Dynamic Games and Applications (2019) 9:884–913
https://doi.org/10.1007/s13235-018-00291-y

Dynamic Games in Cyber-Physical Security: An Overview

S. Rasoul Etesami[1] · **Tamer Başar**[2]

Published online: 1 January 2019

Abstract
Due to complex dependencies between multiple layers and components of emerging cyber-physical systems, security and vulnerability of such systems have become a major challenge in recent years. In this regard, game theory, a powerful tool for modeling strategic interactions between multiple decision makers with conflicting objectives, offers a natural paradigm to address the security-related issues arising in these systems. While there exists substantial amount of work in modeling and analyzing security problems using game-theoretic techniques, most of the existing literature in this area focuses on static game models, ignoring the *dynamic* nature of interactions between the main players (defenders vs. attackers). In this paper, we focus only on dynamic game analysis of cyber-physical security problems and provide a general overview of the existing results and recent advances based on application domains. We also discuss several limitations of the existing models and identify several hitherto unaddressed directions for future research.

Keywords Dynamic game · Cyber-physical security · Network security · Mechanism design · Learning · Security game

1 Introduction

Cyber-physical systems are integration of many components such as physical plants, cyber components (e.g., digital controllers or computing devices), and the communication networks among them. Such systems are inseparable part of our modern world and emerge in wide range of applications such as Internet, smart grids, sensor networks, or even smart transporta-

The research leading to this paper was supported in part by ONR MURI Grant N00014-16-1-2710 and in part by ARO Grant W911NF-16-1-0485.

✉ S. Rasoul Etesami
etesami1@illinois.edu

Tamer Başar
basar1@illinois.edu

[1] Department of Industrial and Enterprise Systems Engineering, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA

[2] Coordinated Science Laboratory, Department of Electrical and Computer Engineering, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA

tion systems. As cyber-physical systems become more involved into our daily life, serious concerns related to their security and vulnerability have been raised. In fact, in certain circumstances, failure of cyber-physical systems can be catastrophic such as terrorist attacks due to failure of national defense organization, political chaos due to hacking of classified information, or information leakage of nations' nuclear infrastructures. While security is arguably one of the most important challenges of our modern world, it is still far from being completely solved. This calls for a systematic study of cyber-physical security in order to prevent *strategic* attacks where the casual defensive methods have shortcomings in identifying them.

In recent years, various models for studying the security of the cyber-physical systems have been introduced. Since security can always be viewed as a conflict between two opposing objectives, it is quite natural to cast it as a multi-agent decision problem in which the defenders (e.g., system authorities) aim to protect the system against potential attackers. However, what makes the analysis of security systems challenging is the fact that most of the recent attacks are highly organized and strategic. In other words, the attackers are quite cautious about the potential consequences of their adversarial actions and hence take strategic actions to hide their identities. In this regard, game theory, a powerful tool for analyzing strategic multi-agent decision systems, proves quite useful to capture the interactions between security players. In fact, one can find a rich literature of survey type since the early 2000 on applications of game theory in security problems, which are mainly based on *static* models [66,76,96,109]. Although static game formulations provide good insights into the behavior of security players (and make quite a bit of sense in certain situations), in most cases they fail to capture the crucial dynamic characteristic of security problems, which is the fact that security players are repeatedly engaged in a multistage game in which the underlying environment can itself change dynamically over the course of interactions. This has motivated researchers in the field to enrich their models by considering more realistic *dynamic* game formulations.

In fact, studying cyber-security problems using dynamic games becomes even more sophisticated once we account for the information structure. This is because in most security problems the defenders and attackers do not have complete information on each other's payoffs, nor are they aware of each others' identities. Furthermore, due to limitation of monitoring devices, it is possible that once an attack has occurred, it is not identified by the system. As a result, the players may only have partial information about each other's past or current strategies. In addition, the existence of heterogeneous security players can completely ruin the symmetric structures of the game which introduces additional challenges for rigorous analysis of realistic cyber-physical security models. As it can be seen, cyber-physical security is a fairly complex problem, offering adversaries a large attack surface and ample room for evasive maneuvers. This requires new insights and novel techniques for modeling the behavior of security players, which not only is scalable to large numbers of players, but also takes into account the dynamic nonstandard information structure of the problem.

In this survey article, we provide an overview of advances in modeling and informational aspects of dynamic games for cyber-physical security problems. Different from most of the earlier game-theoretic security surveys [66,76,93,96,109], we focus here on only *dynamic* game formulations of cyber-physical security and discuss several common methods for analyzing them. In this effort, we group security problems based on both theme and application domains and approach them from diverse perspectives such as network security, mechanism design, learning, and optimal investment. Following a review of some relevant basic concepts of game theory in Sect. 1.1, we organize the rest of the discussion based on security applications and themes. More specifically, in Sect. 2, we address several important issues related to network security such as *risk assessment* and *intrusion detection*. We then look into several important classes of security games such as *signaling*, *deception*, and *Stackel-*

berg security games in Sect. 3. In Sect. 4, we consider decision making in the physical layer and in the presence of *jammers* or *eavesdroppers*. In Sect. 5, we approach cyber-security from the perspective of incentive and mechanism design. Optimal investment strategies for security players with limited resources are discussed in Sect. 6. We review several learning algorithms for finding optimal defense/attack strategies in Sect. 7. We conclude the paper by providing several future directions of research and discussing new emerging issues in the next generation of cyber-physical security problems in Sect. 8.

1.1 An Overview of Game Theory

In this subsection, we briefly review several important concepts and solutions from game theory, as relevant to this paper. We assume that, as appropriate for this journal, the readers have a basic knowledge of game theory, and particularly dynamic game theory; for a broader exposition, we refer the reader to [21].

Game theory provides a mathematical model for studying the conflict and cooperation among intelligent rational decision makers. A game is comprised of several elements: (i) *players*, who are the strategic entities participating in the decision-making process in the game; (ii) *actions*, which are players' decisions taken at each move of the game; (iii) *strategies*, which determine how the players select their actions based on their past and current information at each stage[1]; and (iv) *payoffs*, which are the rewards or punishments received by players as a consequence of their own actions and others'. A game can be either *static*, i.e., a one-shot game in which all players make decisions simultaneously without knowledge of others' strategies, or *dynamic* in which at least one player is allowed to use a strategy that depends on previous actions. In other words, a dynamic game is the one with sequential moves over multiple stages with new revealed information to the players at different stages. For dynamic games, a somewhat more generalized notion of strategy is so-called *policy* which is a sequence of strategies taken by a player at different stages of the game based on the observed information. In this paper, we only focus on dynamic game formulation of cyber-physical security problems (Table 1).

A game can be viewed as belonging to one of four possible groups, based on its underlying information structure. A game is of *complete information* if all the players know completely the structure of the game being played, such as the number of players in the game, payoff functions of the players, the underlying dynamics, the information structure, etc., and it is of *incomplete information*, otherwise. A game is of *perfect information* if all the players know the historical actions of each other at the time of their move, and it is of *imperfect information*, otherwise. Next, we identify several important *nonexclusive* classes of dynamic games which are particularly suitable for addressing security problems. Here, we want to stress the fact that each of the following games can itself be of the form of complete–perfect information, complete–imperfect information, incomplete–perfect information, or incomplete–imperfect information, depending on the description of the game.

- *Zero-sum game* A class of multistage dynamic games in which the sum of players' payoffs at each stage is identical to zero.
- *Stochastic game* A dynamic multistage game where at the beginning of each stage the game is in some state. The players choose their actions and receive payoffs that depend on the current state and their chosen actions. The game then moves to a new state with some transition probability which depends on the previous state and the actions chosen

[1] A strategy can be either pure or mixed, meaning that a player can either choose a particular action with probability 1, or based on a probability distribution over its set of possible actions.

Table 1 A nonexclusive classification of security games based on the theme and the game structure

Theme	Application	Dynamic game structure						
		Zero sum	Stochastic	Repeated	Differential	Stackelberg	Bayesian	Others
Network security	Intrusion detection	[117, IC]	[86, P],[121, P],[73, P]	[84, IP-IC]		[113, P]	[44, P-IC]	[45,72,85,88,94]
		[112, IP],[68]	[6, IP],[117, IC],[73, C]	[5, IP–IC]			[90, P–IC]	[27,71,73,77,101,117,123]
	Risk assessment		[110, P–C],[104, P–IC]			[103, P]		[8,29,64]
Security games	Signaling games		[89, P–AS]	[69], [62]			[89, P–AS]	[31,69]
	Honeypot/deception	[112, IP]	[62],[30, P]	[62]		[119]	[62],[30, P]	[32,62,79,91]
	Cascading games	[60]			[60],[23]			[23,60,115,118]
	Stackelberg security					[9,11,40,58,59,61,87,106]		[9,11,40,58,59,61,87,106]
	CBG/hypergame		[46,53]	[50, C]				[46,50,53]
Physical layer	Jamming and eavesdropping	[13,15,19,20,80]	[2,16]	[68]		[3]	[98, IC], [68]	[17],[97]
		[49,67,74,75]	[67]	[18]		[41]	[99, IC]	[24–26,65]
Incentives	Mechanism design		[42],[70]	[83],[1]	[22]	[22]	[70]	[1,22,33,35,39,42,70,83,116]
Resource allocation	Security investment		[106], [48, AS]	[78], [55, IC]	[12]	[106],[92]	[54]	[108]
			[9]	[28]		[9, IP],[54]	[48, AS]	[56, AS]
Learning in security	Reinforcement learning	[29],[51]	[29],[37],[51]	[107]	[120]			[29,37]
		[120]	[120],[122, IC]			[106]		[51,106]
	Regret-Based learning	[47,107]	[107]	[11]		[11]		[58]
		[87]	[47]	[47]		[87]		[47]

Letters P, IP, C, IC, and AS next to a reference shows the information structure of the game being perfect, imperfect, complete, incomplete, or asymmetric by all the players. The total payoff to a player is typically defined to be the discounted average of his payoff over the course of the game.

- *Repeated game* A game that consists of a number of repetitions of some base static game, usually referred to as the *stage game*.
- *Differential game* A game in which the players' payoffs depend on the evolution of a state and their actions over a finite or infinite time horizon. The state evolution is usually captured by a differential equation which also depends on the players actions throughout the time horizon.
- *Stackelberg game* A hierarchical game that in its simplest form is composed of a leader and a follower. The leader chooses his strategy first. The follower, observing the leader's strategy, maximizes his payoff by best responding to the leader's strategy.
- *Extensive form game* An explicit representation of a dynamic game which displays the evolution of the game by showing important factors such as the information structure, the possible actions for players at every decision stage, and the sequence of players' moves.

In fact, we can also consider a combinations of the above games. For instance, we can talk about zero-sum stochastic differential games which is a zero-sum stochastic game whose state evolution is driven by a probability transition rule or a stochastic differential (or difference) equation which depends on players' actions at each stage [63]. As another example, one can consider repeated zero-sum or differential games with incomplete and asymmetric information, where the players' knowledge about the underlying game is neither complete nor identical [63,105].

An important solution concept in game theory is *equilibrium* which refers to a joint strategy (policy) profile from which no player has a unilateral incentive to change his strategy (policy) within the rules of the game. The concept of equilibrium can also be readily adopted for each of the aforementioned types of dynamic games, which yields the notions of *Nash* equilibrium, *Stackelberg* equilibrium, *stationary* equilibrium, *saddle-point* equilibrium, *Bayesian* equilibrium, etc. However, at this point we will not go further into the details of these notions and we refer the readers to [21] for precise definitions of these equilibrium concepts; see also [4] for a security perspective.

2 Network Security

Since the early 2000s, a large amount of effort has been devoted to analyzing security issues connected to networks. In this section, we review several important topics within the context of network security with analysis based on dynamic games.

2.1 Intrusion Detection and Information Limitations

Intrusion detection (ID) is one of the important ingredients of recent computer networks which compensates the shortcomings of standard prevention methods. The main task of an intrusion detection system (IDS) is to detect intrusions by monitoring the events occurring in the network and report them to a system administrator in order to stop or mitigate the effects of an attack [66].

