2018 European Control Conference (ECC)
June 12-15, 2018. Limassol, Cyprus

Static and Dynamic Informational Incentive Mechanisms for Security Enhancement

Farzaneh Farhadi[1,2], Demosthenis Teneketzis[1], and S. Jamaloddin Golestani[2]

Abstract— **In strategic relationships, the control of information is an important instrument for coordinating and incentivizing actions. In this paper, we consider a game with asymmetric information where a principal privately observes the security state of a system and chooses a signal to disclose his information strategically to an agent, who then takes a strategic action that affects the utility of both. We study the problem of designing an optimal information disclosure mechanism for the principal in both static and dynamic settings. In the dynamic setting, unlike the existing works, we consider non-myopic players who are attempting to maximize their long-run incentives. Using future information revelation as the basis for promises and threats, we design an optimal information disclosure mechanism that maximizes the principal's utility, and provides the agent with strong incentives to obey.**

I. Introduction

In strategic interactions between two parties with asymmetric information and mismatched goals (i.e. games with asymmetric information), the better informed party, call him principal, can use his superior information as a control instrument to persuade the other party, call her agent, to behave as he wants. In many economic and engineering applications, the principal and the agent have different utility functions which depend on the system state, that is known solely to the principal, and on the agent's decision. In these applications, the principal can utilize his superior information about system state to provide informational incentives to the agent by a selective disclosure of information so as to influence the agent's decision, and ultimately to maximize his own utility. Examples of such applications include marketing a new product by advertising, congestion control via real-time traffic signals, demand response management via announcement of time-varying electricity prices, and health and security promotion via timely environmental information disclosure. In this paper, we study the problem of designing optimal information disclosure mechanisms, in both static and dynamic settings, motivated by security games.

We consider a security game where a strategic agent is exposed to attacks from the environment which, at any time, can be in one of the two states: low-risk or high-risk. The agent cannot observe the environment's state; therefore she does not know the level of risk she is facing so as to determine whether it is worthwhile for her to undertake the cost of exerting effort and ultimately protecting herself. There is a principal in the system who privately observes the environment's risk level and sends a message to the agent. The principal's objective is different from that of the agent, as the principal is only concerned with the system's safety and does not care how much establishing security will cost to the agent. Therefore, utilizing his superior information, he attempts to provide informational incentives to the agent so as to persuade her that investment in protection is beneficial. Since the agent is strategic, she utilizes the information she receives from the principal to her own advantage and does not necessarily follow the actions suggested by the principal. Consequently, the principal needs to disclose his information strategically so that the agent's best response to the information she receives maximizes his objective. When the principal chooses an information disclosure mechanism/strategy, he announces it and commits to it. The agent's knowledge of the mechanism shapes her interpretation of the messages she receives.

In this paper, we study the problem of designing an optimal information disclosure mechanism for the principal. We first study the problem in a *static* setting where the environment's state is fixed. In this setting, we formulate the information design problem as an optimization problem and state the characteristic of an optimal informational incentive mechanism. We show that strategic disclosure of information significantly improves the principal's utility compared to the utility he could get by adopting non-strategic policies such as full-information or no-information disclosure mechanisms.

Next, we consider a *dynamic* two-stage setting where the environment's state evolves dynamically over time according to an uncontrolled Markov chain, and the principal and the agent attempt to maximize their long-run utilities. In dynamic settings, using future information revelation as the basis for promises and threats, we design a dynamic information disclosure mechanism that maximizes the principal's objective. We prove that this optimal dynamic information disclosure mechanism is of threshold type with respect to the agent's patience. If the agent is patient enough, the principal puts the priority on the agent's first-period security and maximizes it by promising to provide *near-truth* information in the future. However, if the agent has little patience, the principal gives the same priority to the agent's security in both periods and does not sacrifice one in favor of the other.

A. Review of Related Works

There exists literature on information transmission in principal-agent or equivalently transmitter-receiver settings. A significant part of the existing works have studied the non-strategic case, where the information transmitter and receiver

*This work was supported in part by the NSF grant CNS-1238962, ARO-MURI grant W911NF-13-1-0421, and ARO grant W911NF-17-1-0232.

[1] University of Michigan, [2] Sharif University of Technology
emails: ffarhadi, teneket@umich.edu, golestani@ieee.org

are cooperative and have the same objectives [1]–[3].

The works that address the strategic case, where the transmitter and receiver have misaligned objectives, are classified into two main categories based on the model considered. One set of works consider the *cheap talk* model where the principal/information transmitter has no commitment power [4]–[9]. In this model, the principal is not committed to any disclosure mechanism before hand, and decides on the information message to be sent after seeing the state's realization. The cheap talk model induces a simultaneous game between the principal and the agent. Thus, the main goal of this category's works is to characterize the Nash equilibria of the induced game. Most of the existing work has focused on the *static* setting [4]–[7]. The work of [4] shows that when the state of the system is one-dimensional and both the principal's and the agent's utilities are quadratic, the principal's equilibrium strategy employs quantization. More general models of cheap talk, such as multidimensional sources [5], noisy communication [6] and multiple principals with misaligned objectives [7], have been studied in the literature for static settings. There are a few works that study the dynamic version of the cheap talk communication [8], [9]. These works show that allowing for dynamic information control improves the informativeness of the communication.

Another set of works - including ours - consider the information disclosure model, where the principal has commitment power and hence announces an information disclosure mechanism (recommendation policy) before seeing the system state and then commits to it [10]–[18]. The main focus of this strand of works, which corresponds to Bayesian persuasion problems, is to design an optimal information disclosure mechanism for the principal. The static version of the problem was studied in [10]–[14]. The authors of [10] consider a problem of static information disclosure where the state of the system is Gaussian and the utilities are quadratic, and show that there exists a linear policy for the principal that leads to a Stackelberg equilibrium. In [11], [12], the authors propose a method for deriving an optimal information provision mechanism explicitly. Our approach to the static information design problem is similar to the ones proposed in [11], [12]. The dynamic version of the problem have been addressed in [15]–[17] when the principal and the agent care only about their immediate utilities at each time step and are, hence, myopic decision makers. Under this simplifying assumption, the work of [17] develops a method for designing optimal dynamic information mechanisms by modifying the agents' beliefs so as to maximize the principal's objective. The work of [18] studies the dynamic information design problem for agents with long-run incentives; using numerical simulations it discusses the effect of each piece of information on the optimal mechanism and its qualitative properties. Our work contributes to the existing literature by deriving an optimal dynamic information disclosure mechanism analytically, when both the principal and the agent have long-run utilities.

