Defending Against Stealthy Attacks on Multiple Nodes With Limited Resources: A Game-Theoretic Analysis

Ming Zhang , *Student Member, IEEE*, Zizhan Zheng , *Member, IEEE*, and Ness B. Shroff , *Life Fellow, IEEE*

Abstract—**Stealthy attacks have become a major threat to cybersecurity. Previous works in this direction fail to capture the practical resource constraints and mainly focus on one-node settings. In this article, we propose a two-player game-theoretic model, including a system of multiple independent nodes, a stealthy attacker, and an observable defender. In our model, the attacker can fully observe the defender's behavior and the system state, whereas the defender has zero feedback information. Furthermore, a strict resource constraint is introduced to limit the frequency of the attacks/defenses for both players. We characterize the best responses for both attacker and defender under both nonadaptive and adaptive strategies. We then study the sequential game where the defender first announces its strategy and the attacker then responds accordingly. We have designed an algorithm that finds a nearly optimal strategy for the defender and provides a full analysis of its complexity and performance guarantee.**

Index Terms—**Game theory, resource constraints, stealthy attacks.**

I. INTRODUCTION

INCREASINGLY sophisticated cyberattacks constantly push the evolution of cybersecurity. In recent years, worldwide organizations and IT companies, e.g., United Nation, Google, and Amazon, are facing a significantly increasing number of *advanced persistent threats* (APT) [8]. The APT attack has several distinguishing properties that render traditional defense mechanisms less effective. First, they are often launched by *incentive-driven* entities, including government and competitive companies with specific targets. Second, the APT attack is *persistent*, which usually involves multiple stages and frequent compromises of the system. Based on the work in [1], half of the entities suffering APT attacks experienced another successful compromise within one year. Third, they are highly adaptive and *stealthy*, often operating in a "low-and-slow" fashion [14] in order to maintain a small footprint and avoid being detected. In fact, some of the past APT attacks have been so effective because they have gone undetected for months or longer [9], [13]. Hence, conventional security measures against one-shot attack and known attack types are not sufficient in the face of long-lasting and stealthy attacks. Meanwhile, the objective of APT attacks usually includes key information theft and complete control over the system, resulting in a much bigger loss than traditional cyberattacks.

In this article, we study a two-player non-zero-sum game that explicitly models stealthy attacks with resource constraints, as an extension of the asymmetric version of the FlipIt game considered in [21]. We consider a system with N independent nodes (or components), an attacker, and a defender. Both players compete for the control of the system by attacking or defending each node, subject to an instantaneous move cost per node and a long-term average resource constraint across the entire system. The attacker tries to maximize its benefits by successfully compromising nodes, and the defender aims at minimizing the total defense cost and value loss incurred by losing control of a node.

To model the stealthy attacks, we assume that the defender has no feedback about the node state and the attacker's behavior across the entire game, which is reasonable in many security setups. On the other hand, the attacker is capable of observing the defender's each move as well as the node state, and makes decisions accordingly. In this work, we consider two commonly adopted solution concepts: 1) Nash equilibrium and 2) sequential equilibrium, both of which have been applied to cybersecurity. In the former, the defender and the attacker determine their strategies at the beginning of the game simultaneously, whereas in the latter, the defender acts as the leader of the game and commits to a strategy first, and the attacker as the follower then responds accordingly.

For tractability and simplifying the analysis, we assume that the set of nodes is *independent* in the sense that the proper

TABLE I
MAIN RESULTS

	Attacker	Defender
Best Response	*i.i.d.* attack $\leftrightarrows$ periodic defense	
	Markovian attack $\rightarrow$ periodic defense	
	i.i.d. attack $\not\leftarrow$ Markovian defense	
Nash Equilibrium	A complete characterization of NEs (6 types)	
Sequential Game	Optimal attack under a given defense strategy (16)	A polynomial time algorithm for optimal defense (Algorithm 1)

functioning of one node does not depend on other nodes, which serves as a first-order approximation of the more general setting of interdependent nodes to be considered in our future work. Despite the assumption that each node is independent, the multinode setting, together with the resource constraints, impose significant challenges in characterizing the best responses, Nash equilibria and sequential equilibria of the games.

One example where our game model can be applied is key rotation. For a system with multiple communication links or servers that are protected by different keys, an APT attacker may compromise some of the keys from time to time. A common practice is to periodically generate fresh keys by a trusted key-management service, without knowing when they are compromised. On the other hand, the attacker can easily detect when the key expires with a negligible cost and there is a constraint on the frequency of moves at both sides. There are also other examples where our model can be useful such as password reset and virtual machine refreshing [15], [21], [30].

To help reader better understand our main results, we briefly explain the key concepts ahead. Formal definitions can be found in Sections III and IV.

1) In a *periodic defense* strategy: The defender protects each node periodically. That is, the time interval between two consecutive defenses is fixed for a given node.
2) In an independent identically distributed (i.i.d.) attack strategy: The attacker's waiting time before each attack (modeled as a random variable) is i.i.d. across time.
3) In a *Markovian defense (resp. attack)* strategy: The time interval between two consecutive defenses (resp. the probability of attacking each node) follows a Markovian process.
4) Nearly optimal strategy: For arbitrary small positive number ϵ, we can always find a strategy that the performance difference between this strategy and the theoretical optimal strategy is less than ϵ.

We have made the following contributions in this article with the main results summarized in Table I.[1]

1) We propose a two-player game model with multiple independent nodes, an overt defender, and a stealthy attacker where both players have strict resource constraints.
2) We prove that periodic defense is a best response against i.i.d. attack among *all* defense strategies, and i.i.d. attack is a best response against periodic defense among *all* attack strategies. We further consider Markovian strategies

and prove that periodic defense is still a best response against a Markovian attacking strategy, but i.i.d. attack is not necessarily a best response against a Markovian defending strategy.
3) For the pair of periodic defense and i.i.d. attack strategies, we fully characterize the set of Nash equilibria of our game, and show that there is always one (maybe more) equilibrium, when the attack times are deterministic.
4) We further consider the sequential game with the defender as the leader and the attacker as the follower. We design a dynamic programming based algorithm that identifies a nearly optimal strategy (in the sense of subgame perfect equilibrium) for the defender. We also fully characterize the tradeoff between algorithm performance and its complexity.

This article is the extended version of the work in [34] with the addition of adaptive strategies and Markovian strategies in Section IV and further improvement of Algorithm 1 in Section VI. The remainder of this article is organized as follows. A summary of related work is provided in Section II. We present our game-theoretic model in Section III, and study best-response strategies of both players in Section IV. Analysis of Nash equilibria of the game is provided in Section V, and the sequential game is studied in Section VI. In Section VII, we present numerical results, and conclude the article in Section VIII.

II. RELATED WORK

Game theory has been extensively applied to cybersecurity and network security [11], [19], [24], [29]. However, traditional models mainly focus on known attacks and largely ignore the budget constraints of both the defender and the attacker.

As mentioned in the introduction, our model is inspired by the FlipIt game [15], [30] proposed in response to an APT attack toward RSA Data Security [10], a non-zero-sum dynamic game that explicitly models the stealthy takeover of a single node. In the original model, a player obtains control over a component instantaneously by "flipping" it, and obtains feedback only when it moves. Dominant strategies or strongly dominant strategies are characterized for several classes of periodic and renewable strategies and some simple adaptive strategies. But the full analysis of Nash equilibrium is only provided when both the defender and the attacker employ a periodic strategy with a random starting phase. Several variants of the basic model have been studied [20], [21]. In particular, a multinode extension is considered in [20] where the attacker needs to compromise either all the nodes (AND model) or a single node (OR model) to take over a system. The authors name such a model as "FlipThem." However, only preliminary analytic results are provided. Leslie *et al.* extend the "FlipThem" model in [22] and [23] where the attacker can obtain partial benefits by compromising a certain number (larger than a threshold) of nodes. An asymmetric model similar to ours where the attacker is stealthy while the defender is observable is considered in [21], where a full Nash equilibrium analysis is provided but only for the single-node setting. In [26], Nochenson *et al.* first initiate the effort of adding player's characterization information including gender

---

[1] In Table I $A \rightarrow B$ means that B is a best response against A; $A \not\rightarrow B$ means that B is NOT a best response against A.

and age in the FlipIt game model. In [35], Zheng *et al.* use multiarmed bandit model to investigate the optimal timing of security updates against stealthy attacks. However, none of the previous works considered an explicit resource constraint on the players.

A different type of security game has also been studied in the literature mainly for protecting physical infrastructures [12], [18], [28], [29]. Essentially, a mixed strategy Stackelberg game is considered, where the defender is the leader and the attacker is the follower. The key assumption is that the defender first decides upon a randomized defense policy, and the attacker then observes the randomized policy of the defender but not its realization before taking action. While this is a useful assumption under certain scenarios, it may not hold when the attacker is highly adaptive. In particular, since the attacker may be able to observe the defender's previous actions, it could take an action before the defender changes its policy to obtain greater benefit. Moreover, the two-stage game is insufficient to capture the persistent and stealthy behaviors of advanced attacks. Despite the fundamental differences of the two models, recent work that extends this model to multiple defenders and bounded rationality [17], [25] provides useful insights to our model as well, which will be studied in our future work.

III. GAME MODEL

In this section, we discuss our two-player game model including its information structure, the action spaces of both attacker and defender, and their payoffs. Our game model extends the single-node model in [21] to multiple nodes and includes a resource constraint on each player.

A. Basic Model

In our game-theoretic model, there are two players (the defender and the attacker) and a network of N *independent* nodes.[2] Each node has a value of r_i representing the payoff the attacker can receive per unit time by successfully compromising node i. We consider finite time horizon where the game starts at time $t = 0$ and goes to any time $t = T$. We assume that time is continuous. Every time when the attacker starts an attack for node i, it incurs a cost of C_i^A and takes a random period of time $\alpha_{i,k}$ to succeed. On the other hand, if the defender makes a move to protect node i, the node is immediately recovered and incurring a cost of C_i^D. Furthermore, this information is immediately learned by the attacker. The attacker's strategy is to determine $W_{i,k}$, the waiting time from the defender's kth move to its next attack on node i, for each i and k. On the contrary, the defender's strategy is to determine the time intervals between its $(k-1)$th move and kth move for each node i and k, denoted as $X_{i,k}$. The attacker (defender) can attack (defend) multiple nodes at the same time and maintain their possession until the other player's next move, which may or may not change the node state. It is important to note that when $W_{i,k} \geq X_{i,k}$, the attacker actually gives up its kth attack against node i (this is possible as the attacker can observe when the defender moves).