In [86], a two-player stochastic game between an attacker and a defender over a network of nodes was proposed. Here, the network captures the influence among the nodes. Each state of this stochastic game can be represented by a binary vector with the ith entry being 0 or 1, depending on whether node i is compromised or not. At each state of the game, the attacker decides either to attack one of the network nodes or to do nothing, while the defender decides to either defend one of the nodes or to stay idle. The possible combinations of the players' pure strategies at a given state define a payoff matrix corresponding to that state whose entries are determined based on the effective security assets at that stage as well as transition probabilities of moving to the next state. Moreover, there is always a chance that the game ends at any stage, meaning that either the defender has detected the attacker and stopped him from further intrusion, or the attacker has stopped attacking. The challenge in analyzing such a stochastic game is that due to correlation among the nodes, if a node is compromised, the effective security assets and the supports of the remaining nodes will change and have to be recalculated. For this game, the authors show existence, uniqueness, and structure of the equilibrium solution and justify the model effectiveness through numerical examples.

Stochastic game formulations for security problems are quite useful when there are imperfections in the sensor network or ID monitoring system in which case the transition probabilities between game states map attacker actions to sensor outputs. In particular, when limitations are imposed on information available to players, then the outcome and evolution of the stochastic game highly depend on how successful the players can estimate the state of the security system (and hence the game) which also depends on their learning rate. In this regard, the work in [6] presents a two-player stochastic game with imperfect information and uses a Q-learning method to evaluate how the players learn and optimize their strategies. It is worth noting that imperfect measurements can also be addressed with fixed structure observers/estimators, which are different in spirit from Q-learning. A similar approach has been used in [121] to extend the two-player stochastic game for IDS to multiple attackers and defenders. However, the proposed framework in [121] assumes a perfect information structure for the stochastic game and uses value iteration method to approximate the stationary Nash equilibrium of the game. Other zero-sum or nonzero-sum stochastic game formulations for security of IDS with different assumptions on their information structure can be found in [27,73,77,101,117,123]. For instance, the work in [73] provides a stochastic game with complete and perfect information and illustrates how one can compute the Nash equilibrium using a nonlinear program.

The work in [71] takes a different path by looking into IDS as a two-player repeated Bayesian game. One player is a potential attacking node which can operate in two different modes (types): *regular* or *malicious*. However, the type of the attacker is *unknown* to the other player. The other player is a defending node which has a regular type and is *known* to the attacker. The attacker in its malicious type can either choose to *attack* or *not attack*, while in its regular type can only choose *not attack*. The defending node which always has the regular type can either choose to *monitor* or *not monitor* the network. The players simultaneously make their moves at each round, whose payoffs are determined by a payoff matrix reflecting the cost of damage to the system. The paper provides a Bayesian rule that the defender can consistently update his belief on his opponent's maliciousness as the game progresses. Moreover, the paper provides an energy-efficient monitoring strategy for the defender by extending the setting to a Bayesian hybrid detection approach where the defender can choose between two alternatives: a lightweight monitoring system to estimate his opponent's actions and a heavyweight monitoring system acting as a last resort of defense. We mention here [44] and [90] as two other examples of dynamic Bayesian game formulation for IDS with incomplete but perfect information structures, with the analysis relying on the notion of *perfect Bayesian equilibrium*.[2] In fact, what makes the concept of perfect Bayesian

[2] A perfect Bayesian equilibrium is a set of strategies and beliefs for every player at every information set, so that the beliefs are derived from the strategies and common prior beliefs using Bayes' rule, and the strategies are optimal at every point in the game, given the players' beliefs.

equilibrium suitable in analysis of security systems and in particular IDS is that a defender cannot directly detect the attacker's activities and vice versa. Therefore, in such a situation, if a player wants to evaluate his strategy, he needs to gradually form beliefs about his opponent's private information that is relevant to his objective.

The work in [5] adopts a repeated incomplete–imperfect information game to model a security game between an attacker and IDS where the imperfection of information stems from the fact that in reality the distributed virtual monitoring sensors may have imperfect detection capabilities. It is shown that the single-stage game admits a unique Nash equilibrium and several basic strategies for the players of the repeated game are discussed through numerical analysis. [84] looks at IDS as a repeated two-player nonzero-sum game with incomplete–imperfect information and proposes a *fictitious play*[3] dynamics in which players improve their strategies based on their past observations. In particular, the convergence properties of the proposed dynamics to the Nash equilibrium in the presence of perfect or imperfect observations are studied. We refer to [45,72,85,88,94] for alternative dynamic game formulations for IDS, network security, and other applications of fictitious play. The authors in [113] adopt a slightly different setting to study network security with potential application to IDS. They consider a two-stage stochastic Stackelberg game where the attacker's goal is to interdict a smart network from routing maximum flow from an origin to a destination. In this game, the results of attack actions are stochastic, and the attacker observes outcomes from an initial attack before choosing a second attack. Using a so-called *branch and bound* technique, the paper provides a fast method for solving the optimal strategy for the intruder with increasingly accurate lower bounds.

2.2 Risk and Security Assessment

Risk management and security assessment is a method to obtain and evaluate the current and future security status of a network information system and is quite beneficial in reducing risks and helping the network system to reach a certain security level. In this regard, [104] provides an incomplete information stochastic game framework to estimate the belief of each possible attack pattern and generate the corresponding defensive strategy. The players' payoff functions, as defined in [104], are generally nonlinear, and the corresponding game is hard to be analyzed toward explicit, closed-form solutions. Instead, the paper provides a defensive strategy based on fictitious play for the approximated linearized system and evaluates its effectiveness using simulation results and a developed software.

Within the same framework, a complete–perfect information stochastic game model between a threat agent and a vulnerability agent for risk assessment was proposed in [110]. In this game, the threat agent aims to increase the system risk by spreading the threat, while the vulnerability agent aims to mitigate the risk by repairing the information system. In this game, the state of the game at some time instant $k = 0, 1, \ldots,$ denoted by $s(k)$, is the entry-wise product of two binary vectors $s^{\mathrm{t}}(k)$ and $s^{\mathrm{v}}(k)$, where $s_i^{\mathrm{t}}(k) = 0$ or 1 denote, respectively, *no threat* or *threat* to the ith asset (node) of the network at time step k. Similarly, $s_i^{\mathrm{v}}(k) = 0$ or 1 mean, respectively, that the ith asset is *not vulnerable* or *vulnerable* at time k. In particular, the state of the entire game at time k equals $s(k) = (s_1^{\mathrm{t}}(k)s_1^{\mathrm{v}}(k), \ldots, s_n^{\mathrm{t}}(k)s_n^{\mathrm{v}}(k))$, where n denotes the total number of system assets. At each stage k, an action for the threat agent, denoted by $u^{\mathrm{t}}(k)$, is to spread the threat to one asset, while an action for the vulnerability agent is to repair vulnerability of one of the assets. Given actions of the players at stage k, the

[3] Fictitious play is a learning rule in which at each round, each player best responds to the empirical frequency of play of his opponents.

threat spreads over the assets at the next time stage based on a transition probability given by $\mathbb{P}(s(k+1)|s(k), u^{\mathrm{t}}(k), u^{\mathrm{v}}(k))$, and the system incurs a damage $V(s(k))$ which depends on the current risk of the system, $R(s(k))$, and the possible future damages given by

$$\begin{aligned} V(s(k)) = R(s(k)) \\ + \beta \sum_{s(k+1)} \mathbb{P}(s(k+1)|s(k), u^{\mathrm{t}}(k), u^{\mathrm{v}}(k)) V(s(k+1)), \end{aligned} \tag{1}$$

where $\beta \in (0, 1)$ is a discounting factor. This formulation allows to take into account the future security status, which also has impact on risk assessment. Therefore, according to the value function (1), one can compute the system risk induced from each threat and accumulate them to get the aggregate risk of the system. Using this risk measure, the paper proposes an automatic generated reinforcement scheme for the system administrator in order to mitigate the amount of vulnerabilities utilized by threats and make the system safer. Along the same research thrust, the authors in [100] use a stochastic game formulation to measure the trustworthiness of a security system and to assess how the parameters of the game may change the expected behavior of the attacker.

Different from [104,110], the authors in [103] propose a hierarchical three-level Stackelberg game with perfect information between an attacker and a defender for security assessment of electricity distribution networks. In this game, the defender first chooses a security strategy to secure a subset of network nodes. Then, the attacker compromises a set of vulnerable nodes and injects false generation set points. In the last stage, the defender responds by controlling loads and uncompromised network nodes. Finding the Stackelberg equilibrium for this three-level game can be cast as a nonlinear mixed-integer program which is hard to solve. Instead, the authors in [103] use a practical linear approximation to find the critical nodes and the optimal attack strategy. In particular, they propose an iterative greedy algorithm to compute attacker's/defender's strategies for the original nonlinear problem. In [114], the authors study risk management in security networks using a linear influence directed graph which captures the dependencies between security players. As a distinct feature of their model, the authors allow the network environment to be stochastic and time varying under which they study the stochastic behavior of the best response dynamics. In particular, the paper provides sufficient conditions under which the existence, uniqueness, and convergence of the best response dynamics are guaranteed. We refer to [8,29,64] for other related works on risk assessment and risk management using dynamic game formulation.

3 Security Games

In this section, we review several popular and well-studied game-theoretic models for cyber-physical security problems. While each of these games can be viewed as a special subclass of dynamic games given in Sect. 1.1, their inherent properties make them particularly suitable for studying security problems.

3.1 Signaling Games

In its simplest form, a *signaling* game is a dynamic Bayesian game with two players (a *sender* and a *receiver*) in which the sender has several types which is private to him and determines his payoff function. The receiver, however, has only one type, and hence, his payoff is common knowledge to both players. The sender takes an action first by sending a message. The receiver observes the sender's message and then takes his action. The players then receive some payoffs depending on sender's type and their actions. As in many security problems, the defender (receiver) does not have knowledge about the attacker's (sender's) target list, it seems quite reasonable to model the interactions between the attacker and the defender in a security problem using a signaling game [31,69,89,90].

The authors in [69] develop a repeated two-way signaling game to model the multi-step attack–defense scenario on confidentiality, and show how to find the actions maximizing the expected payoffs through the equilibrium. In their proposed game, both attacker and defender have incomplete information about their opponent and use the received signals to gradually reduce their uncertainty about their opponents. Using a case study, it is shown that upon receiving an attack alert, if the defender analyzes the attack targets through the equilibrium of a basic signaling game, then he can gradually learn and deploy his optimal strategy.

3.2 Honeypot and Deception Games

Deception is a method which can be employed by both attackers and defenders in advanced security problems to make the situation ambiguous, and hence, maximize their payoffs. Using deception in security systems makes the situation more complex as now a player relies less on his opponents' strategies in arriving at his decision [30,32,62,79,91]. As discussed in [32], the information asymmetry in most security problems (e.g., one player may have more information than the other) is one of the main reasons that contributes to players acting deceptively in order to gain advantages. In fact, deception can also be used in active cyber *defense* to manipulate the beliefs of an adversary. Perhaps, *honeypot* is one of the most common deception techniques for cyber defense, which is an effective deception mechanism set by the security system to detect, deflect, or counteract attempts of attackers in an information security system [30,34,62].

In [30], the authors use a deceptive signaling game to study the behavior of attacker and defender in an information security system. The defender can either disguise a normal system as a honeypot or disguise a honeypot as a normal system. For this model, the authors determine the set of perfect Bayesian equilibrium strategies in which no player has any incentive to deviate unilaterally. They also extend the model to a hybrid setting which can accommodate a mix of security systems such as honeypot and normal system, which allows more effective deceptive strategies for the defender. The work in [62] looks at a security problem within the context of Internet of Things (IoT) and formulates it using a repeated signaling game between an attacker and a defender. In this game, the attacker can deceive the defender by employing different types of attacks ranging from a suspicious to a seemingly normal activity, while the defender can deceive and trap the attacker using honeypots. For this game, the authors fully characterize the set of perfect Bayesian Nash equilibrium strategies and use simulation results to show that deploying honeypots by defenders can be quite effective when there is a high concentration of active attacks.

