Pricing and Selection of Channels for Remote State Estimation Using a Stackelberg Game Framework

Yuqing Ni, Alex S. Leong, Daniel E. Quevedo, *Senior Member, IEEE*, and Ling Shi, *Senior Member, IEEE*

***Abstract*—We consider the communication channel pricing and selection problem in a networked control system. To encompass the sequentialized nature of the decision-making process, we use game theory and formulate a Stackelberg game framework, where the server first determines the channel pricing strategy, and the clients then make channel selection decisions. Both single-server-single-client (SSSC) scenario and single-server-multi-client (SSMC) scenario are discussed. The existence of an optimal stationary and deterministic policy for the clients is proved. We show that for the SSSC scenario, the server's optimal pricing strategy in terms of maximizing revenue is to ensure that the client uses the good channel all the time. For the SSMC scenario, it is assumed that the channel price remains invariant. As a consequence, each client has an optimal policy with threshold structure. Some properties of the optimal policy pair for both scenarios are obtained. Simulation results confirm the structure and properties of both the server and clients' optimal strategies.**

***Index Terms*—Networked control system, Stackelberg game, Markov decision process, pricing mechanism.**

I. Introduction

IN VARIOUS control systems, such as manufacturing industries, aerospace vehicles, public transportation and intelligent homes, a real-time network is utilized to exchange information between sensors and estimators, and controllers and actuators [1]–[3]. A control loop is called a networked control system (NCS) when it is closed through wireless communication channels [4]. Compared to traditional wired systems, NCSs are well known for a lower cost in installment, diagnosis, debugging and maintenance.

For the implementation of NCSs, feasible candidate networks are DeviceNet, Ethernet, and FireWire [2]. The performance of NCSs largely depends on the communication networks. For example, packet dropouts, delay and quantization may happen when sensors with sensing and wireless communication abilities transmit their measurements to a remote estimator [5]. Therefore network servers should charge appropriately for the communication services with different transmission qualities they provide.

The Internet-of-Things (IoT) has brought new business markets. For example, Sun *et al.* [6] proposed a new business model for service provision in a wireless sensor network. Three roles exist in their scenario: the sensors, the service providers which gather the sensing data and provide service to users, and end users. In addition, the development of 5G wireless networks leads to many discussions regarding pricing models for resource management, as summarized in the survey [7]. Since 5G involves multiple rational stakeholders and entities that may have different objectives, e.g., utility maximization, cost minimization, and revenue maximization, studies on incorporating economic implications into the solutions, i.e., the pricing mechanism design, are vital for practical applications.

To the best of our knowledge, pricing mechanism design for communication channels in NCSs has not been addressed. One major challenge is how to formulate the problem. In this work, we consider a scenario with a server and some clients in NCSs. The server, which is the communication service provider, should determine the channel price to make the most profits. The clients, which mean the sensors in the NCSs scenario considered in the present work, should ponder between the benefits from selecting good channels to transmit data packets and the resulting costs. The server and clients can thus be regarded as playing a game for the communication resources.

For the scenario where a rational and selfish stakeholder (e.g., a server who profits by providing communication service) and entities (e.g., clients who may increase their utilities by paying for communication service) are involved, a Stackelberg game is useful to model the interactions between such two players with unequal status. Generally, a Stackelberg game is a strategic game where players make decisions sequentially. The first player usually occupies a dominant position and moves first, and the second player makes a choice after observing the first player's action. This class of games has been utilized to analyze pricing and selection problems in state-of-the-art works in various fields. For example, in software-defined networks, Gu *et al.* [8] considered a payoff optimization problem involving an Internet service provider (ISP) and network subscribers. The ISP decides the price according to the real-time network traffic load first, and network subscribers then specify the amount of bandwidth to be reserved based on the price offered. A Stackelberg game was constructed to analyze the interactions between the ISP and subscribers. In heterogeneous wireless access networks, Yun *et al.* [9] constructed a Stackelberg game framework to

Manuscript received December 4, 2017; revised January 7, 2019 and July 24, 2019; accepted July 24, 2019. Date of publication August 1, 2019; date of current version October 22, 2019. The work of Y. Ni and L. Shi was supported by the Hong Kong RGC General Research Fund 16208517. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Alejandro R. Ribeiro. *(Corresponding author: Yuqing Ni.)*

Y. Ni and L. Shi are with the Department of Electronic and Computer Engineering, Hong Kong University of Science and Technology, Kowloon, Hong Kong (e-mail: yniac@ust.hk; eesling@ust.hk).

A. S. Leong and D. E. Quevedo are with the Department of Electrical Engineering (EIM-E), Paderborn University, Paderborn 33098, Germany (e-mail: alex.leong@upb.de; dquevedo@ieee.org).

Digital Object Identifier 10.1109/TSIPN.2019.2932684

address the joint pricing and load distribution problem of multihoming where a service provider and users were involved. The service provider makes a pricing strategy to impose a cost on the users, whereas the users make rate-allocation decisions to maximize their utilities based on the given pricing strategy. In NCSs, a Stackelberg game framework for pricing and selection of "parallel channels" with a remote estimator was formulated in our recent conference contribution [10]. Structural results over a finite time horizon were derived. Thus in our channel pricing and selection problem with infinite time horizon, a Stackelberg game is still used to model the interactions between the channel server and the sensor clients. The channel server aims at maximizing its revenue, while each sensor client aims at minimizing a linear combination of the remote estimation error covariance and the communication cost, in a sequential order.

Our current work is mainly based on [10] and extends the results to the infinite time horizon. Differences are summarized as follows. First, in [10], the optimal strategy pair and the Stackelberg equilibrium exist trivially since the time horizon is finite. However, the infinite time horizon induces a countable but not finite state space, where there may not exist an optimal strategy pair, not to mention an equilibrium [11] since the minimum average cost is not guaranteed to be bounded. The optimization problem with infinite time horizon may be theoretically unsolvable. Although the problem becomes challenging, we succeed in proving the existence of an equilibrium under some assumptions. Second, [10] showed the server's nondecreasing pricing policy by backward induction. In the infinite time horizon case, the backward induction method does not work any more. We overcome this technical difficulty and characterize the properties of the server's optimal pricing policy. Furthermore, a closed-form expression is obtained. Third, single-server-multi-client (SSMC) scenario is considered in our paper. Properties regarding both players' optimal strategies are analyzed. However, [10] only studied the single-server-single-client (SSSC) case. In summary, recalling the drawbacks of the finite horizon problem, e.g., the rapidly increasing computational complexity when time N grows, and the dependence on the initial value, our current work studies the infinite horizon case to have a thorough understanding of this pricing and selection problem.

