Game Theoretic Approaches to Cyber Security: Challenges, Results, and Open Problems

Hamidreza Tavafoghi[1], Yi Ouyang[2], Demosthenis Teneketzis[3], and Michael P. Wellman[4]

[1]Mechanical Engineering, University of California, Berkeley, CA
[2]Industrial Engineering & Operations Research, University of California, Berkeley, CA
[3]Electrical Engineering & Computer Science, University of Michigan, Ann Arbor, MI
[4]Computer Science & Engineering, University of Michigan, Ann Arbor, MI

Abstract

We formulate cyber security problems with many strategic attackers and defenders as stochastic dynamic games with asymmetric information. We discuss solution approaches to stochastic dynamic games with asymmetric information and identify the difficulties/challenges associated with these approaches. We present a solution methodology for stochastic dynamic games with asymmetric information that resolves some of these difficulties. Our main results are based on certain key assumptions about the game model. Therefore, our methodology can solve only specific classes of cyber security problems. We identify classes of cyber security problems that our methodology cannot solve and connect these problems to open problems in game theory.

1 Introduction

The high and continually increasing connectivity of modern cyber networks has resulted in significant improvement in the functionality and efficiency of our networked systems, but has also created new entry points for attackers, thus making these systems more vulnerable to intrusion. As noted by Miehling et al. [23], recent events such as information leakage and theft [7], car hacking [11], and denial of service attacks [6], have highlighted this vulnerability. Such vulnerability has become an issue of great concern because (i) the operation of critical infrastructure is increasingly reliant upon (potentially unreliable) networked systems and (ii) cyber attacks are becoming persistent and increasingly sophisticated. As reported by the Department of Homeland Security's Industrial Control Systems

Cyber Emergency Response Team (ICS-CERT), attacks on the critical infrastructure sectors (such as energy, communication, manufacturing, transportation, and water systems) have remained persistent over the past few years, with 245 in 2014, 295 in 2015, and 290 in 2016 [28], and many of the recent intrusions have had the potential to inflict severe and widespread damage (an increasing number of attacks have reached the system's control system layer [28]). Therefore, it is imperative to detect and mitigate cyber attacks so as to ensure secure operation of society's critical systems.

Cyber security is a complex problem. The complexity of the problem stems primarily from the fact that many individuals/agents (attackers, defenders) with different objectives, and different information about the cyber network's structure/topology and security status interact with one another through the network. At each time instant the cyber network's security status and each agent's information depend, in general, on exogenous random events (e.g., random failures in hosts and connections among hosts) and all the agents' strategies; such strategies are not common knowledge [1, 41] among all agents. Furthermore, the degree to which each agent achieves his objectives depends on his strategy and *all* the other agents' unknown strategies. Agents can use these these features of the cyber security problem to their advantage. For example, an attacker can take undetectable actions, or detectable actions that do not fully reveal intent; a defender can likewise take actions that are not observable by attackers. Under these conditions, the determination of strategic equilibria—configurations of strategies that leave no agent any incentive to deviate—is a formidable problem.

In this chapter we propose and present a game-theoretic approach to the cyber security problem. In Section 2 we explain why the formulation of the cyber security problem as a stochastic dynamic game with asymmetric information provides a reasonable approach to the problem. Then, in the remainder of the chapter we present: the game model that captures the salient features of cyber security problems (Section 3); current approaches to stochastic dynamic games with asymmetric information and the difficulties/challenges associated with them (Section 4); a new approach/methodology that resolves some of these difficulties along with a discussion of the methodology's key results (Section 5); the literature on game theory that is relevant to cyber security problems (Section 6); some open problems in game theory that are tightly connected to cyber security (Section 7).

The literature on stochastic dynamic games with asymmetric information is rich in deep ideas and is very technical, therefore, it is not easily accessible. Our goal in this chapter is to present and explain, in as plain language as possible, the key ideas behind the approaches to stochastic dynamic games with asymmetric information along with the main results of our approach to these games. For this reason we provide an informal presentation of the approaches to and the results on dynamic games of asymmetric information along with references that formally describe these approaches and results. The results presented in this chapter summarize a series of papers describing our work, motivated by issues in cyber security, on dynamic games with asymmetric information [30, 31, 36, 37, 38, 39].

2 The Cyber Security Problem as a Stochastic Dynamic Game with Asymmetric Information

As pointed out in the introduction, cyber security is a multi-agent problem involving attackers and defenders. The salient features of cyber security problems are: (i) Attacks are progressive in nature. Attackers are using their capabilities to attack and capture computers/hosts. Defenders attempt, through their actions, to retake hosts that are under the attackers' control and to limit the attackers' exploits (e.g., by isolating certain hosts from the rest of the network). As a result of the attackers' and defenders' actions, the security status of the network changes/evolves over time. The evolution is also affected by the occurrence of random events (e.g., random failures in hosts and connections between hosts) in the cyber network. (ii) Each agent has different information about the network's security status. For example: each defender knows in part the network's topology, but does not know the hosts/computers that are under the attackers' control and the information of other defenders; each attacker knows the hosts he has captured, but does not know the network's topology and the hosts captured by other attackers. In addition to their private information, attackers and defenders possess, at any time instant, some common information, consisting of events which they all observe, such as an attack to a group of servers that is detected and the detection is common knowledge ([2, 41]) among all attackers and defenders (for examples of events that are common information see [6, 7, 11, 28]). (iii) Attackers and defenders are strategic and self-interested: each agent attempts to optimize his own objective rather than a social/agent-wide objective. (iv) Each agent has different objectives. An attacker's objective is to acquire the information he is looking for and to take control of the cyber network. A defender's objective is to protect the information that is stored in the network without compromising too much the network's integrity and availability (a defender can protect his network by turning off the corresponding section of the network, however, such an action would make that section unavailable to its users).

As a result of the above features, the cyber security problem can be formulated as a dynamic game with asymmetric information where the underlying system is stochastic and dynamic. The attackers' and defenders' different objectives along with their strategic behavior lead to a game. The stochastic and dynamic nature of the game is due to the fact that the attackers' and defenders' information changes over time (it increases over time) and the network's security status evolves randomly over time. The fact that attackers and defenders possess, at every time instant, private information (in addition to their common information) results in an asymmetric information structure, thus a game with asymmetric information.

3 The Game Model

We present a model that captures the salient features of cyber security problems discussed in the previous section. We consider a stochastic dynamic system the evolution of which over a time horizon T is affected by the decisions/actions of N strategic agents along with random events that occur in nature. Such a system is described by a stochastic difference equation. Specifically, the system's state at time $t+1$, X_{t+1}, is a function of its state X_t at time t, the actions A_t^i, $i = 1, 2, \ldots, N$, of the N strategic agents at t, and random events that occur at time t and are statistically independent of the system's state at t and the agents' actions at t. At each time t, each agent has some *noisy/imperfect private information* about the system's evolution up to time t; such information is described by noisy observations of X_t and all the agents' actions $A_{t-1} = (A_{t-1}^i, i = 1, 2, \ldots, N)$ at time $t-1$. Furthermore, all agents have some *noisy/imperfect common information* about the system's evolution up to t; such information is described by common noisy observations of the system's state X_t and all the agents' actions A_{t-1}. We assume that the system's state, all the agents' private and common observations, and all the agents' actions are finite-valued. All agents have *perfect recall*, that is, at any time t, they remember everything they have observed up to t and all the actions they have taken up to $t-1$. Denote by P_t^i agent i's private information at t and by C_t the agents' common information at t, $i = 1, 2, \ldots, N$, $t = 1, 2, \ldots, T$. At each time t, agent i's action, $i = 1, 2, \ldots, N$, is generated by g_t^i, his strategy at t; g_t^i is a function of i's private information P_t^i at t and the common information C_t at t. We denote by g^i agent i's *strategy profile* in the T-horizon game; g^i is the collection of strategies g_t^i, $t = 1, 2, \ldots, T$. We term the collection of the agents' strategy profiles g^i, $i = 1, 2, \ldots, N$, the *strategy profile* g in the T-horizon game. At each t, agent i has an instantaneous utility $U_t^i(X_t, A_t)$ that is a function of the system's state X_t and all the agents' actions A_t^i, $i = 1, 2, \ldots, N$. Each strategic agent's objective is to determine his strategy profile so as to maximize the expected sum of his instantaneous utilities from the beginning (time 1) until the end of the game (time T).

