Designing Incentive Schemes Based on Intervention: The Case of Imperfect Monitoring

Jaeok Park and Mihaela van der Schaar

University of California, Los Angeles (UCLA), Los Angeles CA 90095, USA

Abstract. In this paper, we propose a class of incentive schemes based on intervention. We develop a general game-theoretic framework for the design of intervention schemes under imperfect monitoring. We examine a model of slotted multiaccess communication to illustrate our framework. In this model, an intervention device monitors the behavior of agents for a period called the test phase and takes an intervention action which affects agents for the remaining period called the intervention phase. We analyze the problems of designing an optimal intervention rule given a length of the test phase and choosing an optimal length of the test phase. Intervention schemes can induce cooperative behavior by applying intervention following signals with a high likelihood of deviation. Increasing the length of the test phase has two counteracting effects: It improves the quality of signals, but at the same time it weakens the impact of intervention due to increased delay.

Key words: intervention, incentive schemes, slotted multiaccess communication, game theory

1 Introduction

In a system where selfish agents compete for available resources, it is common that the resources are not utilized optimally from a system-wide point of view. This calls for an incentive scheme to drive selfish agents towards the system objective. In this paper, we propose a class of incentive schemes, called intervention schemes. An intervention scheme can be implemented by augmenting the system with an intervention device that can monitor the behavior of agents and intervene in their resource usage. To analyze interaction under an intervention scheme, we formulate a class of games called intervention games and propose a solution concept called intervention equilibrium. In an intervention game, a system manager specifies an intervention rule used by the intervention device and the actions of agents. The intervention rule prescribes an action taken by the intervention device following each possible signal it can observe regarding the actions of agents. By increasing the intensity of intervention following signals that suggest a deviation from the specified actions, an intervention scheme can provide incentives for agents to follow the actions recommended by the manager. At an intervention equilibrium, the manager chooses a pair of an intervention rule and an action profile that optimizes the system objective while making the action profile self-enforcing given the intervention rule.

The idea of using intervention to provide incentives can be found in [1], [2]. Garg et al. [1] consider a congestion control problem and analyze performance under different scheduling mechanisms, which can be considered as intervention schemes. With a scheduling mechanism that assigns higher priority to flows with a smaller rate, the input rates of users can be restrained voluntarily. In other words, users do not send their traffic excessively in their self-interest because doing so will increase the probability that their packets are dropped. In our previous work [2], we consider a slotted multiaccess communication network with an intervention device that can jam the packets of users. We showed that intervention schemes can successfully regulate the transmission probabilities of selfish users in the case of perfect monitoring, where the intervention device observes the transmission probabilities of users. In the current paper, we provide a general framework for the design of intervention schemes, unlike the two previous papers that focus on specific communication scenarios. Then we illustrate our framework with a model of slotted multiaccess communication, as considered in [2], while relaxing the monitoring requirement. That is, we consider the case of imperfect monitoring, where the intervention device obtains only imperfect information about the transmission probabilities of users.

In an intervention game, the manager chooses an intervention rule before agents take their actions, and thus the manager can be considered as a leader and agents as followers. Such a hierarchical structure in the interaction between the leader and the followers has been analyzed using a Stackelberg game in the literature (see, for example, [3]–[5]). In [3]–[5], the leader uses a Stackelberg strategy, simply taking an action before the followers take theirs. In contrast, in an intervention game, the manager chooses an intervention rule, which is a complete contingent plan for intervention actions to be taken given each possible signal about the actions of agents. Thus, an intervention rule requires more overhead in terms of monitoring capability for implementation than a Stackelberg strategy. However, when the manager does not value his resource usage, intervention schemes have an advantage over Stackelberg strategies. In intervention schemes, intervention actions can be adjusted to the observed behavior of agents, and thus intervention can be applied only when it is necessary. On the contrary, Stackelberg strategies lack such adaptivity.

