A Dynamic Mechanism for Security Management in Multi-Agent Networked Systems

1st Shiva Navabi
Electrical and Computer Engineering Department
University of Southern California
Los Angeles, USA
navabiso@usc.edu

2nd Ashutosh Nayyar
Electrical and Computer Engineering Department
University of Southern California
Los Angeles, USA
ashutosn@usc.edu.

***Abstract*—We study the problem of designing a dynamic mechanism for security management in an interconnected multi-agent system with N strategic agents and one coordinator. The system is modeled as a network of N vertices. Each agent resides in one of the vertices of the network and has a privately known security state that describes its safety level at each time. The evolution of an agent's security state depends on its own state, the states of its neighbors in the network and on actions taken by a network coordinator. Each agent's utility at time instant t depends on its own state, the states of its neighbors in the network and on actions taken by a network coordinator. The objective of the network coordinator is to take security actions in order to maximize the long-term expected social surplus. Since agents are strategic and their security states are private information, the coordinator needs to incentivize agents to reveal their information. This results in a dynamic mechanism design problem for the coordinator. We leverage the inter-temporal correlations between the agents' security states to identify sufficient conditions under which an incentive compatible expected social surplus maximizing mechanism can be constructed. We then identify two special cases of our formulation and describe how the desired mechanism is constructed in these cases.**

***Index Terms*—dynamic mechanism design, interdependent valuations, social surplus maximization, incentive compatibility, network security, multi-agent systems.**

I. Introduction

Networked computing and communication systems are ubiquitous and widely integrated in numerous contexts including electronic commerce, telecommunications, smart grids, cloud-based platforms as well as emerging intelligent transportation systems. While the increasing prevalence of these technologies in various industries enhances efficiency and reliability in their operation, it also exposes the existing infrastructures to new types of security threats in the form of cyber attacks administered by various adversarial entities [1]. Therefore, in tandem with further integration of advanced computing and communication technologies into various establishments, sophisticated security machinery needs to be devised and deployed so as to provide sufficient protection against malicious infiltration.

In this work, we consider a networked cyber-physical system which houses N strategic agents that interact with each other over a discrete infinite time horizon. These agents can be viewed as the operators of computing units in a networked computing environment. Each agent resides in one of the N vertices of the network and has a security state that describes its safety level at each time. For instance, in a computer network the security state of each computing unit may indicate whether or not that computer is infected with a computer virus. The internal security of the network can be sabotaged through cyber-attacks attempted by exogenous adversarial entities. Due to the interconnections between agents in the network, once an agent's security is undermined its effects could spread throughout the network and cause damage to other agents' security as well. There is a network coordinator that has security resources and takes actions to maximize the expected social welfare of all the agents in the long run. In the context of computer networks, the coordinator can be viewed as the network manager who is in charge of preserving the long-range network-wide security. It may have some computer security resources such as a limited number of antivirus software licences, firewall technology with limited netwrok coverage, etc. In this context, the actions available to the coordinator can be its choices of the computers that it can select to invest its limited security resources in so as to improve the network security. The computers that are recipients of these security resources will be more likely to remain secure afterwards and/or recover from cyber attacks that have already happened. To choose the best security enhancing action, the coordinator needs the information about the security states of all the agents at each time. Each agent's security state at each time is *privately* observed by that agent only and is not known to others. The coordinator thus needs to *elicit* the information about the security states of the agents in order to select the optimal action. Since agents are self-interested and strategic, they may not truthfully reveal their private information if they believe that they can enhance their individual utility by misreporting their security state. In order to resolve the potential conflict of incentives that could arise between agents and the coordinator, we borrow tools and principles from the theory of dynamic mechanism design [2] to devise the rules of interaction in the induced dynamic game in a way that the objectives of all the participants get aligned in the emerging equilibrium.

This work was supported by NSF Grant ECCS 1750041 and ARO Award No. W911NF-17-1-0232.

A number of studies have adopted mechanism design principles to model, formulate and address problems in the context of networked systems security as surveyed in [3, Section 5]. Our work is inspired by the study of a network security mangement problem in [4]. Farhadi et al in [4] consider a specific interdependent valuation model and develop an incentive compatible expected social surplus maximizing mechanism for managing network security in a given setup. In the present work, we *generalize* the class of interdependent valuation models for which an incentive compatible expected social surplus maximizing mechanism can be obtained. The valuation model considered in [4] can be viewed as a *special case* of the broader family of the valuation models that is considered in our work. We employ the idea of forming cross-inference signals proposed in [4] to construct payment functions for a truthfully implementable mechanism. While the authors in [4] consider a *particular* cross-inference signal form to specify the desired mechanism, we provide more *general sufficient conditions* under which any proposed set of cross-inference signals can be used to construct an incentive compatible expected social surplus maximizing mechanism. Two special instances of our formulation are provided in this paper to demonstrate the construction of the desired mechanism.

The design of dynamic social surplus maximizing (efficient) mechanisms was undertaken in the well-known studies [5] and [6] for agents with independent *private* types. However, the design of dynamic efficient mechanisms in presence of *interdependencies* between the agents' valuations was first studied in [7] where existence of expected social surplus maximizing, budget balanced and incentive compatible mechanisms was established. The idea of exploiting the inter-temporal correlation between agents' states to design payment functions that would ensure satisfaction of incentive compatibility constraints, was also initially proposed in [7]. These ideas were applied in [4] to design an incentive compatible, ex-ante budget balanced and individually rational mechanism to maximize the long-term expected social surplus in a given network model under certain assumptions about the agents' utility models and state evolution dynamics of the network.

The rest of the paper is organized as follows: we discuss the problem formulation and the mechanism setup in Section II. In Section III, we construct an incentive compatible expected social surplus maximizing mechanism under a generic setup. In Sections IV and V we describe the construction of the desired mechanism for two special classes of problems. Section VI discusses conditions for the satisfaction of ex-ante individual rationality and budget balance constraints for an incentive compatible mechanism. We summarize our findings and briefly point out potential extensions to the current framework in Section VII.

Notations: $\mathbf{1}_{1\times k}$ is the k dimensional all ones vector. $\mathbb{Z}^+$ is the set of positive integers. For a set $\mathcal{A}$, $|\mathcal{A}|$ denotes the cardinality of $\mathcal{A}$. $\mathbb{1}_{\{\cdot\}}$ is the indicator function that equals 1 if the statement in its subscript is true and is 0 otherwise. $\mathbb{E}$ denotes the expectation operator. For a random variable/random vector θ, $\mathbb{E}_\theta$ denotes that the expectation is with respect to the probability distribution of θ. $\mathbb{P}$ denotes the probability measure. $x_{1:n}$ and $y^{1:m}$ are shorthands for vectors $(x_1,\dots,x_n)$ and $(y^1,\dots,y^m)$, respectively. For the vector $y^{1:m}$, y^{-j} is the shorthand for $(y^1,\dots,y^{j-1},y^{j+1},\dots,y^m)$.

II. Problem Formulation

Consider N strategic agents that interact with each other through an interconnected network over time steps $t\in\mathbb{T}:=\{0,1,2,\dots\}$. Let the graph $\mathcal{G}=(\mathcal{N},\mathcal{E})$ denote the network with $\mathcal{N}:=\{1,2,\dots,N\}$ denoting the set of vertices and $\mathcal{E}$ denoting the set of (directed) edges in this graph. Each agent resides in one of the vertices in $\mathcal{G}$. Let $\mathcal{N}^i:=\{j\in\mathcal{N}: j\neq i\ ,\ (j,i)\in\mathcal{E}\}$ denote the set of *input* neighbors of agent i. Let $\mathcal{O}^i:=\{j\in\mathcal{N}: j\neq i\ ,\ (i,j)\in\mathcal{E}\}$ denote the set of *output* neighbors of agent i.