The state X_t represents the system's/network's security status at time t. The progressive nature of cyber attacks is captured by the fact that X_t evolves dynamically over time, and its evolution at any time t is affected by the agents' (attackers' and defenders') actions at t and random events that occur in nature at t, such as network failures, and are independent of the agents' actions and the system's security status. Since cyber security systems are networks consisting of a finite number of computers, each computer's security status can be described by one of a finite number of states and each agent can take at any time t a finite number of actions, it is reasonable to assume that the system's state space along with the observation and action spaces are finite. We assume that all agents have *perfect recall*, that is, every agent remembers everything he has seen and every thing he has done. We will discuss the implication of this assumption in the analysis of dynamic games with asymmetric information later in this chapter. We wish to bring to the reader's attention two important features of the above-described model. (1) The instantaneous utility of each agent (hence his overall utility) depends on all agents' actions that are not all perfectly observable and are generated by their respective strategies. Therefore, each agent's choice of strategy must take into account all the other agents' choice of strategies, thus the agents' strategy choices are inter-dependent. (2) Since the dynamic system's evolution over time depends on all the agents' actions through their strategy choices, at any time t, each agent's information, private and common, depends on all agents' strategies up to $t-1$. These two features of the model result in significant difficulties in the analysis of dynamic games with asymmetric information and in the computation of the appropriate equilibria. We will discuss these difficulties along with ways of overcoming them in the rest of this chapter.

We conclude this section by presenting an example, drawn from [30, 31], that we will use throughout the chapter to illustrate the ideas and results we present. Even though the model of the example does not capture all the essential features of cyber security problems, (the current state of the art on stochastic dynamic games with asymmetric information cannot be used to solve the cyber security problem in its full generality), we hope that it will help the reader to understand and appreciate the difficulties/issues that arise in these games along with the key ideas of our approach.

An Example

Consider a game, played over a time horizon T, with N agents that are split into two groups. Group 1 consists of N_1 agents, group 2 consists of N_2 agents, $N_1 + N_2 = N$. The state of the dynamic system at time t is denoted by $X_t = (X_t^1, X_t^2, \ldots, X_t^{N_1}$). The component X_t^n of the state is privately observed by agent n in group 1. The private state X_t^n has uncontrolled Markovian dynamics with given time-invariant matrix of transition probabilities Q^n, $n = 1, 2, \ldots, N_1$. At the beginning of time t, each agent n in group 1 observes X_t^n, $n = 1, 2, \ldots, N^1$ and takes an action A_t^n. The actions $A_t = A_t^n, n = 1, 2, \ldots, N_1$ are announced to all N agents. After this announcement, agent m, $m = 1, 2, \ldots, N_2$, in group 2 makes a decision $D_t^m = (D_t^m(1), D_t^m(2), \ldots, D_t^m(N_1))$. Let $D_t = (D_t^1, D_t^2, \ldots, D_t^{N_2})$ denote the decisions of the agents in group 2 at time t. The decisions D_t are observed by all N agents. After the decisions D_t are made, all agents receive noisy observations $Y_t^1, Y_t^2, \ldots, Y_t^{N_1}$ of the states $X_t^1, X_t^2, \ldots, X_t^{N_1}$, respectively. The utility of agent n, $n = 1, 2, \ldots, N_1$, in group 1 is given by $U_t^{n,1}(A_t, D_t) = (A_t^n - c)(\sum_{i=1}^{N_2} D_t^i(n))$. The utility of agent m, $m = 1, 2, \ldots, N_2$, in group 2 is given by $u_t^m(Y_t, D_t, A_t) = V_t^m(Y_t, D_t) - (\sum_{i=1}^{N_1} D_t^m(i) A_t^i)$, where $V_t^m(Y_t, D_t)$ is a given function of Y_t and D_t, and $Y_t = (Y_t^1, Y_t^2, \ldots, Y_t^{N_1})$.

We assume that the state X_t of the dynamic system, the agents' actions A_t, D_t and the observations Y_t, $t = 1, 2, \ldots, T$ take values in finite spaces. Furthermore, we assume that the Markov state processes $X_t^n, t = 1, 2, \ldots, T, n = 1, 2, \ldots, , N_1$, describing the evolution of the dynamic system, and the observations Y_t, conditioned on X_t, $t = 1, 2, \ldots, T$, are all mutually independent.

In this game, the private information of agent n in group 1 at time t, before any action/decision is made at t, is $P_t^n = X_{1:t}^n$, where $X_{1:t}^n = X_1^n, X_2^n, \dots, X_t^n, n = 1, 2, \dots, N_1$. Agents in group 2 have no private information at any time. The common information of all N agents at time t is $C_t = A_{1:t-1}, D_{1:t-1}, Y_{1:t-1}$,

The action/decisison of agent n, $n = 1, 2, \dots, N$, at time t is a function g_t^n of his private information P_t^n at t and the common information C_t at t. The functions $g_t^n, t = 1, 2, \dots, T$ define agent n's strategy g^n in the game. The collection of strategies $g^n, n = 1, 2, \dots, N$, define the strategy profile played in the game.

4 Current Approaches to Dynamic Games with Asymmetric Information and the Associated Challenges

We provide an informal presentation of the key ideas underlying current approaches to dynamic games with asymmetric information along with the challenges associated with these approaches. For a formal presentation of current approaches to dynamic games with asymmetric information we refer the reader to [9, 25, 29].

The fundamental difficulties in the analysis of stochastic dynamic games with asymmetric information arise from the fact that the agents involved in the game (e.g. attackers and defenders) are strategic, they possess private information, their private and common information increase over time, their strategy choices are inter-dependent, and each agent's information depends, in general, on the strategies employed by all other agents.

To address these difficulties, classical approaches to dynamic games with asymmetric information proceed by taking into account the following considerations. At every instant of time, each agent has to form: (i) An appraisal about the current state of the (stochastic dynamic) system (e.g., the current security status of the network) and the other agents' private information; such an appraisal is about the history/past of the game. (ii) An appraisal about how other agents will play in the future so as to evaluate the performance of his strategy choices; such an appraisal is about the future of the game. Consider any agent, say agent i; given the other agents' strategies, at each time t, agent i can utilize his information (private and common) at t along with (i) all other agents' past strategies up to time $t-1$ and (ii) all other agents' future strategies from t up to T to form these appraisals about the the history/past and the future of the game, respectively.

Since each agent has his own objective, each agent's strategy is his own private information, thus, it is not known to other agents. Therefore, each agent has to form a prediction about the other agents' strategies. According to Nash's idea/model, all agents have a "common prediction about the strategy profile" played in the game (as stated in Section 3, a strategy profile describes all agents' strategies at all times). Such a prediction strategy profile does not necessarily coincide with the actual strategy profile that is being played in the game. Thus, it is possible that an agent's strategy, say agent i's strategy, is different from the prediction strategy profile. Denote by $g^* := (g^{*1}, g^{*2}, \dots, g^{*N})$ the common prediction about the agents' strategy profile, where g^{*i} is the strategy prediction profile for agent i (his strategy from time 1 up to time T); denote by $g := (g^1, g^2, \dots, g^N)$ the actual strategy profile, that is, the strategy profile that is being played by the agents, i.e. $g^{*i} \neq g^i$, in the game. Below we discuss the implications of agent i's deviation from the prediction strategy profile on all agents' behavior. For that matter, we first consider agent i who may want to deviate from his predicted strategy, then we examine the response of an agent who faces such a deviation, and finally we discuss how an agent determines his optimal strategy at each time t for all realizations of his information.

When agent i chooses a strategy, he needs to know how other agents will play/react for any choice that is different from the prediction strategy profile g^{*i}. Since a deviation from agent i may generate information (*e.g.* an observation) that is not expected when the prediction strategy profile g^* is played by all agents, (that is, information that has zero probability according to g^*), the prediction strategy profile g^* has to determine how agents will act for all possible realizations of their information, even those realizations that have zero probability according to it. Then, using g^*, agent i can form an appraisal about the future of the game for any choice of his own strategy and can evaluate the performance of that strategy.

By the same rationale, when agent i chooses any strategy g^i, he needs to determine that strategy for all possible realizations of his information, even those that have zero probability according to the prediction strategy profile $g^{*1}, g^{*2}, \dots, g^{*(i-1)}, g^{*(i+1)}, \dots, g^{*N}$. This is because some other agent j, different from i, may deviate from the prediction strategy profile g^{*j}, therefore, agent i must foresee such a possible deviation and must determine his response (according to g^i) to these deviations.