Pricing offers an alternative means that the manager can use to provide incentives. In pricing schemes, the manager collects payments from agents based on their resource usage. Due to its foundations in market economies, pricing has received a significant amount of attention in the literature (see, for example, [6], [7]). The main difference between intervention and pricing is that intervention affects the resource usage of agents inside of the system whereas pricing affects the payoffs of agents through an outside instrument. Intervention schemes are robust in the sense that agents cannot avoid intervention as long as they use resources in the system, while pricing is not effective if agents can evade payments while still using resources. Moreover, even in the case where payments can be enforced technologically, there are situations where pricing is undesirable due to policy considerations. For example, it has been argued in a public policy debate that access to the Internet should be provided as a public good by a public authority rather than as a private good in a market [8]. In addition, intervention schemes have informational advantages over pricing schemes. Intervention affects the payoffs of agents through physical quantities (e.g., data rates and delay) that agents care about, and the impacts of intervention actions on such physical quantities are relatively easy to measure. On the contrary, finding out the impact of pricing on agents requires the manager to know agents' subjective valuation of payments relative to physical quantities, which is difficult to measure. Below we summarize related work discussed so far in Table 1.

Table 1. Summary of related work

Intervention	Stackelberg strategy	Pricing
[1]: congestion control [2]: medium access control	[3]: congestion control [4]: medium access control [5]: power control	[6]: congestion control [7]: cognitive radio

In game theory, repeated games have been studied as a method to provide incentives in a long-run relationship [9]. In a repeated game, agents monitor the actions of other agents and choose their actions depending on their past observations. Hence, the burden of monitoring and executing reward and punishment is distributed to agents in a repeated game, while it is imposed solely on the intervention device in an intervention game. In other words, compared to repeated game strategies, intervention schemes require a central infrastructure in the system while reducing a burden on agents. We need repeated interaction among agents to apply a repeated game strategy, while intervention schemes are applicable to systems with a dynamically changing agent population (e.g., mobile networks). In the repeated game literature, reputation schemes are used to overcome limited observation due to infrequent interaction (see, for example, [10]). A reputation scheme usually has a central infrastructure, but its role is limited to collecting observations from agents, processing the collected observations, and disseminating the processed information.

An active area of research that focuses on designing incentive schemes is mechanism design. In a standard mechanism design problem [11, Chap. 23], the manager can decide an outcome that agents care about but has incomplete information about the types of agents. Relying on the revelation principle, standard mechanism design is concerned about incentives for agents to reveal their types truthfully. In contrast, our focus in designing intervention schemes is on providing incentives for agents to choose desirable actions. In this paper, we assume that the manager has complete information, knowing the action spaces and the payoff functions of agents. We can extend our framework to study the design of intervention schemes when the manager has incomplete information. We leave this extension for future research.

The remainder of the paper is organized as follows. In Sect. 2, we develop a game-theoretic framework for the design of intervention schemes. In Sect. 3, we apply the framework to a model of slotted multiaccess communication and analyze optimal intervention schemes. In Sect. 4, we conclude.

2 Framework for the Design of Intervention Schemes

We consider a system (e.g., a communication network) where N agents and an intervention device interact. The set of the agents is denoted by $\mathcal{N} = \{1, \ldots, N\}$. The action space of agent i is denoted by A_i, and an action for agent i is denoted by $a_i \in A_i$, for all $i \in \mathcal{N}$. An action profile is represented by a vector $a = (a_1, \ldots, a_N) \in A := \prod_{i \in \mathcal{N}} A_i$. An action profile of the agents other than agent i is written as $a_{-i} := (a_1, \ldots, a_{i-1}, a_{i+1}, \ldots, a_N)$ so that a can be expressed as $a = (a_i, a_{-i})$. Once an action profile of the agents is chosen, a signal is realized from a finite set of signals, denoted by Y, and is observed by the intervention device. The probability that a signal $y \in Y$ is realized given an action profile $a \in A$ is denoted by $\rho(y|a)$. After observing the realized signal, the intervention device takes an action, called an intervention action. We use a_0 and A_0 to denote an intervention action and the action space of the intervention device, respectively.

Since the intervention device chooses its action after observing a signal, a strategy for it can be represented by a mapping $f : Y \to A_0$, which is called an *intervention rule*. The set of all possible intervention rules is denoted by $\mathcal{F}$. There is a system manager who determines the intervention rule used by the intervention device. We assume that the manager can commit to an intervention rule, for example, by using a protocol embedded in the intervention device. The payoffs of the agents and the manager are determined by the intervention action, the action profile of the agents, and the realized signal. Thus, we denote the payoff function of agent $i \in \mathcal{N}_0 := \mathcal{N} \cup \{0\}$ by $u_i : A_0 \times A \times Y \to \mathbb{R}$. The payoff of the manager, u_0, can be interpreted as a measure of system performance. The game played by the manager and the agents is formulated as an *intervention game*, which is summarized by the data

$$\Gamma = \langle \mathcal{N}_0, (A_i)_{i \in \mathcal{N}_0}, (u_i)_{i \in \mathcal{N}_0}, (Y, \rho) \rangle \ . \tag{1}$$