In the context of computer networks, presence of an edge (i,j) in the network graph can represent a communication link from node i to node j through which node i can send data to node j. Consequently, security of the computer at node j can be impacted by the security of the computer at node i.

A. System Dynamics

Let S_t^i denote the security state of agent i at time t. The realization of S_t^i is privately observed by agent i only and is unknown to others. S_t^i takes values in a discrete set $\mathcal{S}$. For example $\mathcal{S}$ could be the set $\{0,1\}$, where 0 and 1 represent "unsafe" and "safe" security states, respectively. The network has a central coordinator[1] that takes an action A_t at time t. Actions are chosen from a finite set of admissible actions denoted by $\mathcal{A}$. For instance, consider $\mathcal{A}=\mathcal{N}$, where action $A_t=i$ denotes that the coordinator is applying a security measure at node i in the network at time t. Let $S_t^{1:N}:=(S_t^1,\dots,S_t^N)$ denote the network state profile at time t. It evolves according to the following Markovian dynamics:

$$\begin{aligned}&\mathbb{P}(S_{t+1}^{1:N}=s_{t+1}^{1:N}\mid S_t^{1:N},A_t) \qquad (1)\\ &=\prod_{i=1}^{N}\mathbb{P}(S_{t+1}^i=s_{t+1}^i\mid S_t^i,\{S_t^j\}_{j\in\mathcal{N}^i},A_t)\ ,\ s_{t+1}^{1:N}\in\mathcal{S}^N.\end{aligned}$$

Thus, the probability distribution of an agent's state at time $t+1$ depends on its own state, the states of its input neighbors and the coordinator's action at time t. We also assume that the states evolution dynamics are time-invariant, that is for any $t,t'\in\mathbb{T}, x^{1:N},y^{1:N}\in\mathcal{S}^N$ and $a\in\mathcal{A}$,

$$\begin{aligned}&\mathbb{P}(S_{t+1}^{1:N}=x^{1:N}\mid S_t^{1:N}=y^{1:N},A_t=a)\\ &=\mathbb{P}(S_{t'+1}^{1:N}=x^{1:N}\mid S_{t'}^{1:N}=y^{1:N},A_{t'}=a). \qquad (2)\end{aligned}$$

B. Consumer Valuation and Utility Models

An agent's valuation at time t is a function of its own state, the states of its input neighbors and the action taken by the coordinator. Agent i's valuation at time t is given as:

$$v^i(S_t^{1:N},A_t)=\sum_{j\in\mathcal{N}^i}\alpha_i^j(S_t^i,A_t)\ S_t^j\ +\ \beta^i(S_t^i,A_t), \qquad (3)$$

[1]By "central" we mean that the coordinator has some network-wide execution power.

where $\alpha_i^j(\cdot,\cdot)$ and $\beta^i(\cdot,\cdot)$ are publicly known functions. The term $\sum_{j\in\mathcal{N}^i} \alpha_i^j(S_t^i, A_t)\ S_t^j$ is a weighted sum of the states of agent i's input neighbors at time t where the weights depend on agent i's current state S_t^i and the coordinator's action A_t. The term $\beta^i(S_t^i, A_t)$ in (3) reflects the intrinsic benefit to agent i when action A_t is taken and its state is S_t^i. Note that agent i's valuation at time t is affine in its input neighbors' states but can depend on its own state and coordinator's action in a non-linear way through the functions $\alpha_i^j(\cdot,\cdot)$ and $\beta^i(\cdot,\cdot)$.

Assumption 1. *We assume that the valuation functions in (3) are uniformly bounded, i.e., there is a positive number $B < \infty$ such that:*

$$\max_{s^{1:N}\in\mathcal{S}^N, a\in\mathcal{A}} |v^i(s^{1:N}, a)| < B\ ,\ \forall i \in \mathcal{N}. \tag{4}$$

The valuation model in (3) implies that the agents' valuations are *interdependent*: each agent's utility is directly influenced by its input neighbors' states as well as its own state.

Remark 1. *The valuation model considered in [4] (Equations (3)-(4) in [4]) is a special case of the family of models in (3) that results by setting $\alpha_i^j(S_t^i, A_t) = \alpha\, \mathbb{1}_{\{S_t^i=1 \text{ or } A_t=i\}} \frac{l_{ji}}{\sum_{k\in\mathcal{N}^i} l_{ki}}$ and $\beta^i(S_t^i, A_t) = S_t^i$, under $\mathcal{S} = \{0,1\}$ and $\mathcal{A} = \mathcal{N}$. As defined in [4], $l_{ji} \in (0,1]$ denotes the probability of attacks spreading from j to i and $0 < \alpha < 1$ indicates value of a safe neighborhood to each agent residing in the network. We refer the reader to [4, Section 2] for precise definition of the parameters as well as interpretation of the specific model in the setup studied therein.*

Let p_t^i denote the payment charged to agent i at time t. Then, agent i's net utility at time t is:

$$u^i(S_t^{1:N}, A_t, p_t^i) = v^i(S_t^{1:N}, A_t) - p_t^i. \tag{5}$$

We assume that all agents discount the future with a common discount factor $\delta \in (0,1)$. The total utility of agent i across the entire time horizon is given by:

$$U^i := (1-\delta)\sum_{t=0}^{\infty} \delta^t u^i(S_t^{1:N}, A_t, p_t^i). \tag{6}$$

The coordinator's objective is to maximize the expected social surplus $\mathbb{E}\Big\{ (1-\delta)\sum_{t=0}^{\infty}\delta^t \Big[\sum_{i=1}^{N} v^i(S_t^{1:N}, A_t) - c(A_t)\Big] \Big\}$ where, $c(\cdot)$ is a bounded function such that $c(a)$ denotes the cost incurred due to taking action $a \in \mathcal{A}$.

We assume that the network configuration specified by $\mathcal{G} = (\mathcal{N}, \mathcal{E})$, the network state evolution dynamics in (1)-(2) and the probability distribution of the network state profile at time $t = 0$, i.e., $\pi_0(s_0^{1:N}) := \mathbb{P}(S_0^{1:N} = s_0^{1:N}), \forall s_0^{1:N} \in \mathcal{S}^N$ are common knowledge. The security states of the agents however, are their private information. In order to maximize its objective, the coordinator needs to incentivize the agents to reveal their private state information at each time. Therefore, the coordinator's problem can be posed as a mechanism design problem. We describe this problem in the following sections.

C. Mechanism Setup

We consider a dynamic direct mechanism where at each time step $t \in \mathbb{T}$, each agent reports a state from the set $\mathcal{S}$ to the coordinator. Let s_t^i and r_t^i denote agent i's true and reported states at time t, respectively. Each agent can misreport its state and thus, r_t^i may be different from s_t^i.[2] Define $h_t := \{r_{0:t-1}^{1:N}, a_{0:t-1}\}$, where, $r_t^{1:N}$ denotes the reported state profile at time t and a_t is the action taken at time t. We call h_t the public history at time t. Let $\mathcal{H}_t$ denote the set of all possible values of h_t. Let h_t^i denote all the information that is known to agent i at time t *prior to* observing S_t^i. Thus, $h_t^i := \{h_t, s_{0:t-1}^i\}$. Let $\mathcal{H}_t^i$ denote the set of all possible values of h_t^i.

