Dynamic Defense Strategy against Advanced Persistent Threat with Insiders

Pengfei Hu*, Hongxing Li*, Hao Fu*, Derya Cansever† and Prasant Mohapatra*

*Department of Computer Science, University of California, Davis, USA,
Email: {pfhu, honli, haofu, pmohapatra}@ucdavis.edu
†US Army CERDEC, USA
Email: derya.h.cansever.civ@mail.mil

***Abstract*—The landscape of cyber security has been reformed dramatically by the recently emerging *Advanced Persistent Threat* (APT). It is uniquely featured by the *stealthy, continuous, sophisticated* and *well-funded* attack process for *long-term* malicious gain, which render the current defense mechanisms inapplicable. A novel design of defense strategy, continuously combating APT in a long time-span with imperfect/incomplete information on attacker's actions, is urgently needed. The challenge is even more escalated when APT is coupled with the *insider threat* (a major threat in cyber-security), where insiders could trade valuable information to APT attacker for monetary gains. The interplay among the defender, APT attacker and insiders should be judiciously studied to shed insights on a more secure defense system. In this paper, we consider the joint threats from APT attacker and the insiders, and characterize the fore-mentioned interplay as a two-layer game model, *i.e.*, a defense/attack game between defender and APT attacker and an information-trading game among insiders. Through rigorous analysis, we identify the best response strategies for each player and prove the existence of Nash Equilibrium for both games. Extensive numerical study further verifies our analytic results and examines the impact of different system configurations on the achievable security level.**

I. Introduction

We are in a new era of Cyber security with the arising challenge from Advanced Persistent Threats (APT) [1]. Different from the traditional Cyber security threats, APT attackers are capable to adopt any *advanced* actions in a *stealthy* manner with a goal of *long-term* utility gain, instead of any *one-shot* benefit. Hence, these unique properties render the existing security solutions [1] inapplicable for APT, since they are confined by one or more of the following limitations: i) each attacker has a discrete and limited set of actions for one specified type of attacks (*e.g.*, DoS attack and password-based attack), violating the feature of "*advanced*" actions in APT which could include the combination all possible types of attacks; ii) the security game runs in a discrete-time fashion and the defender and attacker take actions either *concurrently* or *alternately* in each time slot, which are far from the real practice for APT since the attacker/defender cannot be accurately coordinated to make a move as the attacker acts *continuously* (not discretely) and *stealthily*[1]; and iii) the security problem is modeled as a *one-shot* static game, which cannot characterize the *persistent* interplay among players for their *long-term* utility gains, or a *repeated* game, whose system status (*e.g.*, how much portion of the system has been compromised) remains static and cannot be impacted by players' behaviors.

To sum up the above, APT calls for a framework which 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 i) a model to accurately capture the continuously evolving process of the system status and how it is influenced by attacker's and defender's actions; and ii) dynamic defense/attack strategies that judiciously and continuously take actions in order to minimize/maximize the long-term system damage without knowing the opponent's behavior.

The challenge is further escalated when we consider the threats from the insiders, which is an inevitable issue for cyber security. The 2013 US State of Cybercrime Survey [2], which is conducted by the U.S. Secret Service, CSO magazine and Carnegie Mellon University Software Engineering Institute's CERT, states that 23% of the electronic crime events are caused by insiders while the damage/cost resulting from insiders (34%) is even more severe than outsiders (31%). It implies that insiders have more threat to the system/organizations than the conventional outsider does.

As most insiders are driven by economic profits while APT attacker is always well funded [1], insiders are prone to be utilized for the APT attacks through information-trading. Existing literature on insider threats merely focuses on the insider-detection mechanism at the defender side [3], without considering the insiders' inherent profit-maximizing objectives and their consequentially proactive and dynamic actions so as to achieve their goals. Current efforts tackle the threats from outside attackers and insiders independently and separately, failing to capture the interconnection between them.[2] To the best of our knowledge, this is the first reported work to investigate the joint interplay among defender, attacker and insiders within one framework, by identifying the best response strategies for each of them to pursue their long-term

[1]There is no way to know the opponent's time to take an action and to react accordingly either *concurrently* or *alternately*.

[2]Without the outside attacker, the insider has no way to sell the information and thus to make profits. Meanwhile, with the help of insiders, the outside attacker could launch more effective attacks.

objectives, respectively, *i.e.*, minimizing (or maximizing) the system damage for defender (or attacker) and maximizing the profit gained from information trading for insiders. However, it is nontrivial to bring insiders into the picture, and to evaluate their impact. Each selfish insider will rationally and independently decide its action *dynamically*, *i.e.*, to which extent the inside information should be traded at each time point, in order to maximize its *long-term* gain from the information trading. An over-aggressive transaction may lead to the risk of exposure to the system defender, which results in a cost of being fired or even sued, while an over-conservative action may hurt its profit gain.

In this paper, we investigate the joint threats from the *APT attacker* and *insiders* over a long time-span within a general framework. We characterize the interplay among defender, attacker and insiders as a two-layer differential game in an open-loop setting: i) the defense/attack game between the defender and the APT attacker; and ii) the information-trading game among multiple insiders based on the attacker's needs. We model the evolving process of the system status as a differential equation defined over the actions by each player, and identify the optimal defense/attack strategies for each player in the defense/attack game for both static and dynamic cases as well as the optimal information-trading strategies for each insider in its dynamic case. Through rigorous analysis, we prove i) the existence of the Nash Equilibrium for the defense/attack game; and ii) the existence and uniqueness of the Nash Equilibrium for the information-trading game among insiders. Those results shed insights on the design of defense strategies towards a securer system. The proposed framework is also evaluated with numerical studies in practical settings.

The contribution of this paper can be summarized as follows,

▷ As a first study in the literature, we consider the joint threats from the APT attacker and insiders, and the impact of the long-term objectives of defender, attacker and insiders on their continuous and dynamic actions.

▷ We propose a general framework of two-layer differential game: one between the defender and the attacker, and the other one among multiple insiders. Optimal response strategies are identified for the defender and attacker in both static and dynamic cases while the optimal information-trading decisions are dynamically made for each insider, so as to optimize their long-term objectives, respectively.

▷ We rigorously analyze the proposed framework, and prove the existence of the Nash Equilibrium for the defense/attack game as well as the existence and uniqueness of the Nash Equilibrium for the information-trading game.

▷ Our numerical study further examines the impact of different system configurations on the achievable security level.

The remainder of the paper is organized as follows. We discuss related work in Sec. II and present the problem model in Sec. III. Detailed algorithm design and performance analysis are presented in Sec. IV. The algorithm performance is evaluated via an empirical study in Sec. VI. Finally, we conclude the paper in Sec. VII.