To determine his optimal strategy for all realizations of his information (those that have positive or zero probability under the prediction strategy profile g^*), at each time t, an agent, say agent i, needs to form an appraisal about the history of the game at t along with an appraisal about the future of the game under the assumption that all other agents follow the prediction strategy profile g^*. To form an appraisal about the history of the game at t, agent i proceeds as follows. For all realizations of his information up to and including t that have positive probability under g^*, he uses Bayes' rule to form this appraisal. For any realization of information up to and including t that has zero probability under g^*, agent i cannot rely on the strategy prediction g^* up to time $t-1$ and use Bayes' rule to form this appraisal. The realization of information of zero probability under g^* tells agent i that his original prediction up to time $t-1$ is not completely correct, consequently, he needs to revise his original strategy prediction up to $t-1$ and to form a revised appraisal about the history of the game at t. Therefore, agent i must determine how to revise/form his appraisal about the history of the game at t for all realizations of his information that have zero probability under the prediction strategy profile g^*.

In the game theoretic literature [9] the above considerations are formalized as follows. Each agent's appraisals, say agent i's appraisals, about the history and future of the game are captured by an *assessment* which consists of a *strategy prediction profile* g^* (that is common to all agents) and a *belief* μ_t^i, for each time t, on the system's state and the other agents' private information at t, based on agent i's information at t; for each t, the realization of such a belief is, in general, agent i's private information. The collection $\mu := (\mu_t^i, i = 1, 2, \dots, N, t = 1, 2, \dots, T)$ of all agents' beliefs at all times is called a *belief system* μ. At any time t and for any realization of agent i's information at t (such a realization may have positive or zero probability under the strategy prediction profile g^*) agent i's belief at t determines his appraisal about the history of the game at t; agent i's appraisal at t about the future of the game is determined by his belief μ_t^i at t and the prediction strategy profile $g_{t:T}^* = (g_t^*, g_{t+1}^*, \dots, g_T^*)$ from t until T (the end of the game).

Based on the definition of assessment, game theorists extended the concept of Nash equilibrium to dynamic games with asymmetric information. An equilibrium of the dynamic game is defined as a common assessment among the agents that satisfies the following conditions under the assumption that agents are rational. (1) Agent i, $i = 1, 2, \dots, N$, chooses his strategy to maximize his total expected utility in all continuation games (*i.e.* the game's continuation from t until T for all $t = 1, 2, \dots, T-1$) given the assessment about the game. Consequently, the prediction about agent i's strategy that other agents hold must be a maximizer of agent i's total expected utility under the assessment about the game. (2) For all times t, any agent i's belief at t that is based on a realization of information that has positive probability under the common assessment must be equal to the conditional probability, under the strategy prediction profile of the common assessment, of the system's state and the other agents' private information at t; this conditional probability for agent i is determined via Bayes' rule when all other agents play according to the common assessment's prediction strategy profile. When the realization of agent i's information at t has zero probability under the prediction strategy profile of the common assessment, his belief at t cannot be determined via Bayes' rule and must be revised. The revised belief must satisfy a certain set of *reasonable* conditions so as to be compatible with agent i's rationality. Game theorists have proposed various sets of conditions (see [9, 25, 29]) to capture the notion of reasonable beliefs that are compatible with the agents' rationality. Different sets of conditions for off-equilibrium beliefs, that is, beliefs along off the equilibrium paths of the game (i.e. paths of zero probability under the common strategy prediction component of the assessment) result in different equilibrium/solution concepts (such as perfect Bayesian equilibria, sequential equilibria, perfect equilibria, proper equilibria, persistent equilibria, *etc*) that have been proposed for dynamic games with asymmetric information (see [9, 25, 29] for the definition and meaning of all these equilibrium concepts).

Perfect Bayesian Equilibrium (PBE) [9, 42], is an equilibrium concept that has been widely accepted as an appropriate solution concept for dynamic games with asymmetric information. A PBE is defined as an assessment (a strategy prediction g^* and a belief system μ) that satisfies the *sequential rationality* and *consistency* conditions. The sequential rationality and consistency conditions for dynamic games with asymmetric information where the underlying system is dynamic are formally defined in [36]. Here we verbally describe these conditions.

Sequential Rationality. *Consider any agent at any time, say agent i at time t.*

Given agent i's information (private and common) at t, his belief at t according to the assessment, and the prediction of all other agents' strategies from t until T according to the assessment, agent i's strategies that maximize his (total) expected utility from t until T are the same as his corresponding strategies from t until T in the assessment.

Sequential rationality requires that the common prediction g^{*i} about agent i's strategy must be an optimal strategy choice for him since it is common knowledge that he is a rational agent. We note that the sequential rationality condition is more restrictive than the optimality condition for Bayesian Nash Equilibrium (BNE) which requires that the optimality condition in italics above should hold only for $t = 1$. By the sequential rationality condition we require the optimality of the strategy prediction g^* even along paths of the game's evolution that are off-equilibrium (*i.e.* paths that have zero probability under g^*), thus, we rule out the possibility of *non-credible threats.* Consider for example an agent who threatens to play an action that is suboptimal for himself upon the realization of a history that has zero probability under the strategy prediction g^* of the assessment. Such a non-credible threat is ruled out by sequential rationality (hence by PBE) but is not ruled out by BNE. Sequential rationality gives rise to a set of conditions that the strategy prediction g^* must satisfy given the belief system μ of the assessment. As discussed above, the belief system μ of the assessment should also be compatible with g^*. The compatibility between the strategy prediction component of the assessment and the belief system component of the assessment is captured by the consistency condition.

Consistency. *The consistency condition requires that along all equilibrium paths (that is, game histories that are realized when all agents play an equilibrium strategy profile) the agents' beliefs should be updated/evolve according to Bayes' rule. Along all other paths of the game's evolution, the consistency condition requires that if the information received by an agent i at time t has zero probability under the assessment, agent i's belief at t must be revised in a "reasonable" manner.*

The work in [36] presents a set of "reasonable" conditions for revising an agent's beliefs along off-equilibrium paths of the game's evolution.

Even though the definition of PBE provides a general formalization of outcomes that are *rationalizable* (that is, consistent with agents' rationality) under some strategy profile and belief system, computation of PBEs is a formidable task. There are two major challenges in computing a PBE (g^*, μ). First, there is an inter-temporal coupling between the agents' strategy prediction g^* and the belief system μ. As discussed before, according to the consistency requirement, the belief system μ must satisfy a set of conditions given a strategy prediction g^* (see [36, 39]). On the other hand, sequential rationality dictates that a strategy prediction g^* must satisfy a set of optimality conditions given the belief system μ (see [36, 39]). Therefore, there is a circular dependency between a strategy prediction g^* and a belief system μ. For example, by sequential rationality, agent i's strategy g_t^{i*} at time t depends on the agents' future strategies $g_{t:T}^*$, (where $t : T$ denotes the time interval $t, t+1, t+2, \dots, T$), and on the agents' past strategies $g_{1:t-1}^*$ indirectly through the consistency condition for μ_t^i. As a result, one has to determine the strategy prediction

g^* and the belief system μ simultaneously for the whole time horizon so as to satisfy all the sequential rationality and consistency conditions; consequently, one cannot sequentially decompose over time the computation of PBEs. Second, since the agents' are assumed to have perfect recall their information (private and common) increases over time, thus, their strategies have a growing domain over time; this feature of the agents' strategies further complicates the computation of PBEs.

We continue discussing the example we introduced at the end of Section 3. Here we illustrate the concepts introduced in this section along with the difficulties that arise in the determination of equilibrium strategies in dynamic games with asymmetric information.

An Example (continued)

An assessment of the game is described by a strategy prediction profile $g^* = (g^{*1}, g^{*2}, \dots, g^{*N})$, that is common to all agents, and a belief system $\mu_t^i, i = 1, 2, \dots, N$, $t = 1, 2, \dots, T$. The component $g^{*i} = (g_1^{*i}, g_2^{*i}, \dots, g_T^{*i}\ )$ describes the strategy player i is predicted to play in the game. The strategy g^{*i} may be different from the actual strategy g^i player i plays in the game. The component μ_t^i describes agent i's belief at time t about the state of the system X_t and the private information P_t^j of all agents j other than i at time t, conditioned on agent i's private information P_t^i and the common information C_t (the private and common information for all agents at all times have been specified at the beginning of this example, at the end of Section 3).