The rest of the paper is organized as follows. We present the static model of our paper in Section II. We formulate the static information design problem and characterize its solution in Section III. We generalize our model to a two-stage dynamic setting in Section IV. In section V, we formulate the dynamic information design problem and provide a step by step solution to it. We present numerical results in Section VI. The proofs of all the results that appear in this paper can be found in [19].

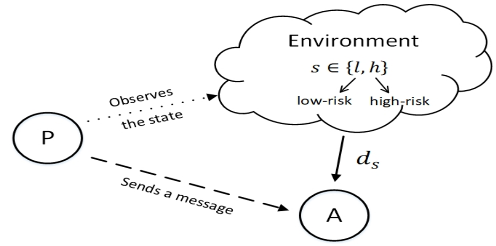


Fig. 1: Model, where A stands for Agent and P for Principal.

II. Static Model

Consider a system consisting of a strategic agent (she) who lives in an unsafe environment and is in danger of being attacked, and a principal (he) who is concerned with the agent's safety (See Fig. 1). The level of risk to which the agent is exposed depends on the safety of the environment she is residing in. We denote the risk level (security state) of the environment by $s \in S := \{l, h\}$, where $s = l$ indicates that the environment is low-risk and $s = h$ indicates that it is high-risk. The agent is more likely to get attacked in a high-risk environment. She may get attacked with probability d_l if the environment is low-risk, and with probability $d_h > d_l$ if it is high-risk. The agent cannot observe the environment's state; therefore she does not know the level of danger she is facing so as to determine whether it is worthwhile to exert effort to protect herself. The principal privately observes the environment's state and sends a message to the agent in order to persuade her that the environment is unsafe so that investment in protection is beneficial. The principal announces and commits to a mechanism/strategy for disclosing information and the agent's knowledge of the mechanism shapes her interpretation of the messages she receives, and ultimately determines her best response action.

Formally, at the beginning of the process, the agent and the principal share a prior belief that the environment is high-risk with probability μ. Then, the principal observes the true state of the environment and sends a signal/message m to the agent. When the agent receives the message, she updates her belief from μ to ν, and based on ν she decides whether or not to exert effort towards her safety. Let $x \in \{0, 1\}$ denote the agent's action, where $x = 0$ indicates the agent exerts no effort for protecting herself and $x = 1$ indicates that she takes full effort to ensure that she remains safe. Exerting full effort protects the agent from any attack but is costly. We denote cost of the agent's effort by $c > 0$ which determines the expenditure by agent required to guarantee her safety.

Based on the agent's action and possible attacks, the agent's security state will be safe $\theta = 1$, or unsafe $\theta = 0$. When the agent chooses action $x \in \{0, 1\}$ and the environment's state is $s \in \{l, h\}$, the probability that the agent will be safe is given by

$$P(\theta = 1|x, s) = \begin{cases} 1, & \text{if} \quad x = 1, \\ 1 - d_s, & \text{if} \quad x = 0. \end{cases} \tag{1}$$

Eq. (1) is interpreted as follows. If the agent invests in her security, i.e. $x = 1$, she will be out of harm's way. However, if she does not exert any effort to secure her safety, she will remain safe if she is not subject to any attack (prob. $1 - d_s$).

The agent has a valuation $r > c$ for her security. Therefore, the agent's total utility when she chooses effort level x and achieves security state θ is

$$u^A(\theta, x) = r\theta - cx, \tag{2}$$

where she gets a reward r if she is safe, and pays a cost c if she decides to take action $x = 1$. Unlike the agent, the principal is only concerned with the agent's safety and does not care how much establishing security will cost. Thus, his utility is 1 if the agent is safe and 0 otherwise, i.e.

$$u^P(\theta) = \theta. \tag{3}$$

According to (3), principal always wishes that the agent exerts effort to protect herself. However, the agent's decision depends on her belief about the environment's unknown state s. We can assume, without loss of generality, that the cost to reward ratio is such that the agent wants to match her action with the environment's state, since otherwise one of the agent's actions is always dominant and hence will be chosen by her, irrespective of the information the principal discloses. This condition is provided by the following assumption:

Assumption 1: The cost to reward ratio is bounded as

$$d_l \le c/r \le d_h. \tag{4}$$

The principal wants to use his superior information to persuade the agent to exert effort. Therefore, he must design a mechanism for disclosing information so that it leads the agent to believe that the environment is more likely to be high-risk. In the next section, we formulate this problem as an information design problem, and design an information disclosure mechanism that maximizes the principal's utility.

III. The Static Information Design Problem

An information disclosure mechanism specifies the set of messages $\mathcal{M}$ that the principal sends to the agent, along with a distribution over $\mathcal{M}$ given the state of the environment (which the principal observes). An information disclosure mechanism could be very complicated since there is no restriction on the set of messages $\mathcal{M}$. However, it is shown in [11] that there is no loss of generality in restricting attention to *direct* information disclosure mechanisms that are *obedient*.

In a direct information disclosure mechanism the principal directly recommends to the agent the action she should take. We denote the set of messages used by the principal by $\mathcal{M} = \{n, e\}$, where n is a recommendation to exert no effort, and e is a recommendation to exert effort. Therefore, the principal's behavior in a direct information disclosure mechanism can be described by a recommendation policy $\boldsymbol{\rho} = (\rho^s, s \in S)$, where ρ^s is the probability according to which the principal sends message e to the agent (i.e. recommends her to exert effort), when he observes the environment's state s. For each state $s \in S$ the principal sends message n to the agent with probability $1-\rho^s$. When the principal designs an information disclosure mechanism, he announces the recommendation policy to the agent and commits to it.