II. Related Work

A. Advanced Persistent Threat

The cyber security domain has been changed dramatically by a new class of threats, which is referred as *Advanced Persistent Threat* by industry. The first well known APT case may be *Stuxnet* [4], which is designed to modify industrial Programmable Logic Controllers and to force them to diverge from the normal behaviors by exploiting a vast majority of security holes and tools. Another famous APT case is *Operation Aurora* [5], which targets at Google and dozens of other companies. The APT attacker can exploit the zero-day vulnerability in the Internet Explorer. Cole [1] introduces the definition of APT and the unique characteristics making it different from traditional security issues. Dijk *et al.* [6] 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. Insider Threat

Insider threat is a major threat bringing severe damage to the cyber security [2]. Martinez-Moyano *et al.* [7] address the insider threat through a dynamic model based on the behavior theory and explain the model with the data related to information technology security violations. Colwill *et al.* [8] investigate several primary issues related to insider threat, which include the nature of the loyalty and betrayal, cultural factors, changing economic and social factors, *etc*. Based on the above, proactive security actions rather than the reactive actions should be taken to deal with the insider threats. A proactive insider threat detection approach which combines Structural Anomaly Detection from social and information networks and Psychological Profiling of individuals is proposed in [9]. Mathew *et al.* [10] present a feature-extraction method rather than the traditional query expression analysis to model user's access pattern which could be applied to detect the insider attacks. A detailed survey on the proposed approaches against the insider threats in the security research literature is summarized in [3].

C. Game Theory in Cyber Security

Game Theory has been widely applied to solve a variety of security and privacy issues in computer and communication networks. A game theoretic model of the interaction between a intruder and the operator of the smart grid is presented in [11]. Alpcan and Buchegger [12] investigate the security issues of the vehicular networks within a game theoretic framework and identify the optimal defensive strategy respect to threats posed by malicious attackers. Alpcan and Tamer [13] address the intrusion detection problem in networks and formulate it as a noncooperative game. They analyze the Nash equilibrium and its implications behind. The security of networked control systems (NCS) is addressed in [14] via integrating it with the economics of security to deal with the interdependence of security-related risks. Anderson and Moore [15] investigate the

economics of information security and show that incentives are becoming as important as technical design in order to achieve dependability.

Different from all the above discussions, this paper is the first in literature to consider the joint threats from APT attacker and insiders, and presents provably optimal response strategies for each player in a two-layer security game.

III. Problem Model

In this section, we present a general model of the joint APT and insider threat, based on which we formulate the interplay among the defender, APT attacker and multiple insiders as a two-layer differential game in Sec. IV. Important notations are summarized Table I.

A. System Model

We consider a system under the joint threats from APT attacker and the insiders. There are four components in the system: the system resource, one APT attacker, one defender, and n insiders. The target of the threats is the system resource, which could include the fire-wall, network, software and operation system etc. Fig. 1 illustrates the interplay among the defender, APT attacker and insiders.

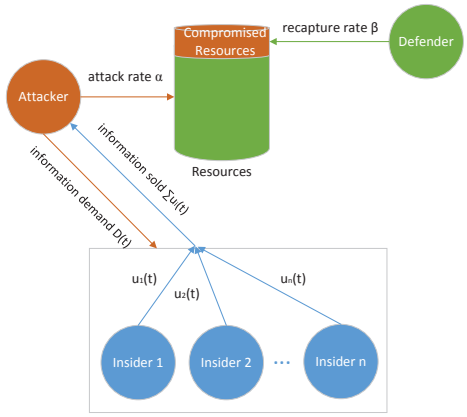


Fig. 1. Illustration of the interplay among the defender, APT attacker and insiders.

▷ The APT attacker aims to obtain malicious gain from the system by launching attacks and compromising partial/all system resource. The cost per attack could decrease if the APT attacker has inside information about the system, which can be purchased from the insiders.

▷ The insiders are selfishly and independently maximizing their individual monetary profits via selling the inside information (which could better assist the APT attacker to compromise the resources) to the APT attacker.

▷ Naturally, the system defender's task is to recapture the compromised resource so as to minimize the damage brought by the APT and insider threats.

Here, we give a general resource model and normalize the total system resource as the value of 1. Let $x(t) \in [0,1]$ denote the fraction of compromised resources at time t, with 0 indicating a system fully under protection while 1 for a completely compromised system. We call $x(t)$ as the *system state* or *system security level* at time t.

TABLE I
IMPORTANT NOTATIONS.

α	Attack rate at which the attacker grabs the system resources
β	Recapture rate of defender to recover the compromised resources
$\dot{x}(t)$	Evolving rate of the system state at each time point
$x(t)$	System state
$c_{\mathcal{A}}(\cdot)$	Instantaneous cost of attacker to launch attack
$c_{\mathcal{D}}(\cdot)$	Instantaneous cost of defender to recapture compromised resources
$Q(t)$	Information demand of attacker at time t
$\mu_i(t)$	Information sold by the insider i at time t
$f(\cdot)$	The function determining the information demand of attacker based on the attack rate
$p(t)$	The information price
$\widehat{p}(t)$	The nominal price of information
$\dot{p}(t)$	The variation of price at each time point
$C(\cdot)$	The risk of being detected by defender when selling information
$\pi_i(\cdot)$	Instantaneous profit of insider

The state of the system is directly driven by the actions of the defender ($\mathcal{D}$) and APT attacker ($\mathcal{A}$), and evolves according to the following dynamics,

$$\dot{x}(t) = \alpha \cdot (1 - x(t)) - \beta \cdot x(t), \text{ and } x(0) = x_0, \quad (1)$$

where $\dot{x}(t)$ is the evolving rate of the system state at each time point, $\alpha \in [0,1]$ is the attack rate at which the attacker grasp the resources, $\beta \in [0,1]$ (which is related with the amount of traded information from the insiders, and to be discussed in Sec. III-C) is the recapture rate of defender to recover the compromised resource. $(1 - x(t))$ is the percentage of resources under the defender's control, thus $\alpha \cdot (1 - x(t))$ denotes the percentage of resources seized by the attacker at time t; $x(t)$ is the fraction of compromised resources, thus $\beta \cdot x(t)$ is the percentage of recaptured resources by defender. The system boots up with an initial state x_0 at time 0.

If the control rates α and β of the attacker and defender stay constant, it indicates that they have no feedback on the system states during the runtime of the system, and cannot adapt their strategies accordingly. We take account of this scenario in the *static case* in Sec. IV.A. The more practical situation is that the control actions of the attacker and defender may vary based on the state of the system at different time t, which can be denoted as $\alpha(t)$ and $\beta(t)$. This scenario will be addressed in the *dynamic case* in Sec. IV.B.

B. Cost Model for Attacker and Defender

After the deployment of the system, the defender and attacker will take a series of actions to minimize their own cost over a long time-span. We model the costs for attacker and defender as follows.

Attacker: For APT attacker, it can launch attacks to compromise the resource. By definition [1], APT attacker intends to gain malicious benefits for a *long term* from the targeted system. Hence, it behaves *stealthily* so as to avoid being caught by the defender's detection. Its instantaneous cost should be

composed of two parts: a) the risk of being detected by the defender, which is related to the attack rate α; and b) the portion of secure resource, the complement of which is the compromised resource or the attacker's utility gain. We use quadratic cost model, which is widely used [16], to characterize the costs. The instantaneous cost of the attacker is as follows,

$$c_{\mathcal{A}}(x(t),\alpha,\beta,t) = r_{\mathcal{A}}(1-x(t))^2 + q_{\mathcal{A}}\alpha^2(1-x(t))^2, \quad (2)$$

where $r_{\mathcal{A}}$ and $q_{\mathcal{A}}$ are unit costs, which are positive constant values for the secure resource and the risk of being caught. The instantaneous cost function of the former one depends on the state of the system, which is depicted by the first part of right hand side of Eq. (2), *i.e.*, $r_{\mathcal{A}}(1-x(t))^2$. And $q_{\mathcal{A}}\alpha^2(1-x(t))^2$ is the expected cost of being detected by the defender when launching attacks at time t.