At any time t,the beliefs μ_t^i, $i = 1, 2, \dots, N$, depend on $g_{1:t-1}$, the strategies of all agents up to time $t-1$. At each time t, when an agent, say agent i, determines his best strategy from t up to time T according to the sequential rationality requirement, he takes into account the strategy prediction profile for all other agents form t up to T and his belief μ_t^i. Therefore, the (equilibrium) strategies and beliefs of all agents are interdependent over time and must all be determined simultaneously for the whole duration of the game. This fact highlights one of the major difficulties associated with the current approaches to stochastic dynamic games with asymmetric information.

As pointed out at the beginning of this example, (end of Section 3), the private information of each agent i in group 1 at time t is $P_t^i = X_{1:t}^i, i = 1, 2, \dots, N_1$. The common information of all agents at time t is $C_t = A_{1:t-1}, D_{1:t-1}, Y_{1:t-1}$. Consequently, the domains of private and common information increase with time, therefore , for large time horizons T, the computation of equilibria becomes a formidable task. This fact highlights another difficulty associated with the current approaches to stochastic dynamic games with asymmetric information.

In the next section we present an approach to dynamic games with asymmetric information that addresses and partly resolves the above challenges.

5 A Sufficient Information Approach to Stochastic Dynamic Games with Asymmetric Information

The definition of PBE requires agents to keep track of all the information they acquire over time and to form beliefs about the private information of all other agents. In this section we show that agents do not need to keep track of all their past information to reach an equilibrium; at any time t, they can take into account only a subset of the information available at t that is *relevant* to the continuation of the game, and ignore the rest of it. Such a selection of the relevant information is motivated by computational and philosophical reasons: the resulting strategies are simpler and the corresponding PBE are easier to compute; furthermore, the simpler strategies proposed in this section offer a more plausible prediction of the outcome of the interactions among strategic agents in cyber security games where the underlying system is dynamic (due to the progressive nature of attacks) and there is significant asymmetry in the information possessed by the agents (attackers and defenders).

The above discussion motivates the approach to dynamic games with asymmetric information that we present in this section. The key steps of our approach are as follows.

1. We present conditions sufficient to guarantee that all the agents involved in the game can compress their private information in a mutually consistent manner. Such a mutually consistent compression leads to the notion/concept of *sufficient private information.*

2. Using the notion of sufficient private information, we present a compression of the common information. Such a compression of the common information leads to the concept of *sufficient common information.*

3. Using the notions of sufficient private information and sufficient common information we define a set of strategies termed *sufficient information based strategies* (SIB strategies), and a set of PBE termed *Sufficient Information Based Perfect Bayesian Equilibria* (SIB-PBE). We show that the set of sufficient information based strategies is closed under the best response correspondence. Thus, we establish the following result: if all agents except one, say agent i, play sufficient information based strategies, then there exists a sufficient information based strategy for agent i that is a best response to those strategies. The implication of this result is that one can restrict attention to SIB strategies to determine SIB-PBE.

4. Using the sufficient common information as an information state we provide a sequential decomposition of stochastic dynamic games with asymmetric information. Such a decomposition leads to an algorithm for the determination of SIB-PBE. We identify instances of games where SIB-PBE exist.

5.1 Sufficient Private Information

Let C_t denote the agents' common information at time t, and let P_t^i denote agent i's private information at t, $i = 1, 2, \ldots, N$.

Sufficient Private Information. We say that the collection $S_t = (S_t^i, i = 1, 2, \ldots, N)$, (where each S_t^i a function of C_t and P_t^i), is sufficient private information for the N agents if the following conditions are satisfied for all agents and for all times: (i) Each S_t^i can be updated recursively, that is, S_t^i can be determined from S_{t-1}^i and the new information agent i acquires at time t. (ii) S_t,C_t and the agents' actions A_t at t provide the information sufficient to statistically determine the sufficient private information S_{t+1} and the common information C_{t+1}. (iii) Agent i's expected utility at t conditioned on his private information P_t^i, the common information C_t, and the agents' actions A_t at t, is the same as his expected utility at t conditioned on his sufficient private information S_t^i, the common information C_t and the agents' actions A_t at t. Furthermore both expected conditional utilities at t are independent of agent i's strategy. (iv) The information provided by agent i's private sufficient information S_t^i and the common information C_t is sufficient for agent i to statistically determine/predict the sufficient private information of all other agents. Furthermore, this statistical determination/prediction is independent of agent i's strategy.

The above discussion provides an informal presentation of conditions (i)-(iv); a formal/mathematical description of these conditions can be found in [36, 38, 39].

We now provide an intuitive interpretation of the above conditions. Condition (ii) requires that that agents' sufficient private information must be rich enough so that, combined with their common information and actions at any time t, it leads to the same prediction of the sufficient private information and the new common information at $t + 1$ as the one that would be obtained if agents used all their information at t. Condition (iii) is similar in spirit to one of the requirements defining an information state in centralized stochastic control [18]. The essence of conditions (ii) and (iii) is that sufficient private information must be a component of a statistic that is sufficient for decision making purposes. Therefore, sufficient private information must be updated recursively (condition (i)). The essence of condition (iv) is the following: the agents' sufficient private information must be defined by a mutually consistent compression of all the agents' private information. Such a compression must not entail any loss of information, as far as an agent's ability to statistically predict the other agents' sufficient private information is concerned; furthermore, such a private information compression must be robust to agents' possible unilateral deviations from the strategy prediction g^*.

We would like to point out that the above conditions do not uniquely determine the agents' sufficient private information. These conditions may lead to many solutions including the trivial one $S_t^i = P_t^i$ for all agents i, $i = 1, 2, \ldots, N$. Therefore, an important question is: is there a minimal sufficient private information for the agents? The existence of a minimal sufficient private information for all agents is currently an important open problem.

5.2 Sufficient Common Information

Based on the characterization of sufficient private information, we introduce the notion of sufficient common information which at any time t is a statistic/compressed version of the common information C_t at t.

Sufficient Common Information. We define the agents' sufficient common information at any time t, denoted by Π_t, to be the agents' belief about the dynamic system's state X_t at t, and all the agents' sufficient private information S_t at t, conditioned on the common information C_t.

We call Π_t the *Sufficient Information Based (SIB) belief at t.* The agents' SIB belief at t is computable by all agents, thus, it is common information ([1, 41]) among all agents. The SIB belief Π_t is recursively updated according to a *SIB update rule* ψ_t. Specifically, Π_{t+1} is determined by Π_t and the common information that becomes available at t, that is $Z_t = C_t \backslash C_{(t-1)}$, according to ψ_t. If the realization of the information Z_t has non-zero probability according to the strategy prediction g^*, then ψ_t updates Π_t according to Bayes' rule; if the realization of Z_t has zero probability according to g^*, then Π_{t+1} is updated according to ψ_t in a *reasonable* manner that is consistent with the agents' rationality (see for example [36]). We denote by $\Pi^{\psi} = (\Pi_t^{\psi}, t = 1, 2, \ldots, T)$ the sequence of SIB beliefs generated by the update rule $\psi := (\psi_t, t = 1, 2, \ldots, T)$.

The above discussion provides an informal presentation of the concept of sufficient common information. For a formal/mathematical description of sufficient common information and its update we refer the reader to [36, 38, 39].

5.3 Sufficient Information-Based Strategies and Sufficient Information-Based Perfect Bayesian Equilibria

The combination of sufficient private information $(S_t^i, i = 1, 2, \ldots, N, t = 1, 2, \ldots, T)$ and sufficient common information Π_t, $t = 1, 2, \ldots, T$, provides a mutually consistent compression of the agents' private and common information, respectively. Using this information compression we define a class of strategies σ_t^i that are based on S_t^i and Π_t for each agent i at each time t. We call σ_t^i a *Sufficient Information Based (SIB) strategy* for agent i at time t. A collection of SIB strategies σ_t^i, $i = 1, 2, \ldots, N$, $t = 1, 2, \ldots, T$, is termed a *SIB strategy profile* σ. We note that at any time t the set of SIB strategies is a subset of all possible strategies agents can choose at t by using all of their private and common information at t. SIB strategies are simpler than general strategies because they have a smaller domain than general strategies as they are based on compressed versions of the agents' private and common information at any time t. We further note that if the dimensionality of S_t^i, agent i's sufficient private information, $i = 1, 2, \ldots, N$, remains fixed over time then the domain of SIB strategies is time-invariant. In Section 5.5 we present instances of dynamic games with asymmetric information where the domain of SIB strategies is time-invariant.