The paper [112] looks at the network interdiction problem using the formulation of a multistage zero-sum game between an attacker and an intelligent network defender. The game is posed as an imperfect information deception game in extensive form where the attacker has imperfect knowledge of the network topology, but it can learn about the topology by monitoring network operations. The network observes the attacker's actions and can choose to avoid using the observed parts of the network in order to disguise information from the attacker. A dynamic programming-like algorithm based on partially observed Markov decision processes (POMDPs) is then developed to obtain a complete set of equilibrium solutions of the game.

3.3 Cascading Games and Robustness

Cascading failure is a phenomenon in security systems where failure of a subsystem can result in failure of successive subsystems. In other words, the failure can propagate over the system once an attack is successful. This issue which can put systems at a serious risk, particularly when there are strong interconnections among the subsystems, has been addressed from a dynamic game-theoretic perspective in recent literature [23,60,115,118].

In [60], a zero-sum differential game for studying the behavior of the malware and the defense in a wireless network which is vulnerable to a cascading failure has been proposed. More specifically, the game is played between two players: a network administrator N and a malware M. At any time t, the state of the game is characterized by a triple $(S(t), I(t), D(t))$ where $S(t)$, $I(t)$, and $D(t)$ denote, respectively, the fraction of network nodes at time $t \in \mathbb{R}^+$ which are *susceptible* to worm, are *infected* by worm, and are *dead*. Here, a susceptible node refers to a node which is not contaminated by the worm, but has potential to be infected. An infected node is already contaminated by the worm and can infect susceptible nodes through communication with them. The worm can kill an infected host which is then declared dead. A functional node that is immune to the worm is referred to as *recovered*.

A strategy for the network N at time t, denoted by $u^N(t) := (u^{N_r}(t), u^{N_i}(t))$, is composed of two components. $u^{N_r}(t)$ determines the communication rate between network nodes at time t (note that this implies that those that are susceptible are transformed to being infective at a rate proportional to $u^{N_r}(t)I(t)S(t)$ as the number of susceptible-infected pairs at time t is proportional to $I(t)S(t)$). Moreover, $u^{N_i}(t)$ determines the rate of communication between network nodes with dispatchers at a cost of using network budget. If the node that receives the patch is a susceptible node, it installs the patch and recovers. Otherwise, if an infective node receives the patch, it is recovered with some probability h, and else, the patch fails to cure and the node remains infective. On the other hand, the malware at an infective host kills the host with rate proportional to $u^M(t)$ at time t. Using the above description, one can capture the evolution of cascade dynamics by

$$\begin{aligned}
\dot{S}(t) &= -\beta_1 u^{N_r}(t)I(t)S(t) - \beta_2 u^{N_i}(t)S(t),\\
\dot{I}(t) &= \beta_1 u^{N_r}(t)I(t)S(t) - \beta_2 u^{N_i}(t)I(t) - u^M(t)I(t),\\
\dot{D}(t) &= u^M(t)I(t),
\end{aligned}$$

where β_1, β_2, and β_3 are nonnegative constants and $\dot{S}(t)$, $\dot{I}(t)$, and $\dot{D}(t)$ are the rates of change in their corresponding quantities. These relations characterize the evolution of the cascade dynamics in the underlying differential game. Finally, the damage over the time horizon $[0, T]$ (or equivalently the payoff to the malware) is given by

$$J(u^N, u^M) = \int_0^T \Big(\alpha_1 I(t)D(t) + \alpha_2 u^{N_i}(t) - \alpha_3 u^{N_r}(t)\Big)\, \mathrm{d}t + \alpha D(T),$$

where $\alpha_1, \alpha_2, \alpha_3$, and α are nonnegative scaling constants. In this differential game, player N wants to minimize $J(\cdot)$ subject to some budget constraints, while the malware wants to maximize this overall damage function.

It was shown in [60] that the above differential game admits a pair of saddle-point strategies (u^{N_*}, u^{M_*}) such that

$$J(u^{N_*}, u^M) \le V := J(u^{N_*}, u^{M_*}) \le J(u^N, u^{M_*}), \quad \forall u^M, u^N.$$

Therefore, a *robust* strategy for the network player is to choose its saddle-point strategy u^{N_*}, in which case irrespective of the strategy of the malware, the damage it incurs is at most V. In particular, it is shown that the structure of the network's saddle-point strategy is of the form of a *threshold policy* in which there are time instances $t_1, t_2 \in [0, T]$ such that for $t \leq t_1$ and $t > t_1$, u^{N*r} attains its minimum value and maximum value, respectively. Similarly, for $t \leq t_2$ and $t > t_2$, u^{N*i} attains its maximum value and minimum value, respectively. This simple structure allows the network player to effectively employ a robust defensive strategy.

Along the same line of work, [23] provides a differential game formulation for the cascade behavior in *Botnets* (computer networks infected with malicious programs) using a modified SIS (susceptible–infectious–susceptible) epidemic model. In particular, the authors are able to establish similar saddle-point strategy characterization based on threshold policies. For other dynamic game approaches to investigating the coupling between cyber-security policy and robust control design in the presence of cascading failures, we refer readers to [115,118] (and the references therein).

3.4 Stackelberg Security Games

One of the widely studied models in security games is known as the *Stackelberg security game* (SSG), where in its simplest form, the defender acts first and assigns resources to potential targets. The attacker observes the defender's strategy and decides to attack one of the targets to maximize his utility. Such games have been the topic of many research papers in security and have been studied under various settings such as repeated SSG under *known* or *unknown* attacker types [9,11,40,58,59,61,87,106]. In fact, as we shall see in Sects. 6 and 7, SSG provides a suitable framework for learning the behavior of the attackers as well as finding optimal allocation strategies when there is a limited amount of security resources. Furthermore, as discussed in [9,11,61], SSGs have many real-world applications in airport security, the Federal Air Marshals, wildlife protection, screening incoming shipments at ports, and patrolling subway systems.

The authors in [61] consider a SSG and provide fast algorithms for computing its optimal strategies, which scale to many resources and targets. The key element in their algorithms is to exploit structural properties of optimal solutions under the payoff restrictions; they consider SSG in a *compact* form due to the fact that, in their formulation, the payoffs depend only on the identity of the attacked target and whether or not it is protected by the defender. As a result, the authors do not need to face a normal form representation of the game which can have exponential size for multiple resources and targets. The work in [59] considers a SSG within a bounded rationality framework and develops a model which takes into account the human behavior decisions. [40] provides a more generalized SSG for wildlife protection by relaxing a standard assumption in SSG that the attacker observes the strategy of the defender before taking his action. Instead, [40] introduces a generalized Green Security Games and provides algorithms to plan effective sequential defender strategies by learning attacker's behavior.

3.5 Other Types of Security Games

In this subsection, we discuss some other forms of dynamic games, which are useful in studying security problems. In [50], a complete information three-player three-stage *Colonel Blotto* game was introduced,[4] with motivation coming from security. In this game, there are two resource-constrained networks of servers and a resource-constrained hacker who wants to

[4] Colonel Blotto game is a multi-dimensional problem on strategic resource allocation. In its classic form, it is a two-player game in which two colonels are tasked with allocating a limited number of troops over multiple

access the servers. In the fist stage of the game, the network operators can invest in additional servers (equivalent to adding battlefields to the Colonel Blotto game). In the second stage, the network operators can share resources among themselves for further securing the servers. In the third stage, the hacker observes the security level of each network and decides on the amount of resource it deploys to hack each of the servers of the two network operators. For a certain range of game parameters, the paper characterizes the subgame-perfect Nash equilibrium of this game and studies whether adding battlefields would be more effective than sharing resources between the network operators. In particular, the paper studies the worst case behavior of the hacker in this security game.

The works in [46,53] look into the cyber-security problem using a *hypergame* formulation. Hypergames are extensions of classical game models, where players may have different understanding of the game due to misperceptions, differences in understanding the game rules, etc. A hypergame contains several different subgames where at each stage possibly one of them is being played by the players. As the security players generally have asymmetric or unknown information about the game, hypergame theory provides many rooms for modeling security type problems. This is because the attackers and defenders can invoke different learning strategies to understand what subgame they are actually playing and to learn how to deceive their opponents about different subgames in order to maximize their own payoffs.

4 Security and Decision Making at the Physical Layer

An important portion of security attacks happen at the physical layer. *Jamming* is an active attack which usually happens at the physical layer of a communication system, where the attacker disrupts the normal communication between a transmitter and a receiver by concurrently transmitting some data to corrupt the flow of information or to mislead the receiver. On the other hand, *eavesdropping* is a passive attack where the attacker overhears the victim's transmitted or received signals which can then be analyzed toward malicious purposes. In fact, in reality, jammers and eavesdroppers can coexist in a communication network, and therefore, having a mathematical model to understand and mitigate, and even stop such adversarial effects in communication systems becomes an important issue. Next we review several dynamic game formulations related to both jamming and eavesdropping.

A prototype for a game-theoretic formulation of jamming in communication systems was introduced in [15] as an extension of the so-called Gaussian test channel, where now there is an intelligent jammer who has access to a possibly noise-corrupted version of a Gaussian random variable transmitted over a Gaussian channel under a quadratic distortion measure, which the transmitter (encoder) and the receiver (decoder) want to jointly minimize while the jammer wants to maximize. More specifically, a Gaussian random variable u with zero mean and unit variance ($u \sim N(0, 1)$) is to be transmitted over a Gaussian channel with input power constraint c^2 and additive Gaussian noise $w = w_1 + w_2$, with zero mean and variance $\xi = \xi_1 + \xi_2$, with $w_i \sim N(0, \xi_i)$, $i = 1, 2$. The transmitter applies a transformation $\gamma(\cdot)$ on u, and the jammer taps the channel and accesses a noisy version of $x := \gamma(u) + w_1$, denoted by $y = x + v$, where $v \sim N(0, \sigma)$ and all random variables (u, w_1, v, w_2) are statistically independent. Using the observed value of y, the jammer feeds back into the channel a second-order random variable, $\nu = \beta(y)$, so that what the receiver receives is the corrupted signal

Footnote 4 continued
battlefields, with the player allocating the most troops to a front being declared the winner, and the overall payoff being proportional to the number of fronts won.

$z = x + v + w_2$, to which it applies a transformation $\delta(\cdot)$ to arrive at an estimate of u under the quadratic distortion measure. The jammer's transformation (strategy) β is allowed to be random, but it has to have a hard-bounded second moment: $E[v^2] \leq k^2$, for some positive constant k; let the associated probability measure for the jammer be denoted by μ. For each fixed triple $(\gamma, \delta; \mu)$, with γ and μ satisfying the given energy constraints, let $R(\gamma, \delta; \mu)$ denote the mean-squared distortion measure, that is,

$$R(\gamma, \delta; \mu) = \int_{-\infty}^{\infty} E\{[\delta(z) - u]^2 | v\}\, d\mu(v),$$

which is to be minimized by the pair (γ, δ) and maximized by μ. More precisely, the paper studies the existence and characterization of a saddle-point solution, $(\gamma^*, \delta^*; \mu^*)$, satisfying the pair of saddle-point inequalities

$$R(\gamma^*, \delta^*; \mu) \leq R(\gamma^*, \delta^*; \mu^*) \leq R(\gamma, \delta; \mu^*)$$

for all permissible $(\gamma, \delta; \mu)$. The paper [15] provides a complete solution to this problem, showing that a saddle point exists (and is essentially unique) for all values of the parameters defining the game, but has different structures and characterizations in three different regions of the parameter space. These regions are determined by the signal-to-noise ratios and relative magnitudes of the noise variances. The best (maximin) policy of the jammer is either to choose a linear function of y, or to choose, in addition, an independent Gaussian random variable (as additive noise), depending on the region where the parameters lie. The optimal (minimax) policy of the transmitter is to scale the input random variable u to the given power level by a linear transformation (that is, γ^* is linear), and that of the receiver is to use a Bayes estimator (which is linear).