Defender: For the defender, we assume that it can continuously scan *part* of the system[3]. Once compromised resources are detected, the defender recaptures the resources, *e.g.*, changing the password, refreshing the virtual machine, etc. The defender's objective is to minimize the damage brought by the compromised resources. There are two components for the defender's instantaneous cost: a) the operational cost to scan and recapture the compromised resources, which is related to the recapture rate β; and b) the damage of the compromised resources. Again, quadratic model is utilized to model the costs. The instantaneous cost of the defender is modeled as follows,

$$c_{\mathcal{D}}(x(t),\alpha,\beta,t) = r_{\mathcal{D}}x(t)^2 + q_{\mathcal{D}}\beta^2 x(t)^2, \quad (3)$$

where $r_{\mathcal{D}}$ and $q_{\mathcal{D}}$ are unit cost, which are constant values. As shown in (3) by the part $r_{\mathcal{D}}x(t)^2$, the cost of the compromised resources $c_{\mathcal{D}}$ is related to the system status. $q_{\mathcal{D}}\beta^2 x(t)^2$ is the cost when the defender takes action to recapture the compromised resources at rate β.

C. Profit Model for Insiders

Under the APT scenario, it is a common practice for the attacker to obtain the foothold inside the system for future attacks [1]. As APT attacker is well-funded and the insiders always pursue higher monetary profits, it is an efficient approach for the attacker to purchase confidential information (such as passwords, etc.) from insiders in order to launch attacks. Let $Q(t)$ be the total information demanded by the attacker at time t. It can be determined by the attacker's attack rate $\alpha(t)$, with $Q(t) = f(\alpha(t))$. Here, we consider a general model of function $f(\cdot)$, which is a non-decreasing function of $\alpha(t)$. Let the information sold by the insider i at time t be $u_i(t) \geq 0$.

According to the linear inverse demand function [17] , the nominal price of information at time t can be evaluated as follows,

$$\widehat{p}(t) = A(Q(t)) - \sum_{i=1}^{n} u_i(t), \quad (4)$$

[3]As the system under APT scenario is commonly huge and complex, it is too expensive for the defender to scan the whole system

where $A(\cdot)$ is a non-decreasing function of $Q(t)$ and n is the number of insiders in the current system. For simplicity, we denote $A(\cdot)$ as A. However, the nominal price is not the current market price, at which the inside information is traded. The reason is that, the market price is commonly sticky [18] and cannot converge to the nominal price immediately with the real-time updates on the information demand of the attacker, *i.e.*, $Q(t)$, and available information at insiders, *i.e.*, $\sum_{i=1}^{n} u_i(t)$. According to [19], the market price evolves with the following dynamics,

$$\dot{p}(t) = s(\widehat{p}(t) - p(t)), \quad (5)$$

where $s \in [0,\infty)$ is a constant value determined by the stickiness of the market, and controls the convergence speed to the nominal price.

As the defender monitors the system continuously, insiders must take the risk, of being detected by the defender when they sell information to the attacker. The instantaneous risk (cost) function [20] of each insider is defined as follows,

$$C(u_i(t)) = cu_i(t) + \frac{1}{2}u_i(t)^2, \ \forall i \in [1,n], \quad (6)$$

where c is the unit risk (cost). c is under the influence of the scan strategy of defender, that means if the defender adopts more active detection strategy, the risk of selling information for the insiders will be larger.

The instantaneous monetary gain for each insider by trading inside information is $p(t)u_i(t)$. Hence, the instantaneous net profit of the insider i is

$$\pi_i(t) = p(t)u_i(t) - (cu_i(t) + \frac{1}{2}u_i(t)^2). \quad (7)$$

IV. Game between Attacker and Defender

In this section, we present the solution to the defense/attack game between the APT attacker and the defender in two settings: a) the actions of each player are static, *i.e.*, each player cannot change their action to adapt the state of the system; and b) the actions of each player are dynamic, which means each player can dynamically change their control action according to the state of the system to take the optimal actions.

A. Static Actions

In this subsection, we address the scenario where the rate of each action is constant. All the rates are pre-configured before the deployment of the system.

▷ For the attacker, it would try its best to grab as much resource as possible with the least cost of attacking over the long term. Based on the instantaneous cost Eq. (2), the cost function of attacker can be modeled as follows,

$$J_{\mathcal{A}}(\alpha,\beta) = \lim_{T\to\infty}\frac{1}{T}\int_0^T c_{\mathcal{A}}(x(t),\alpha,\beta,t)dt. \quad (8)$$

▷ The target of the defender is to protect its resources from the attacker while minimizing the cost of recapturing the

compromised resources. Based on Eq. (3), the cost function of defender can be modeled as follows,

$$J_{\mathcal{D}}(\alpha,\beta)=\lim_{T\to\infty}\frac{1}{T}\int_0^T c_{\mathcal{D}}(x(t),\alpha,\beta,t)dt. \tag{9}$$

Accordingly, we have the definition of the Nash Equilibrium for this game as follows.

Definition 1: Consider the attacker-defender game defined by the system dynamics (1) and the cost functions (8) and (9). A set of strategies $\{\alpha^*,\beta^*\}$ constitutes a Nash Equilibrium if and only if

$$J_{\mathcal{A}}(\alpha^*,\beta^*)\le J_{\mathcal{A}}(\alpha,\beta^*),$$
$$J_{\mathcal{D}}(\alpha^*,\beta^*)\le J_{\mathcal{D}}(\alpha^*,\beta).$$

In the dynamic system, we have assumed that the whole system is fully under defender's control at the beginning, *i.e.*, $x(0)=0$. As Eq. (1) is a first order nonhomogeneous differential equation with constant coefficients, there exists the general solution to it. By solving it, we can derive the state of system at time t as follows,

$$x(t)=\frac{\alpha}{\alpha+\beta}(1-e^{-(\alpha+\beta)t}). \tag{10}$$

By substituting Eq. (10) into Eq. (8), we can get the cost function of the attacker as follows,

$$C_{\mathcal{A}}(\alpha)=(r_{\mathcal{A}}+q_{\mathcal{A}}\cdot\alpha^2)(\frac{\beta}{\alpha+\beta})^2.$$

The aim of the attacker is to minimize its cost function through an optimal action. Hence, the optimization problem for the attacker is

$$\min_{\alpha}\ \ C_{\mathcal{A}}(\alpha)$$
$$s.t.\ \ \alpha\in[0,1]$$

We solve the optimal control for the attacker with the following lemma.