It is clear from (1) and (3) that the ideal information disclosure mechanism for the principal is the one where the agent is always recommended to exert effort, i.e. $\rho^s = 1$, for all $s \in S$; however, if the principal announces this recommendation policy, the agent realizes that the principal's messages are uncorrelated to the environment's state and hence she will disregard them. In other words, since the agent is strategic, she utilizes the information she receives from the principal to her own advantage and does not necessarily follow the actions recommended by the principal. Therefore, the principal must design an information disclosure mechanism that possesses the *obedience* property, that is, it provides the agent with incentives to follow his recommendations.

When the agent receives the principal's recommendation $m \in \{n, e\}$, she updates her belief about the environment's high-risk state to $\nu(m)$ using Bayes rule as follows

$$\nu(m) = P(s=h|m) = \begin{cases} \frac{\mu\rho^h}{\mu\rho^h + (1-\mu)\rho^l}, & m = e, \\ \frac{\mu(1-\rho^h)}{\mu(1-\rho^h)+(1-\mu)(1-\rho^l)}, & m = n. \end{cases} \tag{5}$$

Then, based on her posterior belief ν, she selects an action x so as to maximize her expected utility, that is,

$$x \in \arg\max_x \mathbb{E}_\nu[u^A(\theta, x)] = \arg\max_x \{r\,\mathbb{E}_\nu[\theta|x] - cx\}. \tag{6}$$

Using (1), the expected safety of the agent who has belief ν and takes action x is

$$\mathbb{E}_\nu[\theta|x] = P(\theta = 1|x) = \nu P(\theta = 1|x, s = h) +$$
$$(1-\nu)P(\theta=1|x, s=l) = \begin{cases} 1, & x = 1, \\ 1 - d_l - \nu(d_h - d_l), & x = 0. \end{cases} \tag{7}$$

Substituting (7) into (6) shows that there is a threshold ν^* of the probability the agent assigns to the environment's high-risk state, above which her best response is to exert effort and below which is to do nothing. This threshold is given as

$$\nu^* := \frac{c/r - d_l}{d_h - d_l}. \tag{8}$$

Therefore, the obedience property of the information mechanism $\boldsymbol{\rho}$ is guaranteed by the following constraints,

$$\nu(e) \ge \nu^*, \quad \text{and} \quad \nu(n) \le \nu^*. \tag{9}$$

When the agent follows the recommendation policy $\boldsymbol{\rho}$, the expected utility the principal gets is

$$\mathbb{E}[u^P(\theta)|\boldsymbol{\rho}] = \mathbb{E}[\theta|\rho^l, \rho^h] = \mu P(\theta = 1|s = h, \rho^h) +$$
$$(1 - \mu)P(\theta = 1|s = l, \rho^l) = \mu\Big[\rho^h + (1 - \rho^h)(1 - d_h)\Big] +$$
$$(1 - \mu)\Big[\rho^l + (1 - \rho^l)(1 - d_l)\Big]. \tag{10}$$

Therefore, we can formulate the information design problem for the principal as follows:

$$\begin{aligned}
&\max_{\boldsymbol{\rho}=(\rho^l,\rho^h)} \mathbb{E}[u^P(\theta)|\boldsymbol{\rho}],\\
&\textit{s.t. the obedience constraints (9),}\\
&0 \le \rho^s \le 1,\ \forall s \in S.
\end{aligned} \tag{11}$$

That is, the principal wants to choose a feasible information disclosure mechanism $\boldsymbol{\rho}$ that satisfies the obedience constraints and maximizes his expected utility. In the following theorem, we describe an information disclosure mechanism that solves the principal's problem (11).

Theorem 1: The solution to the static information design problem (11) for the principal is

$$\rho^{*h} = 1, \qquad \rho^{*l} = \min\left\{1, \frac{\mu(1-\nu^*)}{\nu^*(1-\mu)}\right\}. \tag{12}$$

The result of Theorem 1 is intuitive. When the environment is high-risk the principal always recommends the agent to exert effort. However, when the environment is low-risk, the principal plays a mixed strategy so as to introduce uncertainty regarding the precise state of the system.

The following theorem shows that the optimal mechanism's percent of improvement over the naive information disclosure mechanisms that can be used by the principal, namely, the no-information and full-information disclosure mechanisms, is unbounded.

Theorem 2: For any arbitrarily large $\eta > 0$, there is a set of parameters μ, c, r, d_l, d_h, such that the optimal mechanism's percent of improvement over the two naive mechanisms is greater than η.

IV. Dynamic Model

The problem we discussed in Sections II and III is a one-period static problem where the environment's risk level is fixed, and both the principal and the agent optimize their decisions based on their one-time utilities. We now consider a *multi-period dynamic* problem where the environment's state evolves dynamically over time and the principal and the agent maximize their long-run utilities.

In comparison to the static model, dynamics add several interesting dimensions to the information design problem: (i) The agent knows that the state is evolving over time, so even if the principal offers no information, the agent's beliefs will evolve autonomously; (ii) Commitment (by the principal) to disclose certain information to the agent at any time t does not affect her decisions only at that time, but also at times before and after t. The disclosure affects the agent's future decisions through altering the path of her beliefs, and influences her decisions at times before t by changing the continuation utility she expects to achieve. Therefore, as far as the principal is concerned, there is a tradeoff between persuading the agent to take desired actions at each time t and the ability to persuade her at times before and after t.

To investigate dynamic information design problems, we consider a dynamic setting with time horizon $T = 2$, i.e. $t \in \mathcal{T} := \{1, 2\}$, where the environment's state $s_t \in \{l, h\}$, $t \in \mathcal{T}$, evolves over time as an uncontrollable Markov chain with transition probability matrix

$$\mathbb{P} = \begin{array}{c c} & \begin{array}{cc} h & l \end{array} \\ \begin{array}{c} h \\ l \end{array} & \left[\begin{array}{cc} 1 & 0 \\ q & 1-q \end{array}\right] \end{array}.$$

At each time $t \in \mathcal{T}$, the risk level of the environment may switch from low to high with probability q. Once the environment becomes insecure and high-risk, it remains in the high-risk state forever.