Given an intervention rule and an action profile, expected payoffs can be computed by taking expectations of payoffs over signals. The expected payoff function of agent $i \in \mathcal{N}_0$ is denoted by a function $v_i : \mathcal{F} \times A \to \mathbb{R}$, which can be computed as

$$v_i(f, a) = \sum_{y \in Y} \rho(y|a) u_i(f(y), a, y) \ . \tag{2}$$

Once the manager chooses an intervention rule f, the agents play a simultaneous game, whose normal form representation is given by

$$\Gamma_f = \langle \mathcal{N}, (A_i)_{i \in \mathcal{N}}, (v_i(f, \cdot))_{i \in \mathcal{N}} \rangle \ . \tag{3}$$

We predict actions chosen by the agents given an intervention rule f by applying the solution concept of Nash equilibrium [12] to the induced game Γ_f.

Definition 1. *An intervention rule* $f \in \mathcal{F}$ sustains *an action profile* $a^* \in A$ *if* a^* *is a Nash equilibrium of the game* Γ_f*, i.e.,*

$$v_i(f, a_i^*, a_{-i}^*) \geq v_i(f, a_i, a_{-i}^*) \quad \forall a_i \in A_i \ , \ \forall i \in \mathcal{N} \ . \tag{4}$$

Let $\mathcal{E}(f) \subseteq A$ be the set of action profiles sustained by f. The manager aims to maximize his expected payoff by specifying an intervention rule for the intervention device and an action profile for the agents sustained by the intervention rule. The manager's problem leads to the following solution concept for intervention games.

Definition 2. $(f^*, a^*) \in \mathcal{F} \times A$ *is an* intervention equilibrium *if* $a^* \in \mathcal{E}(f^*)$ *and*

$$v_0(f^*, a^*) \geq v_0(f, a) \quad \textit{for all } (f, a) \textit{ such that } a \in \mathcal{E}(f) \ . \tag{5}$$

$f^* \in \mathcal{F}$ *is an* optimal intervention rule *if there exists an action profile* $a^* \in A$ *such that* (f^*, a^*) *is an intervention equilibrium.*

An intervention equilibrium solves the following optimization problem:

$$\max_{f \in \mathcal{F}} \max_{a \in \mathcal{E}(f)} v_0(f, a) \ . \tag{6}$$

An intervention equilibrium can be considered as a subgame perfect equilibrium (or Stackelberg equilibrium) of an intervention game, with an implicit assumption that the manager can induce the agents to choose the best Nash equilibrium for him in case of multiple Nash equilibria. In our interpretation, the manager specifies an intervention equilibrium (f^*, a^*) to the agents so that a^* becomes a self-enforcing operating point for them given the intervention rule f^*. In other words, the manager can make a^* a focal point [12] of the game Γ_{f^*} by recommending it to the agents.

3 Intervention Schemes in a Slotted Multiaccess Communication Network

3.1 Model

There is a communication channel shared by N users. Time is divided into slots of equal length, and packets have a length that can be transmitted within a time slot. In each time slot, a user can attempt to transmit its packet or wait. If there is only one transmission attempt in a slot, the packet is successfully transmitted. If there is more than one transmission attempt in a slot, packets collide and no transmission is successful. For simplicity, we assume that each user can choose one of two transmission probabilities p_l and p_h, where $p_\mathrm{l} = 1/N < p_\mathrm{h} < 1$. Note that each user choosing p_l maximizes the total throughput, defined as the average number of successfully transmitted packets per time slot, assuming that all the users choose the same transmission probability [13].

We consider a period consisting of T consecutive time slots, and analyze interaction in the period without any consideration of past or future periods. We assume that the number of users and their transmission probabilities are fixed throughout a period. Then the action space of user i is given by $A_i = \{p_{\mathrm{l}}, p_{\mathrm{h}}\}$, for all $i \in \mathcal{N}$. The payoff of user i is given by the number of its successfully transmitted packets per time slot. Then the expected payoff of user i is given by the probability of its successful transmission, $a_i \prod_{j \in \mathcal{N} \setminus \{i\}} (1 - a_j)$. It is easy to see that the action p_{h} is a dominant strategy for every user. Hence, $(p_{\mathrm{h}}, \ldots, p_{\mathrm{h}})$ is the unique Nash equilibrium, which yields a lower total throughput than the symmetric social optimum $(p_{\mathrm{l}}, \ldots, p_{\mathrm{l}})$.