The model and results of [15] were then extended in different directions in follow-up works. The model in [19] allows for the jammer to tap the input to the encoder (rather than the input to the channel) and thus be able to directly correlate its input to the channel with the message to be transmitted. In this case, a saddle point exists only if the encoder is allowed to randomize and a side channel of a specific nature is allowed between the transmitter and the receiver; the saddle-point solution again involves only linear maps. If the encoder is not allowed to randomize, however, then there exists no saddle-point solution, but minimax and maximin solutions exist, whose complete characterizations are given in the paper. A follow-up paper [20] extends these results to vector sources and vector channels and obtains secure strategies for the transmitter (and the receiver) within the class of linear policies. Another paper, [13], revisits the model of [15], but this time with the hard constraints on the transmitter and jammer policies replaced by partially or totally soft constraints. Depending on whether the soft constraints are partial or total, the paper shows that in most cases a saddle point exists and is linear for the transmitter–receiver pair and is Gaussian (partially correlated with the message) for the jammer. In one case (specifically, when the jammer is subject to a hard constraint and the transmitter operates under soft power constraint), the maximin policy is not well defined. Another extension has been provided in the recent paper [2], which looks at the general source-channel case (not necessarily Gaussian), and shows that under certain conditions linearity of transmitter–receiver policies prevails, as in a sense the presence of the jammer forces the transmitter–receiver to use linear policies to secure an achievable least upper bound on the quadratic distortion. Another set of results on communication jamming games has been given in [16–18], where the model is now in continuous time, involving transmission of a Gaussian stochastic process being encoded and transmitted over a continuous-time Gaussian channel, which is also under a jamming attack, where the jammer can tap the system either at the input or at the output of the transmitter. For different scenarios, the papers prove the existence of saddle points and provide characterizations as linear Gaussian policies, when there is a clean feedback link from the receiver to the transmitter. In most cases, the receiver has the structure of a Kalman filter.

Another type of a jamming game has been introduced and solved in the paper [98], where the game is among a set of transmitters and jammers, and is played at the medium access control layer. The game is of the form of Bayesian incomplete information in which the users do not have complete information about each other's identities. In particular, using both gradient and fictitious play in the repeated jamming game, the performance loss of jamming attacks with respect to network uncertainty is evaluated. [80] proposes a dynamic zero-sum game to model the interactions between a transmitter and a joint-eavesdropper/jammer who can choose to behave as either a passive eavesdropper or an active jammer. Moreover, the equilibrium strategies for the extensive form of the game where the players move sequentially are derived. A discrete-time two-player zero-sum dynamic game to model the interactions between a communicator and a jammer has been given in [75]. In this game, both transmitter and jammer are subject to temporal energy constraints such that in each time slot, they must choose their respective power levels randomly to be either zero or a positive value in order to maximize their average payoffs over the entire horizon. Depending on the underlying parameters of the players' payoffs, the authors distinguish between two different types of behavior on the players' optimal stationary strategies such that in different regions playing mixed or pure strategies become dominant.

Another representative work that is concerned with security in the presence of eavesdroppers is [97], which has introduced a game-theoretic formulation that enables in multi-hop networks a number of wireless nodes to interact and optimize the security of their uplink transmissions. In the game proposed, the strategy of each node is to choose its preferred path to reach the base station, while optimizing physical layer security-related utilities. The type of adopted utility depends on the knowledge that the nodes have about the eavesdroppers channels. To solve the game, the paper introduces a distributed algorithm that enables the nodes to engage in pairwise negotiation so as to decide on the graph structure that will interconnect them. It is shown that the algorithm converges to a Nash network and leads to significant performance gains in terms of both the average bottleneck secrecy rate per node and the average path qualification probability per node, relative to classical algorithms and the star network.

In a different application, a two-player zero-sum dynamic game between a jammer and a smart grid communication system is considered in [74]. In this game, the jammer aims to interfere real-time communication in order to manipulate the electricity price toward its benefits. Using dynamic programming and extensive form representation of the game, the authors construct the saddle-point equilibrium strategies and provide further demonstration of their optimality. Within the same context, a stochastic zero-sum game between a jammer and a smart grid communication system was formulated in [67] in order to study a denial-of-service for remote state monitoring. In [49], a discrete-time dynamic zero-sum game between a controller of a linear time-invariant plant and a jammer was introduced. In this game, it is assumed that the jammer has only M possibilities of intercepting the communication between the controller and the plant over a horizon N, where $M < N$. At each time $k = 0, \dots, N-1$, the jammer can either choose to jam $\alpha_k = 0$, or to stay inactive $\alpha_k = 1$ and let the control signal reach the plant, where $\sum_{k=0}^{N-1} \alpha_k = N - M$. Denoting the controller's action at time k by u_k, the state of the plant $x_k \in \mathbb{R}$ evolved as

$$x_{k+1} = Ax_k + \alpha_k u_k + w_k, \quad k = 0, \dots, N-1,$$

where $\{w_k\}$ is a discrete-time zero-mean Gaussian white noise process with variance σ_w^2 and x_0 is a zero-mean Gaussian random variable with variance σ^2. The objective function of the jammer is to maximize its expected payoff function given by

$$J := \mathbb{E}\left[\sum_{k=0}^{N-1}(x_k^2 + \alpha_k u_k^2) + x_N^2\right],$$

while the controller aims to minimize this quantity. At each point in time, the controller and jammer have access to the current and past values of the state and have full memory on whether any of the previous control transmissions were intercepted or not. Denoting this information set at time step k by I_k, the jammer and controller seek, respectively, saddle-point strategies $\alpha_k^* = \mu_k^*(I_k) \in \{0, 1\}$, and $u_k^* = \gamma_k^*(I_k) \in \mathbb{R}$ such that $J(\gamma^*, \mu) \leq J(\gamma^*, \mu^*) \leq J(\gamma, \mu^*)$, for all admissible strategies γ and μ of the controller and the jammer. By extending the state space of the above game from x to (x, t, s) to account for the number of remaining game stages (t) and remaining jamming opportunities (s), it was shown in [49] that this jamming game admits a saddle point. In particular, it was shown that for the case of $M = 1$ (i.e., when the jammer can act only once), the optimal saddle-point strategy is a threshold-based policy which is further characterized in the large state limit.

The paper [119] develops a multi-hop multistage dynamic *deception* game framework for data routing in communication networks. In this game, the network nodes aim to send some amount of valid data from a source to a destination, while the network is subject to adversarial jamming effects. In the first stage of the game, the source strategically splits its data into two intermediate routing nodes. One is legitimate and the other one is deceptive to distract the jammers. In the second stage, the nodes search for the optimal multi-hop paths to the destination in response to jamming behaviors from adversaries. The underlying game is defined as a Stackelberg game, and in this context, the paper introduces a number of solution concepts, specifically path Stackelberg equilibrium (PSE), rate Stackelberg equilibrium (RSE), and their counterparts in behavioral and mixed strategies, and has obtains closed-form expressions for PSE and RSE when the utility functions are logarithmic. These solutions lead to an assessment of the benefit that arises from deception.

Paper [3] proposes a game-theoretic framework to model optimal secure communication over a sensor network. In this game of a single Gaussian source observed by multiple sensors after being corrupted by independent additive Gaussian noises, some of the sensors are normal and want to communicate with the receiver under minimum mean-squared error, while others act as adversaries and aim to maximize the distortion in communication. It was shown in [3] that for a certain range of parameters, the underlying game reduces to a zero-sum game while for others the game becomes a Stackelberg game (with normal transmitting sensors and adversarial ones acting as leader and follower, respectively). In both cases, the paper characterizes the structures of the saddle-point or Stackelberg strategies.

A Bayesian game framework was used in [99] to model wireless medium access control in which selfish transmitters (players) decide on their power level or transmission probability. In this game, transmitters may have incentive to act as malicious jammers in order to maximize their own payoffs, which results in an incomplete information game among the selfish and malicious transmitters. For this game and under different degrees of uncertainties, the Bayesian Nash equilibrium strategies and a learning mechanisms for the type belief updates are derived. In [65], the problem of sensor node communication in the presence of a jamming attack is considered. The problem is cast as a discrete-time zero-sum dynamic game with linear time-invariant state process in which both the sensor and the attacker have energy constraints. It is shown that optimal strategies for both players constitute a Nash equilibrium of this game which can be obtained using a *Lagrange multipliers* method in the case of offline setting (i.e., when the players choose their entire policies at the beginning of the game). Moreover, the game is extended to an online setting where players sequentially observe the game outcomes and update their strategies at the next time step, and a recursive algorithm to update both sides decision processes is provided. The paper [68] studies security strategies in two-player zero-sum repeated Bayesian games with application to jamming in sensor networks. In such games, each player has a private type chosen from a known distribution. At every stage, players simultaneously choose their actions which are observed by the public and receive payoffs which depend on players' types and actions, and are not directly observed by any player. The paper develops explicit algorithms to compute the security strategies of the players, i.e., players' optimal strategies in the worst case.

Another class of jamming games arises in the framework of mobile networks, where nodes are under constant motion (such as UAVs or terrestrial vehicles aiming at accomplishing a mission) and they are required to maintain a certain distance with their neighbors for purposes of connectivity (such as UAVs or terrestrial vehicles to stay within their communication ranges). A jammer, which is also mobile (such as a jamming UAV), wants to break the connectivity (by getting close to some of the UAVs or vehicles, and jamming their communication links). This problem can be formulated as a *pursuit–evasion differential game*,[5] where the jammer (pursuer) tries to break connectivity as quickly as possible, whereas the friendly mobile nodes (evaders) want to extend connectivity for as long as possible. The paper [25] has introduced this class of mobile jamming problems and provided optimal saddle-point policies for both sets of players under certain conditions; further results can be found in [24,26]. Finally, in [41], a Stackelberg game for communication between a source and a set of selfish relays is considered. In this game, the source coordinates the relays in order to benefit the relays (followers) in forwarding the signals and defend the system against the eavesdropping attacks. In particular, an algorithm is developed for the relays to find their Nash equilibrium strategies. It is shown that through such cooperative communication, the source can achieve better secure transmission against eavesdroppers.

5 Incentives and Mechanism Design

In real applications, many systems fail because of wrong incentives among their subsystems or agents. As discussed in [10], "the people who guard a system often are not the people who suffer the full costs of failure and as a result they make less effort than would be socially optimal." This shows that an effective incentive mechanism design can play a major role in improving the system security which has been the center of many recent works [1,22,33,35,39,42,70,83,116].

In [42], the authors design a dynamic incentive mechanism for security in networks of interdependent agents. In this model, a system of n strategic agents who are interconnected through a fixed directed network is considered. The agents interact over time instances $t = 0, 1, \ldots$ with their neighbors where the security status of agent i at time t is given by a binary random variable θ_t^i (agent i is safe at time t if $\theta_t^i = 1$, and unsafe if $\theta_t^i = 0$) whose realization is a private information to agent i. There is a network manager whose goal is to maximize the overall security of the network over time against external attacks and/or propagation of internal attacks. At each time, a safe agent i can be attacked externally with probability d_i or

[5] Pursuit–evasion games model many security problems where one or more evaders try to escape a group of pursuing units; see [21].

internally from any of its neighbors j with probability l_{ij} which corresponds to the strength of agent j's influence on agent i in the network. At each time t, the network manager's action a_t is to choose an agent i and apply a security measure on it. If $a_t = i$ and agent i is unsafe (that is, $\theta_t^i = 0$), then it will be safe with a constant probability h while being protected from external attack with the same probability during time instance t (note that network director's action does not affect the internal attacks within the network). As a result of network manager's action as well as external and internal attacks, the security state of the network $\theta_t = (\theta_t^1, \ldots, \theta_t^n)$ evolves based on a Markov chain with a stochastic rule governed by

$$\mathbb{P}(\theta_{t+1} = \boldsymbol{b}|\theta_t, a_t) = \prod_{i=1}^{n} \mathbb{P}(\theta_{t+1}^i = b_i|\theta_t, a_t), \forall \boldsymbol{b} \in \{0, 1\}^n,$$

where $\mathbb{P}(\theta_{t+1}^i = b_i|\theta_t, a_t), i = 1, \ldots, n$, can be computed in a closed form based on the probabilities h, d_i and l_{ji}. For instance, if $a_t = i$ and $\theta_t^i = 0$, then $\mathbb{P}(\theta_{t+1}^i = 1|\theta_t, a_t) = h(1 - d_i(1 - h)) \prod_{j \in N^i : \theta_t^j = 0}(1 - l_{ji})$, where N^i denotes the set of neighbors of node i. This is because if agent i is in the unsafe state and receives the security measure, he will be safe in the next time instance if the security measure is successful with probability h, he is not the subject of new successful external attacks with probability $1 - d_i(1 - h)$, and he is not attacked internally by his unsafe neighbors with probability $\prod_{j \in N^i : \theta_t^j = 0,}(1 - l_{ji})$. Note that the underlying assumption in these derivations is that the external and internal attacks are independent across different agents.