In the absence of any additional information, the agent's belief about the environment's state evolves over time according to

$$\mu_t \triangleq P(s_t = h) = \mu_{t-1} + (1-\mu_{t-1})q \triangleq f(\mu_{t-1}). \tag{13}$$

However, at each time $t \in \mathcal{T}$, the principal observes the current state s_t and sends a message m_t to the agent in order to influence the evolution of her beliefs, hence her future actions $x_\tau \in \{0, 1\}$, $\tau \ge t$. At each time $t \in \mathcal{T}$, if the agent is currently in the safe state, i.e. $\theta_{t-1} = 1$, and decides to undertake the cost c of exerting effort, i.e. $x_t = 1$, she is secure against new attacks at time t. The agent who is already in the unsafe state $\theta_{t-1} = 0$, remains unsafe regardless of her action. The principal can perfectly observe the evolution of the agent's security state θ_t, $t \in \mathcal{T}$, but not the agent's actions. If the agent switches to the unsafe state at time t, the principal finds out that the agent exerted no effort at that time, but if the agent remains safe, the principal cannot determine whether she exerted effort or was just lucky.

The agent's and the principal's instantaneous utilities at each time t are similar to the static case, i.e. $u^A(\theta_t, x_t) = r\theta_t - cx_t$ and $u^P(\theta_t) = \theta_t$. Let $\delta \in (0, 1]$ denote the common discount factor. Then, the agent's and the principal's long-run discounted utilities are

$$U^A = \sum_{t\in\mathcal{T}} \delta^{t-1} u^A(\theta_t, x_t), \quad \text{and} \quad U^P = \sum_{t\in\mathcal{T}} \delta^{t-1} u^P(\theta_t), \tag{14}$$

respectively. Similar to the static case, we assume that the agent's utility function is such that she wishes to match her action with the environment's unknown state. At any time, if the agent is absolutely sure that she is living in a high-risk environment, irrespective of what might happen in the future, she prefers to pay the cost and protect herself; but if she is absolutely sure that the environment is low-risk she finds protection investment unworthy. This condition is provided by the following assumption:

Assumption 2: The cost to reward ratio is bounded as

$$d_l(2 - d_l) \le c/r \le d_h. \tag{15}$$

This assumption is an analogue of Assumption 1 for dynamic environments.

V. The Dynamic Information Design Problem

In dynamic problems, similar to the static ones, we can restrict attention, without loss of generality, to direct information disclosure mechanisms that possess the obedience property. A dynamic direct information mechanism is a recommendation policy $\rho(\cdot) = \{\rho_t(\cdot), t \in \mathcal{T}\}$ that the principal designs and commits to it, where $\rho_t : S \times \mathcal{H}_t \rightarrow [0,1]$ determines the probability according to which the principal recommends to the agent to exert effort at time t based on the current state $s_t \in S$ and the time-t public history $h_t \in \mathcal{H}_t$. The time-t public history contains all the principal's recommendations up to time $t-1$, i.e.,

$$h_t := \{m_\tau, \tau \leq t-1\}, \tag{16}$$

where $m_\tau \in \{n, e\}$ is the principal's message at time τ. We denote the set of all possible public histories at time t by $\mathcal{H}_t$. For ease of notation, hereafter we write $\rho_t^{s_t,h_t}$ for $\rho_t(s_t, h_t)$. Then, a direct information mechanism in our two-period model can be characterized by a recommendation vector

$$\boldsymbol{\rho} = (\rho_1^l, \rho_1^h, \rho_2^{l,n}, \rho_2^{h,n}, \rho_2^{l,e}, \rho_2^{h,e}), \tag{17}$$

where $\rho_1^{s_1}$ is the probability of recommending the agent to exert effort at time $t=1$ when the environment's state is s_1, and $\rho_2^{s_2,m_1}$ is the probability of recommending her to exert effort at time $t=2$ when the state is s_2 and the principal's message at time $t=1$ was $m_1 \in \{n, e\}$.

The principal wants to choose a recommendation policy $\boldsymbol{\rho}$ that possesses the obedience property and maximizes his long-run expected utility. By an argument similar to that given for the static problem, when the environment is high-risk the strategic agent herself is interested to invest in her security. Therefore, in this case there is no point for the principal to mislead the agent. As a result, when the environment's risk level is high, it is always optimal for the principal to recommend to the agent to exert effort; i.e.

$$\rho_1^{*\,h} = \rho_2^{*\,h,n} = \rho_2^{*\,h,e} = 1. \tag{18}$$

Equation (18) determines three components of the optimal recommendation vector described by (17).

As a result of (18), when the agent is recommended to exert no effort she is absolutely sure that the environment is low-risk. The agent always follows this recommendation, because by Assumption 2, the cost to reward ratio is high enough so that when the environment is low-risk exerting effort is unworthy. Therefore, in finding the other three components of the optimal recommendation vector, the principal needs to pay attention to the obedience constraints only when he recommends to the agent to exert effort.

Finding the conditions that guarantee obedience in dynamic settings with long-term optimizing agents is not as straightforward as for static settings. Because in a dynamic problem, when an agent with a long-run incentive receives a recommendation at time t, the decision she takes to obey or disobey is not only based on her interpretation of the message she has received at time t, shaped by the recommendation rule of the present time, but also on the history of past messages she has received as well as the past and future recommendation rules. This can be seen as follows. The long-term optimizing agent obeys the principal's recommendation at time t if the recommended action maximizes her expected continuation utility. The agent's expected continuation utility depends on two factors: (1) the current belief ν_t she has about the environment's risk level, which is constructed from the past messages m_τ she received and the past recommendation rules $\rho_\tau^{s_\tau,h_\tau}$, $\tau \leq t$; and (2) the information she expects to receive in the future, which is determined by future recommendation rules to which the principal commits, i.e. $\rho_\tau^{s_\tau,h_\tau}$, $\tau > t$. The dependence of the agent's decision at each time on the recommendation policy for the whole horizon as well as the history of the past messages makes the derivation of the obedience constraints a formidable task.