The remainder of the paper is organized as follows: Section II introduces the system model and the main problem and Section III presents preliminaries regarding a Stackelberg game. Section IV provides the equilibrium result in the SSSC scenario as well as the perturbation analysis. Section V focuses on the SSMC scenario and proves that the optimal strategy pair for the server and clients is attainable. Section VI provides simulations and interpretations. Section VII draws conclusions.

Notation: For a matrix X, we use $X^\top$, $\mathrm{Tr}\{X\}$ and $\rho(X)$ to denote its transpose, trace and spectral radius, respectively. $\mathbb{S}_+^n$ is the set of $n \times n$ positive semi-definite matrices. When $X \in \mathbb{S}_+^n$, it is written as $X \succeq 0$. For two symmetric matrices X and Y, $X \succeq Y$ means $X - Y \succeq 0$. $\mathbb{R}^n$ is the n-dimensional Euclidean space. $\mathbb{E}[\cdot]$ is the expectation operator and $\mathbb{E}[\cdot|\cdot]$ denotes conditional expectation. The notation $\mathbb{P}[\cdot]$ refers to probability. For functions f, f_1, f_2, $f_1 \circ f_2$ is defined as $f_1 \circ f_2(X) \triangleq f_1(f_2(X))$ and f^k, $k \in \{0, 1, \dots\}$, is defined as $f^k(X) \triangleq \underbrace{f \circ f \cdots f}_{k \text{ times}}(X)$ with $f^0(X) = X$.

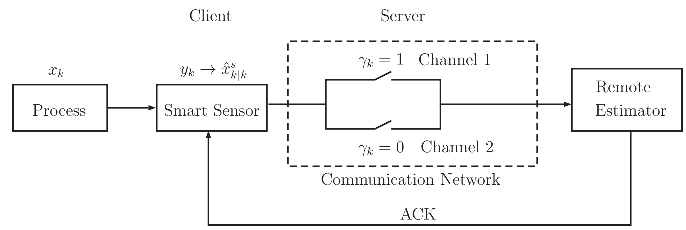


Fig. 1. System block diagram.

II. Remote State Estimation with Channel Pricing and Selection

Consider the system in Fig. 1. The discrete linear time-invariant (LTI) process for $k \geq 0$ is as follows:

$$x_{k+1} = Ax_k + w_k, \tag{1}$$

$$y_k = Cx_k + v_k, \tag{2}$$

where $x_k \in \mathbb{R}^n$ is the process state vector, $w_k \in \mathbb{R}^n$ is the process noise which is i.i.d. zero-mean white Gaussian with covariance $Q \succeq 0$. The measurement collected by the sensor is $y_k \in \mathbb{R}^m$. The measurement noise $v_k \in \mathbb{R}^m$ is i.i.d. zero-mean white Gaussian with covariance $R \succ 0$. The initial state x_0 is a zero-mean Gaussian random variable with covariance $\Pi_0 \succeq 0$, which is uncorrelated with w_k and v_k. The pair (A, C) is assumed to be observable and $(A, \sqrt{Q})$ is controllable.

A. Smart Sensor

The smart sensor in Fig. 1 is capable of running a local Kalman filter. Its minimum mean-squared error state estimate $\hat{x}^s_{k|k}$ and the corresponding error covariance $P^s_{k|k}$ for $k \geq 1$ are denoted as:

$$\hat{x}^s_{k|k} = \mathbb{E}\left[x_k \mid y_1, \dots, y_k\right],$$

$$P^s_{k|k} = \mathbb{E}[(x_k - \hat{x}^s_{k|k})(x_k - \hat{x}^s_{k|k})^\top \mid y_1, \dots, y_k],$$

which are computed via a Kalman filter as follows:

$$\begin{aligned}
\hat{x}^s_{k|k-1} &= A\hat{x}^s_{k-1|k-1}, \\
P^s_{k|k-1} &= AP^s_{k-1|k-1}A^\top + Q, \\
K_k &= P^s_{k|k-1}C^\top(CP^s_{k|k-1}C^\top + R)^{-1}, \\
\hat{x}^s_{k|k} &= \hat{x}^s_{k|k-1} + K_k(y_k - C\hat{x}^s_{k|k-1}), \\
P^s_{k|k} &= P^s_{k|k-1} - K_kCP^s_{k|k-1}.
\end{aligned}$$

The initial states are $\hat{x}^s_{0|0}$ and $P^s_{0|0}$. From [12], the estimation error covariance of the Kalman filter converges to a unique value $\overline{P}$ no matter what the initial value is. Define the Lyapunov operator $h(\cdot) : \mathbb{S}_+^n \to \mathbb{S}_+^n$ as

$$h(X) \triangleq AXA^\top + Q,$$

and Riccati operator $\tilde{g}(\cdot) : \mathbb{S}_+^n \to \mathbb{S}_+^n$ as

$$\tilde{g}(X) \triangleq X - XC^\top \left(CXC^\top + R\right)^{-1} CX.$$

We assume that the error covariance at the smart sensor has already reached steady state and let

$$P_{k|k}^s = \overline{P}, \quad k \geq 0,$$

where $\overline{P}$ is the unique positive semi-definite solution to $\tilde{g} \circ h(X) = X$, as [12] shows.

Lemma 1: The error covariance $\overline{P}$ satisfies:

$$h^{k_1}(\overline{P}) \succeq h^{k_2}(\overline{P}) \tag{3}$$

for any $k_1 \geq k_2 \geq 0$. Furthermore,

$$\text{Tr}\left\{h^{k_1}\left(\overline{P}\right)\right\} \geq \text{Tr}\left\{h^{k_2}\left(\overline{P}\right)\right\}. \tag{4}$$

This well-ordering of the estimation error covariance is helpful for the subsequent analysis. The proof is provided in Appendix A.

B. Channel Pricing and Selection

Fig. 1 portrays a two-player scenario. The process equipped with the smart sensor is the "client" and the communication network belongs to the "server". At each time k, the server decides the price W_k of the high quality channel 1. The price of the low quality channel 2 is normalized to 0. Without loss of generality, we assume $W_k > 0$. Here the "price" represents the cost the client needs to pay when choosing the transmission service. The pricing is to incentivize clients to use channel 1. After a price is decided, it is then the client's turn to decide whether to use channel 1 or channel 2. If using channel 1, the probability that $\hat{x}_{k|k}^s$ arrives error-free at the remote estimator is fixed at λ_1. While using channel 2, the error-free arrival probability is λ_2, where $1 > \lambda_1 > \lambda_2 > 0$.