Further, at each time t, agent i evaluates his safety based on the network security state and the measurement that he receives from the network manager by

$$v^i(\theta_t, a_t) = \theta_t^i + \Big(\frac{\alpha}{|N^i|} \sum_{j \in N^i} \theta_t^j\Big) \mathbf{1}_{\{\theta_t^i = 1 \text{ or } a_t = i\}},$$

where $\mathbf{1}_{\{\cdot\}}$ denotes the indicator function and $\alpha \in (0, 1)$ is a scaling constant. As a result, an agent feels safer if his state is safe or he receives measurement from the network manager. Now, denoting the monetary payment by agent i to the network manager at time t by $p_t^i \in \mathbb{R}$, the payoff received by agent i at time t can be expressed by $u_t^i(\theta_t, a_t, p_t^i) = v^i(\theta_t, a_t) - p_t^i$, and his total payoff equals

$$U^i := (1 - \delta) \sum_{t=1}^{\infty} \delta^t u_t^i(\theta_t, a_t, p_t^i),$$

where $\delta \in (0, 1)$ is a common discounting factor. In particular, the goal of the network manager is to maximize the expected social welfare (safety) of the network given by

$$W := \mathbb{E}\left[(1 - \delta) \sum_{i=1}^{n} \sum_{t=1}^{\infty} \delta^t v^i(\theta_t, a_t)\right].$$

Let $\mathcal{H}_t$ denote all possible histories of agents' safety reports up to time t (note that agents may have incentives to misreport their safety state θ_t^i in order to maximize their own payoffs). The mechanism design problem for the network manager is then to determine a measurement allocation policy $\pi_t(\cdot) : \mathcal{H}_t \to \{1, \ldots, n\}$ (i.e., determine at each time t which agent receives the measurement) together with a monetary payment policy $p_t^i(\cdot) : \mathcal{H}_t \to \mathbb{R}, \forall i$ (i.e., at time t how much each agent must pay to the network manager as a tax) which maximizes

W while being *incentive compatible* and *individually rational*.[6] It was shown in [42] that this mechanism design optimization problem has a solution by first providing an incentive compatible monetary payment rule which aligns the incentives of each agent with the social welfare. Using this monetary payment, agents report their security states truthfully, and thus, the problem of searching for an optimal measurement allocation policy becomes a stochastic control problem with complete information. Therefore, a dynamic programming technique can be used to obtain the optimal measurement allocation policy.

In [83], a repeated game has been introduced in order to understand firms' incentives for disclosing their security information and to find out whether inter-temporal incentives can lead to support of cooperation. It has been shown that even an imperfect rating/monitoring system can be very helpful to design inter-temporal incentives that lead firms to cooperate on sharing their security information. Moreover, similar results have been shown when the monitoring system is replaced by a platform to communicate firms' privately observed beliefs on each others' adherence to the agreement. In [22], a differential Stackelberg game is used to model the interactions between the "West" and "International Terror Organization" (ITO) as two strategic decision makers. In this model, at each time instance t the West acts first by deploying some terror control activities and the ITO acts as a follower by deciding how many attacks to initiate at time t. As a result, the state of the game which is defined to be the number of terrorists at time t evolves according to a differential equation based on the players' actions and the underlying parameters of the game. The equilibrium points of this game have been characterized. In particular, the model has been used to analyze the effect of West's defensive strategies on incentives and strategies of ITO and to explore how changing the underlying parameters can change the ITO incentives in both short or long term.

Using an economics approach, the authors in [1] leverage a repeated game model to explore the incentives of participants to offer and use anonymity services which is an important need in a variety of circumstances, such as using the Internet, surfing the web, or online purchases. As discussed in [1], in economic systems anonymity is particularly important from a security perspective in which usability, efficiency, reliability, and cost become *security* objectives which affect user size and hence the achievable degree of anonymity. Finally, in [70] a general incentive-based framework to model attacker intent, objectives, and strategies, has been developed. Moreover, to infer attacker incentives and objectives, several dynamic game-theoretic formulations such as a repeated Bayesian game or a stochastic game are proposed. In particular, the authors provide some insights into how one should choose the most appropriate game in real-world attack–defense scenarios.

6 Resource Allocation and Optimal Investment

Broadly speaking, resource allocation in security refers to the set of problems in which defenders must allocate limited security resources to protect targets from attack by adversaries. In [106], the authors study the problem of dynamic security resource allocation by developing a stochastic Stackelberg game between a defender and multiple attackers. In this game, the defender acts as a leader and commits first to a patrolling allocation strategy over the security targets. Then, the followers make the strategy selection trying to realize a Nash equilibrium among themselves and obtain payoffs conditioned on the leader's strategy. Depending on defender's and attackers' strategies, the game will then move to a new state assuming that the

[6] Incentive compatibility means that agents have no incentive to misreport their safety states, while individual rationality implies that agents voluntarily have incentives to participate in the mechanism.

underlying stochastic Markov chain is ergodic. A reinforcement learning method is applied to this general setting in which the players can learn the game behavior through trial-and-error interactions with this dynamic environment, and a so-called *Extraproximal approach* is adopted to accelerate the reinforcement learning process. The results are numerically justified for the case of a single defender and a single attacker, showing that the players can efficiently learn their equilibrium strategies.

In [12], a differential game is used to analyze the optimal IT security investment of different firms with similar information assets which are subject to hacking attacks. Such firms have an incentive to deflect hackers toward others by strategically raising their own security levels. It was shown that this competition leads firms to over-invest in IT security while creating proper incentives can result in collaborative security effort among the firms (and hence savings in industry). The authors in [78] develop a repeated best response dynamics to study the security decision making in interdependent organizations (players) whose security levels can affect others' by a linear influence network. It is shown that by adjusting the influence matrix, the equilibrium achieved for the single-stage game can be improved for any two players while maintaining others' levels of investment unchanged. The results are applied to the setting of Web site security with shared passwords where users have the same passwords for different websites, and hence, the decision of one organization may have substantial impact on the security of the others.

As discussed in Sect. 3.4, SSGs provide a suitable framework for studying optimal resource allocation in security systems (see, e.g., [92] for an application of SSG to assist in resource allocation tasks for airport protection). However, the underlying assumption on the players' information can result in considerably different analysis. In its simplest form [9], a two-player SSG is composed of n possible targets $T = \{1, \ldots, n\}$, and m identical resources ($m < n$). The defender has a set of N feasible actions, $\mathcal{A}$, where each $A \in \mathcal{A}$ is a binary coverage vector representing which m targets are protected ($A_i = 1$ if target i is protected in action A and $A_i = 0$, otherwise). The defender can choose to play a mixed (random) strategy x, where x is a probability distribution over the action set $\mathcal{A}$, with x_A denoting the probability that action A is being played. The defender's mixed strategy can also be represented more compactly using a coverage vector $\boldsymbol{c}(x) = (c_1(x), \ldots, c_n(x))$, where $c_i(x) = \sum_{A \in \mathcal{A}} x_A A_i$, which is the overall probability that target i is protected (covered) when defender's mixed strategy is x. On the other hand, the attacker's mixed strategy $\boldsymbol{a} = (a_1, \ldots, a_n)$ is a probability vector with a_i being the probability that the attacker attacks target i. Given the strategy profile $(\boldsymbol{c}, \boldsymbol{a})$ of the players, the utilities of the defender and the attacker denoted by $U^d(\boldsymbol{c}, \boldsymbol{a})$ and $U^a(\boldsymbol{c}, \boldsymbol{a})$, respectively, are given by

$$U^d(\boldsymbol{c}, \boldsymbol{a}) = \sum_{i=1}^{n} a_i \left(c_i R_i^d + (1 - c_i) P_i^d\right)$$

$$U^a(\boldsymbol{c}, \boldsymbol{a}) = \sum_{i=1}^{n} a_i \left(c_i P_i^a + (1 - c_i) R_i^a\right),$$

where R_i^d and R_i^a are the rewards (similarly P_i^d and P_i^a are the penalties) associated with the defender and the attacker given, respectively, whether target i is covered or not. Here, it is assumed that $R_i^d > P_i^d$, and $R_i^a > P_i^a$ to assure that the attacker wants the attack to be successful while the defender wants it to fail. In SSG, the defender acts first by committing to a mixed strategy x (or equivalently $\boldsymbol{c}(x)$). The attacker fully observes the defender's strategy and responds to it optimally by attacking a target which brings him the highest payoff.

Unlike the standard assumption on SSG, where the attacker has *perfect knowledge* of the defender's strategy, in [9], the authors study the optimal resource allocation strategies using a dynamic SSG played over multiple stages and with *limited surveillance* in which the attacker dynamically updates his belief in order to determine the point when to stop surveillance based on observed actions and cost of surveillance. In their model, the attacker is of the form of a Bayesian decision maker who observes defender moves, and optimally decides to either attack at the current stage, or to make another observation of the defender at some fixed cost. The game ends when the attacker attacks a target. One motivation for studying such a model is security concerns at airports. Police deploy checkpoints at the entrance roads to airports according to randomized rules. Attackers typically engage in surveillance such as driving around different airport entrances, but will eventually launch an attack based on a finite number of observations of the checkpoint locations.

The work in [9] formulates the attacker's optimal stopping problem as a finite state space *Markov Decision Process* (MDP) in which the states are the observation vectors. An observation vector $\boldsymbol{o} = (o_A)_{A\in\mathcal{A}}$ is a vector in which o_A denotes the number of times up to now that the defender chooses coverage action (pure strategy) A. Therefore, each transition in the states of this MDP corresponds to moving from an observation vector $\boldsymbol{o}$ to $\boldsymbol{o}'$ where $\boldsymbol{o}'$ differs from $\boldsymbol{o}$ in only one observation A (the most recent observation by the attacker). Now if the attacker decides to attack his best target $\psi(\boldsymbol{o})$ at state with observation vector $\boldsymbol{o}$, his immediate payoff will be

$$W(\boldsymbol{o}) = c_{\psi(\boldsymbol{o})} P^a_{\psi(\boldsymbol{o})} + (1 - c_{\psi(\boldsymbol{o})}) R^a_{\psi(\boldsymbol{o})} - \lambda \Delta(\boldsymbol{o}),$$

where λ is a fixed cost of making observations and $\Delta(\boldsymbol{o}) = \sum_{A\in\mathcal{A}} o_A$ is the length of observation $\boldsymbol{o}$ (i.e., the number of times the attacker has observed the defender's strategy in $\boldsymbol{o}$). Moreover, $c_{\psi(\boldsymbol{o})}$ is the marginal coverage of target $\psi(\boldsymbol{o})$ according to the attacker's belief after observing $\boldsymbol{o}$. Therefore, at each state $\boldsymbol{o}$, the attacker can either attack the best target $\psi(\boldsymbol{o})$ and receive a payoff $W(o)$, or make another observation reaching state $\boldsymbol{o}' = \boldsymbol{o} \cup \{A\}$, with some probability $\mathbb{P}(A|\boldsymbol{o})$ (which can be computed based on posterior belief of the attacker). Letting $V(\boldsymbol{o})$ be the optimal value function for the attacker (i.e., the attacker's expected utility when his observation vector is $\boldsymbol{o}$ and he follows the optimal policy afterward), we have

$$V(\boldsymbol{o}) = \max\left\{W(\boldsymbol{o}), \sum_{A\in\mathcal{A}} \mathbb{P}(A|\boldsymbol{o}) V(\boldsymbol{o} \cup \{A\})\right\}.$$

Solving this MDP faces several challenges including the infinite size of state space. However, it is shown in [9] that without loss of generality one can restrict attention to finitely many states from some time onward, as making new observations for the attacker has negligible effect in changing its posterior belief, and hence improving his expected payoff. In particular, the authors obtain an upper bound on the maximum number of observations the attacker can make before he attacks, and provide a backward induction–forward search technique together with a mixed linear integer program to approximate and compute the exact attack and defense strategies, respectively.