In the next subsection V-A, we first derive explicitly the conditions that guarantee the obedience property of the recommendation policy $\boldsymbol{\rho} = (\rho_1^l, 1, \rho_2^{l,n}, 1, \rho_2^{l,e}, 1)$. Then, using these conditions we formulate the dynamic information design problem for the principal. We provide a solution to this problem in subsection V-B.

A. Obedience Constraints

Based on the discussion in Section V, at each time $t \in \mathcal{T}$, we only need to derive the obedience constraint for the agent who is recommended to exert effort. The agent's decision to obey or disobey the recommendation depends on her belief about the system's state; this belief is constructed from the past messages she has received. Therefore, at each time t, we must investigate the obedience of agents with different histories, separately. At time $t=2$, the agent could have two different histories $h_2 = \{m_1 = n\}$ or $h_2 = \{m_1 = e\}$, while at the first time $t=1$, the agent always has an empty history $h_1 = \emptyset$. As a result, the principal is faced with three obedience constraints which we derive below.

First consider an agent in the last period $t=2$ with history $h_2 = \{m_1\}$. At this time the agent recognizes that there is no possibility of future costs or rewards, thus, similar to the static case, she takes a decision just based on her one-time utility. That is, she follows the recommendation to exert effort if the her posterior belief ν_2 exceeds the threshold ν^* defined by (8). Using Bayes rule, the posterior belief of the agent who has history $h_2 = \{m_1\}$ and is received message $m_2 = e$ at time $t=2$ is

$$\nu_2(m_1, m_2{=}e) = \frac{f(\nu_1(m_1))}{f(\nu_1(m_1)) + (1 - f(\nu_1(m_1)))\rho_2^{l,m_1}}, \tag{19}$$

where $\nu_1(m_1)$ is the agent's belief at time $t=1$ after she receives message m_1, i.e.

$$\nu_1(m_1) = \begin{cases} 0, & m_1 = n, \\ \frac{\mu_1}{\mu_1 + (1-\mu_1)\rho_1^l}, & m_1 = e, \end{cases} \tag{20}$$

and function $f(.)$, defined by (13), reflects the evolution of beliefs from time $t=1$ through the beginning of time $t=2$. Therefore, the obedience conditions for the agent with histories $h_2 = \{m_1 = n\}$ and $h_2 = \{m_1 = e\}$ who is asked to exert effort at time $t=2$ are, respectively,

$$\frac{q}{q + (1-q)\rho_2^{l,n}} \geq \nu^*, \tag{OC1}$$

and

$$\frac{f(\frac{\mu_1}{\mu_1 + (1-\mu_1)\rho_1^l})}{f(\frac{\mu_1}{\mu_1 + (1-\mu_1)\rho_1^l}) + (1 - f(\frac{\mu_1}{\mu_1 + (1-\mu_1)\rho_1^l}))\rho_2^{l,e}} \geq \nu^*. \tag{OC2}$$

Now consider the first period $t = 1$. At time $t = 1$ the agent follows the effort recommendation if exerting effort maximizes her expected continuation utility, i.e.

$$\mathbb{E}[U^A|x_1 = 1, m_1 = e] \geq \mathbb{E}[U^A|x_1 = 0, m_1 = e]. \quad (21)$$

Substituting (14) in (21), we obtain

$$(r+\delta\mathbb{E}[u^A(\theta_2, x_2)|\theta_1=1, \nu_1(e)])(d_l + \nu_1(e)(d_h-d_l))\geq c, \quad (22)$$

as a condition equivalent to (21), where

$$\mathbb{E}[u^A(\theta_2, x_2)|\theta_1 = 1, \nu_1(e)] = f(\nu_1(e))(r - c)+ \quad (23)$$
$$(1 - f(\nu_1(e)))\Big[\rho_2^{l,e}(r - c) + (1 - \rho_2^{l,e})(1 - d_l)r\Big],$$

is the second-period expected utility of the agent who is safe and has belief $\nu_1(e)$ at time $t = 1$, assuming that she follows the second-period recommendation. If the environment is high-risk at time $t = 2$ (prob. $f(\nu_1(e))$), the principal always recommends the agent to exert effort (cf. Eq. (18)), thus the agent pays protection cost c and collects safety reward r. If the environment is low-risk at time $t = 2$ (prob. $1-f(\nu_1(e))$), the agent is either recommended to exert effort (prob. $\rho_2^{l,e}$) and gets utility $r-c$, or is recommended to do nothing (prob. $1 - \rho_2^{l,e}$) and collects reward r if she remains safe (prob. $1 - d_l$). Substituting (20) and (23) into (22), we derive the obedience condition for the agent who is recommended to exert effort at $t = 1$ as follows:

$$(r + \delta\Big[\,f(\frac{\mu_1}{\mu_1 + (1 - \mu_1)\rho_1^l})(r - c)+$$
$$(1-f(\frac{\mu_1}{\mu_1+(1-\mu_1)\rho_1^l}))(\rho_2^{l,e}(r-c) + (1-\rho_2^{l,e})(1-d_l)r)\Big])\times$$
$$(d_l + \frac{\mu_1}{\mu_1 + (1 - \mu_1)\rho_1^l}(d_h - d_l)) \geq c. \quad \text{(OC3)}$$

The constraints OC1-OC3 derived above are sufficient to guarantee the obedience property of recommendation policy $\boldsymbol{\rho} = (\rho_1^l, 1, \rho_2^{l,n}, 1, \rho_2^{l,e}, 1)$. When the agent obeys the recommendation policy $\boldsymbol{\rho}$ the expected utility the principal gets is

$$\mathbb{E}[U^P|\boldsymbol{\rho}] = \sum_{t\in\mathcal{T}} \delta^{t-1}\mathbb{E}[\theta_t|\boldsymbol{\rho}] = \mu_1(1 + \delta)+$$
$$(1 - \mu_1)q\Big[\rho_1^l + (1 - \rho_1^l)(1 - d_l)\Big](1 + \delta)+$$
$$(1 - \mu_1)(1 - q)\Big[\rho_1^l(1 + \delta(\rho_2^{l,e} + (1 - \rho_2^{l,e})(1 - d_l)))+$$
$$(1 - \rho_1^l)(1 - d_l)(1 + \delta(\rho_2^{l,n}+(1 - \rho_2^{l,n})(1 - d_l)))\Big]. \quad (24)$$