Thus the transmission choice of $\hat{x}_{k|k}^s$ in the communication network can be characterized by a binary variable γ_k:

$$\gamma_k = \begin{cases} 1, & \text{if the client uses channel 1,} \\ 0, & \text{if the client uses channel 2.} \end{cases} \tag{5}$$

Furthermore, the arrival of packets at the remote estimator can be characterized by a binary random sequence $\{\delta_k\}$:

$$\delta_k = \begin{cases} 1, & \text{if } \hat{x}_{k|k}^s \text{ arrives error-free at time } k, \\ 0, & \text{otherwise.} \end{cases}$$

Based on the error-free arrival probabilities of the two channels, we have

$$\mathbb{P}\left[\delta_k = 1\right] = \gamma_k \lambda_1 + (1 - \gamma_k)\, \lambda_2.$$

C. Remote Estimator

Since the smart sensor sends the local minimum mean square error estimates instead of raw measurements, the optimal remote estimator can be shown to have the following form [13]. Denote $\hat{x}_k$ and P_k as the state estimate and error covariance at the remote estimator. If the packet $\hat{x}_{k|k}^s$ arrives error-free at time k, the estimator synchronizes $\hat{x}_k$ with $\hat{x}_{k|k}^s$ from the smart sensor; otherwise, it just uses the time update value based on the system model (1). The recursion of $\hat{x}_k$ is

$$\hat{x}_k = \begin{cases} \hat{x}_{k|k}^s, & \text{if } \delta_k = 1, \\ A\hat{x}_{k-1}, & \text{if } \delta_k = 0. \end{cases} \tag{6}$$

Correspondingly, the recursion of P_k is

$$P_k = \begin{cases} \overline{P}, & \text{if } \delta_k = 1, \\ h(P_{k-1}), & \text{if } \delta_k = 0. \end{cases} \tag{7}$$

To simplify the problem, we assume that the initial state $P_0 = \overline{P}$. At each time k, P_k takes a value from the countable set $\{\overline{P}, h(\overline{P}), h^2(\overline{P}), \ldots, h^k(\overline{P})\}$. Note that the server is the owner of the communication network and it has good knowledge of the channel transmission process. The server always knows exactly whether $\hat{x}_{k|k}^s$ arrives at the remote estimator successfully or not. We assume that the remote estimator sends Acknowledgments (ACKs) of receipt to the client. Thus this is a causal perfect-information case for both the server and the client.

In the remainder of this work, we will present the preliminaries regarding a Stackelberg game and the optimal server pricing strategies and client channel selection strategies for both the SSSC and SSMC scenarios.

III. Stackelberg Games

In this section, we introduce the Stackelberg game, which is a key supporting concept for our problem formulation.

Let Θ_1 and Θ_2 be the sets of admissible strategies for players 1 and 2, respectively. Let the cost functions $J_1(\theta_1, \theta_2)$ and $J_2(\theta_1, \theta_2)$ be two functions mapping $\Theta_1 \times \Theta_2 \to \mathbb{R}$ such that player 1 wishes to maximize J_1 and player 2 wishes to minimize J_2. In a Stackelberg game, the player 1 who selects his strategy first is called *the leader*. The player 2 who selects his strategy second is called *the follower*. The definitions of a Stackelberg optimal strategy pair and the Stackelberg game equilibrium are stated as follows, mainly based on [14]–[16].

Definition 1: If there exists a mapping $\Phi : \Theta_1 \to \Theta_2$ such that, for any fixed $\theta_1 \in \Theta_1$, $J_2(\theta_1, \Phi\theta_1) \leq J_2(\theta_1, \theta_2)$ for all $\theta_2 \in \Theta_2$, and if there exists a $\theta_1^\star \in \Theta_1$ such that $J_1(\theta_1^\star, \Phi\theta_1^\star) \geq J_1(\theta_1, \Phi\theta_1)$ for all $\theta_1 \in \Theta_1$, then the pair $(\theta_1^\star, \theta_2^\star) \in \Theta_1 \times \Theta_2$, where $\theta_2^\star = \Phi\theta_1^\star$, is called a Stackelberg optimal strategy pair. An equilibrium in the Stackelberg game is reached under this optimal strategy pair.

IV. Single-Server-Single-Client Game

In the infinite time horizon single-server-single-client (SSSC) case as shown in Fig. 1, for the client sensor, its objective function J_C is a linear combination of the trace of expected estimation error covariance and the cost when using high transmission quality channel 1:

$$\min_{\{\gamma_k\}} J_C \triangleq \limsup_{N\to\infty} \frac{1}{N} \sum_{k=1}^{N} \left[\zeta \text{Tr}\left\{\mathbb{E}\left[P_k\right]\right\} + (1-\zeta)\gamma_k W_k\right], \tag{8}$$

for some weight parameter $\zeta \in (0,1)$. A larger ζ attaches more importance to the error covariance and a smaller ζ attaches more importance to the channel costs. For every policy $\{W_k\}$ which depends on the estimation error covariance and is given by the server, the client needs to determine the optimal action $\{\gamma_k\}$ to minimize its cost J_C. For the server, its objective function J_S is the total revenue:

$$\max_{\{W_k\}} J_S \triangleq \liminf_{N\to\infty} \frac{1}{N}\sum_{k=1}^{N} \gamma_k W_k. \tag{9}$$

The server is at the leading position in this Stackelberg game. At each time $k \geq 1$, it sets the price W_k first and then the follower, the client, decides γ_k sequentially. We define $\mathcal{I}_k^S$ as the information set available to the server up to time k, i.e., $\mathcal{I}_k^S \triangleq \{P_0, P_1, \ldots, P_{k-1}\} \cup \{W_1, W_2, \ldots, W_{k-1}\}$, and $\mathcal{I}_k^C \triangleq \{P_0, P_1, \ldots, P_{k-1}\} \cup \{W_1, W_2, \ldots, W_k\}$ as the information set for the client. Assume both players are rational, which means that they take the optimal actions based on their information sets, respectively.