Based on the concept of SIB strategies we introduce the concept of *Sufficient Infor-*

mation Based-Perfect Bayesian Equilibrium (SIB-PBE) that is informally described as follows.

Sufficient Information Based-Perfect Bayesian Equilibrium (SIB-PBE). A SIB-PBE is a PBE in which all agents play SIB strategies.

For a formal definition of SIB-PBE we need to consider, as in the case of PBE, a SIB assessment that consists of a SIB strategy prediction profile σ and a SIB belief system μ^{ψ}, and to define the sequential rationality and consistency conditions that σ and μ^{ψ} must satisfy so that they should specify a SIB-PBE. A formal definition of SIB-PBE can be found in [36].

The class of SIB assessments needed to formally define a SIB-PBE imposes two additional restrictions/requirements on the agents' strategies and beliefs as compared to the general class of assessments presented in Section 4. First, SIB assessments require that each agent i, $i = 1, 2, \dots, N$, must play a SIB strategy σ^i instead of a general strategy g^i. Second, SIB assessments require that at every time t each agent i must form a belief about the status of the game using only the SIB belief Π_t along with his sufficient private information S_t^i (instead of a general belief μ_t^i that is based on his private information P_t^i and the common information C_t). Such restrictions generate the following strategic concerns. First, a strategic agent does not have to restrict his choice to SIB strategies, he may deviate from a SIB strategy σ^i to a non-SIB strategy g^i if such a deviation is profitable for him. Second, at any time t, a strategic agent does not have to limit himself to forming a belief about the status of the game by using only the SIB belief Π_t and sufficient private information S_t^i; he may want to form a belief using all of his private information and all the common information if such a belief enables him to improve his overall expected utility. These concerns are addressed by the results of our methodology that appear in the next section.

5.4 Main Results

The results we present in this section address the difficulties associated with current approaches to dynamic games with asymmetric information, specifically: the inter-dependence over time between strategies and beliefs (Section 4); the growing domain of the agents' strategies (Section 4); and the strategic concerns arising from restricting attention to SIB strategies and SIB beliefs (Section 5.3). These results have been derived under the following key assumption, the meaning of which we discuss in the following subsection.

Key Assumption. At any time t, $t = 1, 2, \dots, T$, and for any sequence of all the agents' actions up to time $t-1$ the following conditions are satisfied: (C1) Every possible value x_t of the system state X_t can be realized with positive probability. (C2) For every agent, every possible value of his private observations can be realized with positive probability.

We present an informal statement of the four main results of the sufficient information approach to dynamic games with asymmetric information. A formal statement of conditions (C1) and (C2) along with a formal statement of the four main results and their proofs can be found in [36].

Result 1. At any time t, every agent i's private belief about the state of the dynamic system and the the private information of all other agents is independent of his own strategy.

Result 2. If every agent $j \neq i$ plays a SIB strategy σ^j, then there exists a SIB strategy σ^i for agent i that is a best response to the strategies $(\sigma^j, j \neq i)$.

Results 1 and 2 address the strategic concerns created by focusing on SIB strategies and SIB beliefs, and, in part, the growing domain of the agents' strategies. Result 1 shows that no agent can alter his private belief about the state of the dynamic system and all the other agents' private information by deviating from the predicted strategy profile. Thus, when all agents $j \neq i$ play according to a SIB assessment (σ^*, μ^ψ) (where the belief system μ^ψ is determined by the update rule ψ described in Section 5.2), agent i cannot mislead these agents by playing a strategy g^i different from σ^{i*} , thus, creating dual beliefs (one belief that is based on the SIB assessment (σ^{i*}, μ^ψ) the functional form of which is known to all agents, and another belief that is based on his private strategy g^i that is only known to him) which he can use to his advantage. Result 2 shows that when all agents play SIB strategies, no agent can profit by deviating from his SIB strategy to a non-SIB strategy. Therefore, we can restrict attention to SIB strategies and attempt to determine PBE within the class of SIB assessments, i.e. SIB-PBE. As pointed out above, SIB strategies are simpler than general strategies (which, at any time, are functions of all the private and common information available to an agent at that time) because they are based on compressed information. However, SIB strategies do not have, in general, a time-invariant domain. Nevertheless, there are several instances in practice where SIB strategies have a time-invariant domain ([30, 31, 37]).

Using Results 1 and 2 we can obtain a sequential decomposition of stochastic dynamic games with asymmetric information. Such a decomposition is described by the following result.

Result 3. SIB-PBE can be determined by the solution of N coupled dynamic programs (one for each agent). These dynamic programs determine sequentially (moving backwards in time) SIB-PBE via the solution (*i.e.* the Bayesian Nash equilibria) of a series of T static Bayesian games that have the following form. For the game at time T, and for any realization π_t, $s_T^i = (s_T^i, i = 1, 2, \ldots, N)$ of the SIB belief Π_T and the sufficient information $S_T = (S_T^i, i = 1, 2, \ldots, N)$, respectively, agent i's utility is the expectation of his original utility U_T^i (see Section 3) conditioned on the π_T and s_T. For the game at time t, $t = 1, 2, \ldots, T-1$, and for any realization π_t, $s_t = (s_t^i, i = 1, 2, \ldots, N)$ of the SIB belief Π_t and the sufficient private information $S_t = (S_t^i, i = 1, 2, \ldots, N)$, respectively, agent i's utility is the sum of two terms: (i) the expectation of his original utility U_t^i conditioned on π_t and s_t; and (ii) his expected payoff from time t+1 until time T, due to the continuation of the game, conditioned on π_t and s_t. The second term of the above sum is a function of the SIB belief Π_{t+1} (which, according to Section 5.2, is recursively determined form π_t and the new common information Z_{t+1} acquired at $t+1$) and the sufficient private information S_{t+1} (which, according to Section 5.1, is recursively determined from s_t and the new information the agents acquire at $t+1$).

Result 3 shows that for finite horizon stochastic dynamic games with asymmetric information our approach resolves the difficulty due to the inter-dependence over time between strategies and beliefs (discussed in Section 5.4) by providing a systematic method for determining the components of SIB-PBE one at a time, starting at time T and sequentially moving backwards in time. The N coupled dynamic programs provide an algorithm for determining SIB-PBE.

Results 1 and 2 hold for both finite and infinite horizon games. Under certain additional assumptions, Result 3 can be extended to infinite horizon games (see [36, 39]).

Result 4. (i) SIB-PBE exist for zero-sum games. (ii) For nonzero-sum games there exists at least one SIB-PBE (σ^*, μ^ψ), if the following condition is satisfied. There exists sufficient information $S_{1:T}^{1:N}$ such that the update rule ψ is independent of the strategy profile σ^*.

The independence condition of Result 4 is not satisfied for all cyber security games. In [36] we present several classes of dynamic games with asymmetric information where the condition of Result 4(ii) is satisfied.

We illustrate the results of our approach to stochastic dynamic games with asymmetric information through the example introduced in Section 3.

An Example (continued)

For the example introduced in Section 3, at any time t the sufficient private information of agent i in group 1 is $S_t^i = X_t^i$. As pointed out earlier, all agents in Group 2 have no private information, thus, no sufficient private information. The sufficient common information for all agents at time t is the belief on the system state X_t conditioned on the common information $C_t = A_{1:t-1}, D_{1:t-1}, Y_{1:t-1}$. Note that the sufficient private information of each agent, and the sufficient common information for all agents have time invariant domains.

Using the sufficient common information as an information state, we can show [31, 30] that PBE assessments can be determined sequentially in time by a backward induction algorithm.

Therefore, for the game of the example introduced in Section 3, our methodology resolves the key difficulties (discussed in Section 4) that are associated with previous approaches to dynamic games with asymmetric information. That is, it breaks the interdependence over time between strategies and beliefs (through the sequential decomposition of the game) and discovers, for each agent and each time, sufficient private information and sufficient common information that have time invariant domains.