A mechanism needs to specify actions and payments at each time t. Such a mechanism consists of a sequence of action functions $q_t, t \in \mathbb{T}$ and a sequence of payment functions $p_t, t \in \mathbb{T}$. For any $h_t \in \mathcal{H}_t$, and any reported state profile $r_t^{1:N} \in \mathcal{S}^N$ at time t, $q_t(h_t, r_t^{1:N}) \in \mathcal{A}$ and $p_t(h_t, r_t^{1:N}) \in \mathbb{R}^N$. $q_t(h_t, r_t^{1:N})$ gives the action to be taken at time t and the ith component of the payment vector $p_t(h_t, r_t^{1:N})$ denoted by $p_t^i(h_t, r_t^{1:N})$ gives the payment charged to agent i at time t.[3]

D. Incentive Compatibility

In a dynamic incentive compatible (IC) mechanism, truthful reporting of private information (security states at each time in our setup) constitutes an equilibrium of the dynamic game induced by the mechanism. Here we adopt the periodic ex post notion of incentive compatibility defined in [5], [6]. A dynamic direct mechanism $(q_t, p_t)_{t\in\mathbb{T}}$ is *periodic ex post incentive compatible* (p-EPIC) if at each time t every agent i would prefer to report its true state regardless of its private history h_t^i and the states of other agents at that time (s_t^{-i}), provided that all other agents adopt truth-telling strategy. (p-EPIC) constraint is thus described as follows

$$\begin{aligned}
&\mathbb{E}\Big[\ (1-\delta)\sum_{t'=t}^{\infty}\delta^{t'-t}\times \\
&u^i\Big(S_{t'}^{1:N}, q_{t'}(H_{t'}, S_{t'}^{1:N}), p_{t'}^i(H_{t'}, S_{t'}^{1:N})\Big) \mid h_t^i, s_t^{1:N}\ \Big] \\
&\geq \mathbb{E}\Big[\ (1-\delta)\sum_{t'=t}^{\infty}\delta^{t'-t}\times \\
&u^i\Big(S_{t'}^{1:N}, q_{t'}(H_{t'}, S_{t'}^{-i}, \sigma_{t'}^i(H_{t'}^i, S_{t'}^i)), \\
&p_{t'}^i(H_{t'}, S_{t'}^{-i}, \sigma_{t'}^i(H_{t'}^i, S_{t'}^i))\Big) \mid h_t^i, s_t^{1:N}\ \Big], \\
&\forall s_t^{1:N} \in \mathcal{S}^N\ ,\ \forall h_t^i \in \mathcal{H}_t^i\ ,\ \forall \sigma_t^i\ ,\ \forall i \in \mathcal{N}\ ,\ \forall t \in \mathbb{T},
\end{aligned} \tag{7}$$

where, $\sigma_{t'}^i : \mathcal{H}_{t'}^i \times \mathcal{S} \longrightarrow \mathcal{S}$ denotes agent i's reporting strategy at time t', $H_{t'}$ and $H_{t'}^i$ are random vectors describing the public history and agent i's private history at time t'.

[2]For example in a computer network, a strategic operator may down-report its computer security level to selfishly enhance its likelihood of receiving the antivirus software license.

[3]Note that the payment functions can output both positive and negative values, indicating taxes and subsidies, respectively.

E. Coordinator's Problem Formulation

The expected social surplus under the IC mechanism $(q_t, p_t)_{t\in\mathbb{T}}$ is $\mathbb{E}\Big\{ (1-\delta)\sum_{t=0}^{\infty}\delta^t\Big[\sum_{i=1}^{N} v^i\Big(S_t^{1:N}, q_t(H_t, S_t^{1:N})\Big) - c\Big(q_t(H_t, S_t^{1:N})\Big)\Big] \Big\}$ when all the agents adopt the truthful strategy.[4] The expected social surplus maximization problem can thus be formulated as

$$\max_{(q_t,p_t)_{t\in\mathbb{T}}} \mathbb{E}\Big\{ (1-\delta)\sum_{t=0}^{\infty}\delta^t\Big[\sum_{i=1}^{N} v^i\Big(S_t^{1:N}, q_t(H_t, S_t^{1:N})\Big) - c\Big(q_t(H_t, S_t^{1:N})\Big)\Big] \Big\} \text{ subject to } (7). \tag{8}$$

When agents report their states truthfully, the best action policy is given by the solution of an infinite-horizon MDP with state dynamics given by (1) and (2) and the total reward being the social surplus. We formalize this observation in the following lemma.

Lemma 1. *Under an incentive compatible mechanism the optimal action policy is a stationary Markov policy $\rho^* : \mathcal{S}^N \longrightarrow \mathcal{A}$ that selects a maximizing action in the following Bellman equation [8]:*

$$V(s^{1:N}) = \max_{a\in\mathcal{A}} \Big\{ \sum_{i=1}^{N} v^i(s^{1:N}, a) - c(a) + \delta\, \mathbb{E}_{S^{1:N}}[V(S^{1:N}) \mid s^{1:N}, a] \Big\}, \forall s^{1:N}\in\mathcal{S}^N. \tag{9}$$

Proof. The proof follows from classical results for infinite-horizon discounted MDPs [9]. □

Note that there exist well-known solution methods for solving Bellman equations such as linear programming methods [8, Chapter 6]. It thus remains to construct payment functions $p_t, t\in\mathbb{T}$, that would guarantee satisfaction of the IC constraints in (7).

III. Incentive Compatible Expected Social Surplus Maximizing Mechanism

In this section we borrow the approach of using cross-inference signals proposed in [4] to construct payment functions that align agents' incentives with those of the coordinator's. We exploit the correlation between each agent i's state at time t (S_t^i) and its output neighbors' states at time $t+1$ ($S_{t+1}^j, j\in\mathcal{O}^i$) to form the *cross-inference signal* m_t^i from which S_t^i can be *inferred* at time $t+1$. We then propose payment functions of the form $p_{t+1}^i(m_t^i, r_t^{-i}, a_t)$, $i\in\mathcal{N}$ that determine agent i's payments based on the cross-inference signal m_t^i rather than agent i's report r_t^i. We provide sufficient conditions for the cross-inference signals that ensure incentive compatibility. We elaborate on our approach below.

At each time t, m_t^i is defined as

$$m_t^i := f^i(\{r_{t+1}^j\}_{j\in\mathcal{O}^i}, r_t^{-i}, a_t), \tag{10}$$

[4]Note that truthful reporting equilibrium may not be the unique equilibrium of the induced game.

where, $f^i(\cdot), i\in\mathcal{N}$ are some time-invariant mappings that are publicly known. m_t^i as defined in (10) can be viewed as a proxy signal that informs the coordinator about the security state of agent i at time t based on the reports of agent i's output neighbors at time $t+1$ ($\{r_{t+1}^j\}_{j\in\mathcal{O}^i}$) as well as the coordinator's action (a_t) and other agents' reports (r_t^{-i}) at time t.

The following theorem describes conditions under which cross-inference signals can be used to construct IC expected social surplus maximizing mechanisms.

Theorem 1. *Suppose the network $\mathcal{G} = (\mathcal{N}, \mathcal{E})$ is such that $\mathcal{O}^i \neq \emptyset, \forall i\in\mathcal{N}$. Suppose the cross-inference signals in (10) are such that*

$$\mathbb{E}[m_t^i \mid s_t^{1:N}, a_t] = s_t^i\ , \ \forall s_t^{1:N}\in\mathcal{S}^N, \forall a_t\in\mathcal{A}, \forall i\in\mathcal{N}, \tag{11}$$

where the expectation is taken assuming that all agents except i truthfully report their states.

Then, the payment functions and action policies defined below constitute an expected social surplus maximizing incentive compatible mechanism:

$$q_t(h_t, r_t^{1:N}) := \rho^*(r_t^{1:N}) \ \text{ for all } \ h_t, r_t^{1:N}, t, \tag{12}$$

$$p_{t+1}^i(m_t^i, r_t^{-i}, a_t) := -\frac{1}{\delta}\Big[\sum_{j\neq i} v^j(r_t^{-i}, m_t^i, a_t) - c(a_t)\Big],\quad i\in\mathcal{N},\ t\in\mathbb{T}, \tag{13}$$

$$p_0^i := \gamma^i\ ,\ i\in\mathcal{N}, \tag{14}$$

where, $\rho^*(\cdot)$ is the stationary Markov policy characterized in Lemma 1 and γ^i is a constant participation fee charged to agent i at time $t = 0$.