Since there exist few unstable systems in the real case, here we consider an asymptotically stable system, i.e., $\rho(A) < 1$. Thus, even if $\delta_k = 0$ all the time, the recursion $P_k = h(P_{k-1})$ will converge and the trace of the error covariance at the remote estimator will always be upper bounded. Furthermore, define the remote estimator's state set as

$$\begin{aligned}\mathbb{S} &= \left\{\mathrm{Tr}\left\{\overline{P}\right\}, \mathrm{Tr}\left\{h\left(\overline{P}\right)\right\}, \mathrm{Tr}\left\{h^2\left(\overline{P}\right)\right\}, \ldots\right\} \\ &\triangleq \{s_0, s_1, s_2, \ldots\}\end{aligned}$$

where $0 \leq s_0 \leq s_1 \leq s_2 \leq \cdots$ because of Lemma 1. The transition probabilities of the states at the remote estimator are given by:

$$\begin{aligned} &p_k\left(\mathrm{Tr}\left\{P_k\right\} \mid \mathrm{Tr}\left\{P_{k-1}\right\}, \gamma_k\right) \\ &\triangleq \begin{cases} \gamma_k\lambda_1 + (1-\gamma_k)\lambda_2, & \text{if } P_k = \overline{P}, \\ 1 - \gamma_k\lambda_1 - (1-\gamma_k)\lambda_2, & \text{if } P_k = h(P_{k-1}), \\ 0, & \text{otherwise.} \end{cases}\end{aligned} \tag{10}$$

The one-stage cost function for the client at time $k = 1, 2, \ldots$ is

$$c_k\left(\mathrm{Tr}\left\{P_{k-1}\right\}, \gamma_k\right) = \zeta \mathrm{Tr}\left\{\mathbb{E}\left[P_k\right]\right\} + (1-\zeta)\gamma_k W_k. \tag{11}$$

This two-player non-zero-sum stochastic game will be studied under the long-run expected average cost criterion. In our problem, we assume that the server's price setting strategy is stationary and deterministic at the current estimation state; see Assumption 1. A stationary policy only depends on the current state, and a deterministic policy is to choose the action with probability 1 [11].

Assumption 1: The pricing strategy of the server is stationary and deterministic.

In view of Assumption 1 and with a slight abuse of notation, we will henceforth denote the stationary and deterministic pricing decision made by the network server, at time k and given state $\mathrm{Tr}\{P_{k-1}\} = s_i$, by W_i. For the decision made by the client sensor, we shall use a similar notation, i.e., γ_i corresponding to $\mathrm{Tr}\{P_{k-1}\} = s_i$ and $W_k = W_i$, since we will prove that there exists an optimal stationary and deterministic policy for the client under Assumption 1.

Our main result is to find a Stackelberg optimal strategy pair for both the server and the client, under which the equilibrium is achieved.

A. Client's Optimal Policy

In this subsection, we prove that, for any stationary and deterministic policy of the server, the client has an optimal stationary and deterministic policy.

We formulate this optimization problem as a Markov decision process (MDP) problem. An MDP $(\mathbb{S}, \mathbb{A}_C, p(\cdot \mid \cdot,\cdot), c(\cdot,\cdot))$ consists of the state space $\mathbb{S}$, the action space $\mathbb{A}_C$, the stage transition probability $p(\cdot \mid \cdot,\cdot)$, and one-stage cost $c(\cdot,\cdot)$. Let $\Gamma_S(\mathbb{S})$ and $\Gamma_C(\mathbb{S})$ denote the set of stationary and deterministic strategies for the server and client. Due to the Markovian structure of the problem setting, past states and actions are irrelevant for the player to take actions. For any given policy $\{W_i\} \in \Gamma_S(\mathbb{S})$, the client's action set is $\mathbb{A}_C = \{0,1\}$. Here W_i refers to the price determined at the state s_i, and $\{W_i\}$ refers to the set of all prices W_i for $i = 0, 1, 2, \ldots$. The action 0 means using channel 2, while action 1 means using the good channel 1. The action taken at state s_i is γ_i. It is obvious that the transition probabilities $p_k(\cdot \mid \cdot,\cdot)$ in (10) are stationary, which means that they only depend on the states and the actions, thus we can substitute $p(\cdot \mid \cdot,\cdot)$ for them, and write:

$$p\left(s_0 \mid s_i, \gamma_i\right) = \gamma_i\lambda_1 + (1-\gamma_i)\lambda_2, \tag{12}$$

$$p\left(s_{i+1} \mid s_i, \gamma_i\right) = 1 - \gamma_i\lambda_1 - (1-\gamma_i)\lambda_2, \tag{13}$$

for $i = 0, 1, 2, \ldots$. The induced one-stage cost in (11) is stationary, in that the server's policy only depends on the current state. The incurred cost at state s_i is

$$\begin{aligned} c\left(s_i, \gamma_i\right) = \zeta &\left[p\left(s_0|s_i,\gamma_i\right)s_0 + p\left(s_{i+1}|s_i,\gamma_i\right)s_{i+1}\right] \\ &+ (1-\zeta)\gamma_i W_i. \end{aligned} \tag{14}$$

The average cost optimality equation (ACOE) for problem (8) is

$$\begin{aligned} \bar{c} + f(s_i) = \min_{\gamma_i \in \mathbb{A}_C} &\left\{c\left(s_i,\gamma_i\right) + p\left(s_0|s_i,\gamma_i\right) f(s_0)\right. \\ &\left. + p\left(s_{i+1}|s_i,\gamma_i\right) f\left(s_{i+1}\right)\right\} \end{aligned} \tag{15}$$

where $\bar{c}$ is the optimal average cost per stage and $f(\cdot)$ is the relative value function (See Chapter 7 in [11]).

The following existence theorem is the basic result for the client in this subsection. The proof is provided in Appendix B.

Theorem 1: For any given stationary and deterministic policy $\{W_i\} \in \Gamma_S(\mathbb{S})$ of the server, there exists a stationary and deterministic policy which is optimal for the client.

Theorem 1 provides a theoretic basis for further analysis of the server's optimal policy and the equilibrium. Together with Assumption 1, this theorem guarantees that given any server's policy $\{W_i\} \in \Gamma_S(\mathbb{S})$, there exists an optimal client's policy $\{\gamma_i\} \in \Gamma_C(\mathbb{S})$, where $\{\gamma_i\}$ is the set of actions γ_i for $i =$

$0, 1, 2, \ldots$. The latter induces a stationary probability distribution $\{\pi_i\}$ over the states. Under every policy pair $(\{W_i\}, \{\gamma_i\})$, the objective functions of both the client and server can be transformed into

$$J_C = \sum_{i=0}^{\infty} \pi_i \left[\zeta s_i + (1-\zeta)\,\gamma_i W_i\right], \tag{16}$$

$$J_S = \sum_{i=0}^{\infty} \pi_i \gamma_i W_i, \tag{17}$$

where $\sum_{i=0}^{\infty} \pi_i = 1$.