5.5 Discussion of the Main Results

Our main results show that the mutually consistent compression of the agents' information (private and common) leads to SIB strategies, SIB beliefs, and SIB-PBE which have several desirable features. Specifically, SIB strategies are simpler than general strategies, and SIB beliefs, which are common knowledge among all agents, can serve as information states in the sequential decomposition of stochastic dynamic games with asymmetric information. In general, the set of SIB-PBE of a dynamic game is a subset of all PBE of the game. This is because in a dynamic game agents can incorporate their past irrelevant observations into their future decisions so as to create rewards (respectively, punishments) that incentivize them to play (respectively, not to play) specific actions over time. By compressing the agents' private and common information, we do not capture such punishment/reward schemes that are based on past irrelevant observations. An example of such a situation appears in [36, 39] within the context of a repeated game, where the set of PBE that can not be captured as SIB-PBE are the ones that utilize payoff-irrelevant information to create reward/punishment schemes in the continuation game.

We would like to note that in dynamic games where the agents' equilibrium payoffs are unique, we can restrict attention to SIB-PBE because the above-described punishment/reward schemes do not lead to additional equilibrium payoffs. One class of such games is the class of zero-sum games (*e.g.*, attacker-defender games within the context of cyber security where the defender's only concern is the network's security). In a zero-sum game agents have completely opposite interests, therefore, it is not rational for them to cooperate on the formation of such punishment/reward schemes; we refer the interested reader to [36, 39] for more discussion and the proof of existence of SIB-PBEs in zero-sum games.

Even though it is true that, in general, the set of PBE of a dynamic game is larger than the set of SIB-PBE, in our opinion there are reasons on why in a highly dynamic environments, such as the the environment of cyber security problems, SIB-PBE are more plausible to arise as an outcome of the game.

First, we argue that in a highly dynamic environment with significant information asymmetries among agents, the creation/formation of reward/punishment schemes that utilize the agents' payoff-irrelevant information requires prior complex agreements among the agents. These complex agreements are more likely to occur in games where the underlying system is not highly dynamic (as in repeated games [19]) and there is no much information asymmetry among agents. In a highly dynamic environment with significant information asymmetries among agents (as in cyber security games) the formation of such complex agreements becomes less likely for the following reasons. First, in these environments each agent's individual decision making process is described by a complex Partially Observable Markov Decision Process (POMDP); thus, strategic agents are less likely to form a prior common agreement (that depends on the solution of their POMDPs) in addition to solving their individual POMDPs. Second, as the information asymmetry among agents increases, reward/punishment schemes that utilize payoff-irrelevant information require an increasingly complex agreement that is sensitive and not robust to changes in the assumptions on the information structure of the game. An example illustrating the lack of robustness of these agreements to changes in the information structure of the game is provided in [36, 39]. The author of [24] provides a general result on the robustness of the above mentioned reward/punishment schemes in repeated games; he shows that the set of equilibria that are robust to changes in the game's information structure that affect only payoff-irrelevant signals do not include the set of equilibria that utilize the reward/punishment schemes described above.

Second, the proposed solution concept SIB-PBE can be viewed as a generalization/ extension of Markov Perfect Equilibrium (MPE) [21] to dynamic games with asymmetric information. Therefore, a similar set of rationales that support the notion of MPE also applies to the notion of SIB-PBE as follows. First, the the set of SIB assessments, as presented in [36, 39], describes the simplest form of strategies capturing the agents' behavior that is consistent with the agents' rationality. Second, the class of SIB-PBE captures the idea that "bygones are bygones", which also underlies the requirement of subgame perfection in equilibrium concepts for dynamic games. That is, the agents' strategies in two continuation games that differ only in the agents' information about payoff-irrelevant events must be identical. Third, SIB assessments embody the principle that "minor changes in the past should have minor effects". This implies that any perturbation in the specification of the game or in the agents' past strategies that are irrelevant to the continuation game should not change drastically the outcome of the continuation game.

We would like to emphasize that the key assumption of Section 5.4 is essential in establishing the assertions of the main results of the approach presented in this section. Condition (C1) says that there is enough exogenous uncertainty (i.e random uncontrollable events) in the system's evolution so that at each time t all states in the system's state space can be reached with positive probability; condition (C2) says that no agent can infer perfectly another agent's actions based only on his private observations; equivalently, condition (C2) says that any deviation that is detected by a certain agent is also simultaneously detected by all the other agents. We believe that within the context of cyber security problems these conditions are fairly reasonable. For example, when the system/network is heavily used there is a high likelihood that random failures induced by the heavy load can potentially lead to one of many security states. Furthermore, the information agents receive from their own (private) sensors can be very noisy, thus they are not able to perfectly detect other agents' actions. Nevertheless, there are instances of cyber security games with many players/agents (attackers and defendants) where an agent's deviation may be detected by a subset of the rest of the agents (this subset of agents use their private information to detect the deviation). The methodology presented in this section cannot address these instances. We present some ideas on how to address these instances in Section 7.

Even though there are instances of dynamic games with asymmetric information where the domain of SIB strategies is time-invariant, *e.g.* [31, 37], the methodology for information compression presented in this section does not always result in sufficient private information the domain of which is time-invariant. Thus, our methodology does not completely resolve the difficulty arising from the growing domain of the agents' strategies in dynamic games with asymmetric information. In Section 7, we present some ideas on how to address this difficulty.

We conclude our discussion by pointing out that the main results 1-3 can be obtained if we replace the key assumption of Section 5.4 with another one where each agent's actions are always observable by all other agents. However, such an assumption is not realistic for cyber security problems.

6 Relevant Literature

The literature on dynamic games with asymmetric information can be divided into two categories: (1) games where the underlying system is static (repeated games); and (2) games where the underlying system is dynamic. There are significant philosophical differences between the approaches to games in the above categories.

Dynamic games where the underlying system is static (repeated games) arise primarily in economic problems where the environment does not change with time or evolves very slowly over time. Works on (discounted) repeated games study primarily their asymptotic properties, specifically their properties when the horizon is infinite and agents are sufficiently patient (that is the discount factor is close to 1). In repeated games agents play a stage (static) game repeatedly over time. The main objectives of the literature on these games are: (i) to analyze situations where the agents can form self-enforcing punishment/reward mechanisms so as to create additional equilibria that improve the payoffs they obtain by playing an equilibrium of the stage game over time; and (ii) to characterize the payoffs corresponding to all the equilibria of the repeated game.

Dynamic games where the underlying system is dynamic arise in engineering problems where the environment evolves rapidly over time. For example, in cyber security, the progressive nature of cyber attacks results in a rapidly changing environment, this is why the underlying system is modeled by a stochastic difference equation (see Section 3). The work existing on games with asymmetric information where the underlying system is dynamic does not restrict attention only to situations where the horizon is infinite and agents are sufficiently patient. The literature addresses situations where the decision problem for each agent, in the absence of interactions with other agents (*i.e.* assuming fixed strategies for the other agents), is a POMDP. Therefore, the determination of a set of equilibrium strategies is a complex problem. Consequently, it is unlikely that the agents seek equilibria that result from the formation of self-enforcing punishment/reward mechanisms that are similar to those used in infinitely repeated games. Existing approaches to and results on dynamic games with asymmetric information where the underlying system is dynamic demonstrate that the equilibria of these games have the same features as the equilibria determined by our approach (see Section 5.5). For this reason, in this section we will provide a detailed description of the literature on stochastic dynamic games with asymmetric information where the underlying system is dynamic. At the end of the section we will provide a few key references on dynamic games with asymmetric information where the underlying system is static.

Stochastic dynamic games with asymmetric information where the underlying system is dynamic can be classified into two categories, zero-sum and nonzero-sum. Cyber security problems are usually modeled as non zero sum games because the attackers' and defenders' objectives are not exactly the opposite of each other (see Section 3). For this reason, first we will briefly review the literature on zero-sum dynamic games and then we will provide a more detailed discussion of the literature on non zero-sum games.