In order to improve the inefficiency of Nash equilibrium, we introduce an intervention device in the system. The intervention device is capable of monitoring the actions of the users and interfering in the transmission of the users. In each slot, the intervention device can sense the channel to learn whether the channel is idle (i.e., no user attempts to transmit its packet) or busy (i.e., at least one user attempts to transmit its packet). After observing channel states for the first t slots, where $1 \leq t \leq T$, the intervention device chooses its transmission probability, which remains fixed until the end of the period. We assume that, unlike the users, the intervention device can choose any transmission probability in $[0, 1]$. Thus, we have $A_0 = [0, 1]$. The first t slots and the remaining $(T - t)$ slots are called the test phase and the intervention phase, respectively.

Let $S = \{\text{idle}, \text{busy}\}$ be the set of channel states that can be observed in a slot. Then the set of all possible signals that the intervention device can obtain in the test phase is S^t. Since the transmission probabilities of the users do not change in a period, there is no gain for the manager to treat channel states from different slots differently. Hence, we focus on the class of intervention rules that use only the number of idle slots, which allows us to work with a smaller signal space $Y = \{0, 1, \ldots, t\}$ instead of S^t. Channel states are independent across slots, and the probability of an idle state in a slot given an action profile a is given by $q(a) := \prod_{i \in \mathcal{N}} (1 - a_i)$. Thus, the probability that k idle slots arise in the test phase is $\rho(k|a) = \binom{t}{k} q(a)^k (1 - q(a))^{t-k}$, for $k = 0, 1, \ldots, t$. Note that monitoring is imperfect in that the intervention device cannot observe the action profile of the users but obtains only imperfect information about the action profile.

The sequence of events in a period can be listed as follows.

1. At the beginning of the period, the users choose their transmission probabilities $a \in A$, which are used from slot 1 to slot T, knowing the intervention rule f adopted by the intervention device.
2. The intervention device collects observations of channel states from slot 1 to slot t (test phase).
3. The intervention device intervenes using the transmission probability prescribed by the intervention rule f from slot $t + 1$ to slot T (intervention phase).

The payoff of user i given intervention action a_0 and action profile a conditional on signal k (i.e., k idle slots in the test phase) is given by

$$
\begin{aligned}
& u_i(a_0, a, k) && (7)\\
&= \frac{t-k}{T} \frac{a_i \prod_{j \in \mathcal{N} \setminus \{i\}} (1-a_j)}{1-q(a)} + \frac{T-t}{T} a_i (1-a_0) \prod_{j \in \mathcal{N} \setminus \{i\}} (1-a_j) && (8)\\
&= \left[\frac{t-k}{T} \frac{1}{1-q(a)} + \frac{T-t}{T} (1-a_0) \right] a_i \prod_{j \in \mathcal{N} \setminus \{i\}} (1-a_j) \ . && (9)
\end{aligned}
$$

The expected payoff of user i given intervention rule f and action profile a can be computed using (2):

$$
v_i(f, a) = \left[1 - \frac{T-t}{T} \sum_{k=0}^{t} \binom{t}{k} q(a)^k (1-q(a))^{t-k} f(k) \right] a_i \prod_{j \in \mathcal{N} \setminus \{i\}} (1-a_j) \ . \quad (10)
$$

Note that $\sum_{k=0}^{t} \binom{t}{k} q(a)^k (1-q(a))^{t-k} f(k)$ is the expected transmission probability of the intervention device while $(T-t)/T$ is the weight on the intervention phase.[1]

3.2 Formulation of the Design Problem

For notation, let us define

$$
\lambda(k;t) = \binom{t}{k} \left[(1-p_\mathrm{l})^N \right]^k \left[1 - (1-p_\mathrm{l})^N \right]^{t-k} \ , \quad (11)
$$

$$
\mu(k;t) = \binom{t}{k} \left[(1-p_\mathrm{l})^{N-1} (1-p_\mathrm{h}) \right]^k \left[1 - (1-p_\mathrm{l})^{N-1} (1-p_\mathrm{h}) \right]^{t-k} \ , \quad (12)
$$

for $k = 0, 1, \ldots, t$, and let $\tau_\mathrm{c} = p_\mathrm{l} (1-p_\mathrm{l})^{N-1}$ and $\tau_\mathrm{d} = p_\mathrm{h} (1-p_\mathrm{l})^{N-1}$. $\lambda(k;t)$ is the probability of k idle slots arising in the t slots of the test phase when every user cooperates (i.e., chooses p_l), while $\mu(k;t)$ is that when exactly one user defects (i.e., chooses p_h). τ_c is the cooperation throughput that each user obtains when all the users choose p_l, while τ_d is the defection throughput that a user obtains when it deviates to p_h unilaterally. Note that an idle slot is more likely to occur when every user cooperates than when some user defects. Also, note that $\tau_\mathrm{d} > \tau_\mathrm{c}$, which reflects the positive gain from defection when there is no intervention.