B. Server's Optimal Policy

In this subsection, we narrow the scope of feasible optimal strategies for the server, and then provide a method for the computation of an optimal server policy.

Theorem 2 characterizes the properties of the server's optimal policy. Lemma 2 and Assumption 2 are preliminary steps. The proof is provided in Appendix C.

Lemma 2: Let $\{m_j\}$ and $\{n_j\}$ be non-negative real-valued sequences with

$$\sum_{j=0}^{\infty} m_j = \sum_{j=0}^{\infty} n_j = 1, \tag{18}$$

and

$$\sum_{j=k}^{\infty} m_j \geq \sum_{j=k}^{\infty} n_j \tag{19}$$

for all non-negative k. Then for any nondecreasing but bounded sequence $\{s_j\}$, we have

$$\sum_{j=0}^{\infty} m_j s_j \geq \sum_{j=0}^{\infty} n_j s_j. \tag{20}$$

Assumption 2: If at any state s_i, with the price W_i set by the server, the client incurs the same long run average cost J_C whether using channel 1 or channel 2, then the client chooses to use channel 1 and the server makes a profit W_i at the state s_i.

The following theorem is the major result of this subsection. It rules out many varieties of strategies in the sense of the server's optimality and leading position in this Stackelberg game.

Theorem 2: A pricing policy $\{W_i^\star\}$, under which the client's optimal policy is to use channel 1 at all the states, i.e., $\gamma_i = 1$ for $i = 0, 1, 2, \ldots$, and the resulting scheduling cost is a constant $J_C^\star$ as described in (21), is an optimal server's policy.

Proof: We present this proof in three steps. First, we show that under any pricing policy $\{W_i\}$ and the incurred optimal scheduling policy $\{\gamma_i\}$, the client's cost is upper bounded by some constant $J_C^\star$. Furthermore, the server can obtain the maximum revenue when the client's cost reaches this upper bound $J_C^\star$. The relation between the server's revenue and the induced state probability distribution is displayed in (22). Second, we prove that a pricing policy $\{W_i^\star\}$, which incentivizes the client to use channel 1 at all the states, minimizes $\sum_{i=k}^{\infty} \pi_i$ for arbitrary non-negative k. Finally, according to Lemma 2 and Equation (22), the maximum revenue is proved to be attained under the pricing policy $\{W_i^\star\}$, which shows the optimality of $\{W_i^\star\}$.

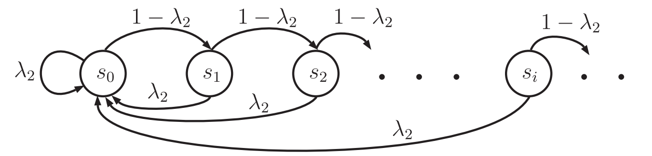


Fig. 2. Markov chain of always using channel 2.

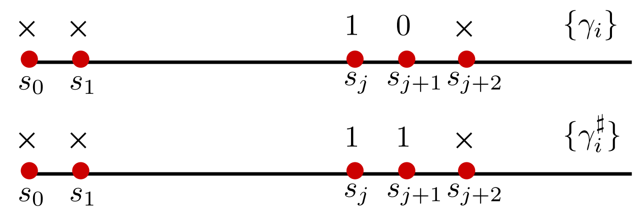


Fig. 3. Client's policies differing only at s_{j+1}.

First, the upper bound of the client's cost is derived as follows. Considering that our system is asymptotically stable, s_i is bounded by some non-negative constant. For some server's policy $\{W_i\}$ where each element is large enough, the client's optimal strategy is never to use channel 1. The resulting Markov chain is in Fig. 2. The induced stationary probability distribution $\{\pi_i^{channel\_2}\}$ is

$$\pi_i^{channel\_2} = \lambda_2 \left(1-\lambda_2\right)^i, \qquad i = 0, 1, \ldots.$$

In this situation, the client's cost is

$$J_C^\star = \zeta \sum_{i=0}^{\infty} \pi_i^{channel\_2} s_i. \tag{21}$$

For any server's policy $\{W_i\}$ and the incurred client's policy $\{\gamma_i\}$, if the induced client's cost $J_C < J_C^\star$, the server can always raise the prices at some states so that the client's policy $\{\gamma_i\}$ remains unchanged but $J_C = J_C^\star$. Due to Assumption 2, the server will profit more by taking this pricing adjustment. However, the server cannot improve these values too much. If so, the server can profit nothing since the client is always willing to choose channel 2. And therefore, $J_C > J_C^\star$ will never happen, which makes $J_C^\star$ an upper bound.

As a consequence, $J_C^\star$ acts as a constraint for the server. According to the "constraint" $J_C^\star$, the server's revenue in this Stackelberg game, with a policy pair $(\{W_i\}, \{\gamma_i\})$ and the induced state probability distribution $\{\pi_i\}$, is denoted as:

$$J_S = \frac{1}{1-\zeta} J_C^\star - \frac{\zeta}{1-\zeta} \sum_{i=0}^{\infty} \pi_i s_i. \tag{22}$$

Second, we show that a pricing policy $\{W_i^\star\}$, under which $\gamma_i = 1$ for $i = 0, 1, \ldots$, minimizes $\sum_{i=k}^{\infty} \pi_i$ for arbitrary non-negative k. We prove this by dividing into two cases. Consider a client's policy $\{\gamma_i\}$ and its corresponding probability distribution $\{\pi_i\}$. As Fig. 3 shows, the policy $\{\gamma_i^\sharp\}$ refers to keeping all the other actions unchanged but only changing at the state s_{j+1},

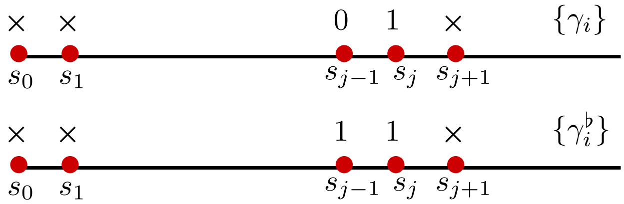


Fig. 4. Client's policies differing only at s_{j-1}.