The works in [3, 10, 16, 17, 32] consider dynamic zero-sum games with asymmetric information. The authors of [3, 32] study two-player games with Markovian dynamics and lack of information on one side (that is, one player/agent who has perfect knowledge of the game that is being played and one player who has partial/incomplete knowledge of the game that is being played). The authors of [10, 16] study two-player zero-sum games with Markovian dynamics and lack of information on both sides (that is, both players possess only partial/incomplete knowledge of the game that is being played). We would like to point out that the authors of [3, 10, 16, 32] consider models with specific Markovian dynamics where each agent observes perfectly a local state that evolves independently of all other local states conditioned on the agents' observable actions. Thus, even if one attempted to formulate cyber security games as dynamic zero-sum games with asymmetric information, the results of the above mentioned papers could not provide any answers or insights because in cyber security games the agents' actions are not, in general, observable, agents have imperfect (noisy) observations of the system's/network's security status, and the game's information structure (who knows what and when) is considerably more complex than that of the above mentioned references. One instance of zero-sum stochastic dynamic games where the agents' actions are not observable is analyzed in [26]. The authors of [26] consider zero-sum games with asymmetric information where the agents, in addition to having private information, share, at each time instant, some common information, and they play pure strategies. They prove that if the set of saddle point equilibria of the above games is non-empty, then the (minmax) value of these games is the same as the value of the (symmetric) games where the agents' only information is their common information. They provide an algorithm for determining the value of the symmetric information games.

The literature on stochastic dynamic non zero-sum games with asymmetric information, where the underlying system is dynamic, addresses mostly situations where, in addition to their private information, all agents have some common information (see [5, 12, 15, 27, 30, 31, 33, 35, 36, 37, 39, 40]). Refernces [5, 15, 35], consider infinite horizon discounted games where the underlying system is a controlled Markov chain. The approach taken in [5, 15, 35] is based on the philosophy and ideas used to analyze infinitely repeated games.

In the work reported in [15] the system's state is perfectly observed by all agents at all times, and each agent's actions are his private information (hidden actions); attention is restricted to *Perfect Pubic Equilibria (PPE)*, that is, equilibria that result in when agents play only common information-based strategies. The authors of [15] characterize, under certain assumptions that appear in [15], the set of the agents' payoffs that correspond to all PPE when all agents are sufficiently patient, that is, the discount factor δ approaches 1. The authors of [5] consider games where at each time all agents observe perfectly each others' actions but each agent has imperfect private information about the system's state. They consider PBE as a solution/equilibrium concept, and characterize, under certain assumptions that are explicitly stated in [5], the set of the agents' payoffs corresponding to all PBE of the game when all agents are sufficiently patient. Sugaya [35] analyzes instances of games where each agent has imperfect private information about the system's state and private monitoring of the other agents' actions; furthermore, he assumes that agents communicate with one another via perfect and public cheap talk. He adopts PBE as the equilibrium/solution concept and characterizes, under certain assumptions that are explicitly stated in [35], the agents' payoffs that correspond to all PBEs of the game when the agents are sufficiently patient. References [12, 27, 30, 31, 33, 36, 37, 39, 40] analyze finite and/or infinite horizon discounted dynamic games. In all of these references, the agents' common information is used as an instrument for coordination of the agents' strategies. In [12, 27, 30, 31, 33, 37, 40], the *Common Information Based (CIB) belief* (the belief on the dynamic state state at time t, and all the agents' private information at t, based on the agents' common information at t, $t = 1, 2, \dots, T$) is an information state/sufficient statistic for decision making for each agent at t. In [36, 38, 39] the *SIB belief* Π_t, $t = 1, 2, \dots, T$, defined in Section 5.2, is an information state for decision making for each agent at time t. In the game instances investigated in [12, 27] the CIB belief is independent of the agents' strategies; in such a situation, assessments (defined in Section 5.4) can be described simply by the agents' strategy prediction (defined in Section 5.4), and an appropriate equilibrium concept is *Common Information Based Markov Perfect Equilibrium* that was introduced in [27]. In the game instances investigated in [30, 31, 33, 40], the CIB beliefs depend on the agents' strategies, the agents' actions are always perfectly observable, and the agents' (private) beliefs (defined in Section 5.4) are common knowledge ([1, 41]) among all agents. An appropriate equilibrium concept for these instances of games is *Common Information Based-Perfect Bayesian Equilibrium (CIB-PBE)* that was introduced in [30, 31]. In the game instances investigated in [36, 37, 39] the agents' SIB beliefs depend on their strategies, the agents' actions are not observable, and the agents' (private) beliefs are their own private information. An appropriate solution concept for these game instances is SIB-PBE that was introduced in [36, 39] and presented in Section 5.3. Since cyber security games have asymmetric information, unobservable actions, and the domain of the agents' strategies' grows with time, the work of [35] along with the methodology and results reported in [36, 39] and informally presented in Section 5 is the literature that is the most relevant to these games.

Infinitely repeated games have been extensively studied, primarily by economists. There is a rich literature available on these games; the book by Mailath and Samuelson,[19], presents the main results on this topic until its publication date. In this chapter we briefly discuss this literature, because some of the ideas and philosophy behind the development of key results for this class of games played a significant role in the development of key results for dynamic games with asymmetric information where the underlying system is dynamic ([5, 15, 35]). Infinitely repeated games can be divided into two categories, zero-sum and non-zero sum.

Infinitely repeated zero-sum games with asymmetric (incomplete) information were initially studied by Aumann et al. ([2]); an excellent survey and discussion of results on this class of games can be found in [44].

Infinitely repeated non-zero sum games with asymmetric information can be classified into three categories: games with perfect public monitoring, in which the agents observe perfectly each others' actions, and Nash equilibrium or perfect equilibrium as a solution concept (see [20, 19] and references therein); games with imperfect public monitoring, in which the agents can observe public noisy signals about the action profile and focus on perfect equilibria where each agent's continuation strategy depends only on past public signals (see [8, 19] and references therein); and games with imperfect private monitoring, in which players observe private noisy signals about other players' actions, and sequential equilibrium as a solution concept (see [22] for two-player games and [34] for many-player games, and references therein). In all of the above categories the authors consider infinitely repeated discounted games and characterize the set of equilibrium payoffs corresponding to all equilibria in the limit as the discount factor approaches one.

7 Conclusion

We have argued that cyber security problems are stochastic dynamic games with asymmetric information where the underlying system is stochastic and dynamic. We presented current approaches to analyzing dynamic games with asymmetric information along with the currently available literature and the challenges/difficulties associated with these approaches. As we pointed out in Section 4, two major difficulties are the interdependence over time between strategy prediction and beliefs, and the increasing domain of the agents' strategies. We presented a "sufficient information approach" (section 5) which breaks the interdependence over time between strategy prediction and beliefs, leads to a sequential decomposition of the dynamic game and specifies an algorithm for determining the SIB-PBE of the game; we also identified instances of games where the sufficient information approach results in a time-invariant domain of the agents' strategies. The results of the sufficient information approach were developed under a key assumption, stated in Section 5.4, which in essence says that any deviation by one agent is either not detected or it is detected simultaneously by all other agents and the detection is based on the agents' common information.

In cyber security problems the domain of the agents' strategies increases, in general, with time. Furthermore, a deviation from one agent may not be detected at all, or it may be detected at different times by different agents. These two features of cyber security games cannot be captured by the approach presented in Section 5. In the rest of this section we present some ideas on how to address them, and we identify open problems in dynamic games with asymmetric information that are tightly connected to cyber security games.

First, consider the situation where the agents' sufficient private information increases with time. In this case assume that each agent has finite memory which he updates at each time instant; specifically, assume that at any time t part of each agent's memory is used to store his private information and another part is used to store his SIB belief about the system state and all the agents' (including himself) private information. At time $t+1$, each agent's private information is determined by an update rule which combines his private information at t and the new information he receives at $t+1$; similarly, the SIB belief at time $t+1$ is formed by an update rule which combines the SIB belief at t and the new common information received at $t+1$. Under these constraints, the objective is to determine decision strategies (that are based on the agents' private information and the SIB belief), private information update rules, and common information update rules that are in equilibrium.

Next, consider the situation where the key assumption of Section 5.4 is relaxed. In this case the challenge is to create public monitoring structures/mechanisms that allow each agent to detect deviations from other agents. Within the context of infinite horizon discounted games (with discount factors close to 1) such monitoring structures are presented (i) in [35] for games where the underlying system is dynamic and is described by a controlled Markov chain, the agents' actions are hidden (unobservable) and the agents' private state observations are imperfect (noisy), and (ii) in [34] for repeated games with an information structure similar to that of [35]. These monitoring mechanisms are described by "review phases" the duration of which is chosen appropriately so that at the end of each phase the law of large numbers should hold with high probability, therefore, allowing agents to detect each others' deviations (see [34]). Such ingenious monitoring structures work well for infinite horizon games but can not be used in finite horizon games. The discovery of monitoring structures that allow agents to detect each others' deviations in finite horizon games where the key assumption of Section 5.4 is relaxed and the information structure is similar to that of [35] is a challenging and important open problem that is closely connected to cyber security games.