Lemma 1: The best strategy of the attacker to the defender is,

$$\alpha^*=\begin{cases}\frac{r_{\mathcal{A}}}{q_{\mathcal{A}}\cdot\beta} & r_{\mathcal{A}}/q_{\mathcal{A}}\le\beta,\\ 1 & r_{\mathcal{A}}/q_{\mathcal{A}}>\beta.\end{cases}$$

Proof: Taking the derivative of $C_{\mathcal{A}}(\alpha)$ w.r.t. α,

$$\frac{dC_{\mathcal{A}}}{d\alpha}=\frac{2\beta^2}{(\alpha+\beta)^3}(q_{\mathcal{A}}\cdot\alpha\cdot\beta-r_{\mathcal{A}})$$

Let $\alpha=0$, $\frac{dC_{\mathcal{A}}}{d\alpha}$ is negative. Thus there are only two situations for $\frac{dC_{\mathcal{A}}}{d\alpha}$: (1) staying negative; or (2) striking the horizontal axis once and only once at the point $\alpha=\frac{r_{\mathcal{A}}}{q_{\mathcal{A}}\cdot\beta}$. For the first situation, it is obvious that $C_{\mathcal{A}}(\alpha)$ is nonincreasing, thus it reaches the minimum at the point $\alpha=1$. For the second situation, if $\frac{r_{\mathcal{A}}}{q_{\mathcal{A}}\cdot\beta}\le 1$, $C_{\mathcal{A}}(\alpha)$ reaches the minimum at the point $\alpha=\frac{r_{\mathcal{A}}}{q_{\mathcal{A}}\cdot\beta}$, otherwise, it arrives at the minimum at $\alpha=1$. ■

Similarly, The optimal control strategy for the defender is found with the following lemma.

Lemma 2: The best strategy of the defender to the attacker is,

$$\beta^*=\begin{cases}\frac{r_{\mathcal{D}}}{q_{\mathcal{D}}\cdot\alpha} & r_{\mathcal{D}}/q_{\mathcal{D}}\le\alpha,\\ 1 & r_{\mathcal{D}}/q_{\mathcal{D}}>\alpha.\end{cases}$$

Based on the Definition 1 and Lemmas 1 and 2, we can derive the existence of Nash Equilibrium in the static case with the following theorem.

Theorem 1: In static case, the Nash Equilibrium will fall into the following four scenarios: (1) $\alpha^*=r_{\mathcal{A}}/q_{\mathcal{A}}$ and $\beta^*=1$, when $r_{\mathcal{A}}/q_{\mathcal{A}}<r_{\mathcal{D}}/q_{\mathcal{D}}<1$; (2) $\alpha^*=1$ and $\beta^*=r_{\mathcal{D}}/q_{\mathcal{D}}$, when $r_{\mathcal{D}}/q_{\mathcal{D}}<r_{\mathcal{A}}/q_{\mathcal{A}}<1$; (3) α^* and β^* are on the curve $\alpha^*\cdot\beta^*=r/q$, when $r_{\mathcal{A}}/q_{\mathcal{A}}=r_{\mathcal{D}}/q_{\mathcal{D}}=r/q<1$; and (4) $\alpha^*=1$ and $\beta^*=1$, when $r_{\mathcal{D}}/q_{\mathcal{D}}>1$ and $r_{\mathcal{A}}/q_{\mathcal{A}}>1$.

Proof: The definition of Nash Equilibrium is the intersection of each player's best response. Thus it is obvious to derive this Theorem from Lemmas 1 and 2 ■

Remarks: The instantaneous cost of each player includes two categories: the cost incurred from uncontrolled resources and the cost of actions. The constant values r_i and q_i (where $i\in\{\mathcal{A},\mathcal{D}\}$) can be regarded as weights of each kind of cost. Hence, $r_{\mathcal{A}}/q_{\mathcal{A}}<r_{\mathcal{D}}/q_{\mathcal{D}}$ indicates that the defender considers the first kind of cost more important than the second compared with that of the attacker. Thus, the defender will tend to decrease the compromised resources through recapturing at a high rate. As its counterpart, $r_{\mathcal{A}}/q_{\mathcal{A}}>r_{\mathcal{D}}/q_{\mathcal{D}}$ will result in a higher attack rate by the APT attacker.

B. Dynamic Actions

In this subsection, we address the more challenging yet practical scenario where the actions of the attacker and defender could change along with the state of the system. In this scenario, the strategies of each player is more flexible as it could change its actions continuously to achieve its long-term goal. Here, the actions of the attacker and defender are denoted as $\alpha(t)$ and $\beta(t)$, respectively, to emphasize that they could change with time.

The cost function of each player in the dynamic case is similar with that of the static case except that we only consider the finite time horizon. The aim of the attacker is to grab as much resource as possible with the minimum cost in the entire time-span. Its cost function could be defined as follows,

$$J_{\mathcal{A}}(\alpha(t),\beta(t))=\int_0^T c_{\mathcal{A}}(x(t),\alpha(t),\beta(t),t)dt. \tag{11}$$

As contrary, the defender's goal is to recapture the compromised resources with the minimum cost in this time-span, *i.e.*, the defender minimizes the following cost function,

$$J_{\mathcal{D}}(\alpha(t),\beta(t))=\int_0^T c_{\mathcal{D}}(x(t),\alpha(t),\beta(t),t)dt. \tag{12}$$

The Nash Equilibrium of this dynamic defense/attack game can be defined as follows.

Definition 2: Consider the defense/attack game defined by the system dynamics (1) and the cost functions (11) and

(12). A set of strategies $\{\alpha^*(t), \beta^*(t)\}$ constitutes a Nash Equilibrium if and only if

$$J_{\mathcal{A}}(\alpha^*(t), \beta^*(t)) \leq J_{\mathcal{A}}(\alpha(t), \beta^*(t)),$$
$$J_{\mathcal{D}}(\alpha^*(t), \beta^*(t)) \leq J_{\mathcal{D}}(\alpha^*(t), \beta(t)).$$

The necessary conditions for the existence of the Nash Equilibrium and its corresponding optimal strategy for each player are derived in the following theorem.

Theorem 2: Consider the defense/attack game defined by the system dynamics (1) and the cost functions (11) and (12). If a set of strategies $\{\alpha^*(t), \beta^*(t)\}$ constitutes a Nash Equilibrium, and $x^*(t), 0 \leq t \leq T$ is the corresponding state trajectory, there exist two costate functions $\lambda_{\mathcal{A}}(t)$ and $\lambda_{\mathcal{D}}(t)$, such that

$$\dot{x}^*(t) = -\frac{\lambda_{\mathcal{A}}(t)}{2q_{\mathcal{A}}} - \frac{\lambda_{\mathcal{D}}(t)}{2q_{\mathcal{D}}}, \quad x^*(0) = x_0, \tag{13}$$

$$\dot{\lambda}_{\mathcal{A}}(t) = 2r_{\mathcal{A}} \cdot (1 - x^*(t)) + \frac{\lambda_{\mathcal{A}}(t) \cdot \lambda_{\mathcal{D}}(t)}{2q_{\mathcal{D}} \cdot x^*(t)},$$
$$\lambda_{\mathcal{A}}(T) = 0, \tag{14}$$

$$\dot{\lambda}_{\mathcal{D}}(t) = -2r_{\mathcal{D}} \cdot x^*(t) - \frac{\lambda_{\mathcal{A}}(t) \cdot \lambda_{\mathcal{D}}(t)}{2q_{\mathcal{A}} \cdot (1 - x^*(t))},$$
$$\lambda_{\mathcal{D}}(T) = 0, \tag{15}$$

and the optimal actions of each player should fulfill

$$\alpha^*(t) = -\frac{\lambda_{\mathcal{A}}(t)}{2q_{\mathcal{A}} \cdot (1 - x^*(t))}, \tag{16}$$

$$\beta^*(t) = \frac{\lambda_{\mathcal{D}}(t)}{2q_{\mathcal{D}} \cdot x^*(t)}. \tag{17}$$