---

[2] The terms "components" and "nodes" are interchangeable in this article.

Symbol	Meaning
T	time horizon
N	number of nodes
r_i	value per unit of time of compromising node i
$\alpha_{i,k}$	attacking time for node i in the k-th move
C_i^A	attacker's move cost for node i
C_i^D	defender's move cost for node i
$W_{i,k}$	attacker's waiting time in its k-th move for node i
$X_{i,k}$	time between the $(k-1)$-th and the k-th defenses for node i
B	budget to the defender, greater than 0
M	budget to the attacker, greater than 0
m_i	frequency of defenses for node i
p_i	probability of immediate attack on node i
L_i	the number of defense moves for node i

It may also happen that the defender makes a defense during the attack, mathematically $W_{i,k} < X_{i,k} < W_{i,k} + \alpha_{i,k}$. Under such circumstances, we assume that the attacker will drop the attack and may launch another attack after waiting time $W_{i,k+1}$.

Both the attacker's and the defender's strategies can be randomized and adaptive in general. We define the strategy space for the attacker (resp. defender) as all possible $\{W_{i,k}\}$ (resp. $\{X_{i,k}\}$) that follow a certain joint distribution. A strategy is defined as *nonadaptive* if the values of $\{W_{i,k}\}$ (resp. $\{X_{i,k}\}$) are either predetermined or follow a fixed probability distribution. An attack strategy is considered *adaptive* when the attacker's decision on $W_{i,k}$ for some i and k depends on the realized values of $X_{j,k'}$ and/or $W_{j,k'}$ for some j and k' with $T_j(k') < T_i(k)$, where $T_i(k)$ refers to the time instance of node i's kth defense. In other words, an attack strategy is adaptive if the waiting time $W_{i,k}$ depends on previous attack or defense moves. On the other hand, since the defender cannot observe the attacker's behavior and node state, it suffices to consider nonadaptive defense strategies. That is, the defender's decisions on $\{X_{i,k}\}$ are independent of the realization of $\{W_{i,k}\}$.

In addition to the move cost, we introduce a strict resource constraint for each player, which is a practical assumption but has been ignored in most prior works on security games. In particular, we place an upper bound on the average amount of resource that is available to each player at any time (to be formally defined ahead). As in typical security games, we assume that r_i, C_i^A, C_i^D, the distribution of $\alpha_{i,k}$, and the budget constraints are all common knowledge of the game, that is, they are known to both players. Without loss of generality, all nodes are assumed to be protected at time $t = 0$. Table II summarizes the notations used in the article and Fig. 1 shows the transition of node state along timeline.

As in [21], we consider an asymmetric feedback model where the attacker's moves are *stealthy*, whereas the defenders' moves are *observable*. More specifically, at any time, the attacker knows the full history of the moves made by the defender (the realization of $\{X_{i,k}\}$), as well as the state of each node, whereas the defender does not know whether a node is compromised or not. This asymmetric information structure is crucial in modeling
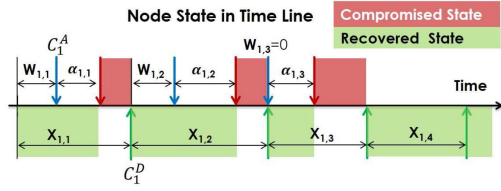
Fig. 1.   Game model.

stealthy attacks in cybersecurity. For clarification, our model is a complete information game setting since we assume that both players are aware of the distribution of the strategies employed by the other player.

B. Defender's Problem

We model the total cost to the defender as the summation of the total time when nodes are compromised and the total move cost. The defender aims at maximizing its payoff, which is defined as the negation of its total loss. Given the attacker's strategy $\{W_{i,k}\}$, the defender faces the following optimization problem:

$$\max_{\{X_{i,k}\},\{L_i\}}$$
$$E\left[\sum_{i=1}^{N}\left(\frac{-\left(T - \sum_{k=1}^{L_i}\min(W_{i,k}+\alpha_{i,k},X_{i,k})\right)\cdot r_i}{T}\right.\right.$$
$$\left.\left. - \frac{L_i C_i^D}{T}\right)\right] \tag{1}$$

and the optimization variable $X_{i,k}$ and L_i satisfy the following two constraints:

$$\sum_{i=1}^{N}\frac{L_i}{T} \le B \text{ with probability } 1$$

$$\sum_{k=1}^{L_i} X_{i,k} \le T \text{ with probability } 1 \; \forall i \tag{2}$$

where L_i (a random variable) is the total number of defense applied to node i during time T. In (1), $T - \sum_{k=1}^{L_i}\min(W_{i,k}+\alpha_{i,k},X_{i,k})$ refers to the total time when node i is compromised and $L_i C_i^D$ is the overall move cost. The first constraint defines an upper bound B of the average number of nodes that can be protected at any time. The second constraint in (2) defines the feasible set of $X_{i,k}$. It is worth mentioning that the actual decision variables of (1) are the joint distribution of $X_{i,k}$ and L_i for all i and k. For the convenience of notation, we simply use $X_{i,k}$ and L_i as decision variables. The expectation is with respect to $X_{i,k}$, L_i, $W_{i,k}$, and $\alpha_{i,k}$.

C. Attacker's Problem

Given the defender's strategy $\{X_{i,k}\}$, the total cost of attacking node i is then $(\sum_{k=1}^{L_i}\phi(W_{i,k},X_{i,k}))\cdot C_i^A$, where $\phi(W_{i,k},X_{i,k})=1$ if $W_{i,k}<X_{i,k}$ and $\phi(W_{i,k},X_{i,k})=0$ otherwise. The attacker's problem can be formulated as follows,

where M is an upper bound on the average number of nodes that the attacker can attack at any time instance

$$\max_{\{W_{i,k}\}} E\left[\sum_{i=1}^{N}\frac{(T - \sum_{k=1}^{L_i}\min(W_{i,k}+\alpha_{i,k},X_{i,k}))\cdot r_i}{T}\right.$$
$$\left. - \frac{(\sum_{k=1}^{L_i}\phi(W_{i,k},X_{i,k}))\cdot C_i^A}{T}\right] \tag{3}$$

and the attacker needs to satisfy the following constraint:

$$E\left[\sum_{i=1}^{N}\frac{1}{T}\int_0^T v_i(t)dt\right] \le M \tag{4}$$

where $v_i(t)=1$ if the attacker is attacking node i at time t and $v_i(t)=0$ otherwise. Note that we make the assumption that the attacker has to keep consuming resources when the attack is in progress. We further have the following equation:

$$\int_0^T v_i(t)dt$$
$$= \sum_{k=1}^{L_i}\left(\min(W_{i,k}+\alpha_{i,k},X_{i,k}) - \min(W_{i,k},X_{i,k})\right). \tag{5}$$

Putting (5) into (3) and (4), and moving the expectation inside, the attacker's problem becomes

$$\max_{\{W_{i,k}\}}\sum_{i=1}^{N}\left[\frac{T\cdot r_i - E[\sum_{k=1}^{L_i}\min(W_{i,k}+\alpha_{i,k},X_{i,k})]\cdot r_i}{T}\right.$$
$$\left. - \frac{E[\sum_{k=1}^{L_i}P(W_{i,k}<X_{i,k})]\cdot C_i^A}{T}\right] \tag{6}$$

with resource constraints

$$\sum_{i=1}^{N}\left[\frac{E[\sum_{k=1}^{L_i}(\min(W_{i,k}+\alpha_{i,k},X_{i,k})-\min(W_{i,k},X_{i,k}))]}{T}\right]$$
$$\le M. \tag{7}$$

IV. BEST RESPONSES

In this section, we analyze the best-response strategies for both players. Our main result is that when the attacker employs an i.i.d. strategy; a periodic strategy is a best response for the defender, and *vice versa*. To prove this result, however, we have provided characterization of best responses in more general settings.

A. Defender's Best Response

We first show that an optimal deterministic defense strategy is always optimal in general for (1). We then prove that the periodic defense is optimal against i.i.d. attacks.

Lemma IV.1: Suppose $x_{i,k}^{\star}$ and $l_i^{\star}$ are the optimal solutions of (1) among all deterministic strategies, then they are also optimal among all the strategies (including both adaptive and nonadaptive strategies).

Proof sketch: For a general defense strategy $X_{i,k}$, we can achieve the same expected payoff for defender by assigning

$X_{i,k} = x^{\star}_{i,k}$ with probability 1 and the same for L_i. Thus, the deterministic strategy is also optimal among all random strategies. For detailed proof, see [33].

According to the lemma, it suffices to consider defender's strategies where both $X_{i,k}$ and $L_{i,k}$ are deterministic. It is also worth mentioning that the order in which nodes are defended makes no difference since the nodes are independent of each other. We then define the set of i.i.d. attack strategies and show that periodic defense is a best response against i.i.d. attacks.

Definition IV.1: An attack strategy is called *i.i.d.* if it is nonadaptive, and $W_{i,k}$ is independent across i and is *i.i.d.* across k.

Theorem IV.1: Assume that for any node i, the attacking time $\alpha_{i,k}$ is *i.i.d.* across k. Then periodic defense is a best response among all defense strategies if the attacker employs an *i.i.d.* strategy.

To prove this result, we need the following definition.

Definition IV.2: For a given L_i, we define a set $\mathcal{X}_i$ that includes all deterministic defense strategies for node i with the following properties:
1) $\sum_{k=1}^{L_i} X_{i,k} = T$;
2) $F_{W_{i,k}+\alpha_{i,k}}(X_{i,k}) = F_{W_{i,j}+\alpha_{i,j}}(X_{i,j}) \ \forall k, j$

where $F_{W_{i,k}+\alpha_{i,k}}(\cdot)$ is the marginal cumulative distribution function of $W_{i,k} + \alpha_{i,k}$. Let $\mathcal{X}$ denote the set of defense strategies where for each node i, a strategy in $\mathcal{X}_i$ is adopted.

Note that 1) $\mathcal{X}_i$ can be an empty set in general due to the randomness of $W_{i,k} + \alpha_{i,k}$; 2) for deterministic $X_{i,k}$, $W_{i,k}$ is independent of any $X_{j,\tau}$ s.t. $T_j(\tau) \geq T_i(k)$. The following lemma shows that when $\mathcal{X}_i$ is nonempty for all i, any strategy that belongs to $\mathcal{X}$ is a defender's best deterministic strategy against a nonadaptive attacker.