Suppose that the objective of the manager is to maximize the sum of the payoffs (i.e., total throughput) while sustaining cooperation among the users. Formally, the payoff function of the manager can be written as

$$
u_0(a_0, a, k) = \begin{cases} \sum_{i \in \mathcal{N}} u_i(a_0, a, k) \ , & \text{if } a_i = p_\mathrm{l} \ \forall i \in \mathcal{N} \ , \\ -\infty \ , & \text{otherwise} \ . \end{cases} \quad (13)
$$

[1] For simplicity, we assume that the users value a successful transmission equally across slots. Introducing time discounting in the model will make the weight on the intervention phase smaller, since a successful transmission in later slots yields less value than that in earlier slots.

Provided that other users cooperate, the payoff to a cooperating user is given by

$$\left[1 - \frac{T-t}{T} \sum_{k=0}^{t} \lambda(k;t) f(k)\right] \tau_{\mathrm{c}} \ , \tag{14}$$

while that to a defecting user is

$$\left[1 - \frac{T-t}{T} \sum_{k=0}^{t} \mu(k;t) f(k)\right] \tau_{\mathrm{d}} \ . \tag{15}$$

Hence, the incentive constraint to sustain cooperation can be written as

$$\left[1 - \frac{T-t}{T} \sum_{k=0}^{t} \lambda(k;t) f(k)\right] \tau_{\mathrm{c}} \geq \left[1 - \frac{T-t}{T} \sum_{k=0}^{t} \mu(k;t) f(k)\right] \tau_{\mathrm{d}} \ , \tag{16}$$

and the problem of designing an optimal intervention rule can be expressed as

$$\max_f N \left[1 - \frac{T-t}{T} \sum_{k=0}^{t} \lambda(k;t) f(k)\right] \tau_{\mathrm{c}} \tag{17}$$

subject to

$$\left[1 - \frac{T-t}{T} \sum_{k=0}^{t} \lambda(k;t) f(k)\right] \tau_{\mathrm{c}} \geq \left[1 - \frac{T-t}{T} \sum_{k=0}^{t} \mu(k;t) f(k)\right] \tau_{\mathrm{d}} \ , \tag{18}$$

$$0 \leq f(k) \leq 1 \quad \forall k = 0, \ldots, t \ . \tag{19}$$

3.3 Analysis of the Design Problem

The design problem (17)–(19) can be rewritten as a linear programming (LP) problem as follows:

$$\min_f \sum_{k=0}^{t} \lambda(k;t) f(k) \tag{20}$$

$$\text{subject to } \frac{T-t}{T} \sum_{k=0}^{t} \left[\tau_{\mathrm{d}} \mu(k;t) - \tau_{\mathrm{c}} \lambda(k;t)\right] f(k) \geq \tau_{\mathrm{d}} - \tau_{\mathrm{c}} \ , \tag{21}$$

$$0 \leq f(k) \leq 1 \quad \forall k = 0, \ldots, t \ . \tag{22}$$

The LP problem (20)–(22) is to minimize the expected transmission probability of the intervention device while satisfying the incentive constraint and the probability constraints. Exerting intervention following some signals is necessary to punish a deviation, but at the same time intervention incurs efficiency loss under imperfect monitoring. Thus, the manager wants to use the minimum possible intervention level while providing the incentive for cooperation. The left-hand side of the incentive constraint (21) is the expected loss from deviation due to the change in the probability distribution of signals induced by deviation, while the right-hand side is the gain from deviation.