Different from [9] and in order to model attacks on a cyber-physical system, a dynamic multiplayer nonzero-sum game with *asymmetric* information has been considered in [48]. In this game, it is assumed that players have limited resource constraints and do not necessarily acquire the same information as the system operators. Under certain conditions on the information structure of the game (e.g., variables affecting the amount of resource invested by players become a part of the common information at the next time step), a virtual game of *symmetric* information is introduced, with an equivalence between its set of Nash equilibria and that of the original asymmetric game established. This virtual symmetric game is then used to compute the Nash equilibrium points of the original asymmetric game. We refer the readers to [56] for other dynamic game formulations of asymmetric information for studying optimal resource allocation in cyber-physical security.

Using a repeated game formulation, it was shown in [28] that for interdependent security networks, such as the Internet, optimal selfish investment of individual users on their own security does not necessarily reduce the overall risk of the network. Moreover, beside optimal investment, the effect of network topology, users' preferences, and their mutual influence on the overall network security has been studied. In particular, it is shown that under a repeated game framework, the best equilibrium yields much better performance as users have more incentive to cooperate for their long-term interests. From an economic perspective, [55] considers a repeated game between a defender with incomplete information and an attacker, who repeatedly targets the weakest link[7] in order to model security investment of an information system and derive optimal security investment over multiple rounds. As opposed to the case of a single-stage game, the authors show that for the repeated game, the defender's strategies may be quite different so that he initially protects fewer assets and waits until the attacker *identifies* the weakest links. This explains why underinvestment in security is rational until threats are realized. Finally, the authors in [108] look into computational aspects of resource allocation in security games by relaxing some of the standard assumptions, such as additivity of the players' payoff functions, or that the attacker can attack only one target. Using a unified framework, they show that computing the equilibrium strategies in security games is essentially a combinatorial optimization problem which can be used to study the complexity of computing equilibrium points and optimal attack/defense strategies. We also refer to [54] for other computational analysis related to equilibrium and optimal resource allocation in Bayesian Stackelberg security games.

7 Learning in Security Games

Security is often an iterative process where the defenders commit to some strategies in order to protect the assets. In return, the attackers decide what strategies to adopt in order to maximize their payoffs. From each failure or success of an attack, both attackers and defenders gain new information which can be used to improve their future strategies. This opens an exciting research area of learning in cyber-physical security, which has been addressed in the past decade. In this section, we mainly focus on two important classes of learning algorithms applied to security games: (i) *Reinforcement learning* (or other variants of it such as *Q-learning*) in which the interaction environment between security players is typically formulated as a Markov decision process. Many reinforcement learning algorithms in this context utilize the notion of value function to represent the quality and expected reward of the players' decisions and use dynamic programming techniques to learn the optimal strategies. Reinforcement learning methods are often model-free and can be applied to large class of stochastic dynamic games (although their performance highly depends on the structure of the game) [29,37,51,106,112,120]. ii) *Online regret-based* learning in which new information becomes available in a sequential order and during the course of the game between security players. The performance metric here is usually based on some notion of *regret* [11,47,58,87].

[7] In an information system, the system's overall security usually depends on its weakest link.

7.1 Reinforcement Learning in Security

In [29], the authors develop a zero-sum stochastic game to model security risk in interdependent organizations and utilize a reinforcement Q-learning to analyze the behavior of the players when the parameters of the game are not known. In [37], a stochastic game for modeling the decision-making process of cyber-security monitoring is proposed. Different variants of Q-learning are used which react automatically to the adversarial behavior of suspicious users. The efficiency of these methods under different environments is also evaluated through simulation results. A zero-sum stochastic game between an attacker and a legitimate system in an abstract form was formulated in [51]. At each stage of this game, $k = 0, 1, \ldots$, both the legitimate system and the attacker observe the current state s_k and take actions o_k, a_k based on their own learned policies and receive immediate rewards $R(s_k, a_k, o_k)$ and $-R(s_k, a_k, o_k)$, respectively. After that, the state of the game (e.g., channel communication state) changes from $s_k \in \mathcal{S}$ to $s_{k+1} \in \mathcal{S}$ with some probability $\mathbb{P}(s_{k+1}|s_k, a_k, o_k)$ which is often unknown to the players. The objective of the legitimate system (attacker) is to learn an optimal stationary policy $\pi := \{\pi_s\}_{s\in\mathcal{S}}$ ($\pi^O := \{\pi_s^O\}_{s\in\mathcal{S}}$) in order to maximize (minimize) the discounted average reward given by $\mathbb{E}[\sum_{k=0}^{\infty} \beta^k R_k(s_k, \pi^O(s_k), \pi(s_k))]$. Here $\pi_s(\cdot)$ (similarly $\pi_s^O(\cdot)$) is a mixed strategy over the action set of the legitimate system (attacker) such that $\pi_s(a)$ determines the probability of choosing action a at a given state $s \in \mathcal{S}$.

A conventional minimax-Q learning method for learning the optimal stationary policies in this security game is to define two functions: an optimal *quality function* $Q_*(s, a, o)$ for each pair of state actions and an optimal *value function* $V(s)$ defined over all the states. $Q_*(s, a, o)$ can be thought of the total expected discounted reward attained by the legitimate system taking action a, given current state s and attacker action o, and then following the optimal policy from then on, i.e.,

$$Q_*(s, a, o) := \mathbb{E}[R(s, a, o) + \beta V_*(S')],$$

where S' is a random variable denoting the next state of the game, and

$$V_*(s) := \max_{\pi_s} \min_{o} \sum_{a} \pi_s(a) Q_*(s, a, o).$$

Based on these functions, one can define a conservative *minimax* optimal policy for the legitimate system given by $\pi_*(s) := \text{argmax}_{\pi_s} \min_o \sum_a \pi_s(a) Q_*(s, a, o)$. The corresponding minimax optimal strategy for the attacker is defined similarly. Now, one can define a sequence of iterated functions $\{Q_k(\cdot)\}$, $\{V_k(\cdot)\}$, and $\{\pi_k(\cdot)\}$, recursively in terms of the above expressions, such that in the limit as $k \to \infty$ these sequences converge to their corresponding functions $Q_*(\cdot)$, $V_*(\cdot)$, and $\pi_*(\cdot)$. However, it is possible that the obtained optimal minimax policy $\pi_*(\cdot)$ is *not* the actual optimal policy for the legitimate system. To resolve this issue, the authors in [51] develop different variants of the above minimax-Q learning algorithm that under certain conditions guarantee convergence of the iterated policies to the actual optimal policy. In particular, it is shown that for unknown dynamic environments, with extra partial information the modified algorithms converge faster to the actual optimal policies.

The paper [120] considers a zero-sum stochastic security game with heterogeneous players in which the players aim to learn their optimal strategies by adopting possibly different distributed reinforcement learning algorithms while they have incomplete information about the game. The paper uses stochastic approximation techniques to show that the heterogeneous learning schemes can be studied in terms of their deterministic ordinary differential equation counterparts. The results are applied to a class of security games in which the attacker and the defender adopt different learning schemes due to differences in their rationality levels and the information they acquire. The book chapter [122] extends the model in [120] to nonzero-sum stochastic security games with incomplete information, and develops fully distributed reinforcement learning algorithms, which require for each player a minimal amount of information regarding the other player. At each time, a player can be in an active mode or in a sleep mode. If a player is in an active mode, she updates her strategy and estimates of unknown quantities using a specific pure or hybrid learning pattern. The players' intelligence and rationality are captured by a weighted linear combination of different learning patterns. As in [120], it has been shown that the pure or hybrid learning schemes with random updates can be studied using their deterministic ordinary differential equation counterparts. Finally, the work [106] applies a reinforcement learning method to a Stackelberg security game and provides an efficient framework for the attackers and defenders to adapt their strategies in a dynamic environment.

7.2 Regret-Based Online Learning

Regret-based online learning algorithms have become quite popular in recent years with interesting applications in studying security games. Given a multistage security game, the regret associated with a player following an online learning algorithm is simply the difference between his payoff for the *best-in-hindsight* fixed strategy[8] and the expected payoff that he can obtain by following the online algorithm. In [11], the authors consider a slightly more general version of the repeated SSG introduced in Sect. 6 and devise an online algorithm for the defender whose regret is sublinear in the number of time steps, and polynomial in the parameters of the game. As a result, the average regret of the defender by following such an online algorithm approaches zero as the number of game stages increases. More precisely, a repeated SSG is considered in [11] in which one defender (leader) must protect a set of n targets against a sequence of attackers (followers). At each step, the defender commits to a randomized allocation strategy over the targets and an attacker (which can be chosen adversarially from a set of k unknown different attackers) observes this randomized allocation and attacks the target with the best expected payoff. Therefore, each stage of the game is an SSG between the defender and a newly arrived attacker. Here, the defender's goal is to maximize his payoff over a period of time, even when the sequence of attackers is unknown.

For the full observation setting (i.e., when the defender can observe the attacker's type once an attack has occurred), it is shown in [11] that there is an online algorithm for the defender such that his regret will be bounded above by $O(\sqrt{Tn^2k\log(nk)})$, where T is the number of game stages. For the partial observation setting (i.e., when the defender can only observe which target was attacked at each round), it is shown that the defender can still follow an algorithm with sublinear regret but with slightly worse upper bound $O(T^{\frac{2}{3}}nk\log^{\frac{1}{3}}(nk))$. For another application of regret-based minimization in security games, we refer to [87] in which the authors provide an algorithm for computing minimax regret-based strategies for SSGs under uncertainty.

The work in [47] considers an online regret-based learning game within the context of *expert advice*. In this problem, there is a learner (e.g., a recommendation system) which has to select at each time $t = 1, 2, \ldots$, one expert from a pool of k experts and follow his advice. At every time t, an adversary sets a gain $g_{it} \in [0, 1]$ for each expert i. Simultaneously, the learner observes all the gains from all previous steps except t and has to choose which expert to follow at stage t. Given that the learner follows expert $j(t)$ at time t, he receives an instant

[8] This is a strategy that yields the highest total payoff for that player if he knows the entire sequence of attacks *a priori*.

payoff of $g_{j(t)t}$. After the learner makes his decision at time t, the gains associated with all experts are revealed to him, and the adversary can also observe the learner's choice of expert. Given that this game is repeated T times, the learner's objective is to select the experts in a manner such that he can achieve a cumulative gain close to that when choosing the best expert in hindsight. In other words, the learner's goal is to minimize his regret defined by

$$R_T := \max_{1 \le i \le k} \sum_{t=1}^{T} g_{it} - \sum_{t=1}^{T} g_{j(t)t},$$

while the adversary aims to maximize this quantity. This defines a dynamic zero-sum game between the learner and the adversary where the players must play in an online fashion. The authors in [47] fully characterize the optimal online policies for the learner and the adversary in the case of $k = 2$ or $k = 3$ experts and provide some general insights into how to design optimal algorithm for the learner and the adversary for an arbitrary number of experts. Along the same line of work, the authors in [107] consider a security system in the form of expert advice setting in which the experts can themselves be malicious and report false predictions to deceive the learning system. Using a dynamic programming approach, the authors characterize the optimal online policies for the malicious experts for certain special cases. We refer to [58] for other possibilities of learning algorithms in security games with rationally bounded players [68].

8 Conclusions and Discussion

In this survey, we have discussed recent advances and applications of dynamic games in cyber-physical security problems. We have categorized these problems based on their application domain as well as the topical area. In particular, we have illustrated several common yet important game-theoretic methodologies in modeling and analyzing security problems. There are many other important security topics which can be effectively captured and analyzed using dynamic games. We mention here the issues related to information leakage, anonymity, or privacy [38,43,95,111], bounded rationality and humans' behaviors [36], and patrolling, pursuit-evasion, or models of crime [7,14,52,81].

While the existing body of literature in dynamic games successfully models the behavior of attackers/defenders in a variety of security circumstances, there are still many limitations in applying them to real-world applications. In particular, the emerging security problems are becoming more sophisticated so that analyzing them might require a combination of the earlier tools and methods. In the following, we sketch briefly some of the current limitations and some future directions of research.

