IEEE Control Syst.* **2015**, *35*, 46–65.
31. Manshaei, M.; Zhu, Q.; Alpcan, T.; Tamer, B.; Jean-Pierre, H. Game Theory Meets Network Security and Privacy. *ACM Comput. Surv.* **2013**, *45*, 1–39. [CrossRef]
32. Liu, R.; Zhai, F. Model Identification of Risk Management System. In Proceedings of the 2008 4th International Conference on Wireless Communications, Networking and Mobile Computing, Dalian, China, 12–14 October 2008; pp. 1–4.
33. Yeung, D.W.K.; Petrosyan, L.A. *Cooperative Stochastic Differential Games*; Springer: New York, NY, USA, 2006; p. 256.
34. Pontryagin, L.S.; Boltyanskii, V.G.; Gamkrelidze, R.V.; Mishchenko, E.F. *The Mathematical Theory of Optimal Processes*; Interscience Publishers: New York, NY, USA, 1962.