Lemma 1. *Suppose that an optimal solution to the LP problem (20)–(22) exists. Then the incentive constraint (21) is satisfied with equality at the optimal solution.*

Proof. Let f^* be an optimal solution. Suppose that $[(T-t)/T]\sum_{k=0}^{t}[\tau_{\mathrm{d}}\mu(k;t)-\tau_{\mathrm{c}}\lambda(k;t)]f^*(k) > \tau_{\mathrm{d}}-\tau_{\mathrm{c}}$. Since $\tau_{\mathrm{d}} > \tau_{\mathrm{c}}$, there exists k' such that $\tau_{\mathrm{d}}\mu(k';t)-\tau_{\mathrm{c}}\lambda(k';t) > 0$ and $f^*(k') > 0$. Then we can reduce $f^*(k')$ while satisfying the incentive constraint and the probability constraint for k', which decreases the objective value since $\lambda(k;t) > 0$ for all k. This contradicts the optimality of f^*. □

Lemma 1 validates the intuition that the manager wants to use a punishment just enough to prevent deviation. The following proposition provides a necessary and sufficient condition for the LP problem to have a feasible solution, and the structure of an optimal solution.

Proposition 1. *Let* $k_0 = \max\{k : \tau_{\mathrm{d}}\mu(k;t)-\tau_{\mathrm{c}}\lambda(k;t) > 0\}$. *Then the LP problem has a feasible solution if and only if*

$$\frac{T-t}{T}\sum_{k\leq k_0}[\tau_{\mathrm{d}}\mu(k;t)-\tau_{\mathrm{c}}\lambda(k;t)] \geq \tau_{\mathrm{d}}-\tau_{\mathrm{c}}\ . \tag{23}$$

Moreover, if the LP problem has a feasible solution, then there exists a unique optimal solution f^* described by

$$f^*(k) = \begin{cases} 1\ , & \textit{if } k < \bar{k}\ , \\ \frac{1}{\tau_{\mathrm{d}}\mu(\bar{k};t)-\tau_{\mathrm{c}}\lambda(\bar{k};t)}\left[\frac{T}{T-t}(\tau_{\mathrm{d}}-\tau_{\mathrm{c}})-\sum_{k=0}^{\bar{k}-1}[\tau_{\mathrm{d}}\mu(k;t)-\tau_{\mathrm{c}}\lambda(k;t)]\right]\ , & \textit{if } k = \bar{k}\ , \\ 0\ , & \textit{if } k > \bar{k}\ , \end{cases} \tag{24}$$

where

$$\bar{k} = \min\left\{k' : \frac{T-t}{T}\sum_{k\leq k'}[\tau_{\mathrm{d}}\mu(k;t)-\tau_{\mathrm{c}}\lambda(k;t)] \geq \tau_{\mathrm{d}}-\tau_{\mathrm{c}}\right\}\ . \tag{25}$$

Proof. Define the likelihood ratio of signal k by

$$L(k;t) = \frac{\mu(k;t)}{\lambda(k;t)} = \left(\frac{1-p_{\mathrm{h}}}{1-p_{\mathrm{l}}}\right)^{k}\left(\frac{1-(1-p_{\mathrm{l}})^{N-1}(1-p_{\mathrm{h}})}{1-(1-p_{\mathrm{l}})^{N}}\right)^{t-k}\ . \tag{26}$$

It is easy to see that $L(0;t) > 1$, $L(t;t) < 1$, and $L(k;t)$ is monotonically decreasing in k. Note that $\tau_{\mathrm{d}}\mu(k;t)-\tau_{\mathrm{c}}\lambda(k;t) > 0$ if and only if $L(k;t) > p_{\mathrm{l}}/p_{\mathrm{h}}$. Hence, k_0 is well-defined, and $\tau_{\mathrm{d}}\mu(k;t)-\tau_{\mathrm{c}}\lambda(k;t) > 0$ if and only if $k \leq k_0$. If (23) is satisfied, then $\tilde{f}$ defined by $\tilde{f}(k) = 1$ for all $k \leq k_0$ and $\tilde{f}(k) = 0$ for all $k > k_0$ is a feasible solution. To prove the converse, suppose that a feasible solution, say f, exists. Then we have

$$\frac{T-t}{T}\sum_{k\le k_0}[\tau_{\rm d}\mu(k;t)-\tau_{\rm c}\lambda(k;t)] \ge \frac{T-t}{T}\sum_{k=0}^{t}[\tau_{\rm d}\mu(k;t)-\tau_{\rm c}\lambda(k;t)]f(k) \tag{27}$$

and

$$\frac{T-t}{T}\sum_{k=0}^{t}[\tau_{\rm d}\mu(k;t)-\tau_{\rm c}\lambda(k;t)]f(k) \ge \tau_{\rm d}-\tau_{\rm c}\ , \tag{28}$$

and combining the two yields (23).