where $\gamma_j = 1$ and $\gamma_{j+1} = 0$. The corresponding probability distribution is denoted as $\{\pi_i^\sharp\}$. In the figure, "$\times$" means the action is either 0 or 1. We want to prove that $\sum_{i=k}^{\infty} \pi_i \geq \sum_{i=k}^{\infty} \pi_i^\sharp$ for any arbitrary non-negative k. Assume that

$$\pi_{j+1} = \mu, \quad \pi_{j+1}^\sharp = \mu^\sharp.$$

Then

$$\pi_{j+2} = \mu\left(1-\lambda_2\right), \quad \pi_{j+2}^\sharp = \mu^\sharp\left(1-\lambda_1\right).$$

The actions at the other states remain the same and we have

$$\sum_{i=0}^{j+1} \pi_i = \mu\sigma_1, \quad \sum_{i=0}^{j+1} \pi_i^\sharp = \mu^\sharp\sigma_1,$$

where σ_1 is a positive constant. Similarly,

$$\sum_{i=j+3}^{\infty} \pi_i = \mu\left(1-\lambda_2\right)\sigma_2, \quad \sum_{i=j+3}^{\infty} \pi_i^\sharp = \mu^\sharp\left(1-\lambda_1\right)\sigma_2,$$

where σ_2 is also a positive constant. The following equations hold for the stationary probability distributions:

$$\mu\sigma_1 + \mu\left(1-\lambda_2\right) + \mu\left(1-\lambda_2\right)\sigma_2 = 1, \tag{23}$$

$$\mu^\sharp\sigma_1 + \mu^\sharp\left(1-\lambda_1\right) + \mu^\sharp\left(1-\lambda_1\right)\sigma_2 = 1. \tag{24}$$

Conditioned on Equation (23) and Equation (24), the channel transmission quality condition $1 > \lambda_1 > \lambda_2 > 0$ indicates that

$$\mu < \mu^\sharp, \tag{25}$$

$$\mu\left(1-\lambda_2\right) > \mu^\sharp\left(1-\lambda_1\right). \tag{26}$$

Then,

$$\sum_{i=k}^{\infty} \pi_i \geq \sum_{i=k}^{\infty} \pi_i^\sharp$$

can be derived from these two inequalities (25) and (26) for any non-negative k.

For the case shown in Fig. 4, we can similarly derive that $\sum_{i=k}^{\infty} \pi_i \geq \sum_{i=k}^{\infty} \pi_i^\flat$ for any non-negative k, where $\gamma_{j-1} = 0$, $\gamma_j = 1$ and $\gamma_{j-1}^\flat = 1$.

According to the above two cases, we conclude that using channel 1 at all the states, which means $\gamma_i = 1$ for $i = 0, 1, \ldots$, minimizes $\sum_{i=k}^{\infty} \pi_i$. We denote this policy as $\{\gamma_i^\star\}$, and the induced probability distribution as $\{\pi_i^\star\}$. A price setting policy which induces $\{\gamma_i^\star\}$ is marked as $\{W_i^\star\}$, and the server's revenue is $J_S^\star$. For any server policy $\{W_i\}$ and the corresponding client's optimal policy $\{\gamma_i\}$ with the induced probability distribution

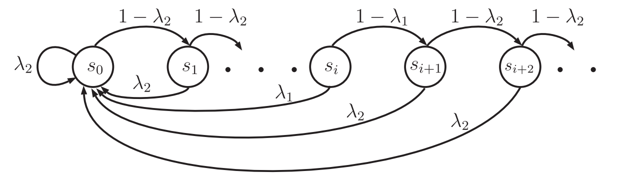


Fig. 5. Markov chain of only using channel 1 at s_i.

$\{\pi_i\}$ and server's revenue J_S, the following inequality holds:

$$\begin{aligned} J_S^\star - J_S &= \left(\frac{1}{1-\zeta}J_C^\star - \frac{\zeta}{1-\zeta}\sum_{i=0}^{\infty}\pi_i^\star s_i\right) \\ &\quad - \left(\frac{1}{1-\zeta}J_C^\star - \frac{\zeta}{1-\zeta}\sum_{i=0}^{\infty}\pi_i s_i\right) \\ &= \frac{\zeta}{1-\zeta}\left(\sum_{i=0}^{\infty}\pi_i s_i - \sum_{i=0}^{\infty}\pi_i^\star s_i\right) \\ &\geq 0. \end{aligned} \tag{27}$$

The last inequality holds because of Lemma 2. Finally, we finish the proof by comparing $J_S^\star$ and J_S. $\{W_i^\star\}$ maximizes the server's revenue and therefore, this is an optimal policy. ■

As for the server, Theorem 2 shows that an optimal $\{W_i^\star\}$ should be made to incentivize the client sensor such that the client's optimal policy $\{\gamma_i^\star\}$ is always to use channel 1 and the resulting cost is $J_C^\star$. Under this optimal policy pair $(\{W_i^\star\}, \{\gamma_i^\star\})$, the probability distribution $\{\pi_i^\star\}$ is

$$\pi_i^\star = \lambda_1\left(1-\lambda_1\right)^i, \quad i = 0, 1, \ldots. \tag{28}$$

The server's optimal revenue is

$$\begin{aligned} J_S^\star &= \frac{1}{1-\zeta}J_C^\star - \frac{\zeta}{1-\zeta}\sum_{i=0}^{\infty}\pi_i^\star s_i \\ &= \frac{\zeta}{1-\zeta}\sum_{i=0}^{\infty}[\lambda_2\left(1-\lambda_2\right)^i - \lambda_1\left(1-\lambda_1\right)^i]s_i. \end{aligned} \tag{29}$$

After characterizing the properties of an optimal server's strategy, we provide a method to find such an optimal strategy. Notice that under an optimal $\{W_i^\star\}$, whether the client uses channel 1 at all the states does not change the client's average cost $J_C^\star$. Then the optimal price at each state i can be calculated assuming that the client's optimal scheduling policy is to use channel 1 only at the state i. For notation simplicity, we define a function $\phi(\cdot)$ as

$$\phi(i) = \frac{\lambda_2}{1-\left(\lambda_1-\lambda_2\right)\left(1-\lambda_2\right)^i}.$$

For a state s_i, we assume that the client's optimal scheduling policy is $\gamma_i = 1$ while $\gamma_j = 0$ for $j \neq i$, and the induced Markov chain is shown in Fig. 5. Consequently, the induced stationary probability distribution $\{\bar{\pi}_i\}$ is shown in Table I. It is trivial to verify that the summation of the stationary probabilities in the second column is equal to 1. Thus the designed optimal price