Proof: Using the Pontryagin minimum principle, the Hamiltonian function for each player is defined as follows,

$$H_{\mathcal{A}}(x(t), \alpha(t), \beta(t), \lambda_{\mathcal{A}}(t)) = r_{\mathcal{A}}(1 - x(t))^2 \tag{18}$$
$$+ q_{\mathcal{A}}\alpha(t)^2(1 - x(t))^2 + \lambda_{\mathcal{A}}(t)[\alpha(t)(1 - x(t)) - \beta(t)x(t)],$$
$$H_{\mathcal{D}}(x(t), \alpha(t), \beta(t), \lambda_{\mathcal{D}}(t)) = r_{\mathcal{D}}x(t)^2 + q_{\mathcal{D}}\beta(t)^2x(t)^2$$
$$+ \lambda_{\mathcal{D}}(t)[\alpha(t)(1 - x(t)) - \beta(t)x(t)]. \tag{19}$$

Once the Hamiltonian function has been constructed, the optimal action of each player should satisfy

$$\alpha(t) = arg\min H_{\mathcal{A}}(x(t), \alpha(t), \beta(t), \lambda_{\mathcal{A}}(t)),$$
$$\beta(t) = arg\min H_{\mathcal{D}}(x(t), \alpha(t), \beta(t), \lambda_{\mathcal{D}}(t)).$$

To solve the above problem, we take the second order partial derivative of Eqs. (18) and (19) w.r.t. $\alpha(t)$ and $\beta(t)$, then we get $\frac{\partial^2 H_{\mathcal{A}}}{\partial \alpha^2} = 2q_{\mathcal{A}}(1 - x(t))^2 \geq 0$ and $\frac{\partial^2 H_{\mathcal{D}}}{\partial \beta^2} = 2q_{\mathcal{D}}x(t)^2 \geq 0$. Due to the convexity, take the partial derivative Eqs. (18) and (19) w.r.t. $\alpha(t)$ and $\beta(t)$ and let $\frac{\partial H_{\mathcal{A}}}{\partial \alpha} = 0$ and $\frac{\partial H_{\mathcal{D}}}{\partial \beta} = 0$, then we could obtain the unique solutions (16) and (17). According to the Pontryagin minimum principle, the costate function should satisfy

$$\dot{\lambda}_i(t) = -\frac{\partial}{\partial x} H_i(x(t), \alpha(t), \beta(t), \lambda_i(t)) \tag{20}$$

where $i \in \{\mathcal{A}, \mathcal{D}\}$. By substituting Eqs. (16) and (17) into Eqs. (1) and (20), we could derive Eqs. (13)-(15). ■

Remarks: According to Theorem 2, the optimal actions of the attacker and defender are only related to the system state $x(t)$ and system parameters. Each player does not need to know its opponent's action. Hence, our solution can model the unique feature of *stealthy* behavior in APT.

A series of conditions in Theorem 2 are the necessary conditions that optimal actions of the attacker and defender must satisfy. They could be applied to generate the candidate solutions. Further, we require there are no conjugate points in the set of Eqs. (13), (14) and (15). Next, we prove the existence of the Nash Equilibrium for our dynamic game as follows.

Theorem 3: Nash Equilibrium exists in the attacker-defender game.

Proof: Substituting the candidate solutions (16) and (17) into Eqs. (18) and (19) respectively, we could obtain the Hamiltonian functions $H_{\mathcal{A}}(x(t), \lambda_{\mathcal{A}}(t), \lambda_{\mathcal{D}}(t))$ and $H_{\mathcal{D}}(x(t), \lambda_{\mathcal{A}}(t), \lambda_{\mathcal{D}}(t))$ which do not contain actions of attacker and defender. Taking second order partial derivative of the Hamiltonian functions w.r.t. $x(t)$, we can obtain $\frac{\partial^2 H_{\mathcal{A}}}{\partial x^2} = 2r_{\mathcal{A}} > 0$ and $\frac{\partial^2 H_{\mathcal{D}}}{\partial x^2} = 2r_{\mathcal{D}} > 0$. Since the Hamiltonian functions are convex w.r.t. $x(t)$, the candidate solution constitutes the Nash Equilibrium [21]. ■

Even though Theorems 2 and 3 provide the necessary and sufficient conditions for the solution of game between attacker and defender, the optimal trajectory of system state $x^*(t)$ and the optimal actions of attacker and defender, *i.e.*, $\alpha^*(t)$ and $\beta^*(t)$, cannot be solved explicitly since the differential equations of state and costate are unsolvable. Here, we provide an iterative numerical solution which is based on the *steepest descent* method to solve the game between the attacker and defender.

We first divide the time interval $[0, T]$ into N subintervals $[t_1, t_2, \cdots, t_N]$. Based on Theorem 2 and its proof, we can find that the optimal actions of attacker and defender, *i.e.*, α^* and β^*, are obtained when both partial derivatives $\frac{\partial H_{\mathcal{A}}}{\partial \alpha} = 0$ and $\frac{\partial H_{\mathcal{D}}}{\partial \beta} = 0$ because of the convexity. Now supposing the partial derivatives are not equal to 0, *i.e.* the actions are not optimal, we should update each action at the direction of its steepest descent in each iteration, which means the i^{th} round of action should be updated to the $(i+1)^{\text{th}}$ round as follows,

$$\alpha^{(i+1)}(t_k) = \alpha^{(i)}(t_k) - \tau \cdot \frac{\partial H_{\mathcal{A}}}{\partial \alpha}, \tag{21}$$

$$\beta^{(i+1)}(t_k) = \beta^{(i)}(t_k) - \tau \cdot \frac{\partial H_{\mathcal{D}}}{\partial \beta}, \tag{22}$$

where τ is the step size and $k = [1, 2, \cdots, N]$. The iteration terminates when $|\frac{\partial H_{\mathcal{A}}}{\partial \alpha}| < \epsilon$ and $|\frac{\partial H_{\mathcal{D}}}{\partial \beta}| < \epsilon$, where ϵ is the error tolerance. In order to update the action of attacker and defender, the state and costates trajectory $\{x^{(i)}(t_k)\}$, $\{\lambda_{\mathcal{A}}^{(i)}(t_k)\}$ and $\{\lambda_{\mathcal{D}}^{(i)}(t_k)\}$ in the i^{th} round should be calculated first. We can numerically integrate the state differential equations (1) from time t_1 to t_N, with the initial state $x^{(i)}(0) = x_0$. As the costates at time $t = 0$ are unknown, the differential equation of costates (20) should be integrated backward from time t_N to t_1. The initial costates $\lambda_{\mathcal{A}}^{(i)}(t_N)$ and $\lambda_{\mathcal{D}}^{(i)}(t_N)$ can be calculated

according to Eqs. (16) and (17). Algorithm 1 summarizes this numerical approach.