To prove the remaining result, suppose that the LP problem has a feasible solution. Then there exists a feasible solution, say f, that satisfies the incentive constraint with equality. Define the likelihood ratio of f by

$$l(f)=\frac{\sum_k \mu(k;t)f(k)}{\sum_k \lambda(k;t)f(k)}\ . \tag{29}$$

Then the objective value in (20) at f can be expressed as

$$\frac{T}{T-t}\frac{\tau_{\rm d}-\tau_{\rm c}}{\tau_{\rm d}l(f)-\tau_{\rm c}}\ . \tag{30}$$

Hence, the objective value decreases as f has a larger likelihood ratio. To optimize the objective function, f should put the probabilities on the signals starting from signal 0 to signal 1, and so on, until the incentive constraint is satisfied with equality. Thus, we obtain $\bar{k}$ in (25), where $0 \le \bar{k} \le k_0$, associated with the unique optimal solution. □

Since a smaller number of idle slots gives a higher likelihood ratio, an intervention rule yields a smaller efficiency loss when intervention is exerted following a smaller number of idle slots. Put differently, signal k provides a stronger indication of defection as k is smaller. However, using only signal 0 may not be sufficient to provide the incentive for cooperation, in which case other signals need to be used as well. Using signal k with $k \le k_0$ contributes to the incentive for cooperation, although the "quality" of the signal decreases as k increases. Hence, it is optimal for the manager to use signals with smaller k primarily, which yields a threshold $\bar{k}$.

Timing of Intervention. So far we have analyzed the problem of designing an optimal intervention rule given the length of the test phase, t. Now we consider a scenario where the manager can choose a length of the test phase as well as an intervention rule. In this scenario, there are two counteracting effects of increasing t. First, note that the objective function in (17) can be expressed as

$$N\left[1-\frac{\tau_{\rm d}-\tau_{\rm c}}{\tau_{\rm d}l(f)-\tau_{\rm c}}\right]\tau_{\rm c}\ , \tag{31}$$

which shows that increasing t affects the objective value only through f. Since $L(k;t_1) > L(k;t_2)$ for all t_1, t_2 such that $t_1 > t_2$ and for all $k \le t_2$, increasing t increases the likelihood ratios of existing signals. At the same time, it adds new signals available for the manager. Thus, a larger likelihood ratio $l(f)$ can be achieved with larger t. In other words, as the intervention device collects more observations, the information about deviation becomes more accurate (quality effect). On the other hand, increasing t decreases the weight given on the intervention phase, which makes the impact of intervention weaker and the incentive constraint harder to satisfy (delay effect).

Let $\tau^*(t)$ be the optimal value of the design problem (17)–(19) with the length of the test phase t, where we set $\tau^*(t) = Np_{\mathrm{h}}(1-p_{\mathrm{h}})^{N-1}$ if there is no feasible solution with t. The problem of finding an optimal length of the test phase can be written as $\max_{t\in\{1,\ldots,T\}} \tau^*(t)$. In general, $\tau^*(t)$ is a non-monotonic function of t, and we provide a numerical example to illustrate non-monotonicity. We consider system parameters $N = 5$, $p_{\mathrm{l}} = 1/N = 0.2$, $p_{\mathrm{h}} = 0.8$, and $T = 100$. Then we have $\tau_{\mathrm{c}} = 0.08$ and $\tau_{\mathrm{d}} = 0.33$. The numerical results show that the LP problem is infeasible for $t = 1$ and $t \geq 21$. With $t = 1$, there is not sufficient information based on which intervention can provide the incentive for cooperation. With $t \geq 21$, the delay effect is too strong to have the incentive constraint satisfied. Figure 1 plots $\tau^*(t)$ for $t = 2,\ldots,20$. We can see that $\tau^*(t)$ is non-monotonic while reaching the maximum at $t = 18$ with $\tau^*(18) = 0.37$. In the plot, the dotted line represents the total throughput at $(p_{\mathrm{l}},\ldots,p_{\mathrm{l}})$, $N\tau_{\mathrm{c}}$. The difference between $\tau^*(t)$ and $N\tau_{\mathrm{c}}$ can be interpreted as the efficiency loss due to imperfect monitoring.[2] Lastly, we note that $\bar{k}$ in Proposition 1 is non-decreasing in t, with $\bar{k} = 1$ for $t = 2,\ldots,7$, $\bar{k} = 2$ for $t = 8,\ldots,13$, $\bar{k} = 3$ for $t = 14,\ldots,18$, and $\bar{k} = 4$ for $t = 19, 20$.