Proof. See Appendix A. □

The key implication of Theorem 1 is the following: if we can construct cross-inference signals of the form in (10) such that (11) holds, then Theorem 1 provides an optimal mechanism for the coordinator (given by (12)-(14)). Thus, the coordinator's mechanism design problem has been reduced to the problem of constructing appropriate cross-inference signals.

Using (10) in (11) and evaluating the expectation on the left-hand side results in the following equations:

$$\sum_{\{y^j\}_{j\in\mathcal{O}^i}\in\mathcal{S}^{|\mathcal{O}^i|}} f^i(\{y^j\}_{j\in\mathcal{O}^i}, s^{-i}, a)\times \mathbb{P}\Big(\{S_{t+1}^j\}_{j\in\mathcal{O}^i} = \{y^j\}_{j\in\mathcal{O}^i} \mid S_t^{-i} = s^{-i}, S_t^i = s^i, A_t = a\Big) = s^i\ ,\quad s^i\in\mathcal{S}. \tag{15}$$

Note that for each $(s^{-i}, a)\in\mathcal{S}^{N-1}\times\mathcal{A}$, (15) is a linear system of equations in the variables $f^i(\cdot, s^{-i}, a)$. Thus, if these linear systems of equations have a solution, then the solution gives us the desired cross-inference signal. The cross-inference signals computed as such can then be used to determine the payments through the payment form in (13).

We next consider two special families of states evolution dynamics under which the linear systems of equations described by (15) have solutions. Therefore, for these special families of problems, an optimal mechanism is given by Theorem 1.

IV. Binary Security States

In this section we restrict our attention to the case where each agent's security state can take only two possible values: *safe* state, indicated by 1 and *unsafe* state, indicated by 0; that is $\mathcal{S} = \{0, 1\}$. In the following lemma we identify a condition on the state evolution dynamics under which an IC expected social surplus maximizing mechanism can be obtained.

Lemma 2. *Consider the network $\mathcal{G} = (\mathcal{N}, \mathcal{E})$ in which $\mathcal{O}^i \neq \emptyset, \forall i \in \mathcal{N}$. Suppose for each agent $i \in \mathcal{N}$ and for each profile $s_t^{-i} \in \{0,1\}^{N-1}$ and each action $a_t \in \mathcal{A}$, there exists some profile $\{s_{t+1}^j\}_{j\in\mathcal{O}^i}$ for which the following condition holds true:*

$$\mathbb{P}\Big(\{S_{t+1}^j\}_{j\in\mathcal{O}^i} = \{s_{t+1}^j\}_{j\in\mathcal{O}^i} \mid s_t^{-i}, S_t^i = 0, a_t\Big) \neq \mathbb{P}\Big(\{S_{t+1}^j\}_{j\in\mathcal{O}^i} = \{s_{t+1}^j\}_{j\in\mathcal{O}^i} \mid s_t^{-i}, S_t^i = 1, a_t\Big), \quad (16)$$

that is, $\mathbb{P}\Big(\{S_{t+1}^j\}_{j\in\mathcal{O}^i} \mid s_t^{-i}, S_t^i = 0, a_t\Big)$ and $\mathbb{P}\Big(\{S_{t+1}^j\}_{j\in\mathcal{O}^i} \mid s_t^{-i}, S_t^i = 1, a_t\Big)$ are two distinct probability mass functions (PMFs). Then, there always exist inference signals $m_t^i = f^i(\{r_{t+1}^j\}_{j\in\mathcal{O}^i}, r_t^{-i}, A_t)$ such that Equation (11) of Theorem 1 is satisfied.

Proof. See Appendix C. □

Example 1. *Consider a network $\mathcal{G} = (\mathcal{N}, \mathcal{E})$ in which $\mathcal{O}^i \neq \emptyset, \forall i \in \mathcal{N}$, $\mathcal{S} = \{0, 1\}$ and $\mathcal{A} := \mathcal{N}$. The state dynamics are given as:*

$$\mathbb{P}(S_{t+1}^i = 1 | S_t^i, \{S_t^j\}_{j\in\mathcal{N}^i}, A_t) = \begin{cases} \tilde{h}(1-d_i) \prod_{j\in\mathcal{N}^i: S_t^j = 0} (1 - l_{ji}) & \text{if } S_t^i = 0\ ,\ A_t \neq i \\ h(1-d_i(1-h)) \prod_{j\in\mathcal{N}^i: S_t^j = 0} (1 - l_{ji}) & \text{if } S_t^i = 0\ ,\ A_t = i \\ (1-d_i) \prod_{j\in\mathcal{N}^i: S_t^j = 0} (1 - l_{ji}) & \text{if } S_t^i = 1\ ,\ A_t \neq i \\ (1-d_i(1-h)) \prod_{j\in\mathcal{N}^i: S_t^j = 0} (1 - l_{ji}) & \text{if } S_t^i = 1\ ,\ A_t = i \end{cases} \quad (17)$$

and $\mathbb{P}(S_{t+1}^i = 0 | S_t^i, \{S_t^j\}_{j\in\mathcal{N}^i}, A_t) = 1 - \mathbb{P}(S_{t+1}^i = 1 | S_t^i, \{S_t^j\}_{j\in\mathcal{N}^i}, A_t)$. The above dynamics model a network where each agent's security state can change due to direct external attacks or through indirect attacks from unsafe neighbors. The action $A_t = i$ denotes that the coordinator is applying a security measure at node i in the network. The parameters of the model are explained below.

1) *$h \in (0, 1)$ is the probability with which an agent's safety can be restored when it is in unsafe state and is the recipient of the security measure. h is also the probability with which an agent's safety is preserved after it is subject to an external attack, provided that the security measure is applied to this agent at the time the attack takes place. These objectives are achieved independently for an agent that is the recipient of the security measure at a given time.*
2) *$d_i \in (0, 1)$ is the probability of agent i being the target of external attack when it is in a safe state.*
3) *$l_{ji} \in (0, 1)$ denotes the probability of attacks spreading from agent $j \in \mathcal{N}^i$ to agent i.*
4) *$\tilde{h} \in (0, h)$ is the probability with which an agent can restore its safety when it is in unsafe state and the security measure is not applied to it.*

The above model is similar to the network security model described in [4]. The key difference in our model and the model in [4] is that if agent i is in unsafe state at time t, i.e., $S_t^i = 0$, and does not receive the security measure at that time, i.e., $A_t \neq i$, under the dynamics in (17) it can still restore its safety with probability $\tilde{h} \in (0, h)$ through some internal protection mechanism that each agent is endowed with.