Algorithm 1 Steepest Descent based Algorithm

Require: initial system status x_0.
Ensure: $\alpha^*(t), \beta^*(t), x^*(t)$.

1: Dividing the time interval $[0, T]$ into N subintervals. Randomly generating the initial controls of attacker and defender at each time slot: $\{\alpha^0(t_k)\}$ and $\{\beta^0(t_k)\}$, where $k = 1, 2, \cdots, N$.
2: **while** $|\frac{\partial H_{\mathcal{A}}}{\partial \alpha}| > \epsilon$ **or** $|\frac{\partial H_{\mathcal{D}}}{\partial \beta}| > \epsilon$ **do**
3: Integrating the system state dynamics from 0 to T according to **Eq**. 1 with the initial status $x^{(i)}(0) = x_0$ and store the the trajectory $\{x^{(i)}(t_k)\}$, where $k = 1, 2, \cdots, N$.
4: Integrating the costates backward according to Eq. (20) with the initial value $\lambda_{\mathcal{A}}^{(i)}(t_N) = -2q_{\mathcal{A}} \cdot \alpha^{(i)}(t_N)[1 - x^{(i)}(t_N)]$ and $\lambda_{\mathcal{D}}^{(i)}(t_N) = 2q_{\mathcal{D}} \cdot \beta^{(i)}(t_N)x^i(t_N)$, storing the costates $\{\lambda_{\mathcal{A}}^{(i)}(t_k)\}$ and $\{\lambda_{\mathcal{D}}^{(i)}(t_k)\}$, where $k = 1, 2, \cdots, N$.
5: Updating the controls of attacker and defender according to Eqs. (21) and (22).
6: **end while**

Theorem 4: Once the the system becomes steady, the best strategies of the attacker and defender in dynamic the case are as follows,

$$\alpha_s^* = \begin{cases} \frac{r_{\mathcal{A}}}{q_{\mathcal{A}} \cdot \beta_s} & r_{\mathcal{A}}/q_{\mathcal{A}} \leq \beta_s, \\ 1 & r_{\mathcal{A}}/q_{\mathcal{A}} > \beta_s. \end{cases} \tag{23}$$

$$\beta_s^* = \begin{cases} \frac{r_{\mathcal{D}}}{q_{\mathcal{D}} \cdot \alpha_s} & r_{\mathcal{D}}/q_{\mathcal{D}} \leq \alpha_s, \\ 1 & r_{\mathcal{D}}/q_{\mathcal{D}} > \alpha_s. \end{cases} \tag{24}$$

where α_s and β_s are the actions in the steady state of system. The stable system state is as follows,

$$x_s = \frac{\alpha_s}{\alpha_s + \beta_s}, \tag{25}$$

Proof: When the system is under steady state, the fraction of compromised resources will not change any more, *i.e.* $\dot{x}(t) = 0$. According to Eq. (1), $\alpha_s \cdot (1 - x_s(t)) - \beta_s \cdot x_s(t) = 0$. Hence, we can obtain $x_s = \alpha_s/(\alpha_s + \beta_s)$.

In the steady state, no one changes its strategy, for the defender, *i.e.*, $\dot{\beta}_s = 0$. By differentiating Eq. (17) w.r.t. t and let it be 0,

$$\dot{\beta}_s = \frac{\dot{\lambda_{\mathcal{D}}} \cdot x_s(t) - \lambda_{\mathcal{D}} \cdot \dot{x}_s(t)}{2q_{\mathcal{D}} \cdot x_s^2(t)} = 0.$$

Substituting Eqs. (15), (1) and $\lambda_{\mathcal{D}} = 2q_{\mathcal{D}}\beta_s x_s(t)$ (according to Eq. (17)) into the above equation, and after simplification, we could obtain

$$-r_{\mathcal{D}} \cdot x_s(t) + q_{\mathcal{D}} \cdot \beta_s^2 \cdot x_s(t) + 2q_{\mathcal{D}} \cdot \alpha_s \cdot \beta_s \cdot x_s(t) - q_d \cdot \alpha_s \cdot \beta_s = 0.$$

■

After substituting $x_s = \alpha_s/(\alpha_s + \beta_s)$ into the above equation and necessary simplifications, we can obtain $-r_{\mathcal{D}} \cdot \alpha_s + q_{\mathcal{D}} \cdot \alpha_s \cdot \beta_s = 0$. Hence, we can arrive at Eq. (24). Similarly, we could prove Eq. (23).

V. Game Among Insiders

In this section we present a model of the interaction among insiders and their best responses to obtain the maximum individual profits over the long term.

Base on the instantaneous profit Eq. (7) of the insider i, its long term profit should be

$$J_i(u_i(t), \boldsymbol{u}_{-i}(t)) = \int_0^T e^{-\rho t}[p(t) \cdot u_i(t) - c \cdot u_i(t) - \frac{1}{2}u_i(t)^2]dt,$$

where ρ is a discount factor with constant value and $\boldsymbol{u}_{-i}$ denotes the actions of all the other insiders.

The aim of the insider i is to maximize its overall profit,

$$\begin{aligned} \max_{u_i} \quad & J_i(u_i(t), \boldsymbol{u}_{-i}(t)) \\ s.t. \quad & \dot{p}(t) = s \cdot (\widehat{p}(t) - p(t)) \\ & p(0) = p_0 \end{aligned}$$

The best strategy to this maximization problem in the game should be a Nash Equilibrium, where no one could increase his profit unilaterally without impairing the others, *i.e.* the amount of information sold by each insider is the best response to the others.

In order to reach the Nash Equilibrium, each insider should find its best response with respect to the others with the following lemma.

Lemma 3: The best response of each insider to the others is as follows,

$$u_i(t) = \begin{cases} p(t) - c - \lambda_i(t) \cdot s & p(t) \geq c + \lambda_i(t) \cdot s, \\ 0, & p(t) < c + \lambda_i(t) \cdot s. \end{cases} \tag{26}$$

Proof: In order to find best strategy of the insider i, we should at first construct the Hamiltonian,

$$\begin{aligned} H_i(p(t), u_i(t), \boldsymbol{u}_{-1}(t), \lambda_i(t)) = e^{-\rho t}[p(t) \cdot u_i(t) - c \cdot u_i(t) \\ - \frac{1}{2}u_i(t)^2 + \lambda_i(t) \cdot s \cdot (A(t) - \sum_{j=1}^{n} u_j(t) - p(t))], \end{aligned}$$

where $\lambda_i(t) = \mu_i(t)e^{\rho t}$ and $\mu_i(t)$ is the costate. Taking the partial derivative of H_i w.r.t. $u_i(t)$, we have that

$$\frac{\partial H_i}{\partial u_i} = e^{-\rho t}(p(t) - c - u_i(t) - \lambda_i(t) \cdot s).$$

If we take the second order partial derivative of H_i w.r.t. $u_i(t)$, we could easily find that it is concave as $\frac{\partial^2 H_i}{\partial u_i^2} < 0$. Thus, let the partial derivative of H_i be 0, we could get the optimal action for the insider i. ■

Lemma 3 provides the necessary conditions that the optimal strategy of each insider i should satisfy. However, the best response of the insider i cannot be determined according to Eq. (26) since $\lambda_i(t)$ remains unknown currently. Thus we should try to either eliminate or determine the value of $\lambda_i(t)$. As $\lambda_i(t)$ keeps changing along with costate $\mu_i(t)$, it is really

hard to get the analytical solution of $\lambda_i(t)$. Thus we try to eliminate it.