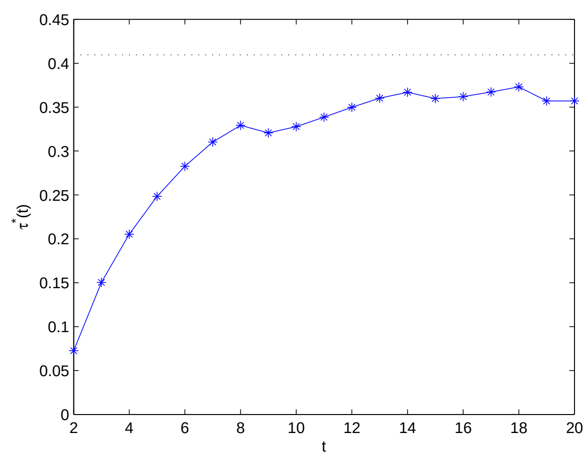

Fig. 1. Plot of $\tau^*(t)$ for $t = 2,\ldots,20$

[2] If the intervention device can observe the actions of the users immediately (i.e., perfect monitoring), it can sustain cooperation without incurring an efficiency loss by using the threat of transmitting with probability 1 when a deviation is detected.

4 Conclusion

In this paper, we have proposed a class of incentive schemes, called intervention schemes. We have presented a general game-theoretic framework for the design of intervention schemes under imperfect monitoring. In order to illustrate our framework and obtain concrete results, we have analyzed a simple model of slotted multiaccess communication. Our results suggest that we can design an intervention scheme that sustains an action profile from which a deviation yields a sufficiently distinct distribution of signals. When the manager cares about efficiency, it is optimal to use a punishment just enough to prevent deviation, in order to minimize the efficiency loss due to imperfect monitoring. Also, in a scenario where the manager can decide the timing of intervention, we have identified the two counteracting effects of having a longer test phase. Our framework of intervention schemes can be potentially applied to any application scenario in which individual objectives are in conflict with the system objective and some monitoring is possible. Investigating intervention schemes in various settings will provide us with insights into the properties of optimal intervention schemes as well as the capabilities and limitations of intervention schemes.

References

1. Garg R., Kamra, A., Khurana, V.: A game-theoretic approach towards congestion control in communication networks. Comput. Commun. Rev. **32** (2002) 47–61
2. Park, J., van der Schaar, M.: Stackelberg contention games in multiuser networks. EURASIP J. Adv. Sig. Process. Article ID 305978 (2009) 15 pages
3. Korilis, Y. A. , Lazar, A. A., Orda, A.: Achieving network optima using Stackelberg routing strategies. IEEE/ACM Trans. Netw. **5** (1997) 161–173
4. Ma, R. T., Misra, V., Rubenstein, D.: An analysis of generalized slotted-Aloha protocols. IEEE/ACM Trans. Netw. **17** (2009) 936–949
5. Su, Y., van der Schaar, M.: A new perspective on multi-user power control games in interference channels. IEEE Trans. Wireless Commun. **8** (2009) 2910–2919
6. Başar, T., Srikant, R.: A Stackelberg network game with a large number of followers. J. Opt. Theory Applicat. **115** (2002) 479–490
7. Bloem, M., Alpcan, T., Başar, T.: A Stackelberg game for power control and channel allocation in cognitive radio networks. In: Proc. ValueTools (2007)
8. Hallgren, M. M., McAdams, A. K.: The economic efficiency of Internet public goods. In: McKnight, L. W., Bailey, J. P. (eds.) Internet Economics. MIT Press, Cambridge (1997) 455–478
9. Mailath, G., Samuelson, L.: Repeated Games and Reputations: Long-run Relationships. Oxford Univ. Press, Oxford (2006)
10. Kandori, M.: Social norms and community enforcement. Rev. Econ. Stud. **59** (1992) 63–80
11. Mas-Colell, A., Whinston, M. D., Green, J. R.: Microeconomic Theory. Oxford Univ. Press, Oxford (1995)
12. Fudenberg, D., Tirole, J.: Game Theory. MIT Press, Cambridge (1991)
13. Massey, J. L., Mathys, P.: The collision channel without feedback. IEEE Trans. Inf. Theory **31** (1985) 192–204