Lemma IV.2: Consider a nonadaptive attack strategy. For any given set of $\{L_i\}$ with $\sum_{i=1}^{N} \frac{L_i}{T} \leq B$, if $\mathcal{X}_i \neq \emptyset$ for any i, then any strategy in $\mathcal{X}$ is a best deterministic strategy for the defender.

Proof: We first define the defender's payoff for node i as

$$U_i^D(\{X_{i,k}\})$$
$$= \frac{-\left(T - \sum_{k=1}^{L_i} E[\min(W_{i,k} + \alpha_{i,k}, X_{i,k})]\right) \cdot r_i}{T}$$
$$- \frac{L_i C_i^D}{T}. \tag{8}$$

Since $\{L_i\}$ are fixed, Problem (1) can be divided into N independent subproblems as follows:

$$\max_{\{X_{i,k}\}} U_i^D(\{X_{i,k}\})$$
$$\text{s.t. } \sum_{k=1}^{L_i} X_{i,k} \leq T. \tag{9}$$

We first assume that $F_{W_{i,k}+\alpha_{i,k}}(\cdot)$ is continuous for any i and k. Since any deterministic $X_{i,k}$ is independent of $W_{i,k}$, we can prove that the objective function is concave by showing that the Hessian matrix of $U_i^D(\{X_{i,k}\})$, with respect to $X_{i,k}$ $(1 \leq k \leq L_i)$, is negative semidefinite. We note that even

when $F_{W_{i,k}+\alpha_{i,k}}(\cdot)$ is not continuous, the concavity can still be proved using the subgradient concept. The details are omitted to save space.

Since $U_i^D(\{X_{i,k}\})$ is concave and continuously differentiable, the optimal solution $\{X_{i,k}\}$ should satisfy the KKT conditions, that is, $\nu^{\star}(\sum_{k=1}^{L_i} X_{i,k} - T) = 0$ and $F_{W_{i,k}+\alpha_{i,k}}(X_{i,k}) = F_{W_{i,j}+\alpha_{i,j}}(X_{i,j}) \ \forall k, j$ by taking the derivative of Lagrangian function and setting it equal to zero, where $\nu^{\star}$ is the optimal Lagrangian multiplier. It is clear that $U_i^D(\{X_{i,k}\})$ is maximized when the constraint is tight, that is, $\sum_{k=1}^{L_i} X_{i,k} = T$. Note that there may exist a set of $X_{i,k}$ with $\sum_{k=1}^{L_i} X_{i,k} < T$ that is also optimal for (9). Thus, the two conditions in Definition IV.2 are sufficient but not necessary. ∎

We now prove Theorem IV.1.

Proof: For any fixed $\{L_i\}$, let $X_i \triangleq [\frac{T}{L_i} \frac{T}{L_i} \cdots \frac{T}{L_i}]$. It is easy to check that $\{X_i\}$ satisfies the first property in Definition IV.2 and will satisfy the second property if $\alpha_{i,k}$ is i.i.d. with respect to k. According to Lemma IV.2, $\{X_i\}$ is an optimal (deterministic) solution given $\{L_i\}$. It follows that if we let $\{L_i^{\star}\}$ denote the optimal solution of

$$\max_{\{L_i\}} \sum_{i=1}^{N}$$
$$\frac{-\left(T - \sum_{k=1}^{L_i} E[\min(W_{i,k} + \alpha_{i,k}, \frac{T}{L_i})]\right) \cdot r_i - L_i C_i^D}{T}$$

with resource constraint $\sum_{i=1}^{N} \frac{L_i}{T} \leq B$. Then $X_i^{\star} \triangleq [\frac{T}{L_i^{\star}} \frac{T}{L_i^{\star}} \cdots \frac{T}{L_i^{\star}}]$ is an optimal solution to the defender's problem. Hence, a periodic strategy with periods of $X_i^{\star}$ for all i is a best-response strategy for the defender. ∎

According to Theorem IV.1, the defender uses periodic strategy to keep the system stable, in the sense of the same total loss between two defenses. Since the distribution of attacker's waiting time $W_{i,k}$ does not change with time, a fixed defense interval provides the same expected payoff between every two consecutive moves. Moreover, the convexity of the defender's optimization problem guarantees an optimal solution under a given attack strategy.

B. Attacker's Best Response

We first analyze the attacker's best response against any deterministic defense strategy, then show that the i.i.d. strategy is the best response against periodic defense.

Definition IV.3: An attack strategy is called independent nonadaptive if it is nonadaptive, and $W_{i,k}$ are independent across i and k.

Lemma IV.3: When the defense strategy is deterministic and $\alpha_{i,k}$ is *i.i.d.* across k, for any attacking strategy (adaptive or nonadaptive), there always exists an independent nonadaptive strategy that gives the attacker the same payoff.

Proof: When the defense strategies are deterministic, we can move the expectation in (6) after the summation over k and the expectation is with respect to $\alpha_{i,k}$ and $W_{i,k}$. The same for constraint (7). Then, the proof is done as long as we can construct

an independent nonadaptive strategy $W'_{i,k}$ such that for all i and k, we have the following:

1) $E[\min(W_{i,k} + \alpha_{i,k}, X_{i,k})] = E[\min(W'_{i,k} + \alpha_{i,k}, X_{i,k})]$;
2) $E[\min(W_{i,k}, X_{i,k})] = E[\min(W'_{i,k}, X_{i,k})]$;
3) $P(W_{i,k} < X_{i,k}) = P(W'_{i,k} < X_{i,k})$.

Since $X_{i,k}$ is deterministic and $\alpha_{i,k}$ is i.i.d. across k, the expectation above is with respect to the marginal distribution of $W_{i,k}$ only. Thus, we can construct $W'_{i,k}$ whose distribution is the same as $W_{i,k}$'s marginal distribution, which does not depend on any realization of $X_{j,\tau}$ and $W_{j,\tau}$ s.t. $T_j(\tau) < T_i(k)$. Meanwhile, $W'_{i,k}$ is independent across i and k. ∎

According to Lemma IV.3, it suffices to consider independent nonadaptive strategies when the defender uses deterministic strategies.

Lemma IV.4: When the defense strategy is deterministic, the attacker's best response (among nonadaptive strategies) must satisfy the following condition:

$$W^\star_{i,k} = \begin{cases} 0, & w.p.\ p_{i,k} \\ \geq X_{i,k}, & w.p.\ 1 - p_{i,k} \end{cases}. \tag{10}$$

Refer to the proof in Section VIII-A. Lemma IV.4 implies that for each node i, the attacker's best strategy is to either attack node i immediately after it realizes the node's recovery, or gives up the attack until the defender's next move. There is no incentive for the attacker to wait a small amount of time to attack a node before the defender's next move. The constraint M actually determines the probability that the attacker will attack immediately. If M is large enough, the attacker will never wait after defender's each move. We then find the attacker's best response when the defender employs the periodic strategy.

Theorem IV.2: Assume that for any i, the attacking times $\alpha_{i,k}$'s are *i.i.d.* across k. When the defender employs a periodic strategy, the *i.i.d.* strategy is the attacker's best response among all strategies.

Proof: Suppose that the defender uses a periodic strategy, where for any i, $X_{i,k} = \frac{T}{L_i}$ for any k. With (10), the attacker's problem (6) can be simplified to a fractional knapsack problem with decision variables $\{p_{i,k}\}$. For any given node i, $p_{i,k}$'s unit rewards (payoff in the target function divided by weight in the constraint) across k are all equal when $\alpha_{i,k}$'s are i.i.d. across k. Thus, setting all $p_{i,k}$ in (10) equal is one of the optimal solutions. Therefore, the i.i.d. strategy is a best solution for attacker when the defender uses a periodic strategy. ∎

C. Simplified Optimization Problems

We put particular emphasis on the case where the defender employs a periodic strategy and the attacker uses an i.i.d. strategy. We assume that all the attacking times $\alpha_{i,k}$ are i.i.d. across k and omit the subscript k in $\alpha_{i,k}$. According to Theorems IV.1 and IV.2, periodic defense and i.i.d. attack can form a pair of best-response strategies with respect to each other. Consider such pairs of strategies. Let $m_i \triangleq \frac{L_i}{T} = \frac{1}{X_{i,k}}$, and let p_i denote the probability that $W_{i,k} = 0$ for all k. The optimization problems to the defender and the attacker can then be simplified as follows.

Defender's problem

$$\max_{\{m_i\}} \sum_{i=1}^{N} \left[\left(E\left[\min\left(\alpha_i, \frac{1}{m_i}\right)\right] p_i r_i - C_i^D \right) \cdot m_i - p_i r_i \right]$$

$$\text{s.t. } \sum_{i=1}^{N} m_i \leq B. \tag{11}$$

Attacker's problem

$$\max_{\{p_i\}} \sum_{i=1}^{N} p_i \cdot \left(r_i \left(1 - E\left[\min\left(\alpha_i, \frac{1}{m_i}\right)\right] \cdot m_i \right) - C_i^A m_i \right)$$

$$\text{s.t. } \sum_{i=1}^{N} E\left[\min\left(\alpha_i, \frac{1}{m_i}\right)\right] \cdot m_i \cdot p_i \leq M. \tag{12}$$

We observe that the defender's problem is a continuous convex optimization problem, whereas the attacker's problem is a fractional knapsack problem. Therefore, the best response strategy of each side can be easily determined. Also, the time period T disappears in both problems. It is worth mentioning that finding the Nash equilibrium of (11)–(12) is very challenging since the constraint of (12) is nonconvex with respect to m_i, thus the strategy space of this generalized Nash equilibrium problem is not jointly convex.

D. Markovian Strategies

Based on Theorems IV.1 and IV.2, the defender's periodic strategy and attacker's i.i.d. strategy form a Nash equilibrium among all adaptive and nonadaptive strategies. However, it remains unclear what is the best response if one of the players uses an adaptive strategy. To the best of our knowledge, there has been virtually no discussion about adaptive strategies in the field of stealthy attacks. Furthermore, even though a deterministic strategy is always optimal for the defender based on Lemma IV.1, there may still exist nondeterministic strategies that are also optimal. Meanwhile, Nash equilibria under more general strategies from both players may exist. In this section, we provide some preliminary results in this direction by considering Markovian strategies from both the defender's and the attacker's perspectives. We assume that the attacker's waiting times $W_{i,k}$ follow (10) and define a Markovian attacking strategy as follows.