We observe that under the dynamics in (17) the following inequalities hold true:

$$\mathbb{P}(S_{t+1}^j = 1 | s_t^{-i}, S_t^i = 1, a_t) > \mathbb{P}(S_{t+1}^j = 1 | s_t^{-i}, S_t^i = 0, a_t)\ , \forall s_t^{-i} \in \mathcal{S}^{N-1}, \forall a_t \in \mathcal{A},\ \forall j \in \mathcal{O}^i.$$

Considering some profile $s_t^{-i} \in \mathcal{S}^{N-1}$ and some action $a_t \in \mathcal{A}$, from the above inequalities it follows that

$$\begin{aligned} &\mathbb{P}\Big(\{S_{t+1}^j\}_{j\in\mathcal{O}^i} = \mathbf{1}_{1\times|\mathcal{O}^i|} \mid s_t^{-i}, S_t^i = 1, a_t\Big) \\ &= \prod_{j\in\mathcal{O}^i} \mathbb{P}(S_{t+1}^j = 1 | s_t^{-i}, S_t^i = 1, a_t) \\ &> \prod_{j\in\mathcal{O}^i} \mathbb{P}(S_{t+1}^j = 1 | s_t^{-i}, S_t^i = 0, a_t) \\ &= \mathbb{P}\Big(\{S_{t+1}^j\}_{j\in\mathcal{O}^i} = \mathbf{1}_{1\times|\mathcal{O}^i|} \mid s_t^{-i}, S_t^i = 0, a_t\Big). \end{aligned}$$

Since the profile s_t^{-i} and the action a_t were picked arbitrarily, this observation implies that 'the PMFs $\mathbb{P}(\{S_{t+1}^j\}_{j\in\mathcal{O}^i} \mid s_t^{-i}, S_t^i = 0, a_t)$ and $\mathbb{P}(\{S_{t+1}^j\}_{j\in\mathcal{O}^i} \mid s_t^{-i}, S_t^i = 1, a_t)$ are indeed distinct for all $s_t^{-i} \in \mathcal{S}^{N-1}, a_t \in \mathcal{A}$. Hence, the dynamics in (17) satisfy the condition in (16). Therefore, based on the results of Lemma 2 and Theorem 1 we conclude that under the dynamics in (17) an IC expected social surplus maximizing mechanism can be obtained. In particular, for each agent i and each profile $(s^{-i}, a) \in \{0,1\}^{N-1} \times \mathcal{A}$ the cross-inference signals $f^i(\cdot, s^{-i}, a)$ are given by the solution to the systems of linear equations in (15). The cross-inference signals computed as such can then be used in Theorem 1 to obtain an IC expected social surplus maximizing mechanism (as specified by (12)-(14)).

V. Affine Dynamics Model

In this section we consider the case where $\mathcal{S} \subset \mathbb{Z}$ and the state dynamics are an affine function of the neighbors' states. Positive and negative state values indicate safety and insecurity, respectively. Larger positive numbers indicate increasingly safer states while smaller negative numbers are indicators of increasingly poorer security states. Zero suggests a neutral security state. The security state of agent i evolves according to the following dynamics:

$$S_{t+1}^i = \sum_{l\in\mathcal{N}^i} w_l^i(S_t^i, A_t)S_t^l + \mu^i(S_t^i, A_t) + D^i(S_t^i, A_t), \quad (18)$$

where A_t is the action taken at time t. The term $\sum_{l\in\mathcal{N}^i} w_l^i(S_t^i, A_t)S_t^l$ is a weighted sum of the states of agent i's input neighbors at time t where the weights depend on agent i's current state S_t^i and the coordinator's action A_t. The weights $w_l^i(\cdot,\cdot)$ are integer-valued quantities. We assume that $w_l^i(\cdot,\cdot) \neq 0$ for all $l \in \mathcal{N}^i$. $D^i(S_t^i, A_t)$ is the disturbance term that takes values in the set $\{0, \pm 1, \pm 2, \dots, \pm m\}, m \in \mathbb{Z}^+$ according to the PMF $\bar{g}(S_t^i, A_t)$ such that $\mathbb{E}[D^i(s_t^i, a_t)] = 0$ for all $s_t^i \in \mathcal{S}, a_t \in \mathcal{A}$. We can see that under the dynamics specified in (18), agent i's state at time $t+1$ is in *affine* relation to the states of its input neighbors at time t but can depend on its own state and the action at time t in an arbitrary fashion.

In the following lemma, we construct cross-inference signals of the form $m_t^i := f^i(\{r_{t+1}^j\}_{j\in\mathcal{O}^i}, r_t^{-i}, a_t)$ which can be used in Theorem 1 to construct an IC expected social surplus maximizing mechanism.

Lemma 3. *For the affine dynamics model in (18), define the cross-inference signal m_t^i as follows:*

$$m_t^i := \frac{1}{|\mathcal{O}^i|} \sum_{j\in\mathcal{O}^i} \frac{1}{w_i^j(r_t^j, a_t)} \times \Big(r_{t+1}^j - \sum_{\substack{l\in\mathcal{N}^j \\ l\neq i}} w_l^j(r_t^j, a_t)\, r_t^l - \mu^j(r_t^j, a_t) \Big), i \in \mathcal{N}. \quad (19)$$

Then, the above cross-inference signals satisfy Equation (11) of Theorem 1.

Proof. The proof is omitted due to space limitation. □

With the cross-inference signals defined in Lemma 3, Theorem 1 provides the IC expected social surplus maximizing mechanism.

VI. Participation and Budget Constraints

In this section we provide further conditions under which a dynamic mechanism satisfies individual rationality (IR) and budget balance (BB) constraints.

Assumption 2. *We make the following assumptions:*

(i) Once each agent decides whether or not they want to participate in the mechanism at time $t = 0$, they remain committed to that decision and will not revise it in future.

(ii) If any of the agents opts out of the mechanism at time $t = 0$, the coordinator will default to a neutral action $\bar{a} \in \mathcal{A}$ from time $t = 0$ onward, which indicates the absence of any regulatory efforts exerted by the coordinator to preserve or enhance the network security state. Under this mode of the mechanism's operation, no security resources are allocated to the agents and no payment is charged to any of the agents across the entire time horizon.

A. Individual Rationality

Individual rationality (IR) constraint ensures that each agent's expected total utility at the truthful reporting equilibrium weakly exceeds the expected reservation utility that this agent would obtain if it unilaterally opts out of the mechanism. The ex-ante[5] IR constraint can be described as follows:

$$\mathbb{E}\Big[(1-\delta)\sum_{t=0}^{\infty} \delta^t u^i\Big(S_t^{1:N}, q_t(H_t, S_t^{1:N}), p_t^i(H_t, S_t^{1:N})\Big) \Big] \geq \mathbb{E}\Big[(1-\delta)\sum_{t=0}^{\infty} \delta^t u^i\Big(S_t^{1:N}, \bar{a}, 0\Big) \Big], \forall i \in \mathcal{N}, \quad (20)$$

where the expression on the right-hand side denotes the expected total utility that agent i would obtain when the mechanism operates under the neutral action $\bar{a}$ after agent i opts out at time $t = 0$. We say that a mechanism $(q_t, p_t)_{t\in\mathbb{T}}$ is individually rational if it satisfies the ex-ante IR constraints in (20).

B. Budget Balance

Budget balance (BB) constraint guarantees that at the truthful reporting equilibrium, the expected value of all the monetary payments collected from the agents minus the expected value of all the monetary subsidies paid to the agents across the entire time horizon is equal to zero. Ex-ante budget balance constraint would thus ensure that at the truthful reporting equilibrium, in expectation, the monetary transfers get accumulated neither on the agents' side nor on the coordinator's side, but rather are in constant circulation to ensure proper functioning of the entire network. Ex-ante budget balance constraint is expressed as follows:

$$\mathbb{E}\Big[(1-\delta) \sum_{t=0}^{\infty} \delta^t \sum_{i=1}^{N} p_t^i(H_t, S_t^{1:N}) \Big] = 0. \quad (21)$$

We say that a mechanism $(q_t, p_t)_{t\in\mathbb{T}}$ is budget balanced if it satisfies the ex-ante BB constraint in (21).

In the following lemmas we provide sufficient conditions that an IC mechanism needs to satisfy in order to respect the BB and IR constraints in (21) and (20), respectively.