Thus, the information design problem for the principal is

$$\max_{\rho_1^l,\rho_2^{l,n},\rho_2^{l,e}} \mathbb{E}[U^P|\boldsymbol{\rho} = (\rho_1^l, 1, \rho_2^{l,n}, 1, \rho_2^{l,e}, 1)],$$
$$\text{s.t.} \quad \textit{obedience constraints OC1-OC3}$$
$$0 \leq \rho^{l,h_t} \leq 1,\ \forall h_t \in \mathcal{H}_t, t \in \mathcal{T}. \quad (25)$$

The information incentive design problem (25) is a dynamic persuasion problem for the principal when both the principal and the agent have long-run incentives. In this problem, the recommendation rules at different time instants are coupled through the objective function and the obedience constraints. As a result, determination of the recommendation rules at different times must be done in a coordinated manner. We provide a step by step solution to the problem (25) below.

B. Specification of the optimal mechanism

We design a dynamic recommendation policy $\boldsymbol{\rho}^*$ that solves problem (25) and hence maximizes the principal's long-run utility. In problem (25), the decision variables ρ_1^l and $\rho_2^{l,e}$ are linked together by constraints OC2-OC3, however, the variable $\rho_2^{l,n}$ appears in just one constraint OC1 that is independent of the other variables. Therefore, using some algebra, we are able to decompose the principal's problem into two sub-problems that can be solved sequentially. The first problem

$$\max \rho_2^{l,n},$$
$$\textit{s.t. constraint OC1},$$
$$0 \leq \rho_2^{l,n} \leq 1, \quad \text{(SP1)}$$

maximizes the decision variable $\rho_2^{l,n}$ subject to the obedience constraint OC1 in which it is involved. The second problem

$$\max_{\rho_1^l,\rho_2^{l,e}} \rho_1^l(1+\delta q+\delta(1-q)(1-d_l)(1-\rho_2^{*\,l,n}))+\rho_1^l\rho_2^{l,e}\delta(1-q),$$
$$\textit{s.t. constraints OC2-OC3},$$
$$0 \leq \rho_1^l, \rho_2^{l,e} \leq 1, \quad \text{(SP2)}$$

uses SP1's optimal solution $\rho_2^{*\,l,n}$ to determine decision variables ρ_1^l and $\rho_2^{l,e}$ so as to maximize the principal's expected utility subject to obedience constraints OC2-OC3. In the rest of this section, we derive the optimal recommendation policy for the principal by solving these two sub-problems.

Sub-Problem SP1. The following lemma establishes the optimal solution to problem SP1.

Theorem 3: The optimal solution of SP1 is given by

$$\rho_2^{*\,l,n} = \min\left\{1,\ \frac{q(1 - \nu^*)}{\nu^*(1 - q)}\right\}. \quad (26)$$

Comparing the result of Theorem 3 with that of Theorem 1, we see that it is efficient for the principal to follow the history $m_1 = n$ with the optimal static recommendation policy at time $t{=}2$ where the agent has prior belief $\mu_2 = f(\nu_1(n)) = q$. This happens because the agent obeys the recommendation of exerting no effort at time $t{=}1$ irrespective of what happens in the future, therefore, $\rho_2^{l,n}$ has no effect on the agent's behavior in the first period and can be designed to maximize just the principal's second-period utility.

Sub-Problem SP2. In this problem, the principal uses the solution of SP1 to find the optimal values of ρ_1^l and $\rho_2^{l,e}$. The interaction of the decision variables ρ_1^l and $\rho_2^{l,e}$ through the obedience constraints OC2-OC3 is such that increasing one of them requires decreasing the other one. Therefore, the principal needs to find a balance between these two parameters so as to maximize his long-run expected utility.

It is clear that if the prior probability μ_1 the agent assigns to the environment's high-risk state is high enough to convince her to exert effort at both periods, the best strategy for the principal is to provide her with no information. The following theorem describes the situations when this trivial case happens.

Theorem 4: The principal's optimal strategy is to provide the agent with no-information (i.e. $\rho_1^{*l} = \rho_2^{*l,e} = 1$), if and only if the agent's prior belief μ_1 exceeds

$$\hat{\mu} := \max\left\{\frac{\nu^* - q}{1-q}, \frac{c/r(1+\delta d_l) - d_l(1+\delta)}{(d_h - d_l)(1+\delta(1-c/r))}\right\}. \quad (27)$$

In the rest of this section, we restrict attention to the non-trivial case where $\mu_1 < \hat{\mu}$. It can be shown that in this case, the feasible region of problem SP2 is one of the four areas shown in Fig. 2, where the dashed and solid lines represent the boundaries below which constraints OC2 and OC3 are satisfied, respectively, and the shaded area represents the feasible region. None of these areas is convex. Therefore, the conventional nonlinear programming algorithms cannot be used to find the solution. We begin deriving the optimal solution of problem SP2 by stating the following lemma.

Lemma 1: The optimal solution of problem SP2 lies on the boundary of its feasible region. Specifically, at least one of the constraints OC2 and OC3 are binding at optimality.

As a result of Lemma 1, to find the optimal solution of problem SP2, we restrict attention to the points at which at least one of the constraints OC2 and OC3 are binding. We study the behavior of the objective function of problem SP2 when each of these two constraints are binding in the two following lemmas: Lemmas 2 and 3.

Lemma 2: The objective function of SP2 is increasing in terms of ρ_1^l when constraint OC2 is binding.

Constraint OC2 is the obedience condition for the agent who is recommended to exert effort at time $t=2$ and has received $m_1 = e$ before. Lemma 2 shows that it is not optimal for the principal to sacrifice the agent's first-period security by reducing ρ_1^l so as to increase her trust in him and make it easier to convince her to exert effort at time $t=2$. This is intuitively expected, because the agent's safety in the first period is a prerequisite for her safety in the second period.