To alleviate the difficulties arising when the key assumption of Section 5.4 is relaxed and public monitoring mechanisms are not in place we can focus on *belief-free equilibria.* An equilibrium is belief-free if, after each history profile, each agent's continuation strategy is optimal independently of his beliefs' of the other agents' history profiles. Game theorists have analyzed and solved repeated infinite horizon discounted games with private imperfect state information, observable actions, and belief-free equilibrium as the solution concept (see [4, 13, 14, 43] and references therein). The analysis and solution of games where the underlying system is dynamic, the agents' private state observations are imperfect (noisy), actions are hidden, and the solution concept is belief-free equilibrium, is an important class of open problems. Such problems are tightly connected to cyber security as they capture several important key features of cyber security games.

References

[1] R. Aumann. Agreeing to disagree. *The annals of statistics*, pages 1236–1239, 1976.

[2] R. Aumann, M. Maschler, and R. Stearns. *Repeated games with incomplete information.* MIT press, 1995.

[3] P. Cardaliaguet, C. Rainer, D. Rosenberg, and N. Vieille. Markov games with frequent actions and incomplete information-the limit case. *Mathematics of Operations Research*, 41(1):49–71, 2015.

[4] J. Ely, J. Hörner, and W. Olszewski. Belief-free equilibria in repeated games. *Econometrica*, 73(2):377–415, 2005.

[5] J. Escobar and J. Toikka. Efficiency in games with Markovian private information. *Econometrica*, 81(5):1887–1934, 2013.

[6] D. Etherington and K. Conger. Large DDos attacks cause outages at twitter, spotify, and other sites. *TechCrunch. Np*, 21, 2016.

[7] J. Finkle and D. Skariachan. Target cyber breach hits 40 million payment cards at holiday peak. accessed: 2016-09-09. [Online]. Available: http://www.reuters.com/article/us-target-breach-idUSBRE9BH1GX20131219.

[8] D. Fudenberg, D. Levine, and E. Maskin. The folk theorem with imperfect public information. *Econometrica (1986-1998)*, 62(5):997, 1994.

[9] D. Fudenberg and J. Tirole. Game theory. *Cambridge, Massachusetts*, 393(12):80, 1991.

[10] F. Gensbittel and J. Renault. The value of Markov chain games with incomplete information on both sides. *Mathematics of Operations Research*, 40(4):820–841, 2015.

[11] Andy Greenberg. Hackers remotely kill a jeep on the highway?with me in it. *Wired*. accessed: 2016-12-15. [Online]. Available: https://www.wired.com/2015/07/hackers-remotely-kill-jeep-highway.

[12] A. Gupta, A. Nayyar, C. Langbort, and T. Başar. Common information based Markov perfect equilibria for linear-Gaussian games with asymmetric information. *SIAM J. Control Optim.*, 52(5):3228–3260, 2014.

[13] J. Hörner and S. Lovo. Belief-free equilibria in games with incomplete information. *Econometrica*, 77(2):453–487, 2009.

[14] J. Hörner, S. Lovo, and T. Tomala. Belief-free equilibria in games with incomplete information: Characterization and existence. *Journal of Economic Theory*, 146(5):1770–1795, 2011.

[15] J. Hörner, T. Sugaya, S. Takahashi, and N. Vieille. Recursive methods in discounted stochastic games: An algorithm for $\delta \to 1$ and a folk theorem. *Econometrica*, 79(4):1277–1318, 2011.

[16] L. Li, C. Langbort, and J. Shamma. Solving two-player zero-sum repeated Bayesian games. *arXiv preprint arXiv:1703.01957*, 2017.

[17] L. Li and J. Shamma. Lp formulation of asymmetric zero-sum stochastic games. In *Decision and Control (CDC), 2014 IEEE 53rd Annual Conference on*, pages 1930–1935. IEEE, 2014.

[18] A. Mahajan and M. Mannan. Decentralized stochastic control. *Annals of Operations Research*, 241(1-2):109–126, 2016.

[19] G. Mailath and L. Samuelson. *Repeated Games and Reputations.* Oxford university press Oxford, 2006.

[20] E. Maskin and D. Fudenberg. The folk theorem in repeated games with discounting or with incomplete information. *Econometrica*, 53(3), 1986.

[21] E. Maskin and J. Tirole. Markov perfect equilibrium: I. observable actions. *Journal of Econonomic Theory*, 100(2):191–219, 2001.

[22] H. Matsushima. Repeated games with private monitoring: Two players. *Econometrica*, 72(3):823–852, 2004.

[23] E. Miehling, M. Rasouli, and D. Teneketzis. A pomdp approach to the dynamic defense of large-scale cyber networks. *IEEE Transactions on Information Forensics and Security*, 13(10):2490–2505, 2018.

[24] D. Miller. Robust collusion with private information. *The Review of Economic Studies*, 79(2):778–811, 2012.

[25] R. Myerson. *Game theory.* Harvard university press, 2013.

[26] A. Nayyar and A. Gupta. Information structures and values in zero-sum stochastic games. In *American Control Conference (ACC), 2017*, pages 3658–3663. IEEE, 2017.

[27] A. Nayyar, A. Gupta, C. Langbort, and T. Başar. Common information based Markov perfect equilibria for stochastic games with asymmetric information: Finite games. *IEEE Transactions on Automatic Control*, 59(3):555–570, March 2014.

[28] Department of Homeland Security. Industrial control systems cyber emergency response team (ICS-CERT). Available: https://ics-cert.us-cert.gov.

[29] M. Osborne and A. Rubinstein. *A course in game theory.* MIT press, 1994.

[30] Y. Ouyang, H. Tavafoghi, and D. Teneketzis. Dynamic oligopoly games with private Markovian dynamics. In *54th IEEE Conference on Decision and Control (CDC)*, 2015.

[31] Y. Ouyang, H. Tavafoghi, and D. Teneketzis. Dynamic games with asymmetric information: Common information based perfect bayesian equilibria and sequential decomposition. *IEEE Transactions on Automatic Control*, 62(1):222–237, 2017.

[32] J. Renault. The value of Markov chain games with lack of information on one side. *Mathematics of Operations Research*, 31(3):490–512, 2006.

[33] A. Sinha and A. Anastasopoulos. Structured perfect Bayesian equilibrium in infinite horizon dynamic games with asymmetric information. *American Control Conference*, 2016.

[34] T. Sugaya. Folk theorem in repeated games with private monitoring. working paper.

[35] T. Sugaya. Folk theorem in stochastic games with private state and private monitoring. working paper.

[36] H. Tavafoghi. *On design and analysis of cyber-physical systems with strategic agents.* PhD thesis, University of Michigan, 2017.

[37] H. Tavafoghi, Y. Ouyang, and D. Teneketzis. On stochastic dynamic games with delayed sharing information structure. In *Decision and Control (CDC), 2016 IEEE 55th Conference on*, pages 7002–7009. IEEE, 2016.

[38] H. Tavafoghi, Y. Ouyang, and D. Teneketzis. A unified approach to dynamic decision problems with asymmetric information-part i: Non-strategic agents. submitted to *IEEE Transactions on Automatic Control*, available on arXiv:1812.01130, 2018.

[39] H. Tavafoghi, Y. Ouyang, and D. Teneketzis. A unified approach to dynamic decision problems with asymmetric information-part ii: Strategic agents. submitted to *IEEE Transactions on Automatic Control*, available on arXiv:1812.01132, 2018.

[40] D. Vasal and A. Anastasopoulos. Signaling equilibria for dynamic LQG games with asymmetric information. In *55th IEEE Conference on Decision and Control (CDC)*, pages 6901–6908. IEEE, 2016.

[41] RB Washburn and D Teneketzis. Asymptotic agreement among communicating decisionmakers. *Stochastics: An International Journal of Probability and Stochastic Processes*, 13(1-2):103–129, 1984.

[42] J. Watson. Perfect Bayesian equilibrium: general definitions and illustrations. *working paper*, 2016.

[43] Y. Yamamoto. A limit characterization of belief-free equilibrium payoffs in repeated games. *Journal of Economic Theory*, 144(2):802–824, 2009.

[44] S. Zamir. Repeated games of incomplete information: Zero-sum. *Handbook of Game Theory*, 1:109–154, 1992.