- Most of the earlier literature on security games views information, rather than the computation, as the main bottleneck. However, as we saw earlier in the case of SSG with multiple resources and a target, the computational complexity of finding optimal defending strategies becomes a critical barrier which hinders their applicability in realistic scenarios. Therefore, devising computationally efficient algorithms for computing (or approximating) the optimal strategies for large-scale dynamic security games is an important problem.
- Humans are mainly the decision makers in security games. While most of the earlier literature assumes rational behavior of the decision makers, it has been observed extensively that in reality humans have subjective views on uncertain events (e.g., risk averse or risk seeking). Hence, human behavioral decisions can substantially affect the anticipated

outcomes. Therefore, incorporating human behavioral decisions into the game structure (e.g., using *prospect theory* [57]) and identifying its implications and deviations from conventional game theory is another important issue which has received limited attention so far. For some very recent work in that direction, we refer to [82,102].

- There are many circumstances in which the structure of the security system is not predefined. For example, in social security the connections between attackers (e.g., terrorists) and the defenders might change based on their incentives. Therefore, developing dynamic game models to understand how the underlying connections forms and evolves as a result of players' interactions is an interesting problem.
- Finally, in many practical situations such as relationships between countries, there are underlying adversarial and defender groups with potential weak links between members of opposing groups which are effectively used by both groups to infiltrate the other. In other words, an agent may possibly belong to multiple groups, which themselves may be connected by multilayered networks. Providing comprehensive dynamic security games which can capture the interactions of the agents in such a multilayered multi-scale environment is another interesting research direction.

References

1. Acquisti A, Dingledine R, Syverson P (2003) On the economics of anonymity. In: International conference on financial cryptography. Springer, pp 84–102
2. Akyol E, Rose K, Başar T (2015) Optimal zero-delay jamming over an additive noise channel. IEEE Trans Inf Theory IT–61:4331–4344
3. Akyol E, Rose K, Başar T (2013) On communication over Gaussian sensor networks with adversaries: further results. In: Proceedings of international conference on decision and game theory for security (GameSec). Springer, pp 1–9
4. Alpcan T, Başar T (2011) Network security: a decision and game theoretic approach. Cambridge University Press, Cambridge
5. Alpcan T, Başar T (2004) A game theoretic analysis of intrusion detection in access control systems. In: 43rd IEEE conference on decision and control (CDC), vol 2, pp 1568–1573
6. Alpcan T, Başar T (2006) An intrusion detection game with limited observations. In: 12th international symposium on dynamic games and applications, vol 26. Sophia Antipolis, France
7. Alpern S, Lidbetter T, Morton A, Papadaki K (2016) Patrolling a pipeline. In: Proceedings of international conference on decision and game theory for security (GameSec). Springer, pp 129–138
8. Amin S, Schwartz GA, Hussain A (2013) In quest of benchmarking security risks to cyber-physical systems. IEEE Netw 27(1):19–24
9. An B, Brown M, Vorobeychik Y, Tambe M (2013) Security games with surveillance cost and optimal timing of attack execution In: Proceedings of the 2013 international conference on autonomous agents and multi-agent systems, pp 223–230
10. Anderson R, Moore T, Nagaraja S, Ozment A (2007) Incentives and information security. Algorithmic Game Theory, pp 633–649
11. Balcan M-F, Blum A, Haghtalab N, Procaccia AD (2015) Commitment without regrets: online learning in Stackelberg security games. In: Proceedings of the sixteenth ACM conference on economics and computation, pp 61–78
12. Bandyopadhyay T, Liu D, Mookerjee VS, Wilhite AW (2014) Dynamic competition in IT security: a differential games approach. Inf Syst Front 16(4):643–661
13. Bansal R, Başar T (1989) Communication games with partially soft power constraints. J Optim Theory Appl 61(3):329–346
14. Basak A, Fang F, Nguyen TH, Kiekintveld C (2016) Combining graph contraction and strategy generation for green security games. In: Proceedings of international conference on decision and game theory for security (GameSec). Springer, pp 251–271
15. Başar T (1983) The Gaussian test channel with an intelligent jammer. IEEE Trans Inf Theory IT–29(1):152–157

16. Başar TÜ, Başar T (1982) Optimum coding and decoding schemes for the transmission of a stochastic process over a continuous-time stochastic channel with partially unknown statistics. Stochastics 8(3):213–237
17. Başar T, Başar TÜ (1984) A bandwidth expanding scheme for communication channels with noiseless feedback and in the presence of unknown jamming noise. J Frankl Inst 317(2):73–88
18. Başar TÜ, Başar T (1989) Optimum linear causal coding schemes for Gaussian stochastic processes in the presence of correlated jamming. IEEE Trans Inf Theory 35(1):199–202
19. Başar T, Wu YW (1985) A complete characterization of minimax and maximin encoder-decoder policies for communication channels with incomplete statistical description. IEEE Trans Inf Theory IT–31(4):482–489
20. Başar T, Wu YW (1986) Solutions to a class of minimax decision problems arising in communication systems. J Optim Theory Appl 51(3):375–404
21. Başar T, Olsder GJ (1999) Dynamic noncooperative game theory. Series in classics in applied mathematics, SIAM, Philadelphia
22. Behrens DA, Caulkins JP, Feichtinger G, Tragler G (2007) Incentive Stackelberg strategies for a dynamic game on terrorism. Adv Dyn Game Theory, pp 459–486
23. Bensoussan A, Kantarcioglu M, Hoe S (2010) A game-theoretical approach for finding optimal strategies in a Botnet defense model. In: Proceedings of international conference on decision and game theory for security (GameSec), pp 135–148
24. Bhattacharya S, Gupta A, Başar T (2013) Jamming in mobile networks: a game-theoretic approach. J. Numer Algebra Control Optim 3(1):1–30
25. Bhattacharya S, Başar T (2010) Game-theoretic analysis of an aerial jamming attack on a UAV communication network. In: American control conference (ACC). IEEE, pp 818–823
26. Bhattacharya S, Başar T (2011) Spatial approaches to broadband jamming in heterogeneous mobile networks: a game-theoretic approach. J Autonomous Robots
27. Bloem M, Alpcan T, Başar T (2006) Intrusion response as a resource allocation problem. In: 45th IEEE conference on decision and control, pp 6283–6288
28. Böhme R, Moore T (2016) The "iterated weakest link" model of adaptive security investment. J Inf Secur 7(02):81
29. Bommannavar P, Alpcan T, Bambos N (2011) Security risk management via dynamic games with learning. In: IEEE international conference on communications (ICC), pp 1–6
30. Carroll TE, Grosu D (2011) A game theoretic investigation of deception in network security. Secur Commun Netw 4(10):1162–1172
31. Casey W, Morales JA, Nguyen T, Spring J, Weaver R, Wright E, Metcalf L, Mishra B (2014) Cyber security via signaling games: toward a science of cyber security. In: International conference on distributed computing and internet technology. Springer, pp 34–42
32. Casey W, Weaver R, Metcalf L, Morales JA, Wright E, Mishra B (2014) Cyber security via minority games with epistatic signaling. In: Proceedings of the 8th international conference on bioinspired information and communications technologies, pp 133–140
33. Cavallo R (2008) Efficiency and redistribution in dynamic mechanism design. In: Proceedings of the 9th ACM conference on electronic commerce, pp 220–229
34. Çeker H, Zhuang J, Upadhyaya S, La QD, Soong B-H (2016) Deception-based game theoretical approach to mitigate DoS attacks. In: International conference on decision and game theory for security. Springer, pp 18–38
35. Chen J, Zhu Q (2016) Optimal contract design under asymmetric information for cloud-enabled internet of controlled things. In: Proceedings of international conference on decision and game theory for security (GameSec). Springer, pp 329–348
36. Chicoisne R, Ordóñez F (2016) Risk averse Stackelberg security games with quantal response. In: Proceedings of international conference on decision and game theory for security (GameSec). Springer, pp 83–100
37. Chung K, Kamhoua CA, Kwiat KA, Kalbarczyk ZT, Iyer RK (2016) Game theory with learning for cyber security monitoring. In: 17th IEEE international symposium on high assurance systems engineering (HASE), pp 1–8
38. Culnane C, Teague V (2016) Strategies for voter-initiated election audits. In: Proceedings of international conference on decision and game theory for security (GameSec). Springer, pp 235–247
39. Dziubiński M, Sankowski P, Zhang Q (2016) Network elicitation in adversarial environment. In: Proceedings of international conference on decision and game theory for security (GameSec). Springer, pp 397–414
40. Fang F, Stone P, Tambe M (2015) When security games go green: designing defender strategies to prevent poaching and illegal fishing. In: IJCAI, pp 2589–2595

41. Fang H, Xu L, Wang X (2017) Coordinated multiple-relays based physical-layer security improvement: a single-leader multiple-followers Stackelberg game scheme. In: IEEE transactions on information forensics and security, pp 75–80
42. Farhadi F, Tavafoghi H, Teneketzis D, Golestani J (2017) A dynamic incentive mechanism for security in networks of interdependent agents. In: 7th EAI international conference on game theory for networks
43. Farhang S, Grossklags J (2016) Flipleakage: a game-theoretic approach to protect against stealthy attackers in the presence of information leakage. In: Proceedings of international conference on decision and game theory for security (GameSec). Springer, pp 195–214
44. Farhang S, Manshaei MH, Esfahani MN, Zhu Q (2014) A dynamic Bayesian security game framework for strategic defense mechanism design. In: International conference on decision and game theory for security. Springer, pp 319–328
45. Ghafouri A, Abbas W, Laszka A, Vorobeychik Y, Koutsoukos X (2016) Optimal thresholds for anomaly-based intrusion detection in dynamical environments. In: International conference on decision and game theory for security. Springer, pp 415–434
46. Gibson AS (2013) Applied hypergame theory for network defense. Air Force Inst of Tech WRIGHT-PATTERSON AFB OH Graduate School of Engineering Management. Tech. Rep
47. Gravin N, Peres Y, Sivan B (2016) Towards optimal algorithms for prediction with expert advice. In: Proceedings of the twenty-seventh annual ACM-SIAM symposium on discrete algorithms, pp 528–547
48. Gupta A, Langbort C, Başar T (2017) Dynamic games with asymmetric information and resource constrained players with applications to security of cyberphysical systems. IEEE Trans Control Netw Syst 4(1):71–81
49. Gupta A, Langbort C, Başar T (2010) Optimal control in the presence of an intelligent jammer with limited actions. In: 49th IEEE conference on decision and control (CDC), pp 1096–1101
50. Gupta A, Schwartz G, Langbort C, Sastry SS, Başar T (2014) A three-stage Colonel Blotto game with applications to cyber-physical security. In: Proceedings of IEEE American control conference (ACC), pp 3820–3825
51. He X, Dai H, Ning P (2015) Improving learning and adaptation in security games by exploiting information asymmetry. In: IEEE conference on computer communications (INFOCOM), pp 1787–1795
52. Horák K, Bošanskỳ B (2016) A point-based approximate algorithm for one-sided partially observable pursuit-evasion games. In: Proceedings of international conference on decision and game theory for security (GameSec). Springer, pp 435–454
53. House JT, Cybenko G (2010) Hypergame theory applied to cyber attack and defense. Sens Command Control Commun Intell (C3I) Technol Homel Secur Homel Def IX 7666:766604–766611
54. Jiang AX, Yin Z, Zhang C, Tambe M, Kraus S (2013) Game-theoretic randomization for security patrolling with dynamic execution uncertainty. In: Proceedings of the 2013 international conference on autonomous agents and multi-agent systems, pp 207–214
55. Jiang L, Anantharam V, Walrand J (2011) How bad are selfish investments in network security? IEEE/ACM Trans Netw 19(2):549–560
56. Jones MG (2013) Asymmetric information games and cyber security. Ph.D. dissertation, Georgia Institute of Technology
57. Kahneman D, Tversky A (1979) Prospect theory: an analysis of decision under risk. Econometrica 47:263–291
58. Kar D, Fang F, Delle Fave FM, Sintov N, Tambe M, Lyet A (2016) Comparing human behavior models in repeated Stackelberg security games: an extended study. Artif Intell 240:65–103
59. Kar D, Fang F, Delle Fave F, Sintov N, Tambe M (2015) A game of thrones: when human behavior models compete in repeated Stackelberg security games. In: Proceedings of the 2015 international conference on autonomous agents and multiagent systems, pp 1381–1390
60. Khouzani M, Sarkar S, Altman E (2012) Saddle-point strategies in malware attack. IEEE J Sel Areas Commun 30(1):31–43
61. Kiekintveld C, Jain M, Tsai J, Pita J, Ordóñez F, Tambe M (2009) Computing optimal randomized resource allocations for massive security games. In: Proceedings of the 8th international conference on autonomous agents and multiagent systems, vol 1, pp 689–696
62. La QD, Quek TQ, Lee J (2016) A game theoretic model for enabling honeypots in IoT networks. In: IEEE international conference on communications (ICC), pp 1–6
63. Laraki R, Sorin S (2015) Advances in zero-sum dynamic games. Handbook of game theory with economic applications 4:27–93
64. Law YW, Alpcan T, Palaniswami M (2015) Security games for risk minimization in automatic generation control. IEEE Trans Power Syst 30(1):223–232
65. Li Y, Shi L, Cheng P, Chen J, Quevedo DE (2015) Jamming attacks on remote state estimation in cyber-physical systems: a game-theoretic approach. IEEE Trans Autom Control 60(10):2831–2836