However, when the binding constraint is the agent's obedience at time $t = 1$, i.e. OC3, whether or not the principal benefits from sacrificing the second-period security by providing more honest information at time $t = 2$ (i.e. reducing $\rho_2^{l,e}$) so as to make the agent more hopeful about the future at time $t = 1$, depends on the agent's patience. If the agent has little patience, her decision at time $t = 1$ gives very little importance/weight to her utility at $t = 2$; thus, it is not advantageous for the principal to lower the second-period security to only slightly increase the chance of protection at time $t = 1$ (i.e. ρ_1^l). However, this sacrifice is the right thing to do if the agent is patient enough.

Lemma 3: When constraint OC3 is binding, there are two threshold levels $\underline{\delta}$ and $\bar{\delta}$ such that the objective function of SP2 is decreasing in terms of ρ_1^l when δ is below the lower threshold level $\underline{\delta}$, and is increasing in ρ_1^l when δ exceeds the higher threshold level $\bar{\delta}$.

Combining the results of Lemmas 2 and 3, with that of Lemma 1, leads to the following theorem.

Theorem 5: The optimal dynamic information disclosure mechanism is of threshold-type with respect to the discount factor δ with two threshold levels.

1) when δ is below the lower threshold level $\underline{\delta}$, the principal considers the security of both periods equally important and selects the intersection of the constraints OC2 and OC3 which results in point A of figures 2a-2d.
2) when δ exceeds the higher threshold level $\bar{\delta}$, the principal puts the priority on the agent's first-period security and maximizes ρ_1^l under the constraints which results in point B of the figures.

Theorem 5 establishes some qualitative properties of the optimal dynamic information disclosure mechanism. Using these properties, the principal can derive the optimal mechanism analytically. The analytical solution to the dynamic persuasion problem (25) is described in the Appendix.

VI. Numerical Results

In Theorem 5 we prove that the optimal mechanism is of threshold-type with two threshold levels $\underline{\delta}$ and $\bar{\delta}$ such that for $\delta < \underline{\delta}$ the principal gives the same priority to the agent's security in both periods and hence chooses point A, while for $\delta > \bar{\delta}$ he puts the whole priority on the agent's security in the first period and hence chooses point B. To investigate this result, we define a priority ratio R as follows

$$R = \frac{\textit{Priority of agent's second-period security}}{\textit{Priority of agent's first-period security}} = \frac{\rho_{B,1}^l - \rho_1^{*l}}{\rho_{B,1}^l - \rho_{A,1}^l},$$

where $\rho_{A,1}^l$ and $\rho_{B,1}^l$ are the values of ρ_1^l in points A and B, respectively. The priority ratio R is 1 if the principal selects point A and is 0 if he chooses B. In Fig. 3, the priority ratio R is plotted versus the discount factor δ for three sets of parameters $\mu_1, c/r, q, d_l, d_h$. This figure shows that in each case, there is a lower threshold below which $R = 1$ and hence the optimal point is A, while there is another threshold above which $R = 0$ and hence the optimal point is B.

In Fig. 4 we investigate the efficiency of the proposed dynamic information disclosure mechanism compared to no-information and full-information disclosure mechanisms, when the value of parameters is $\mu_1 = 0.1, c/r = 0.65, q = 0.1, d_l = 0.4, d_h = 0.82$. We observe that while the full-information and no-information mechanisms outperform each other in different regions, the optimal mechanism proposed in this paper always outperforms both.

References

[1] H. S. Witsenhausen, "On the structure of real-time source coders," *The Bell System Technical Journal*, vol. 58, no. 6, July 1979.

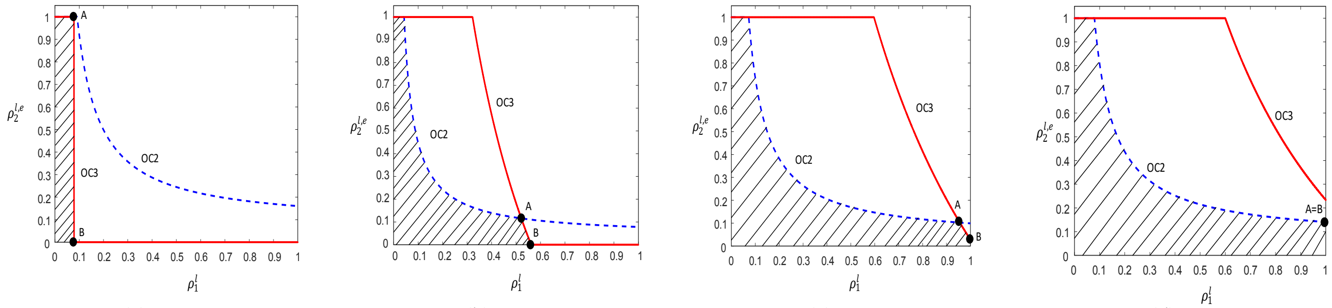
[2] J. Walrand and P. Varaiya, "Optimal causal coding - decoding problems," *IEEE Transactions on Information Theory*, Nov. 1983.

[3] A. Mahajan and D. Teneketzis, "Optimal design of sequential real-time communication systems," *IEEE Transactions on Information Theory*, vol. 55, no. 11, Nov 2009.

[4] V. P. Crawford and J. Sobel, "Strategic information transmission," *Econometrica*, vol. 50, no. 6, 1982.

[5] M. Battaglini, "Multiple referrals and multidimensional cheap talk," *Econometrica*, vol. 70, no. 4, 2002.

[6] S. Saritas, S. Yuksel, and S. Gezici, "On multi-dimensional and noisy quadratic signaling games and affine equilibria," in *ACC*, 2015.