Theorem 5: In the stable information market, the amount of information sold by each insider i is

$$u_i^*(t) = \frac{(s+\rho)(A - \sum_{j\neq i} u_j(t) - c)}{3s + 2\rho} \tag{27}$$

Proof: According to the Pontryagin maximum principle, the costate is as follows:

$$\dot{\mu}_i(t) = -\frac{\partial H_i}{\partial p(t)} = e^{-\rho t}[\lambda_i(t) \cdot s - u_i(t)]. \tag{28}$$

By differentiating $\lambda_i(t) = \mu_i(t)e^{\rho t}$, we have that

$$\dot{\lambda}_i(t) = \dot{\mu}_i(t) \cdot e^{\rho t} + \rho \cdot \mu_i(t) \cdot e^{\rho t}. \tag{29}$$

Substituting Eq. (28) into Eq. (29), we have that

$$\dot{\lambda}_i(t) = (s+\rho) \cdot \lambda_i(t) - u_i(t). \tag{30}$$

By differentiating Eq. (26), substituting $\dot{p}_i(t)$ and $\dot{\lambda}_i(t)$ with Eqs. (5) and (30), respectively, substituting $s \cdot \lambda_i(t)$ with $(p(t) - c - u_i(t))$ according to Eq. (26), we have that

$$\begin{aligned}\dot{u}_i(t) &= s[A - \sum_{j=1}^{n} u_j(t) - p(t)] - s[(s+\rho) \cdot \lambda_i - u_i(t)] \\ &= s[A - \sum_{j=1}^{n} u_j(t) - p(t)] - (s+\rho)(p(t) - c - u_i(t)) \\ &\quad + s \cdot u_i(t).\end{aligned}$$

Since the amount of information sold by the insider i will not change at the stationary status, *i.e.* $\dot{u}_i(t) = 0$. Thus we could obtain

$$u_i^*(t) = \frac{(s+\rho)[p(t) - c] - s[A - \sum_{j\neq i} u_j(t) - p(t)]}{s+\rho}. \tag{31}$$

As the price will not change at the stationary status, *i.e.* $\dot{p}(t) = 0$, we could obtain $p(t) = \widehat{p}(t)$. Substituting $p(t)$ in Eq. (31) by $\widehat{p}(t)$ in Eq. (4), we can obtain the optimal amount of information sold by each insider. ■

If we look into the insiders' game, the solution structure of the insiders' game is symmetric. By applying the symmetry $u_i^*(t) = u_j^*(t)$, it is easy to obtain the Nash Equilibrium for the insiders' game according to the Theorem 5.

Corollary 1: The optimal amount pf information that each insider should sell to attacker at the Nash Equilibrium is

$$u_i^*(t) = \frac{(s+\rho)(A-c)}{(N+2)s + (N+1)\rho}. \tag{32}$$

Remarks: According to Eq. (32), it is obvious that the action of each insider is closely related to the information demand of attacker A and the unit risk cost c. The information demand is determined by the attack rate α, which means if the attacker wants to achieve a high attack rate, it should purchase a large amount of confidential information from the insiders to support its attack, *i.e.* large A. The unit risk cost c is determined by the detection strategy of the defender. If the defender scans the system at a higher rate, the insider must take more risk to sell confidential information, *i.e.* c is larger.

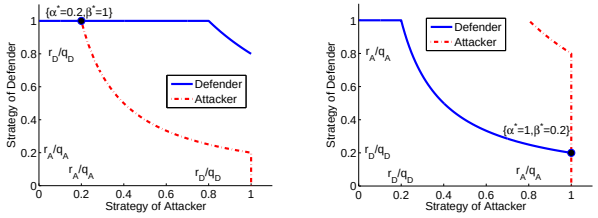
VI. Numerical Study

In this section, we examine our proposed framework with numerical study under different settings of system configurations.

A. Defense-Attack Game in Static Case

For static case, we identify the Nash Equilibria of the defense/attack game in two representative system configurations as follows,

- **Configuration 1**: $r_{\mathcal{A}}/q_{\mathcal{A}} < r_{\mathcal{D}}/q_{\mathcal{D}}$, *e.g.*, $r_{\mathcal{A}} = 2, r_{\mathcal{D}} = 8, q_{\mathcal{A}} = q_{\mathcal{D}} = 10$. With Fig. 2(a), we can see that the attacker and defender can reach the unique equilibrium $\{\alpha^* = 0.2, \beta^* = 1\}$ according to the Theorem 1. In this scenario, the defender achieves full speed of scanning and recapturing. This result echoes our remark in Sec. IV that the defender is more actively taking actions to recapture the compromised resources in this setting.
- **Configuration 2**: $r_{\mathcal{A}}/q_{\mathcal{A}} > r_{\mathcal{D}}/q_{\mathcal{D}}$, *e.g.*, $r_{\mathcal{A}} = 8, r_{\mathcal{D}} = 2, q_{\mathcal{A}} = q_{\mathcal{D}} = 10$. As shown in Fig. 2(b), the attacker and defender can reach the unique equilibrium $\{\alpha^* = 1, \beta^* = 0.2\}$ according to Theorem 1. In this scenario, the attacker launches attacks at its full speed. This result matches our remark in Sec. IV that the attacker will actively compromise the resources in this setting.



(a) $r_{\mathcal{A}}/q_{\mathcal{A}} < r_{\mathcal{D}}/q_{\mathcal{D}}$ (b) $r_{\mathcal{A}}/q_{\mathcal{A}} > r_{\mathcal{D}}/q_{\mathcal{D}}$

Fig. 2. Nash Equilibrium in Static Case

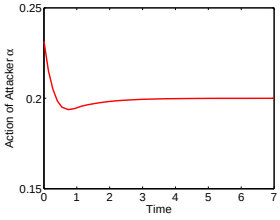
B. Defense-Attack Game in Dynamic Case

As proved in Theorem 4, the Nash Equilibrium under steady state in dynamic case is identical with that in static case. Hence, we selectively present the evolution of both system state and optimal actions for each player in the setting of $r_{\mathcal{A}}/q_{\mathcal{A}} < r_{\mathcal{D}}/q_{\mathcal{D}}$, due to limited space.