Definition IV.4: An attacking strategy is a Markovian strategy if the attack probabilities $\{p_{i,k}\}$ for node i follow a discrete Markov chain over K states $v_1, v_2, \ldots, v_K$ with transition matrix M_i^A. That is, $\Pr(p_{i,k+1} = v_s | p_{i,k} = v_t) = M_i^A(s, t)$ for any s and t.

A Markovian defense strategy is defined similarly by considering $\{X_{i,k}\}$ instead of $\{p_{i,k}\}$. For tractability, we only consider the expected payoffs for the defender in a steady state. We show our main results about Markovian strategies in the following.

Theorem IV.3: If the attacker employs an ergodic Markovian strategy, the periodic strategy is the defender's best response.

The proof can be found in Section VIII-B. Theorem IV.3 tells us that the defender still prefers using a periodic strategy when the attacker's strategy space includes Markovian strategies. Consequently, the pair of periodic strategy and i.i.d. strategy

naturally forms the Nash equilibrium in this case. However, the i.i.d. attack strategy may not be optimal against a Markovian defending strategy as shown in the following theorem.

Theorem IV.4: If the defender employs a Markovian strategy, the *i.i.d.* attack strategy is not optimal in general.

Proof sketch: The main idea is to construct a counter example where the defender uses a Markovian strategy to defend a single node and the defense interval X_k (we omit i since there is only one node) takes one of the two values x_1 and x_2. The probability of $X_k = x_1$ only depends on the value of X_{k-1}. We then find the steady-state probability of the node being compromised. The total expected payoff for the attacker can then be written in a similar form as (12) with decision variables p_1 and p_2, referring to the probabilities of launching an attack when $X_{k-1} = x_1$ and $X_{k-1} = x_2$, respectively. Since the optimal values of p_1 and p_2 for solving the knapsack problem are not equal in general, the i.i.d. attack is not optimal. A detailed proof is given in [33].

Theorem IV.4 tells us that the attacker may use an adaptive strategy against the Markovian defending strategy. Compared to the defender, since the attacker is able to observe the defending periods and the node states, the attacking strategy may become state-dependent. Therefore, Nash equilibria beyond periodic defense and i.i.d. attack can exist in the space of both adaptive and nonadaptive strategies.

E. Discussion on Security Games in Networks

In this work, we focus on protecting a set of independent nodes where the payoff functions are additive, that is, the total payoff to a player is a weighted summation of the payoffs from each node. Even in this case, finding the equilibrium solutions of the game (11)–(12) is already very challenging as we mentioned in Section IV-C. Solving a security game in a general network setting that yields nonadditive utility is even harder. Because of that, existing security game work typically assumes additive utility as we did.

To extend our solutions discussed in Sections V and VI to a network setting, a promising direction is to introduce non-additive payoff functions to the defender and the attacker to capture the dependencies of node values. There are several recent works [16], [31], [32] that consider security games in network settings. In particular, Wang *et al.* [32] developed a general framework to convert a security game with nonadditive utility to a combinatorial optimization problem over a set system, and characterized the complexity of finding the Nash equilibrium. However, efficient algorithms are only known for some special cases and none of them apply to our setting directly. Furthermore, most previous work on security games including [32] considers a static setting (or the steady state in a repeated setting) where the game is played only once, which cannot faithfully model the joint spatial and temporal decisions in dynamic stealthy games as we consider in this article.

V. NASH EQUILIBRIA

In this section, we study the set of Nash equilibria of the game where the defender employs a periodic strategy, and the attacker employs an i.i.d. strategy. For tractability, we further assume that the attacking time $\alpha_{i,k}$ is deterministic for all i and

we omit the subscript k. We show that this game always has a Nash equilibrium and may have multiple equilibria of different values.

We first observe that for deterministic α_i, when $m_i \geq \frac{1}{\alpha_i}$, the defender's payoff becomes $-m_i C_i^D$, which is maximized when $m_i = \frac{1}{\alpha_i}$. Therefore, it suffices to consider $m_i \leq \frac{1}{\alpha_i}$. Thus, the optimization problems to the defender (11) and the attacker (12) can be simplified as follows.

For a given p, the defender aims at maximizing its payoff

$$\max_{\{m_i\}} \sum_{i=1}^{N} [m_i(r_i\alpha_i p_i - C_i^D) - p_i r_i]$$
$$\text{s.t.} \sum_{i=1}^{N} m_i \leq B \tag{13}$$
$$0 \leq m_i \leq \frac{1}{\alpha_i} \quad \forall i.$$

On the other hand, for a given m, the attacker aims at maximizing its payoff

$$\max_{\{p_i\}} \sum_{i=1}^{N} p_i[r_i - m_i(r_i\alpha_i + C_i^A)]$$
$$\text{s.t.} \sum_{i=1}^{N} m_i\alpha_i p_i \leq M \tag{14}$$
$$0 \leq p_i \leq 1 \quad \forall i.$$

For a pair of strategies (m, p), the payoff to the defender is $U_d(m, p) = \sum_{i=1}^{N} [m_i(p_i r_i \alpha_i - C_i^D) - p_i r_i]$, whereas the payoff to the attacker is $U_a(m, p) = \sum_{i=1}^{N} p_i[r_i - m_i(r_i\alpha_i + C_i^A)]$. A pair of strategies (m^*, p^*) is called a (pure strategy) *Nash equilibrium (*NE*)* if for any pair of strategies (m, p), we have $U_d(m^*, p^*) \geq U_d(m, p^*)$ and $U_a(m^*, p^*) \geq U_a(m^*, p)$.

Let $\mu_i(p) \triangleq p_i r_i \alpha_i - C_i^D$ denote the coefficient of m_i in U_d, and $\rho_i(m) \triangleq \frac{r_i - m_i(r_i\alpha_i + C_i^A)}{m_i\alpha_i}$. Note that for a given p, the defender tends to protect more a component with higher $\mu_i(p)$, whereas for a given m, the attacker will attack a component more frequently with higher $\rho_i(m)$. When m and p are clear from the context, we simply let μ_i and ρ_i denote $\mu_i(p)$ and $\rho_i(m)$, respectively. For a given strategy (m, p), we define $\mu^*(p) \triangleq \max_i \mu_i(p)$, $\rho^*(m) \triangleq \min_i \rho_i(m)$, $F(p) \triangleq \{i : \mu_i(p) = \mu^*(p)\}$, and $D(m, p) \triangleq \{i \in F : \rho_i(m) = \rho^*(m)\}$. We omit m and p when they are clear from the context.

Theorem V.1: Any pair of strategies (m, p) with $F(p) = F$ and $D(m, p) = D$ is an NE iff it is a solution to one of the following sets of constraints:
1) $\sum_{i \in F} m_i = B$; $\rho^* = 0$;
2) $\sum_{i \in F} m_i = B$; $\rho^* > 0$; $\sum_{i \in F} m_i\alpha_i p_i = M$;
3) $\sum_{i \in F} m_i = B$; $\rho^* > 0$; $p_i = 1 \quad \forall i \in F$;
4) $\sum_{i \in F} m_i < B$; $\mu^* = 0$; $F = N$; $\rho^* = 0$;
5) $\sum_{i \in F} m_i < B$; $\quad \mu^* = 0$; $\quad F = N$; $\quad \rho^* > 0$; $\sum_{i \in F} m_i\alpha_i p_i = M$;
6) $\sum_{i \in F} m_i < B$; $\mu^* = 0$; $F = N$; $\rho^* > 0$; $p_i = 1$ $\forall i \in F$.

In the following, NEs that fall into each of the six cases considered earlier are named as Type 1–Type 6 NEs, respectively. Among the six types of NE, Types 1–3 character the NEs that the defending budget B is fully utilized. Type 1 NE indicates that the attacker's benefit from all the nodes are zero due to full defense (all nodes are protected in the way that the expected payoffs of each node are zero). Type 2 NE characterizes the case that both attacker and defender fully use their budget, whereas Type 3 NE refers to a special case that the attacker attacks all the nodes in F with probability 1. Cases 4–6 then follow from a similar pattern for Cases 1–3 with underutilized defending budget. Since B is not fully used, the defender has no incentive to use more budget only if $\mu^\star = 0$, meaning that protecting more nodes will not bring more benefits.

The next theorem shows that our game has at least one equilibrium and may have more than one NE.

Theorem V.2: The attacker–defender game always has a pure strategy Nash equilibrium, and may have more than one NE of different payoffs to the defender.

The detailed proofs of Theorems V.1 and V.2 can be found in our technical report [33].

VI. SEQUENTIAL GAME

In this section, we study the subgame perfect equilibrium [27] of the Stackelberg game when the defender employs a periodic strategy and the attacker employs an i.i.d. strategy. In the sequential game, the defender first commits to a strategy and makes it public, the attacker then responds accordingly. We assume that at $t = 0$, the leader (defender) has determined its strategy and the follower (attacker) has learned the defender's strategy and determined its own strategy in response. In addition, the players do not change their strategies thereafter. Our objective is to identify the best sequential strategy for the defender. As mentioned in Section III, we consider the sequential game under full information. We adopt the same assumption in Section V and then define the subgame perfect equilibrium as follows.

Definition VI.1: A pair of strategies $(m^\star, p^\star)$ is a subgame perfect equilibrium of the sequential game if $m^\star$ is the optimal solution of

$$\max_{\{m_i\}} \sum_{i=1}^{N} [m_i(r_i\alpha_i p_i^\star - C_i^D) - p_i^\star r_i]$$

$$\text{s.t. } \sum_{i=1}^{N} m_i \leq B \tag{15}$$

$$0 \leq m_i \leq \frac{1}{\alpha_i} \quad \forall i$$

where $p^\star$ is the optimal solution of

$$\max_{\{p_i\}} \sum_{i=1}^{N} p_i[r_i - m_i(r_i\alpha_i + C_i^A)]$$

$$\text{s.t. } \sum_{i=1}^{N} m_i\alpha_i p_i \leq M \tag{16}$$

$$0 \leq p_i \leq 1 \quad \forall i.$$

Note that in a subgame perfect equilibrium $(m^\star, p^\star)$, $p^\star$ is the optimal solution of (16) with respect to $m^\star$, but the defender's best strategy $m^\star$ is not necessarily optimal in (15) with respect to a fixed set of $p^\star$. Due to the multinode setting and the resource constraints, it is very challenging to identify an exact subgame perfect equilibrium strategy for the defender. We first establish several properties about the optimal defense strategy and then propose a dynamic programming-based algorithm that finds a nearly optimal defense strategy.