TABLE I
STATIONARY PROBABILITY DISTRIBUTION $\{\bar{\pi}_i\}$

State	Stationary probability
s_0	$\bar{\pi}_0 = \phi(i)$
s_1	$\bar{\pi}_1 = \phi(i)(1-\lambda_2)$
$\dots$	
s_i	$\bar{\pi}_i = \phi(i)(1-\lambda_2)^i$
s_{i+1}	$\bar{\pi}_{i+1} = \phi(i)(1-\lambda_2)^i(1-\lambda_1)$
s_{i+2}	$\bar{\pi}_{i+2} = \phi(i)(1-\lambda_2)^{i+1}(1-\lambda_1)$
$\dots$	

$W_i^\star$ at state s_i can be calculated as follows:

$$W_i^\star = \frac{1}{\bar{\pi}_i(1-\zeta)}\left(J_C^\star - \zeta\sum_{j=0}^{\infty}\bar{\pi}_j s_j\right). \tag{30}$$

Interestingly, this optimal pricing policy features some monotone structures as stated in the following theorem:

Theorem 3: In the SSSC scenario, the server's optimal strategy $\{W_i^\star\}$ is nondecreasing in the state s_i.

Proof: For any $i \geq 0$, it can be calculated by combining (21) and (30):

$$W_i^\star = \frac{\zeta}{1-\zeta}\frac{\lambda_1-\lambda_2}{(1-\lambda_2)^i}\left[-\sum_{j=0}^{i}(1-\lambda_2)^{j+i}s_j + \sum_{j=i+1}^{\infty}(1-\lambda_2)^{j-1}[1-(1-\lambda_2)^{i+1}]s_j\right].$$

Then we make the comparison between $W_i^\star$ and $W_{i+1}^\star$:

$$\begin{aligned} W_{i+1}^\star - W_i^\star &= \frac{\zeta}{1-\zeta}\frac{\lambda_1-\lambda_2}{(1-\lambda_2)^{i+1}}\left[\sum_{j=i+2}^{\infty}\lambda_2(1-\lambda_2)^{j-1}s_j - (1-\lambda_2)^{i+1}s_{i+1}\right] \\ &\geq \frac{\zeta}{1-\zeta}\frac{\lambda_1-\lambda_2}{(1-\lambda_2)^{i+1}}s_{i+1}\left[\sum_{j=i+2}^{\infty}\lambda_2(1-\lambda_2)^{j-1} - (1-\lambda_2)^{i+1}\right] \\ &= 0. \end{aligned} \tag{31}$$

The inequality holds due to the well-ordering of s_i. Thus an optimal strategy designed using this mechanism is nondecreasing in the states. ■

So far, we have shown that the server's optimal strategy is to drive the client to use the good channel 1 all the time. This is also an equilibrium of this Stackelberg game. An optimal strategy of the server can be computed, and the optimal price is nondecreasing in the state. This indicates that when the remote estimation error covariance is larger, the server can set higher prices and profit more. However, the price cannot be set too high. It must guarantee the existence of the "constraint" $J_C^\star$.

Remark 1: This is a Stackelberg game model, where the server has priority over the client. For both the server and the client, they make decisions based on the current state. In the first period, the server determines the channel pricing strategy $\{W_i\}$. This decision is irreversible. In the second period, the client makes the channel scheduling strategy $\{\gamma_i\}$ after observing the pricing strategy given by the server. From the above analysis, we can find that the server's optimal strategy restricts the actions of the client. The client has no choice but to pay $J_C^\star$ all the time. The server is taking advantage of being the "first mover" [17]. This case is similar to a monopoly.

C. Perturbation Analysis

In this subsection, we mainly focus on how $J_C^\star$ and $J_S^\star$ change when perturbing the weight parameter ζ. Recall that the client's objective value under the Stackelberg optimal strategy pair $(\{W_i^\star\},\{\gamma_i^\star\})$ is

$$J_C^\star = \zeta\sum_{i=0}^{\infty}\lambda_2(1-\lambda_2)^i s_i,$$

and the server's maximum revenue is given by

$$J_S^\star = \frac{\zeta}{1-\zeta}\sum_{i=0}^{\infty}\left[\lambda_2(1-\lambda_2)^i - \lambda_1(1-\lambda_1)^i\right]s_i,$$

where $\zeta \in (0,1)$. It can be concluded that both $J_C^\star$ and $J_S^\star$ are strictly increasing in the parameter ζ, which means that the increase of ζ brings more benefits to the server and squeezes more out of the client. A larger ζ implies that the client attaches more importance to the estimation quality. Consequently, the client needs to pay more for such concerns since this is actually a monopoly market held by the server. On the contrary, if the client does not pay much attention to the estimation quality and is easily contented, the server can do nothing but earn a meager profit. The analysis with regard to the weight parameter ζ again emphasizes the server's monopoly position in this Stackelberg game.

V. EXTENSION TO MULTIPLE PROCESSES: SINGLE-SERVER-MULTI-CLIENT GAME

In this extension section, consider a single-server-multi-client (SSMC) scenario as portrayed in Fig. 6. There exist M independent processes, which are all asymptotically stable, i.e., $\rho(A_\ell) < 1$. For each client ℓ, its LTI process is

$$x_{k+1}^\ell = A_\ell x_k^\ell + w_k^\ell, \tag{32}$$

$$y_k^\ell = C_\ell x_k^\ell + v_k^\ell. \tag{33}$$

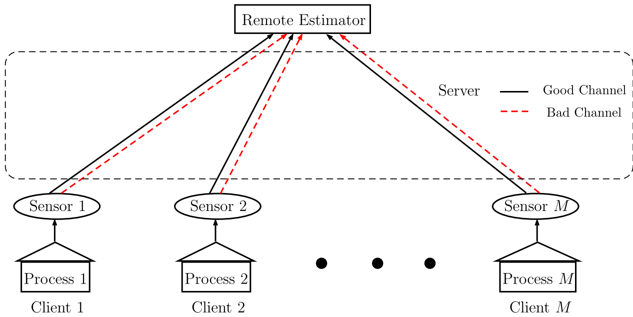


Fig. 6. SSMC scenario.