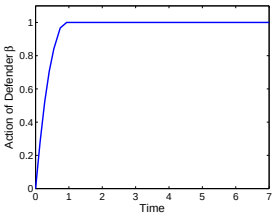
We set $r_{\mathcal{A}} = 2, r_{\mathcal{D}} = 8, q_{\mathcal{A}} = q_{\mathcal{D}} = 10$. The attacker and defender will arrive the unique equilibrium $\{\alpha^* = 0.2, \beta^* = 1\}$ and the system will reach its stable state at $x_s = 0.1667$ according to Theorem 4. We apply the Algorithm 1 to calculate the system state trajectory and the optimal actions of attacker and defender at each time point. We set the step size of action update τ to be 0.001 in each iteration and the error tolerance $\epsilon = 0.01$.

Fig. 3 shows the actions of attacker and defender at each time point by our Algorithm 1. We can find that the action on stable status of attacker, *i.e.*, α^*, is 0.2 and that of defender, *i.e.*, β^*, is 1, which are consistent with the result derived from Theorem 4. Fig. 4 shows how the system state evolves to its

stable status. The stable state is 0.1667, which also matches the result derived from Theorem 4.



(a) Action of attacker



(b) Action of defender

Fig. 3. The actions of attacker and defender

C. Information-Trading Game

As discussed in Section V, the optimal amount of information sold by each insider is affected by the attacker's demand and the defender's detection strategy. We set $A(D(t)) = 500\alpha(t)^2$ since $A(\cdot)$ is an non-decreasing function with respect to $\alpha(t)$, the number of insider $n = 2$ and the unit risk cost $c = 10$. If the converge speed $s \to \infty$ (the current market price converges to the nominal price immediately), Eq. (32) will be $u^*(t) = \frac{A-c}{N+2}$. Fig. 5 shows how the amount of information sold by the insider changes with the attacker's attack rate (as shown in Fig. 3(a)). When the attack rate decreases, the information sold by insiders decreases accordingly.

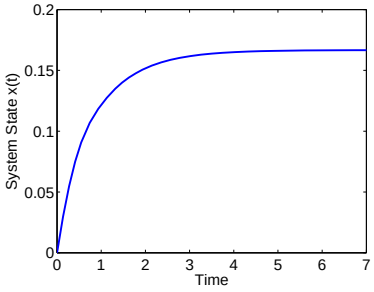


Fig. 4. The evolution of the system state $x(t)$

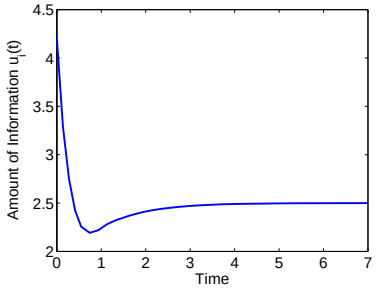


Fig. 5. Insider's best response

VII. Conclusion

This is the first reported literature that investigates the joint threats from APT attacker and insiders. We characterize the interplay among defender, APT attacker and insiders with a two-layer differential game framework, *i.e.*, a defense/attack game between the defender and APT attacker and an information-trading game among the insiders. Through rigorous analysis, we identify the best response strategies for each player via optimizing their long-term objectives, respectively, and prove the existence of the Nash Equilibrium for each game. Extensive numerical study further evaluates the impact of different system configurations on the achievable security level. The results in this paper can shade insights on practical system design for higher security levels facing the joint APT and insider threats.

References

[1] E. Cole, *Advanced Persistent Threat: Understanding the Danger and How to Protect Your Organization*. Newnes, 2012.
[2] "2013 us state of cybercrime survey."
[3] M. B. Salem, S. Hershkop, and S. J. Stolfo, "A survey of insider attack detection research," in *Insider Attack and Cyber Security*. Springer, 2008, pp. 69–90.
[4] N. Falliere, L. O. Murchu, and E. Chien, "W32. stuxnet dossier," *White paper, Symantec Corp., Security Response*, 2011.
[5] D. Alperovitch, *Revealed: operation shady RAT*. McAfee, 2011.
[6] M. van Dijk, A. Juels, A. Oprea, and R. L. Rivest, "Flipit: The game of stealthy takeover," *Journal of Cryptology*, vol. 26, no. 4, pp. 655–713, 2013.
[7] I. J. Martinez-Moyano, E. Rich, S. Conrad, D. F. Andersen, and T. R. Stewart, "A behavioral theory of insider-threat risks: A system dynamics approach," *ACM Transactions on Modeling and Computer Simulation (TOMACS)*, vol. 18, no. 2, p. 7, 2008.
[8] C. Colwill, "Human factors in information security: The insider threat–who can you trust these days?" *Information security technical report*, vol. 14, no. 4, pp. 186–196, 2009.
[9] O. Brdiczka, J. Liu, B. Price, J. Shen, A. Patil, R. Chow, E. Bart, and N. Ducheneaut, "Proactive insider threat detection through graph learning and psychological context," in *Security and Privacy Workshops (SPW), 2012 IEEE Symposium on*. IEEE, 2012, pp. 142–149.
[10] S. Mathew, M. Petropoulos, H. Q. Ngo, and S. Upadhyaya, "A data-centric approach to insider attack detection in database systems," in *Recent Advances in Intrusion Detection*. Springer, 2010, pp. 382–401.
[11] S. Backhaus, R. Bent, J. Bono, R. Lee, B. Tracey, D. Wolpert, D. Xie, and Y. Yildiz, "Cyber-physical security: A game theory model of humans interacting over control systems," *Smart Grid, IEEE Transactions on*, vol. 4, no. 4, pp. 2320–2327, 2013.
[12] T. Alpcan and S. Buchegger, "Security games for vehicular networks," *Mobile Computing, IEEE Transactions on*, vol. 10, no. 2, pp. 280–290, 2011.

[13] T. Alpcan and T. Basar, "A game theoretic approach to decision and analysis in network intrusion detection," in *Decision and Control, 2003. Proceedings. 42nd IEEE Conference on*, vol. 3. IEEE, 2003, pp. 2595–2600.

[14] S. Amin, G. A. Schwartz, and S. Shankar Sastry, "Security of interdependent and identical networked control systems," *Automatica*, vol. 49, no. 1, pp. 186–192, 2013.

[15] R. Anderson and T. Moore, "The economics of information security," *Science*, vol. 314, no. 5799, pp. 610–613, 2006.

[16] T. Basar, G. J. Olsder, G. Clsder, T. Basar, T. Baser, and G. J. Olsder, *Dynamic noncooperative game theory*. SIAM, 1995, vol. 200.

[17] S. G. M. William F. Samuelson, *Managerial Economics, 8th Edition*. Wiley, 2002.

[18] M. Obstfeld, K. S. Rogoff, and S. Wren-lewis, *Foundations of international macroeconomics*. MIT press Cambridge, MA, 1996, vol. 30.

[19] M. Simaan and T. Takayama, "Game theory applied to dynamic duopoly problems with production constraints," *Automatica*, vol. 14, no. 2, pp. 161–166, 1978.

[20] C. Fershtman and M. I. Kamien, "Dynamic duopolistic competition with sticky prices," *Econometrica: Journal of the Econometric Society*, pp. 1151–1164, 1987.

[21] Q. Zhu, L. Bushnell, and T. Basar, "Game-theoretic analysis of node capture and cloning attack with multiple attackers in wireless sensor networks," in *Decision and Control (CDC), 2012 IEEE 51st Annual Conference on*, Dec 2012, pp. 3404–3411.