Lemma 4. *Consider the IC mechanism $(\rho^*, p_t)_{t\in\mathbb{T}}$ where ρ^* is the stationary Markov policy characterized in Lemma 1 and payment functions $p_t, t \in \mathbb{T}$, are of the form $p_{t+1}^i(m_t^i, r_t^{-i}, a_t)$, $i \in \mathcal{N}$, with m_t^i being as defined in* (10)*. Suppose the participation fees charged to the agents at time $t = 0$ are of the following form:*

$$p_0^i := -\mathbb{E}\Big[\sum_{t=0}^{\infty} \delta^{t+1}\, p_{t+1}^i(m_t^i, S_t^{-i}, a_t) \Big], \; i \in \mathcal{N}. \quad (22)$$

where, $a_t := \rho^*(S_t^{1:N})$ is the action taken at time t and m_t^i is the inference signal defined in (10)*. The expectation in (22) is taken with respect to the evolution dynamics of the network*

[5]For the distinction between the various phases (*ex-ante*, *interim* and *ex-post*) in the timeline of a game induced under a given mechanism, we refer the reader to [10, Section 3.2.2]

state profile given in (1)-(2) and the initial distribution of the network state profile, i.e., $\pi_0(\cdot)$.

Then the mechanism $(\rho^*, p_t)_{t\in\mathbb{T}}$ *is budget balanced.*

Proof. See Appendix D. □

In order to guarantee satisfaction of the IR constraints in (20) we make the following assumption.

Assumption 3. *We assume that the valuation functions in (3) are such that*

$$\mathbb{E}\Big[\sum_{t=0}^{\infty}\delta^t\ v^i(S_t^{1:N}, \rho^*(S_t^{1:N}))\Big] \geq \mathbb{E}\Big[\ \sum_{t=0}^{\infty}\delta^t v^i\Big(S_t^{1:N}, \bar{a}\Big)\ \Big],$$
$$\forall t\ ,\ \forall i\in\mathcal{N}, \tag{23}$$

where, ρ^* *is the expected social surplus maximizing action policy characterized in Lemma 1.*

The above assumption essentially means that in expectation, each agent's accumulated utilities over the entire time horizon when the coordinator takes actions given by the expected social surplus maximizing policy ρ^* (see Lemma 1), exceeds that under the coordinator's inaction ($\bar{a}$).

The following lemma provides sufficient conditions that an IC mechanism needs to satisfy in order to respect the IR constraints in (20).

Lemma 5. *Consider the IC mechanism* $(\rho^*, p_t)_{t\in\mathbb{T}}$ *where* ρ^* *is the stationary Markov policy characterized in Lemma 1 and payment functions* $p_t, t\in\mathbb{T}$*, are of the form* $p^i_{t+1}(m^i_t, r^{-i}_t, a_t)$*,* $i\in\mathcal{N}$*, with* m^i_t *being as defined in* (10)*. Suppose the participation fees charged to the agents at time* $t=0$ *are of the form given in (22). Then, under Assumption 3 the mechanism* $(\rho^*, p_t)_{t\in\mathbb{T}}$ *is individually rational.*

Proof. See Appendix E. □

VII. Conclusion

The problem of designing a dynamic mechanism for security enhancement in an interconnected multi-agent system with N strategic agents and one coordinator was considered. We modeled the system as a network with N vertices. Each agent lives in one of the vertices of the network and has a security state that describes its safety level at each time and is privately observed by that agent only. In our setup each agent's security state evolution dynamics follows a Markovian model and depends on its own state, the states of its neighbors in the network and on actions taken by the network coordinator. The agents have interdependent valuations in the sense that each agent's utility at each time is directly influenced by its own state as well as the states of its neighbors in the network. The network coordinator chooses security actions so as to maximize the expected social surplus in the long run. The coordinator therefore needs to elicit agents' private security state information to choose the best action at each time. Since agents are strategic and self-interested they need to be *incentivized* to truthfully reveal their private information. The coordinator's problem was therefore posed as a dynamic mechanism design problem. We used the inter-temporal correlations between the agents' security states to specify sufficient conditions under which an incentive compatible expected social surplus maximizing mechanism can be obtained. We then described construction of the desired mechanism for two special classes of problems.

Extending these results under richer valuation models that can incorporate more sophisticated forms of interdependencies is an interesting future direction. Moreover, it would be interesting to study the design of dynamic security enhancing mechanisms that also exhibit desirable revenue properties for a coordinator that seeks to earn profits through providing security preserving services to the network.

Appendix A
Proof of Theorem 1

We first prove the following lemma which will be used to establish Theorem 1.

Lemma 6. *Consider the mechanism* $(q_t, p_t)_{t\in\mathbb{T}}$ *described in Theorem 1. Suppose this mechanism satisfies the following condition:*

$$\mathbb{E}[\ p^i_{t+1}(m^i_t, S^{-i}_t, A_t)\mid s^{1:N}_t, a_t\ ] = -\frac{1}{\delta}\Big[\sum_{j\neq i} v^j(s^{1:N}_t, a_t) - c(a_t)\Big]\ ,$$
$$\forall s^{1:N}_t\in\mathcal{S}^N, \forall a_t\in\mathcal{A}, \forall i\in\mathcal{N}, \forall t\in\mathbb{T}, \tag{24}$$

where m^i_t *is given by* (10) *and the expectation is taken under truthful reporting strategies for all agents other than* i*. Then, the mechanism of Theorem 1 is an incentive compatible expected social surplus maximizing mechanism.*

Proof. See Appendix B. □

We will now show that condition (11) of Theorem 1 implies condition (24) of Lemma 6. This would mean that the mechanism of Theorem 1 is an incentive compatible expected social surplus maximizing mechanism. If all agents except i are truthful, we can use the definition of payment functions in (13) to write

$$\begin{aligned}
&\mathbb{E}\Big[\ p^i_{t+1}(m^i_t, S^{-i}_t, A_t)\ \mid\ s^{1:N}_t, a_t\ \Big]\\
&= -\frac{1}{\delta}\ \mathbb{E}\Big[\ \sum_{j\neq i} v^j(S^{-i}_t, m^i_t, A_t) - c(A_t)\mid s^{1:N}_t, a_t\ \Big]\\
&= -\frac{1}{\delta}\Big\{\ \sum_{\substack{j\neq i\\ j:i\in\mathcal{N}^j}}\mathbb{E}\Big[\ \alpha^i_j(S^j_t, A_t)m^i_t + \sum_{\substack{l\in\mathcal{N}^j\\ l\neq i}}\alpha^l_j(S^j_t, A_t)S^l_t\\
&+ \beta^j(S^j_t, A_t)\mid s^{1:N}_t, a_t\ \Big]\\
&+ \sum_{\substack{j\neq i\\ j:i\notin\mathcal{N}^j}}\mathbb{E}\Big[\ \sum_{l\in\mathcal{N}^j}\alpha^l_j(S^j_t, A_t)S^l_t + \beta^j(S^j_t, A_t)\mid s^{1:N}_t, a_t\ \Big]\\
&- c(a_t)\ \Big\}
\end{aligned}$$

$$= -\frac{1}{\delta}\Big\{ \sum_{\substack{j\neq i\\ j:i\in\mathcal{N}^j}} \Big(\alpha_j^i(s_t^j,a_t)\underbrace{\mathbb{E}[\, m_t^i | s_t^{1:N}, a_t\,]}_{s_t^i} + \sum_{\substack{l\in\mathcal{N}^j\\ l\neq i}} \alpha_j^l(s_t^j,a_t)s_t^l + \beta^j(s_t^j,a_t) \Big) + \sum_{\substack{j\neq i\\ j:i\notin\mathcal{N}^j}} \Big(\sum_{l\in\mathcal{N}^j} \alpha_j^l(s_t^j,a_t)s_t^l + \beta^j(s_t^j,a_t) \Big) - c(a_t) \Big\}$$
$$= -\frac{1}{\delta}\Big[\sum_{j\neq i}\Big(\sum_{l\in\mathcal{N}^j} \alpha_j^l(s_t^j,a_t)s_t^l + \beta^j(s_t^j,a_t) \Big) - c(a_t)\Big]$$
$$= -\frac{1}{\delta}\Big[\sum_{j\neq i} v^j(s_t^{1:N},a_t) - c(a_t)\Big].$$

Thus, condition (24) of Lemma 6 is true and hence the mechanism of Theorem 1 is an incentive compatible expected social surplus maximizing mechanism.