(a) (b) (c) (d)

Fig. 2: Feasible region of problem SP2 (shaded area)

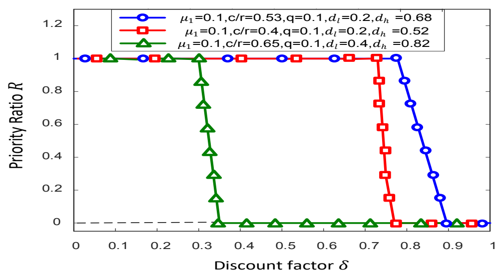


Fig. 3: Priority ratio R versus discount factor δ

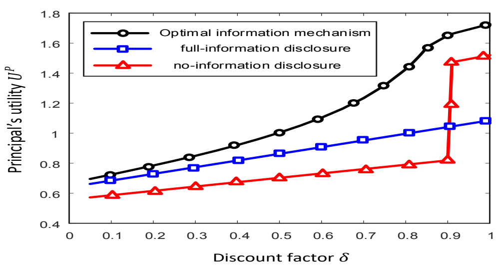


Fig. 4: Principal's utility U^P versus discount factor δ

[7] A. Ambrus and S. Takahashi, "Multi-sender cheap talk with restricted state spaces," *Theoretical Economics*, vol. 3, no. 1, 2008.

[8] M. Ivanov, "Dynamic information revelation in cheap talk," *The B.E. Journal of Theoretical Economics*, vol. 15, no. 2, July 2015.

[9] ——, "Dynamic learning and strategic communication," *International Journal of Game Theory*, vol. 45, no. 3, Aug 2016.

[10] E. Akyol, C. Langbort, and T. Basar, "Information-theoretic approach to strategic communication as a hierarchical game," *Proceedings of the IEEE*, vol. 105, no. 2, Feb 2017.

[11] E. Kamenica and M. Gentzkow, "Bayesian persuasion," *American Economic Review*, vol. 101, no. 6, October 2011.

[12] D. Bergemann and S. Morris, "Information design: A unified perspective," *Cowles Foundation Discussion Paper No. 2075R.*, 2017.

[13] W. Tamura, "A Theory of Multidimensional Information Disclosure," Institute of Social and Economic Research, Osaka University, ISER Discussion Paper 0828, Jan. 2012.

[14] H. Tavafoghi and D. Teneketzis, "Informational incentives in congestion games," in *Proc. of the Allerton*, Oct. 2017.

[15] F. Farokhi, A. M. H. Teixeira, and C. Langbort, "Estimation with strategic sensors," *IEEE Trans. on Automatic Control*, vol. 62, 2017.

[16] M. O. Sayin, E. Akyol, and T. Basar, "Strategic control of a tracking system," in *Proc. of the CDC*, Dec 2016.

[17] J. C. Ely, "Beeps," *American Economic Review*, Jan. 2017.

[18] H. Tavafoghi, "On design and analysis of cyber-physical systems with strategic agents," Ph.D. dissertation, University of Michigan, 2017.

[19] F. Farhadi, D. Teneketzis, and J. Golestani, "Static and dynamic informational incentive mechanisms for security enhancement," available at http://www-personal.umich.edu/%7Effarhadi/Information.pdf.

VII. APPENDIX

Define

$$g(x) := \frac{q(1-\nu^*)}{\nu^*(1-q)} + \frac{\mu_1(1-\nu^*)}{\nu^*(1-\mu_1)(1-q)x}, \tag{28}$$

and

$$h(x) := \frac{1-\delta q(c/r-d_l)+\delta(1-d_l)}{\delta(1-q)(c/r-d_l)} + \frac{\mu_1(1+\delta(1-c/r))}{\delta(1-q)(1-\mu_1)(c/r-d_l)x} - \frac{c/r(\mu_1+(1-\mu_1)x)^2}{\delta(1-q)(1-\mu_1)(c/r-d_l)x(d_l(1-\mu_1)x+\mu_1 d_h)}, \tag{29}$$

where $\rho_2^{l,e} = g(\rho_1^l)$ and $\rho_2^{l,e} = h(\rho_1^l)$ are the equations satisfied when constraints OC2 and OC3, respectively, are binding. By some algebra, it can be shown that each of the functions $h(x)$ and $h(x)-1$ has exactly one positive root. Let x^* and $\bar{x}$ denote the positive roots of functions $h(x)$ and $h(x)-1$, respectively. Moreover, it can be shown that functions $g(x)$ and $h(x)$ have a unique positive intersection, denoted as $\hat{x}$.

According to Theorem 5, the optimal solution of problem SP2 is a point between points A and B. Therefore, at first we characterize these two points based on the above-defined functions and parameters.

Point A**.**

$$(\rho_{A,1}^l, \rho_{A,2}^{l,e}) = \begin{cases} (\hat{x}, g(\hat{x})), & \hat{x} \le 1, g(\hat{x}) \le 1, \\ (\bar{x}, 1), & \hat{x} \le 1, g(\hat{x}) > 1, \\ (1, \min\{1, g(1)\}), & \hat{x} > 1. \end{cases} \tag{30}$$

Point B**.**

$$\rho_{B,1}^l = \min\{1, x^*\},\ \rho_{B,2}^{l,e} = \begin{cases} 0, & x^* < 1, \\ \min\{1, g(x^*), h(x^*)\}, & x^* \ge 1. \end{cases} \tag{31}$$

Now, we describe the solution $(\rho_1^{*l}, \rho_2^{*l,e})$ of problem SP2;

$$(\rho_1^{*l}, \rho_2^{*l,e}) = \begin{cases} (\rho_{A,1}^l, \rho_{A,2}^{l,e}), & \hat{\rho} \le \rho_{A,1}^l, \\ (\rho_{B,1}^l, \rho_{B,2}^{l,e}), & \hat{\rho} \ge \rho_{B,1}^l, \\ (\hat{\rho}, h(\hat{\rho})), & \rho_{A,1}^l < \hat{\rho} < \rho_{B,1}^l, \end{cases} \tag{32}$$

where

$$\hat{\rho} = -\frac{\mu_1 d_h}{d_l(1-\mu_1)} + \frac{(d_h-d_l)\mu_1}{d_l(1-d_l)(1-\mu_1)} \sqrt{\frac{c/r(1-d_l)}{(c/r-d_l)(1-\delta(1-q)(1-\rho_2^{*1,n}))-\delta}}, \tag{33}$$

if (33) returns a real value; otherwise $\hat{\rho} = 1$. Indeed, $\hat{\rho}$ is the point where the maximum of the objective function of SP2 is attained when the constraint OC3 is binding. This point is the optimum if it is between A and B; otherwise, one of the points A or B is optimal (See (32)).