To clearly state the properties, we partition all the nodes into four disjoint sets defined in the following:
1) $F = \{i | m_i > 0, \ p_i = 1\}$;
2) $D = \{i | m_i > 0, \ 0 < p_i < 1\}$;
3) $E = \{i | m_i > 0, \ p_i = 0\}$;
4) $G = \{i | m_i = 0, \ p_i = 1\}$.

We observe that the set D has at most one element since (16) is a fractional knapsack problem. Let $\rho_i(m_i) \triangleq \frac{r_i - m_i(r_i\alpha_i + C_i^A)}{m_i\alpha_i}$ represent the attacker's reward normalized by $m_i\alpha_i$ [the coefficient of p_i in the constraint of (16)] from attacking node i, which is a key factor in finding the optimal solution for the knapsack problem. We use m_d to represent $m_i, \ i \in D$ for simplicity and denote $\rho_d = \rho_d(m_d)$. If D is empty, we pick any node i in F with minimum ρ_i and treat it as a node in D.

Lemma VI.1: For all optimal solutions of (15)–(16), we always have $\rho_d \geq 0$.

Lemma VI.2: For any given non-negative ρ_d, an optimal solution for (15)–(16) satisfies the following properties:
1) $r_i\alpha_i - C_i^D > 0 \quad \forall i \in F \cup E \cup D$;
2) $m_i \leq \overline{m}_i \quad \forall i \in F$;
3) $m_j = \overline{m}_j \quad \forall j \in E$;
4) $\overline{m}_i \leq \frac{1}{\alpha_i} \quad \forall i$;
5) $B - \sum_{i \in E} \overline{m}_i - m_d > 0$
where $\overline{m}_i = \rho_i^{-1}(\rho_d)$.

Lemma VI.3: For any non-negative ρ_d, there exists an optimal solution for (15)–(16) such that $\forall i \in F$, there is at most ONE $m_i < \overline{m}_i$ and all the other $m_i = \overline{m}_i$.

The proof of Lemma VI.3 is in Section VIII-C. Lemmas VI.1–VI.3 establish the foundation for the following key result about the optimal defense strategy of (15)–(16).

Proposition VI.1: For any non-negative ρ_d, there exists an optimal solution $\{m_i\}_{i=1}^{n}$ such that the following statements hold:
1) $\forall i \in F$, there is at most one $m_i < \overline{m}_i$ and all the other $m_i = \overline{m}_i$;
2) $m_d = \overline{m}_d$
3) $\forall i \in E, m_i = \overline{m}_i$;
4) $\forall i \in G, m_i = 0$.

We denote the node whose $m_i < \overline{m}_i$ in the first property of Proposition VI.1 as node f and its defending frequency as m_f. Based on Proposition VI.1, we can easily compute the value of m_i for each node (except m_f) after the set allocation is fixed. Also, we can explicitly list the defender's payoff, defender's budget usage, and attacker's budget usage by putting each node into different sets, as shown in Table III. For the fractional node, its m_i can be computed using linear programming when all the other m_i have been determined. We use dynamic programming to determine the optimal set allocation.

TABLE III
NODES IN DIFFERENT SETS WITH GIVEN ρ_d

	F	E	G
Defender's payoff	$\overline{m}_i(r_i\alpha_i - C_i^D) - r_i$	$-\overline{m}_i C_i^D$	$-r_i$
Defender's budget usage	$\overline{m}_i$	$\overline{m}_i$	0
Attacker's budget usage	$\overline{m}_i\alpha_i$	0	0

From the earlier discussion, we propose the following algorithm to the defender's problem (see Algorithm 1). The algorithm iterates over all possible node d in set D and all possible node f with fractional assignment in set F. We first compute a special case when set G is empty (line 2). In this case, the defender's optimal strategy can be obtained by solving (17) based on Proposition VI.1

$$
\begin{aligned}
\text{Val}(d, f) = \max_{\rho_d, \{p_i\}, m_f} & \sum_{i\neq f} \overline{m}_i(p_i r_i \alpha_i - C_i^D) \\
& - p_i r_i + m_f(r_f\alpha_f - C_f^D) - r_f \\
\text{s.t.} \sum_{i\neq f} \overline{m}_i + m_f \leq B, & \quad \sum_{i\neq f} p_i \overline{m}_i \alpha_i + m_f\alpha_f \leq M \cdot \\
\rho_d = \frac{r_i - \overline{m}_i(r_i\alpha_i + C_i^A)}{\overline{m}_i\alpha_i} & \geq 0, \; 0 \leq p_i \leq 1, \; m_f \geq 0
\end{aligned}
$$
(17)

The algorithm then iterates over non-negative ρ_d with a step size ρ_{step} (line 4). Given ρ_d, d, f, the best set allocation (together with m_i for all i) are determined using dynamic programming as explained before.

For any given ρ_d, d and f, we compute $\overline{m}_i$ for all i (line 5). Let $\text{SEQ}(i, b, m, d, f, \text{ind})$ denote the maximum payoff of the defender considering only node 1 to node i (excluding nodes d and f), for a given defender's budget $b \in [0, B]$ and an attacker's budget $m \in [0, M]$. The parameter ind is a Boolean variable that indicates whether we can put nodes in set E arbitrarily. If ind is $True$, any node (except nodes d and f) can be in set E. Otherwise, a node i can be allocated to set E only if $r_i - \overline{m}_i(r_i\alpha_i + C_i^A) \leq 0$. The value of $\text{SEQ}(i, b, m, d, f, \text{ind})$ is determined recursively. If node i is either d or f, we simply set $\text{SEQ}(i, b, m, d, f, \text{ind}) = \text{SEQ}(i - 1, b, m, d, f, \text{ind})$. Otherwise, we have the following recurrence equation, where the three cases refer to putting node i in sets F, E, and G, respectively:

$$
\begin{aligned}
& \text{SEQ}(i, b, m, d, f, \text{ind}) \\
& = \max
\begin{cases}
\text{SEQ}(i - 1, b - \overline{m}_i, m - \alpha_i\overline{m}_i, d, f, \text{ind}) \\
\qquad + \overline{m}_i(r_i\alpha_i - C_i^D) - r_i \\
\text{SEQ}(i - 1, b - \overline{m}_i, m, d, f, \text{ind}) - \overline{m}_i C_i^D \\
\text{SEQ}(i - 1, b, m, d, f, \text{ind}) - r_i.
\end{cases}
\end{aligned}
$$
(18)

We have the following boundary conditions:
1) The recursion SEQ will return $-\infty$ when $i > 0$ and i) $m < 0$, or ii) $b < 0$, or iii) $m = 0$ and ind = False;

Algorithm 1: Sequential Strategy for Defender.

1: **for** $d, f \leftarrow 1\ to\ n$ **do**
2: $\quad$ Solve (17) to obtain $\text{Val}(d, f)$
3: $\quad$ $\rho_{\max} \leftarrow \rho : \sum_{i=1}^n \alpha_i m_i(\rho) = M$
4: $\quad$ **for** $\rho_d \leftarrow 0$ to $\rho_{\max}$ with step size ρ_{step} **do**
5: $\quad\quad$ $\overline{m}_i \leftarrow m_i(\rho_d)$ for all i
6: $\quad\quad$ $\text{val}'_{d,f,\rho_d} \leftarrow \text{SEQ}(n, B, M, d, f, \text{True})$
7: $\quad\quad$ $\text{val}''_{d,f,\rho_d} \leftarrow \text{SEQ}(n, B, M, d, f, \text{False})$
8: $\quad$ **end for**
9: $\quad$ $P_{dp}(d, f) \leftarrow \max_\rho \{\text{val}'_{d,f,\rho_d}, \text{val}''_{d,f,\rho_d}\}$
10: **end for**
11: $P_{\text{alg}} \leftarrow \max_{d,f}\{P_{dp}(d, f), \text{Val}(d, f)\}$

Algorithm 1 computes the optimal solution by searching over all combinations of d, f, and ρ_d. For any given combination, the dynamic program actually finds all the solutions that satisfy Proposition VI.1, meaning that $P_{dp}(d, f)$ returns the optimal defense strategy under given d, f, and ρ_d (line 9). Therefore, P_{alg} is the maximum payoff that the defender can achieve (line 11). For the dynamic program, we round the input before running $\text{SEQ}(n, B, M, d, f, \text{ind})$, since the recursion may never stop without rounding. Denote δ as the rounding parameter, we have $\overline{m}_i \leftarrow \lfloor \frac{\overline{m}_i}{\delta} \rfloor$, $\alpha_i \leftarrow \lfloor \frac{\alpha_i}{\delta} \rfloor$ for all i and $B \leftarrow \lfloor \frac{B}{\delta} \rfloor$, $M \leftarrow \lfloor \frac{M}{\delta} \rfloor$. By setting δ small enough, Algorithm 1 can find a strategy that is arbitrarily close to the subgame perfect equilibrium strategy of the defender. Formally, we can establish the following result.

Theorem VI.1: Let $|P_{alg}|$ denote the defender's cost obtained by Algorithm 1 and $|P^\star|$ the optimal cost. Given ρ_{step} and the rounding parameter δ, we have $\frac{|P_{alg}|}{|P^\star|} \leq 1 + (\rho_{\text{step}} + \delta)O(N)$.