Appendix B
Proof of Lemma 6

In order to prove that the described mechanism is IC in the periodic ex post sense expressed in (7), we need to argue that at each time t, agent i can maximize its expected continuation utility by truthfully reporting its state at that time regardless of its private history h_t^i and the states of other agents s_t^{-i} at that time, provided that all other agents adopt truth-telling strategy. According to the one-shot deviation principle [6], [5] it suffices to argue that agent i is deterred from deviating from truth-telling strategy in a single time step and then reverting to truthful reporting afterwards. That is, we need to show that reporting s_t^i is the best response to agent i's problem of maximizing its expected continuation utility at time t as given below:[6]

$$\max_{r_t^i\in\mathcal{S}} \Big\{ v^i(s_t^{1:N}, q_t(h_t, r_t^i, s_t^{-i})) - \underbrace{p_t^i(m_{t-1}^i, r_{t-1}^{-i}, a_{t-1})}_{\dagger} + \delta\,\mathbb{E}[\, v^i(S_{t+1}^{1:N}, q_{t+1}(H_{t+1}, S_{t+1}^{1:N})) - p_{t+1}^i(m_t^i, s_t^{-i}, q_t(h_t, r_t^i, s_t^{-i})) \mid s_t^{1:N}, h_t^i\,] + \mathbb{E}\Big[\sum_{t'=t+2}^{\infty} \delta^{t'-t}\Big(v^i(S_{t'}^{1:N}, q_{t'}(H_{t'}, S_{t'}^{1:N})) - p_{t'}^i(m_{t'-1}^i, S_{t'-1}^{-i}, q_{t'-1}(H_{t'-1}, S_{t'-1}^{1:N})) \Big) \mid s_t^{1:N}, h_t^i \Big] \Big\}. \tag{25}$$

[6]Note that the participation fee charged to agent i at time 0 as given in (14), is independent of agent i's future reports and thus can be dismissed in agent i's best response optimization problem in (25).

Noting that the term † in (25) does not depend on r_t^i and using (12), the objective in (25) can be rewritten as

$$\max_{r_t^i\in\mathcal{S}} \Big\{ v^i(s_t^{1:N}, \rho^*(r_t^i, s_t^{-i})) + \delta\,\mathbb{E}[\, v^i(S_{t+1}^{1:N}, \rho^*(S_{t+1}^{1:N})) - p_{t+1}^i(m_t^i, s_t^{-i}, \rho^*(r_t^i, s_t^{-i})) \mid s_t^{1:N}\,] + \mathbb{E}\Big[\sum_{t'=t+2}^{\infty} \delta^{t'-t}\Big(v^i(S_{t'}^{1:N}, \rho^*(S_{t'}^{1:N})) - p_{t'}^i(m_{t'-1}^i, S_{t'-1}^{-i}, \rho^*(S_{t'-1}^{1:N})) \Big) \mid s_t^{1:N} \Big] \Big\}. \tag{26}$$

From (24) it follows that:

$$\mathbb{E}[p_{t+1}^i(m_t^i, s_t^{-i}, \rho^*(r_t^i, s_t^{-i})) \mid s_t^{1:N}\,] = -\frac{1}{\delta}\Big[\sum_{j\neq i} v^j(s_t^{1:N}, \rho^*(r_t^i, s_t^{-i})) - c(\rho^*(r_t^i, s_t^{-i}))\Big],$$

and that,

$$\begin{aligned}
&\mathbb{E}[\, p_{t+2}^i(m_{t+1}^i, S_{t+1}^{-i}, \rho^*(S_{t+1}^{1:N})) \mid s_t^{1:N}\,] \\
&= \mathbb{E}[\, \mathbb{E}[p_{t+2}^i(m_{t+1}^i, S_{t+1}^{-i}, \rho^*(S_{t+1}^{1:N})) \mid s_t^{1:N}, S_{t+1}^{1:N}\,] \mid s_t^{1:N}\,] \\
&= \mathbb{E}[\, \mathbb{E}[p_{t+2}^i(m_{t+1}^i, S_{t+1}^{-i}, \rho^*(S_{t+1}^{1:N})) \mid S_{t+1}^{1:N}\,] \mid s_t^{1:N}\,] \\
&= \mathbb{E}[-\frac{1}{\delta}\Big[\sum_{j\neq i} v^j(S_{t+1}^{1:N}, \rho^*(S_{t+1}^{1:N})) - c(\rho^*(S_{t+1}^{1:N}))\Big] \mid s_t^{1:N}\,] \\
&= -\frac{1}{\delta}\,\mathbb{E}\Big[\sum_{j\neq i} v^j(S_{t+1}^{1:N}, \rho^*(S_{t+1}^{1:N})) - c(\rho^*(S_{t+1}^{1:N})) \mid s_t^{1:N}\Big].
\end{aligned} \tag{27}$$

Using similar steps as in (27), it can be shown that:

$$\mathbb{E}[\, p_{t+k+1}^i(m_{t+k}^i, S_{t+k}^{-i}, \rho^*(S_{t+k}^{1:N})) \mid s_t^{1:N}\,] = -\frac{1}{\delta}\,\mathbb{E}\Big[\sum_{j\neq i} v^j(S_{t+k}^{1:N}, \rho^*(S_{t+k}^{1:N})) - c(\rho^*(S_{t+k}^{1:N})) \mid s_t^{1:N}\Big],$$
for $k > 1$. (28)

Using the above equations, (26) can be simplified and written as:

$$\max_{r_t^i\in\mathcal{S}} \Big\{ \sum_{j=1}^{N} v^j(s_t^{1:N}, \rho^*(r_t^i, s_t^{-i})) - c(\rho^*(r_t^i, s_t^{-i})) + \mathbb{E}\Big[\sum_{t'=t+1}^{\infty} \delta^{t'-t}\Big(\sum_{j=1}^{N} v^j(S_{t'}^{1:N}, \rho^*(S_{t'}^{1:N})) - c(\rho^*(S_{t'}^{1:N}))\Big) \mid s_t^{1:N} \Big] \Big\}. \tag{29}$$

From (29) we observe that agent i's expected continuation utility at time t is in alignment with expected continuation social surplus at time t. Since $\rho^*(s_t^i, s_t^{-i})$ is the expected social surplus maximizing action[7] when the network state profile is (s_t^i, s_t^{-i}), it follows that

$$\begin{aligned}
&\sum_{j=1}^{N} v^j(s_t^{1:N}, \rho^*(s_t^i, s_t^{-i})) - c(\rho^*(s_t^i, s_t^{-i})) \\
&+ \mathbb{E}\Big[\sum_{t'=t+1}^{\infty} \delta^{t'-t} \Big(\sum_{j=1}^{N} v^j(S_{t'}^{1:N}, \rho^*(S_{t'}^{1:N})) \\
&- c(\rho^*(S_{t'}^{1:N})) \Big) \mid s_t^{1:N} \Big] \\
&\geq \sum_{j=1}^{N} v^j(s_t^{1:N}, \rho^*(r_t^i, s_t^{-i})) - c(\rho^*(r_t^i, s_t^{-i})) \\
&+ \mathbb{E}\Big[\sum_{t'=t+1}^{\infty} \delta^{t'-t} \Big(\sum_{j=1}^{N} v^j(S_{t'}^{1:N}, \rho^*(S_{t'}^{1:N})) \\
&- c(\rho^*(S_{t'}^{1:N})) \Big) \mid s_t^{1:N} \Big], \\
&\forall r_t^i \in \mathcal{S}. \qquad (30)
\end{aligned}$$

That is, the truthful report s_t^i is the best response to agent i's problem of maximizing its expected continuation utility at time t (see (25)) under the mechanism described in Lemma 6. This in turn implies that indeed under the described mechanism the expected social surplus maximizing action $\rho^*(s_t^{1:N})$ corresponding to the true state profile $s_t^{1:N}$ will be selected at each time t. Therefore, provided that the conditions in (24) are satisfied, the mechanism described in Theorem 1 is incentive compatible and expected social surplus maximizing. This completes the proof.