66. Liang X, Xiao Y (2013) Game theory for network security. IEEE Commun Surv Tutor 15(1):472–486
67. Li H, Lai L, Qiu RC (2011) A denial-of-service jamming game for remote state monitoring in smart grid. In: 45th annual conference on information sciences and systems (CISS). IEEE, pp 1–6
68. Li L, Langbort C, Shamma J (2017) Computing security strategies in finite horizon repeated bayesian games. In: American control conference (ACC), pp 3664–3669
69. Lin J, Liu P, Jing J (2012) Using signaling games to model the multi-step attack-defense scenarios on confidentiality. In: Proceedings of conference on decision and game theory for security (GameSec). Springer, pp 118–137
70. Liu P, Zang W, Yu M (2005) Incentive-based modeling and inference of attacker intent, objectives, and strategies. ACM Trans Inf Syst Secur (TISSEC) 8(1):78–118
71. Liu Y, Comaniciu C, Man H (2006) A Bayesian game approach for intrusion detection in wireless ad hoc networks. In: Proceedings of the 2006 workshop on game theory for communications and networks. ACM, p 4
72. Luo Y, Szidarovszky F, Al-Nashif Y, Hariri S (2010) Game theory based network security. Executive Editor in Chief
73. Lye K-W, Wing JM (2005) Game strategies in network security. Int J Inf Secur 4(1–2):71–86
74. Ma J, Liu Y, Song L, Han Z (2015) Multiact dynamic game strategy for jamming attack in electricity market. IEEE Trans Smart Grid 6(5):2273–2282
75. Mallik RK, Scholtz RA, Papavassilopoulos GP (2000) Analysis of an on–off jamming situation as a dynamic game. IEEE Trans Commun 48(8):1360–1373
76. Manshaei MH, Zhu Q, Alpcan T, Başar T, Hubaux J-P (2013) Game theory meets network security and privacy. ACM Comput Surv (CSUR) 45(3):25:1–25:39
77. Miao F, Pajic M, Pappas GJ (2013) Stochastic game approach for replay attack detection. In: IEEE 52nd annual conference on decision and control (CDC), pp 1854–1859
78. Miura-Ko RA, Yolken B, Bambos N, Mitchell J (2008) Security investment games of interdependent organizations. In: 46th annual Allerton conference on communication, control, and computing. IEEE, pp 252–260
79. Mohammadi A, Manshaei MH, Moghaddam MM, Zhu Q (2016) A game-theoretic analysis of deception over social networks using fake avatars. In: International conference on decision and game theory for security. Springer, pp 382–394
80. Mukherjee A, Swindlehurst AL (2013) Jamming games in the mimo wiretap channel with an active eavesdropper. IEEE Trans Sig Process 61(1):82–91
81. Mukhopadhyay A, Zhang C, Vorobeychik Y, Tambe M, Pence K, Speer P (2016) Optimal allocation of police patrol resources using a continuous-time crime model. In: Proceedings of international conference on decision and game theory for security (GameSec). Springer, pp 139–158
82. Nadendla VSS, Akyol E, Langbort C, Başar T (2017) Strategic communication between prospect theoretic agents over a Gaussian test channel. In: Proceedings of MILCOM 2017, Baltimore, MD, Oct 23–25 (to appear)
83. Naghizadeh P, Liu M (2016) Inter-temporal incentives in security information sharing agreements. In: Information theory and applications workshop (ITA). IEEE, pp 1–8
84. Nguyen KC, Alpcan T, Başar T (2008) Fictitious play with imperfect observations for network intrusion detection. Preprints of the 13th international symposium dynamic games and applications (ISDGA), June 30–July 3, Wroclaw, Poland
85. Nguyen KC, Alpcan T, Başar T (2009) Security games with incomplete information. In: IEEE international conference on communications, ICC'09, pp 1–6
86. Nguyen KC, Alpcan T, Başar T (2009) Stochastic games for security in networks with interdependent nodes. In: International conference on game theory for networks. IEEE, pp 697–703
87. Nguyen TH, Yadav A, An B, Tambe M, Boutilier C (2014) Regret-based optimization and preference elicitation for Stackelberg security games with uncertainty. In: AAAI, pp 756–762
88. Noureddine MA, Fawaz A, Sanders WH, Başar T (2016) A game-theoretic approach to respond to attacker lateral movement. In: International conference on decision and game theory for security. Springer, pp 294–313
89. Ouyang Y, Tavafoghi H, Teneketzis D (2017) Dynamic games with asymmetric information: common information based perfect bayesian equilibria and sequential decomposition. IEEE Trans Autom Control 62(1):222–237
90. Patcha A, Park J-M (2006) A game theoretic formulation for intrusion detection in mobile ad hoc networks. IJ Netw Secur 2(2):131–137
91. Pawlick J, Zhu Q (2015) Deception by design: evidence-based signaling games for network defense. arXiv preprint arXiv:1503.05458

92. Pita J, Tambe M, Kiekintveld C, Cullen S, Steigerwald E (2011) GUARDS: game theoretic security allocation on a national scale. In: The 10th international conference on autonomous agents and multiagent systems, vol 1, pp 37–44
93. Ramachandran K, Stefanova Z (2016) Dynamic game theories in cyber security. In: Proceedings of dynamic systems and applications, pp 1–8
94. Raya M, Manshaei MH, Félegyházi M, Hubaux J-P (2008) Revocation games in ephemeral networks. In: Proceedings of the 15th ACM conference on computer and communications security, pp 199–210
95. Raya M, Shokri R, Hubaux J-P (2010) "On the tradeoff between trust and privacy in wireless ad hoc networks. In: Proceedings of the third ACM conference on wireless network security, pp 75–80
96. Roy S, Ellis C, Shiva S, Dasgupta D, Shandilya V, Wu Q (2010) A survey of game theory as applied to network security. In: 43rd Hawaii international conference on system sciences (HICSS). IEEE, pp 1–10
97. Saad W, Zhou X, Maham B, Başar T, Poor HV (2012) Tree formation with physical layer security considerations in wireless multi-hop networks. IEEE Trans Wirel Commun 11(11):3980–3991
98. Sagduyu YE, Berry RA, Ephremides A (2011) Jamming games in wireless networks with incomplete information. IEEE Commun Mag, 49(8)
99. Sagduyu YE, Berry R, Ephremides A (2009) MAC games for distributed wireless network security with incomplete information of selfish and malicious user types. In: Proceedings of international conference on game theory for networks. IEEE, pp 130–139
100. Sallhammar K, Helvik BE, Knapskog SJ (2006) On stochastic modeling for integrated security and dependability evaluation. JNW 1(5):31–42
101. Sallhammar K, Knapskog SJ, Helvik BE (2005) Using stochastic game theory to compute the expected behavior of attackers. In: 2005 symposium on applications and the internet workshops, saint workshops. IEEE, pp 102–105
102. Sanjab A, Saad W, Başar T (2017) Prospect theory for enhanced cyber-physical security of drone delivery systems: a network interdiction game. In: Proceedings of IEEE ICC 2017 communication and information security symposium (CISS 2017), Paris, France, May 21–25
103. Shelar D, Amin S (2017) Security assessment of electricity distribution networks under DER node compromises. IEEE Trans Control Netw Syst 4(1):23–36
104. Shen D, Chen G, Blasch E, Tadda G (2007) Adaptive Markov game theoretic data fusion approach for cyber network defense. In: Military communications conference (MILCOM). IEEE, pp 1–7
105. Sorin S (2011) Zero-sum repeated games: recent advances and new links with differential games. Dynamic games and applications 1(1):172–207
106. Trejo KK, Clempner JB, Poznyak AS (2016) Adapting strategies to dynamic environments in controllable Stackelberg security games. In: Proceedings of IEEE 55th conference on decision and control (CDC), pp 5484–5489
107. Truong A, Etesami SR, Etesami J, Kiyavash N (2018) Optimal attack strategies against predictors-learning from expert advice. IEEE Transactions on Information Forensics and Security
108. Wang S, Shroff N (2017) Security game with non-additive utilities and multiple attacker resources. Proc ACM Measur Anal Comput Syst 1(1):13
109. Wang Y, Wang Y, Liu J, Huang Z, Xie P (2016) A survey of game theoretic methods for cyber security. In: IEEE international conference on data science in cyberspace (DSC), pp 631–636
110. Xiaolin C, Xiaobin T, Yong Z, Hongsheng X (2008) A Markov game theory-based risk assessment model for network information system. In: International conference on computer science and software engineering, vol 3. IEEE, pp 1057–1061
111. Zhang N, Yu W, Fu X, Das SK (2010) gPath: a game-theoretic path selection algorithm to protect Tors anonymity. In: Proceedings of international conference on decision and game theory for security (GameSec). Springer, pp 58–71
112. Zheng J, Castañón DA (2012) Dynamic network interdiction games with imperfect information and deception. In: IEEE 51st annual conference on decision and control (CDC), pp 7758–7763
113. Zheng J, Castanón DA (2012) Stochastic dynamic network interdiction games. In: Proceedings of IEEE American control conference (ACC), pp 1838–1844
114. Zhou Z, Bambos N, Glynn P (2016) Dynamics on linear influence network games under stochastic environments. In: Proceedings of international conference on decision and game theory for security (GameSec). Springer, pp 114–126
115. Zhu Q, Başar T (2015) Game-theoretic methods for robustness, security, and resilience of cyberphysical control systems: games-in-games principle for optimal cross-layer resilient control systems. IEEE Control Syst 35(1):46–65
116. Zhu Q, Fung C, Boutaba R, Başar T (2012) GUIDEX: a game-theoretic incentive-based mechanism for intrusion detection networks. IEEE J Select Areas Commun (JSAC) Spec Iss Econ Commun Netw Syst (SI-NetEcon) 30(11):2220–2230

117. Zhu Q, Başar T (2009) Dynamic policy-based IDS configuration. In: Proceedings of the 48th IEEE conference on decision and control (CDC), pp 8600–8605
118. Zhu Q, Başar T (2012) A dynamic game-theoretic approach to resilient control system design for cascading failures. In: Proceedings of the 1st international conference on high confidence networked systems. ACM, pp 41–46
119. Zhu Q, Clark A, Poovendran R, Başar T (2012) Deceptive routing games. In: IEEE 51st annual conference on decision and control (CDC), pp 2704–2711
120. Zhu Q, Tembine H, Başar T (2010) Heterogeneous learning in zero-sum stochastic games with incomplete information. In: 49th IEEE conference on decision and control (CDC), pp 219–224
121. Zhu Q, Tembine H, Başar T (2010) Network security configurations: a nonzero-sum stochastic game approach. In: IEEE American control conference (ACC), pp 1059–1064
122. Zhu Q, Tembine H, Başar T (2013) Hybrid learning in stochastic games and its applications in network security. In: Lewis FL, Liu D (eds) Reinforcement Learning and Approximate Dynamic Programming for Feedback Control, Series on Computational Intelligence, IEEE Press/Wiley, chapter 14, pp 305–329
123. Zonouz SA, Khurana H, Sanders WH, Yardley TM (2014) RRE: a game-theoretic intrusion response and recovery engine. IEEE Trans Parallel Distrib Syst 25(2):395–406