Proof sketch: If the set G is empty in the optimal solution $P^\star$, Algorithm 1 computes the optimal payoffs for the defender by solving (17). Then, we have $\max_{d,f} \text{Val}(d, f) = P^\star$. Therefore, $\frac{|P_{alg}|}{|P^\star|} = 1$. If the set G is not empty in the optimal solution $P^\star$, we first bound the performance loss due to ρ_{step}. Denote $\rho^\star$ as the optimal ρ_d for computing $P^\star$ and ρ' the first ρ_d that is greater than $\rho^\star$ in Algorithm 1. Let $\overline{m}_i^\star = m_i(\rho^\star)$ and $\overline{m}_i' = m_i(\rho')$. Let $|P_{\rho'}|$ refer to the total cost when $\rho^\star$ increases to ρ' for the optimal solution $P^\star$. By increasing $\rho^\star$ to ρ', each $\overline{m}_i^\star$ decreases to $\overline{m}_i'$ and the total cost increases in two parts. The first part is due to the decrease of $\overline{m}_i^\star$ for all i in F. The second part comes from sets E and D. Since $\overline{m}_i^\star$ decreases, the attacker has extra budget to attack the nodes in sets E and D, moving these nodes to sets F and D. For all sets F, D, E, and G above, we refer to the set allocation in optimal solution $P^\star$. Let H_F and H_E denote the increase of total cost from the two parts, respectively. We can bound these cost differences as $H_F = \sum_{i \in F} \triangle\overline{m}_i(r_i\alpha_i - C_i^D)$ and

$$
H_E \leq \sum_{i=1}^N \triangle\overline{m}_i\alpha_i \cdot \max_i \left\{ \frac{r_i(1 - \alpha_i\overline{m}_i')}{\alpha_i\overline{m}_i'} \right\}.
$$

Then, we can bound the total cost as $\frac{|P_{alg}| - |P^\star|}{|P^\star|} \leq \frac{|P_{\rho'}| - |P^\star|}{|P^\star|} = \frac{H_F + H_E}{|P^\star|} \leq \rho_{\text{step}}O(N)$. Find the detailed proof in [33].
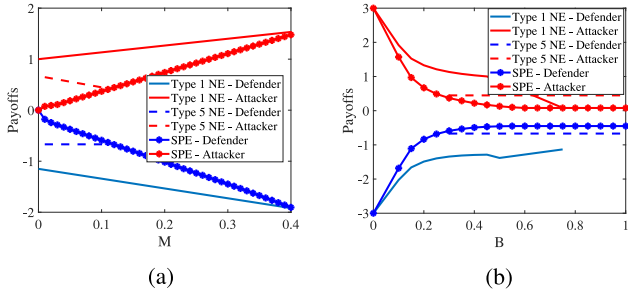
Fig. 2. Effects of varying resource constraints on payoffs. In both figures, $r_1 = 2, r_2 = 1, w_1 = 1.7, w_2 = 1.6, C_1^D = 0.5, C_2^D = 0.6, C_1^A = 1, C_2^A = 1.5, B = 0.3$ in (a), and $M = 0.1$ in (b). (a) Payoffs with varying M. (b) Payoffs with varying B.

Theorem VI.1 provides the performance guarantee of Algorithm 1 showing the tradeoff between performance and the time complexity. Based on Theorem VI.1, we have the following corollary.

Corollary VI.1: By setting both ρ_{step} and δ with $O(\frac{1}{N})$, Algorithm 1 can achieve a near-optimal solution and its complexity is $O(N^5 BM)$.

VII. NUMERICAL RESULTS

In this section, we present numerical results for our game models. For the illustrations, we assume that all the attacking times α_i are deterministic as in Section VI. We study the payoffs of both the attacker and the defender and their strategies in both Nash equilibrium (two-node setting) and subgame perfect equilibrium (both two-node and five-node settings), and study the impact of various parameters including resource constraints B, M, and the unit value r_i.

A. Simulations With Selected Parameters

We first study the impact of the resource constraints M and B on the player's payoffs in a two-node setting. The results are given in Fig. 2, where we have plotted both Type 1 and Type 5 NEs[3] and subgame perfect equilibria. A Type 5 NE only occurs when M is small as shown in Fig. 2(a), whereas Type 1 NE appears when B is small as shown in Fig. 2(b), which is expected since B is fully utilized in a Type 1 NE while M is fully utilized in a Type 5 NE. When the defense budget B becomes large, the summation of m_i does not necessarily equal to B and thus Type 1 NEs disappear. Similarly, Type 5 NEs disappear for large attack budget M. In both figures, the subgame perfect equilibria always bring the defender higher payoffs compared with Nash equilibria, which is expected.

B. Simulations With Real-World Data

To have a better understanding of the performance of Algorithm 1, we consider a five-node setting and use real-word data from the National Vulnerability Database [2]. We pick five vulnerability incidents about IoT devices revealed by the database.
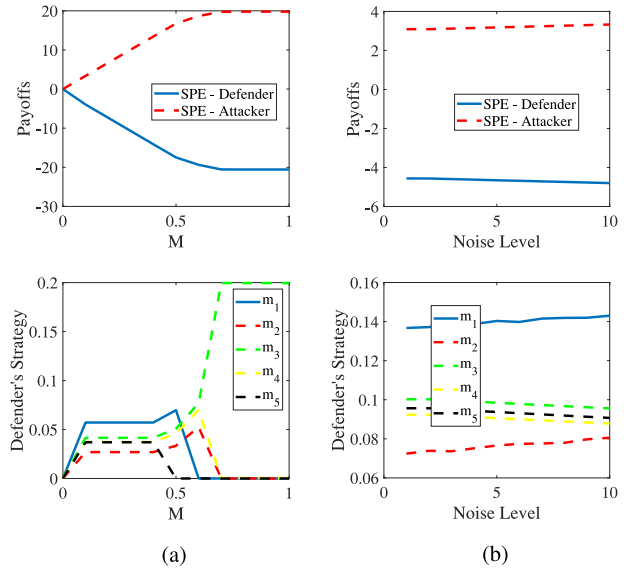

Fig. 3. Effects of varying resource constraint M and unit value r_i, where $B = 0.2$ in (a), $B = 0.5$ and $M = 0.3$ in (b). In (b), a random noise level is added to r_1 and r_2. (a) Payoffs and strategies with varying M. (b) Payoffs and strategies with varying noise level.

For each incident, we use their impact score (the potential impact of the vulnerability), exploitability score (how vulnerable the thing itself is to attack), vulnerability base score (how critical the vulnerability is), and attack complexity (low or high) [3]–[7] as an approximation of the node value, attacking time, defending cost, and attacking cost, respectively. Specifically, we set node values as $r = [5.9\ 3.6\ 5.9\ 5.2\ 3.6]$. For the attacking times, since higher exploitability score means easier attack, we take the reciprocal and set $\alpha = [10/3.9\ 10/2.8\ 10/2.8\ 10/2.8\ 10/3.9]$ where the constant 10 is used for normalization. The vulnerability base score is utilized to approximate the defending cost by setting $C^D = [9.8\ 6.5\ 8.8\ 8.1\ 7.5]/3$, whereas the attacking cost is set to 2 if the attack complexity is high and 1 otherwise. We study the effects of varying M and r in Fig. 3(a).

In Fig. 3(a), the attacker's budget M varies from 0 to 1 and the defending budget $B = 0.2$. When $M = 0$, the defender can set m_i for all i to arbitrary small (but positive) values, so that the attacker is unable to attack any node, leading to a zero payoff for both players. As M becomes larger, the attacker's payoff increases, whereas the defender's payoff decreases, and the defender tends to defend the nodes with higher values more frequently, as shown in Fig. 3(a)(lower). The defender gradually stop protecting low-value nodes and move all the resources to defend node 3. Note that the defending frequency for node 3 is smaller than that for node 1 at the beginning. This is because when M is small, the attacker attacks each node with a very small probability, thus the defender can protect all the nodes at the same time to prevent big loss. Since $r_1 = r_3$ but $\alpha_1 < \alpha_3$, the defender protects node 1 more frequently. However, when the attacker has enough resources to attack each node with a much higher probability, it is not beneficial for the defender to protect other nodes except node 3 since it has the highest node value and attacking time.

---

[3] There are also Type 2 NEs, which are omitted for the sake of clarify.

In Fig. 3(b), we fixed r_3 through r_5 and increase r_1 and r_2 by adding a random noise uniformly distributed between [noise_level − 1, noise_level] ∗ 0.1. We vary the noise_level from 1 to 10. As shown in the figure, m_1 and m_2 keep increasing when the noise level becomes larger, whereas the defending frequencies for all other three nodes decrease due to limited defending resources, which indicates that the defender should protect the nodes with higher values more frequently in the subgame perfect equilibrium. Table IV compares the running time of Algorithm 1 and that of the corresponding algorithm in our conference paper [34]. All experiments are conducted on a desktop with 4-Core Intel i5-4670 K CPU @ 3.40 GHz and MATLAB R2019a. The same simulation setting, as shown in Fig. 3(a), is applied with fixed $M = 0.2$. We observe that Algorithm 1 is much faster than the original algorithm in our conference paper and the improvement is more significant in a larger setting.

TABLE IV
RUNNING TIME IMPROVEMENT

No. of nodes	Algorithm 1	Algorithm 1 in [35]
2	12.1 sec	31.4 sec
3	131.5 sec	410.8 sec
4	1036.3 sec	3710.8 sec
5	2117.9 sec	9261.3 sec
6	3.45 hours	24 - 26 hours

VIII. CONCLUSION

In this article, we proposed a two-player non-zero-sum game for protecting a system of multiple components against a stealthy attacker where the defender's behavior is fully observable and both players have strict resource constraints. We proved that periodic defense and nonadaptive i.i.d. attack are a pair of best-response strategies with respect to each other in the space of both adaptive and nonadaptive strategies. For this pair of strategies, we characterize the set of Nash equilibria of the game, and show that there is always one (and maybe more) equilibrium, for the case when the attack times are deterministic. We further study the sequential game where the defender first publicly announces its strategy and design an algorithm that can identify a strategy that is arbitrarily close to the subgame perfect equilibrium strategy for the defender. We also provide a full analysis of the algorithm performance and its complexity guarantee.

APPENDIX

A. Proof of Lemma IV.4

Proof: In order to get the attacker's best responses against any defender's deterministic strategies, we can divide (6) into $N * L$ suboptimization problems

$$\min_{W_{i,k}} \frac{E[\min(W_{i,k} + \alpha_{i,k}, X_{i,k})]r_i + P(W_{i,k} < X_{i,k})C_i^A}{T}$$

$$\text{s.t.} \quad \frac{E[\min(W_{i,k} + \alpha_{i,k}, X_{i,k})] - E[\min(W_{i,k}, X_{i,k})]}{T}$$

$$\leq M_{i,k} \tag{19}$$

where $\sum_{i=1}^{N} \sum_{k=1}^{L_i} M_{i,k} = M$ and $M_{i,k}$ can be arbitrary positive number. Note that we consider the equivalent minimization problem by taking the negative of the target function of (3) and omitting the constant part. We claim that the optimal solution to (19) is to allocate as much budget as possible to $P(W_{i,k} = 0)$, that is

$$W_{i,k}^* = \begin{cases} 0, & \text{w.p. } p_{i,k}^* \\ \geq X_{i,k}, & \text{w.p. } 1 - p_{i,k}^* \end{cases} \tag{20}$$

where $p_{i,k}^*$ equals $\min\left(1, \frac{M_{i,k}T}{E[\min(\alpha_{i,k}, X_{i,k})]}\right)$ if $r_i(E[\min(\alpha_{i,k}, X_{i,k})] - X_{i,k}) + C_i^A < 0$, and 0 otherwise.