Appendix C
Proof of Lemma 2

Given that $s_t^i \in \{0,1\}$, for each agent i and each $s_t^{-i} \in \{0,1\}^{N-1}, a_t \in \mathcal{A}$, the conditions in (11) result in the following system of linear equations:

$$\begin{aligned}
&\sum_{\{s_{t+1}^j\}_{j\in\mathcal{O}^i} \in \{0,1\}^{|\mathcal{O}^i|}} f^i(\{s_{t+1}^j\}_{j\in\mathcal{O}^i}, s_t^{-i}, a_t) \\
&\mathbb{P}\Big(\{S_{t+1}^j\}_{j\in\mathcal{O}^i} = \{s_{t+1}^j\}_{j\in\mathcal{O}^i} \mid s_t^{-i}, S_t^i = x, a_t\Big) \quad = x \,, \\
&x \in \{0,1\}, \qquad (31)
\end{aligned}$$

where the unknowns are agent i's cross-inference signal values $f^i(\{s_{t+1}^j\}_{j\in\mathcal{O}^i}, s_t^{-i}, a_t)$ corresponding to each of the profiles $\{s_{t+1}^j\}_{j\in\mathcal{O}^i}$, given the profile s_t^{-i} and the action a_t. The condition in (31) represents a system of linear equations with a $2\times 2^{|\mathcal{O}^i|}$ dimensional coefficients matrix. This system admits at least one solution if its coefficients matrix is full rank, i.e., if its rows are linearly independent. It is straightforward to verify that two distinct PMFs give rise to two linearly independent vectors. Therefore, provided the conditions in (16) are satisfied, there always exist cross-inference signals $m_t^i = f^i(\{r_{t+1}^j\}_{j\in\mathcal{O}^i}, r_t^{-i}, A_t), \forall i \in \mathcal{N}, \forall t \in \mathbb{T}$ such that Equation (11) of Theorem 1 holds true. These signals are given by a solution to the system of linear equations in (31). This completes the proof.

[7] See characterization of $\rho^*(\cdot)$ in Lemma 1.

Appendix D
Proof of Lemma 4

We need to argue that the described mechanism satisfies the BB constraint given in (21). Let us rewrite the left-hand side in (21) as:

$$\begin{aligned}
&(1-\delta)\mathbb{E}\Big[\sum_{i=1}^{N} p_0^i + \sum_{t=0}^{\infty} \delta^{t+1} \sum_{i=1}^{N} p_{t+1}^i(m_t^i, S_t^{-i}, a_t) \Big] \\
&= (1-\delta)\mathbb{E}\Big[\sum_{i=1}^{N} p_0^i + \sum_{i=1}^{N}\sum_{t=0}^{\infty} \delta^{t+1} p_{t+1}^i(m_t^i, S_t^{-i}, a_t) \Big] \\
&= (1-\delta)\mathbb{E}\Big[\sum_{i=1}^{N} \Big(p_0^i + \sum_{t=0}^{\infty} \delta^{t+1} p_{t+1}^i(m_t^i, S_t^{-i}, a_t) \Big) \Big] \\
&= (1-\delta) \sum_{i=1}^{N} \mathbb{E}\Big[p_0^i + \sum_{t=0}^{\infty} \delta^{t+1} p_{t+1}^i(m_t^i, S_t^{-i}, a_t) \Big].
\end{aligned}$$

Using the value of p_0^i given in (22), it is straightforward to see that the above expression is equal to 0. This completes the proof.

Appendix E
Proof of Lemma 5

Given the mechanism $(\rho^*, p_t)_{t\in\mathbb{T}}$ is IC and the participation fees $p_0^i, i \in \mathcal{N}$, are of the form given in (22), agent i's maximum expected total utility is given by:

$$\begin{aligned}
&\mathbb{E}\Big[(1-\delta) \sum_{t=0}^{\infty} \delta^t u^i\Big(S_t^{1:N}, \rho^*(S_t^{1:N}), p_t^i\Big) \Big] \\
&= \mathbb{E}\Big\{ (1-\delta) \Big[v^i(S_0^{1:N}, \rho^*(S_0^{1:N})) - p_0^i \\
&+ \sum_{t=0}^{\infty} \delta^{t+1} \Big(v^i(S_{t+1}^{1:N}, \rho^*(S_{t+1}^{1:N})) - p_{t+1}^i(m_t^i, S_t^{-i}, \rho^*(S_t^{1:N})) \Big) \Big] \Big\} \\
&= \mathbb{E}\Big\{(1-\delta) \sum_{t=0}^{\infty} \delta^t \, v^i(S_t^{1:N}, \rho^*(S_t^{1:N}))\Big\} \\
&\underbrace{- \mathbb{E}\Big\{(1-\delta)\Big[p_0^i + \sum_{t=0}^{\infty} \delta^{t+1} \, p_{t+1}^i(m_t^i, S_t^{-i}, \rho^*(S_t^{1:N}))\Big]\Big\}}_{=0} \qquad (32) \\
&= \mathbb{E}\Big[(1-\delta) \sum_{t=0}^{\infty} \delta^t \, v^i(S_t^{1:N}, \rho^*(S_t^{1:N}))\Big] \\
&\geq \mathbb{E}\Big[(1-\delta) \sum_{t=0}^{\infty} \delta^t v^i\Big(S_t^{1:N}, \bar{a}\Big) \Big] \qquad (33) \\
&= \mathbb{E}\Big[(1-\delta) \sum_{t=0}^{\infty} \delta^t u^i\Big(S_t^{1:N}, \bar{a}, 0\Big) \Big].
\end{aligned}$$

Note that the second term in (32) is zeroed out because p_0^i is of the form given in (22). The inequality in (33) followed from Assumption 3. This establishes satisfaction of the IR constraints in (20).

References

[1] T. Alpcan and T. Başar, *Network security: A decision and game-theoretic approach*. Cambridge University Press, 2010.

[2] D. Bergemann and J. Valimaki, "Dynamic mechanism design: An introduction," 2018.

[3] S. R. Etesami and T. Başar, "Dynamic games in cyber-physical security: An overview," *Dynamic Games and Applications*, pp. 1–30, 2019.

[4] F. Farhadi, H. Tavafoghi, D. Teneketzis, and S. J. Golestani, "An efficient dynamic allocation mechanism for security in networks of interdependent strategic agents," *Dynamic Games and Applications*, pp. 1–28, 2018.

[5] D. Bergemann and J. Välimäki, "The dynamic pivot mechanism," *Econometrica*, vol. 78, no. 2, pp. 771–789, 2010.

[6] S. Athey and I. Segal, "An efficient dynamic mechanism," *Econometrica*, vol. 81, no. 6, pp. 2463–2485, 2013.

[7] H. Liu, "Efficient dynamic mechanisms in environments with interdependent valuations: The role of contingent transfers," *Theoretical Economics*, vol. 13, no. 2, pp. 795–829, 2018.

[8] M. L. Puterman, *Markov decision processes: discrete stochastic dynamic programming*. John Wiley & Sons, 2014.

[9] P. R. Kumar and P. Varaiya, *Stochastic systems: Estimation, identification, and adaptive control*. SIAM, 2015, vol. 75.

[10] T. Borgers, R. Strausz, and D. Krahmer, *An introduction to the theory of mechanism design*. Oxford University Press, USA, 2015.