Since $M_{i,k}$ is any number such that $\sum_{i=1}^{N} \sum_{k=1}^{L_i} M_{i,k} = M$, the optimal solution of (6) also satisfies the same structure of (20). We then prove our claim. For simplicity, we assume that $W_{i,k}$ is a discrete random variable (r.v.), and without loss of generality, it has the following probability mass function (p.m.f.):

$$W_{i,k} = \begin{cases} 0, & \text{with probability (w.p.) } p_0 \\ v_i, & \text{w.p. } p_i, \ i = 1 \cdots n \\ \geq X_{i,k}, & \text{w.p. } 1 - \sum_{j=0}^{n} p_j \end{cases} \tag{21}$$

where $n \in \mathbb{N}$ such that $0 < v_1 < v_2 < \cdots < v_n < X_{i,k}$. The following proof can be adapted to the continuous $W_{i,k}$ as well by replacing sums with integrals and p.m.f with probability density function.

Putting (21) into (19), attacker's problem can then be converted to the following form:

$$\min \sum_{j=0}^{n} p_j(r_i[E[\min(v_j + \alpha_{i,k}, X_{i,k})] - X_{i,k}] + C_i^A) + X_{i,k}r_i \tag{22}$$

with two constraints: $\sum_{j=0}^{n} p_j E[\min(\alpha_{i,k}, X_{i,k} - v_i)] \leq M_{i,k}T$ and $\sum_{j=0}^{n} p_j \leq 1$. where $v_0 = 0$.

Let $J(\{p_0, \ldots, p_n\})$ denote the objective function in (22). Since $r_i(E[\min(\alpha_{i,k}, X_{i,k})] - X_{i,k}) + C_i^A < r_i(E[\min(v_j + \alpha_{i,k}, X_{i,k})] - X_{i,k}) + C_i^A$, if $r_i(E[\min(\alpha_{i,k}, X_{i,k})] - X_{i,k}) + C_i^A \geq 0$, $J(\{p_0, \ldots, p_n\})$ is minimized by setting $p_j = 0, \forall j = 0, \ldots, n$, which implies $W_{i,k} \geq X_{i,k}$ w.p.1. Such condition describes the case that even if the attacker attacks the node immediately after it is recovered, its reward is still less than 0. Therefore, the attacker never attacks. If $r_i(E[\min(\alpha_{i,k}, X_{i,k})] - X_{i,k}) + C_i^A < 0$, we claim that the optimal solution is to allocate as much budget $M_{i,k}T$ as possible to p_0, that is, we set all $p_j = 0$, $1 \leq j \leq n$, and $p_0 = \min(1, \frac{M_{i,k}T}{E[\min(\alpha_{i,k}, X_{i,k})]})$. This is clearly true if $r_i(E[\min(v_j + \alpha_{i,k}, X_{i,k})] - X_{i,k}) + C_i^A \geq 0$. Therefore, it suffices to consider the case when $r_i(E[\min(\alpha_{i,k}, X_{i,k})] - X_{i,k}) + C_i^A < r_i(E[\min(v_j + \alpha_{i,k}, X_{i,k})] - X_{i,k}) + C_i^A < 0$.

To prove the claim, consider an optimal solution $\{p_0, p_1, \ldots, p_n\}$ to (22). We show that if $p_0 < \min(1, \frac{M_{i,k}T}{E[\min(\alpha_{i,k}, X_{i,k})]})$, then we can find another optimal solution $\{p'_0, p'_1, \ldots, p'_n\}$ such that $p'_0 > p_0$. We distinguish the following two cases.

Case 1: $p_0 E[\min(\alpha_{i,k}, X_{i,k})] + \sum_{j=1}^{n} p_j E[\min(\alpha_{i,k}, X_{i,k} - v_i)] < M_{i,k}T$. Then by the optimality of $\{p_0, p_1, \ldots, p_n\}$

and the assumption that $r_i(E[\min(v_j + \alpha_{i,k}, X_{i,k})] - X_{i,k}) + C_i^A < 0$, we must have $\sum_{j=0}^n p_j = 1$. Let $j \geq 1$ denote an index such that $p_j > 0$. Then there must exist a small amount $\triangle p > 0$ such that $p_0' = p_0 + \triangle p, p_j' = p_j' - \triangle p, p_k' = p_k, \forall k \neq 0$ and $k \neq j$ is again a feasible solution to (22). We further have

$$J(\{p_0, \ldots, p_n\}) - J(\{p_0', \ldots, p_n'\})$$

$$= \triangle p(r_i[E[\min(v_j + \alpha_{i,k}, X_{i,k})] - X_{i,k}] + C_i^A)$$

$$- \triangle p(r_i[E[\min(\alpha_i, X_{i,k})] - X_{i,k}] + C_i^A)$$

$$= \triangle p r_i(E[\min(v_j + \alpha_{i,k}, X_{i,k})] - E[\min(\alpha_{i,k}, X_{i,k})]) \geq 0.$$

Case 2: $p_0 E[\min(\alpha_{i,k}, X_{i,k})] + \sum_{j=1}^n p_j E[\min(\alpha_{i,k}, X_{i,k} - v_i)] = M_{i,k} T$. Again let $j \geq 1$ denote an index such that $p_j > 0$. Then there must exist a small amount $\triangle M > 0$ such that $p_0' = p_0 + \frac{\triangle M}{E[\min(\alpha_{i,k}, X_{i,k})]}, p_j' = p_j - \frac{\triangle M}{E[\min(\alpha_{i,k}, X_{i,k} - v_j)]}, p_k' = p_k, \forall k \neq 0$ and $k \neq j$ is a feasible solution to (22). We further have

$$J(\{p_0, \ldots, p_n\}) - J(\{p_0', \ldots, p_n'\})$$

$$= \frac{\triangle M(r_i[E[\min(v_j + \alpha_{i,k}, X_{i,k})] - X_{i,k}] + C_i^A)}{E[\min(\alpha_{i,k}, X_{i,k} - v_j)]}$$

$$- \frac{\triangle M(r_i[E[\min(\alpha_{i,k}, X_{i,k})] - X_{i,k}] + C_i^A)}{E[\min(\alpha_{i,k}, X_{i,k})]}$$

$$= \frac{\triangle M}{E[\min(\alpha_{i,k}, X_{i,k} - v_j)]}(r_i v_j - r_i X_{i,k} + C_i^A)$$

$$- \frac{\triangle M}{E[\min(\alpha_{i,k}, X_{i,k})]}(-r_i X_{i,k} + C_i^A) \geq 0.$$

∎

B. Proof of Theorem IV.3

Proof: When the attacker's strategy is an ergodic Markov chain, the $p_{i,k}$'s time-average distribution is the same as its steady-state distribution. Therefore, the defender's problem (1) can be transferred to the following:

$$\max_{\{X_{i,k}\}, L_i} \lim_{T \to \infty} E\left[\sum_{i=1}^N \left(-\frac{L_i C_i^D + T r_i}{T} + \right. \right.$$

$$\left. \left. \frac{(\sum_{k=1}^{L_i} E[p_{i,k}] \min(\alpha_{i,k}, X_{i,k}) + (1 - E[p_{i,k}]) X_{i,k}) \cdot r_i}{T} \right) \right]$$

(23)

with the same resource constraint in (2) where the expectation in the numerator is with respect to the steady-state distribution of $p_{i,k}$. We find that (23) is the same as (1) if we set $W_{i,k}^\star = 0$ with probability $E[p_{i,k}]$ and $W_{i,k} = \infty$ elsewhere. Here, $E[p_{i,k}]$ is the expected value of $p_{i,k}$'s steady-state distribution. Therefore, based on Lemma IV.1 and Theorem IV.1, we know that the periodic strategy is defender's best response. ∎

C. Proof of Lemma VI.3

Proof: Suppose the set allocation and ρ_d are fixed, which means m_d and $\overline{m}_i \forall i$ are also fixed. According to Lemma VI.2,

we can now convert (15)–(16) to the following problem:

$$\max_{m_i, i \in F} \sum_{i \in F} [m_i(r_i \alpha_i - C_i^D) - r_i] - \sum_{i \in G} r_i - \sum_{i \in E} \overline{m}_i C_i^D$$

$$+ m_d(pr_d \alpha_d - C_i^D) - pr_d$$

(24)

with constraints: $\sum_{i \in F} m_i \leq B - \sum_{i \in E} \overline{m}_i - m_d$, $\sum_{i \in F} \alpha_i m_i + p\alpha_d m_d \leq M$ and $0 \leq m_i \leq \overline{m}_i \ \forall i \in F$, where $p = \min\{1, \frac{M - \sum_{i \in F} \alpha_i m_i}{\alpha_d m_d}\}$.

Case 1: If $\frac{M - \sum_{i \in F} \alpha_i m_i}{\alpha_d m_d} \leq 1$, we put $p = \frac{M - \sum_{i \in F} \alpha_i m_i}{\alpha_d m_d}$ back into the target function of (24) and convert it to

$$\max_{m_i, i \in F} \sum_{i \in F} [m_i(r_i \alpha_i - C_i^D) - r_i] - \sum_{i \in G} r_i - \sum_{i \in E} \overline{m}_i C_i^D$$

$$+ \frac{M - \sum_{i \in F} \alpha_i m_i}{\alpha_d m_d} r_d(\alpha_d m_d - 1) - m_d C_d^D$$

(25)

with constraints: $\sum_{i \in F} m_i \leq B - \sum_{i \in E} \overline{m}_i - m_d$ and $0 \leq m_i \leq \overline{m}_i \ \forall i \in F$. It is easy to see that (25) is a fractional knapsack problem. Thus, there is at most one fractional variable, which means at most one $m_i < \overline{m}_i$.

Case 2: If $\frac{M - \sum_{i \in F} \alpha_i m_i}{\alpha_d m_d} > 1$, the attacker's budget is not fully utilized and all $p_i^\star$ in (15) equals 1. Thus, the sets D and E are empty. Now suppose there exist two nodes j and k in F with $m_j < \overline{m}_j$ and $m_k < \overline{m}_k$. Without loss of generality, by assuming $r_j \alpha_j - C_j^D \geq r_k \alpha_k - C_k^D$, we can always increase the defender's payoff by decreasing m_k and increasing m_j until either $m_j = \overline{m}_j$ or $m_k = 0$. If $m_k = 0$, node k is in set G. If the attacker's budget is fully utilized (as in Case 1), we cannot guarantee the new payoff by decreasing m_k and increasing m_j is always bigger, since α_k may be much smaller than α_j, making the increase of m_j is very small due to limited attacker's budget.

Above all, we can claim that there exists an optimal solution with at most one node in set F with $m_i < \overline{m}_i$. ∎

REFERENCES

[1] 2018. [Online]. Available: https://www.fireeye.com/content/dam/collateral/en/mtrends-2018.pdf

[2] [Online]. Available: https://nvd.nist.gov

[3] 2019. [Online]. Available: https://nvd.nist.gov/vuln/detail/CVE-2019-10891

[4] 2019. [Online]. Available: https://nvd.nist.gov/vuln/detail/CVE-2019-9461

[5] 2019. [Online]. Available: https://nvd.nist.gov/vuln/detail/CVE-2019-13267

[6] 2019. [Online]. Available: https://nvd.nist.gov/vuln/detail/CVE-2019-11061

[7] 2019. [Online]. Available: https://nvd.nist.gov/vuln/detail/CVE-2019-16159

[8] "Advanced persistent threat." [Online]. Available: http://en.wikipedia.org/wiki/Advanced_persistent_threat

[9] *ESET and Sucuri Uncover Linux/Cdorked.A: The Most Sophisticated Apache Backdoor*, 2013. [Online]. Available: http://www.eset.com/int/about/press/articles/article/eset-and-sucuri-uncover-linuxcdorkeda-apache-webserver-backdoor-the-most-sophisticated-ever-affecting-thousands-of-web-sites/

[10] A. Coviello, "Open letter to RSA customers," Mar. 17, 2011. [Online]. Available: http://www.rsa.com/node.aspx?id=3872

[11] T. Alpcan and T. Başar, *Network Security: A Decision and Game-Theoretic Approach*. Cambridge, U.K.: Cambridge Univ. Press, 2010.

[12] B. An, M. Brown, Y. Vorobeychik, and M. Tambe, "Security games with surveillance cost and optimal timing of attack execution," in *Proc. Int. Conf. Auton. Agents Multiagent Syst.*, 2013, pp. 1–9.

[13] B. Bencsáth, G. Pék, L. Buttyán, and M. Félegyházi, "The cousins of Stuxnet: Duqu, Flame, and Gauss," *Future Internet*, vol. 4, pp. 971–1003, 2012.
[14] K. D. Bowers, M. E. V. Dijk, A. Juels, A. M. Oprea, R. L. Rivest, and N. Triandopoulos, "Graph-based approach to deterring persistent security threats, U.S. Patent 8813234, 2014.
[15] K. D. Bowers *et al.*, "Defending against the unknown enemy: Applying FlipIt to system security," in *Proc. Conf. Decis. Game Theory Secur.*, 2012, pp. 248–263.
[16] A. Gueye, V. Marbukh, and J. C. Walrand, "Towards a metric for communication network vulnerability to attacks: A game theoretic approach," in *Proc. Int. Conf. Game Theory Netw.*, Springer, 2012, pp. 259–274.
[17] A. X. Jiang, A. D. Procaccia, Y. Qian, N. Shah, and M. Tambe, "Defender (Mis)coordination in security games," in *Proc. Int. Joint Conf. Artif. Intell.*, 2013.
[18] D. Korzhyk, Z. Yin, C. Kiekintveld, V. Conitzer, and M. Tambe, "Stackelberg vs. Nash in security games: An extended investigation of interchangeability, equivalence, and uniqueness," *J. Artif. Intell. Res.*, vol. 41, pp. 297–327, 2011.
[19] H. Kunreuther and G. Heal, "Interdependent security," *J. Risk Uncertainty*, vol. 26, no. 2/3, pp. 231–249, 2003.
[20] A. Laszka, G. Horvath, M. Felegyhazi, and L. Buttyán, "FlipThem: Modeling targeted attacks with FlipIt for multiple resources," in *Proc. Conf. Decis. Game Theory Secur.*, 2014, pp. 175–194.
[21] A. Laszka, B. Johnson, and J. Grossklags, "Mitigating covert compromises: A game-theoretic model of targeted and non-targeted covert attacks," in *Proc. Conf. Web Internet Econ.*, 2013, pp. 319–332.
[22] D. Leslie, C. Sherfield, and N. P. Smart, "Threshold FlipThem: When the winner does not need to take all," in *Proc. Conf. Decis. Game Theory Secur.*, 2015, pp. 74–92.
[23] D. Leslie, C. Sherfield, and N. P. Smart, "Multi-rate threshold FlipThem," in *Proc. Eur. Symp. Res. Comput. Secur.*, 2017, pp. 174–190.
[24] M. H. Manshaei, Q. Zhu, T. Alpcan, and T. Başar, "Game theory meets network security and privacy," *ACM Comput. Surv.*, vol. 45, 2012, Art. no. 25.
[25] T. H. Nguyen, R. Yang, A. Azaria, S. Kraus, and M. Tambe, "Analyzing the effectiveness of adversary modeling in security games," in *Proc. AAAI Conf. Artif. Intell.*, 2013, pp. 718–724.
[26] A. Nochenson *et al.*, "A behavioral investigation of the FlipIt game," in *Proc. 12th Workshop Econ. Inf. Secur.*, 2013, p. 93.
[27] M. J. Osborne and A. Rubinstein, *A Course in Game Theory*. Cambridge, MA, USA: MIT Press, 1994.
[28] P. Paruchuri, J. P. Pearce, J. Marecki, M. Tambe, F. Ordóñez, and S. Kraus, "Playing games for security: An efficient exact algorithm for solving Bayesian Stackelberg games," in *Proc. Int. Conf. Auton. Agents Multiagent Syst.*, 2008, pp. 895–902.
[29] M. Tambe, *Security and Game Theory: Algorithms, Deployed Systems, Lessons Learned*. Cambridge, U.K.: Cambridge Univ. Press, 2011.
[30] M. van Dijk, A. Juels, A. Oprea, and R. L. Rivest, "FlipIt: The Game of 'Stealthy Takeover'," *J. Cryptol.*, vol. 26, no. 4, pp. 655–713, 2013.
[31] S. Wang, F. Liu, and N. Shroff, "Non-additive security games," in *Proc. 31st AAAI Conf. Artif. Intell.*, 2017, pp. 728–735.
[32] S. Wang and N. Shroff, "Security game with non-additive utilities and multiple attacker resources," *Proc. ACM Meas. Anal. Comput. Syst.*, vol. 1, no. 1, 2017, Art. no. 13.
[33] M. Zhang, Z. Zheng, and N. B. Shroff, "Defending against stealthy attacks on multiple nodes with limited resources: A game-theoretic analysis," Tech. Rep. [Online]. Available: http://arxiv.org/abs/1508.01950
[34] M. Zhang, Z. Zheng, and N. B. Shroff, "A game theoretic model for defending against stealthy attacks with limited resources," in *Proc. Int. Conf. Decis. Game Theory Secur.*, Springer, 2015, pp. 93–112.
[35] Z. Zheng, N. B. Shroff, and P. Mohapatra, "When to reset your keys: Optimal timing of security updates via learning," in *Proc. AAAI Conf. Artif. Intell.*, 2017, pp. 3679–3685.
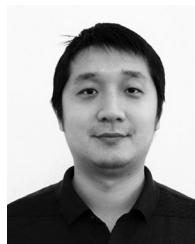
Ming Zhang (Student Member, IEEE) received the B.S. degree in electrical engineering from Shanghai JiaoTong University, Shanghai, China, in 2012, the M.S. degree in electrical and computer engineering and the Ph.D. degree in computer science engineering from The Ohio State University, Columbus, OH, USA, in 2014 and 2018, respectively.

His research interests and thesis focused on game theoretical analysis and algorithm design in the field of network security and smart grid.

Zizhan Zheng (Member, IEEE) received the M.S. degree in computer science from Peking University, Beijing, China, in 2005, and the Ph.D. degree in computer science and engineering from The Ohio State University, Columbus, OH, USA, in 2010.

He was a Postdoctoral Researcher with the Electrical and Computer Engineering Department, The Ohio State University, from 2010 to 2014, and an Associate Specialist with the Computer Science (CS) Department, University of California, Davis, from 2014 to 2016. He is currently an Assistant Professor with the CS Department, Tulane University, New Orleans, LA, USA. His research interests include network optimization, machine learning, and cybersecurity.

Ness B. Shroff (Life Fellow, IEEE) received the Ph.D. degree in electrical engineering from Columbia University, New York, NY, USA, in 1994.

He joined Purdue University, West Lafayette, IN, USA, immediately thereafter as an Assistant Professor with the School of Electrical and Computer Engineering (ECE). At Purdue, he became a Full Professor of ECE and the Director of a university-wide center on wireless systems and applications in 2004. In 2007, he joined The Ohio State University, Columbus, OH, USA, where he holds the Ohio Eminent Scholar Endowed Chair in networking and communications, in the Departments of ECE and computer science engineering. He holds or has held Visiting (Chaired) Professor positions with Tsinghua University, Beijing, China; Shanghai Jiaotong University, Shanghai, China; and IIT Bombay, Mumbai, India.

Prof. Shroff is the recipient of numerous best paper awards for his research and is listed in Thomson Reuters on *The World's Most Influential Scientific Minds*, and noted as a Highly Cited Researcher by Thomson Reuters. He is also the recipient of the IEEE INFOCOM Achievement Award for seminal contributions to scheduling and resource allocation in wireless networks. He is currently the Steering Committee Chair for ACM Mobihoc and an Editor-at-Large of the IEEE/ACM Transactions on Networking.