In the infinite time horizon case, each client ℓ has its own objective function J_C^ℓ, which is still a linear combination of the trace of expected estimation error covariance and the cost when using channel 1:

$$\min_{\{\gamma_k^\ell\}} J_C^\ell \triangleq \limsup_{N\to\infty} \frac{1}{N}\sum_{k=1}^{N}\left[\zeta^\ell \mathrm{Tr}\left\{\mathbb{E}\left[P_k^\ell\right]\right\} + (1-\zeta^\ell)\gamma_k^\ell W_k\right], \tag{34}$$

for some weight parameter $\zeta^\ell \in (0,1)$. For every policy $\{W_k\}$ given by the server, each client ℓ needs to determine the optimal action $\{\gamma_k^\ell\}$ to minimize its J_C^ℓ. While for the server, its objective function J_S is the total revenue from all the M clients:

$$\max_{\{W_k\}} J_S \triangleq \liminf_{N\to\infty}\frac{1}{N}\sum_{k=1}^{N}\sum_{\ell=1}^{M}\gamma_k^\ell W_k. \tag{35}$$

Assumption 3: The server's pricing strategy is a "constant" one, i.e., $\{W\}$.

Remark 2: In this SSMC scenario, if the server can set the price $\{W_k^\ell\}$ independently for each client, then the pricing and selection method in the SSSC scenario can be applied here. However, such kind of pricing policy may undermine the market fairness since the server provides different charging policies for different clients. Besides, a general time-varying pricing policy $\{W_k\}$ makes the problem intractable since the state space $\mathbb{S}^1 \times \mathbb{S}^2 \times \cdots \times \mathbb{S}^M$ grows quickly with the number M of clients. The server needs to take into account the states of all clients when making a decision. It becomes quite difficult for the server to set optimal prices based on such a state space whose size increases exponentially. To get some insights regarding the pricing and selection strategies in the SSMC scenario, we assume that once the price W_k is determined, it cannot be changed at other times or other states, which implies a constant pricing policy. Thus the server needs to design the variable W to maximize its benefits. While for the clients, once the channel 1 price W is determined, they are totally independent of each other. Although this assumption seems to be a little bit strong, it makes the analysis about optimal strategy pair possible. Some theoretic results regarding this Stackelberg game equilibrium are presented in the following subsections.

A. Client's Optimal Policy

In this subsection, we prove that for any given W, each client has an optimal stationary and deterministic policy, which is nondecreasing in its state.

Define the remote estimator's state set for the client ℓ as

$$\mathbb{S}^\ell \triangleq \left\{s_0^\ell, s_1^\ell, s_2^\ell, \dots\right\}$$

where $0 \le s_0^\ell \le s_1^\ell \le s_2^\ell \dots$ according to Lemma 1. The action set is $\mathbb{A}_{C^\ell} = \{0,1\}$, where the action 0 means using channel 2 while action 1 means using the good channel 1. The action taken at state s_i^ℓ is γ_i^ℓ. The transition probabilities of the states at the remote estimator are given by:

$$\begin{aligned} p\left(s_0^\ell \mid s_i^\ell, \gamma_i^\ell\right) &= \gamma_i^\ell\lambda_1 + \left(1-\gamma_i^\ell\right)\lambda_2, \\ p\left(s_{i+1}^\ell \mid s_i^\ell, \gamma_i^\ell\right) &= 1 - \gamma_i^\ell\lambda_1 - \left(1-\gamma_i^\ell\right)\lambda_2, \end{aligned}$$

for $i = 0,1,2,\dots$. The one-stage cost function for the client ℓ at state s_i^ℓ is

$$\begin{aligned} c\left(s_i^\ell, \gamma_i^\ell\right) = \zeta\left[p\left(s_0^\ell | s_i^\ell, \gamma_i^\ell\right)s_0^\ell + p\left(s_{i+1}^\ell | s_i^\ell, \gamma_i^\ell\right)s_{i+1}^\ell\right] \\ + (1-\zeta)\gamma_i^\ell W. \end{aligned}$$

For each client ℓ, Theorem 1 holds under any given strategy W by the server in this SSMC scenario. That means there always exists an optimal stationary and deterministic policy for each client. Furthermore, we will show that the optimal strategy is nondecreasing in the state. The following Definition 2 is necessary for the proof of Theorem 4.

Definition 2: For the well ordered sets X and Y and a real-valued function $g(x,y)$ on $X \times Y$, g is *superadditive* if for $x^+ \ge x^-$ in X and $y^+ \ge y^-$ in Y,

$$g(x^+, y^+) + g(x^-, y^-) \ge g(x^+, y^-) + g(x^-, y^+).$$

If the reverse inequality holds, g is *subadditive* [18].

Theorem 4: Each client $\ell's$ optimal stationary and deterministic policy has a threshold structure under each fixed W given by the server. This means only when the remote estimation error covariance is beyond some threshold, the client sensor will choose to use channel 1.

Proof: Here we prove that client $\ell's$ optimal strategy is nondecreasing in the state s_i^ℓ.

Theorem 8.11.4 in [18] states that if the one stage cost and the transition probability satisfy the following conditions in the countable state space model under the average reward criteria, then there exists an optimal monotone policy.

1) $c\left(s_i^\ell, \gamma_i^\ell\right)$ is nondecreasing in s_i^ℓ for all $\gamma_i^\ell \in \mathbb{A}_{C^\ell}$,
2) $\sum_{j=k}^{\infty} p\left(s_j^\ell \mid s_i^\ell, \gamma_i^\ell\right)$ is nondecreasing in s_i^ℓ for all non-negative integer k and $\gamma_i^\ell \in \mathbb{A}_{C^\ell}$;
3) $c\left(s_i^\ell, \gamma_i^\ell\right)$ is a subadditive function on $\mathbb{S} \times \mathbb{A}_{C^\ell}$; and
4) $\sum_{j=0}^{\infty} p\left(s_j^\ell \mid s_i^\ell, \gamma_i^\ell\right) u\left(s_j^\ell\right)$ is a superadditive function on $\mathbb{S} \times \mathbb{A}_{C^\ell}$ for nonincreasing u.

We verify these conditions point by point: Condition (1) holds trivially for any given W. As for condition (2),

$$\sum_{j=k}^{\infty} p\left(s_j^\ell \mid s_i^\ell, \gamma_i^\ell\right) = \begin{cases} 1, & k = 0, \\ 1 - \gamma_i^\ell\lambda_1 - (1-\gamma_i^\ell)\lambda_2, & 1 \le k \le i+1, \\ 0, & k > i+1, \end{cases}$$

is invariant in s_i^ℓ. For the one stage cost function, we have

$$\begin{aligned}
&\left[c\left(s_{i+1}^\ell,1\right)+c\left(s_i^\ell,0\right)\right]-\left[c\left(s_{i+1}^\ell,0\right)+c\left(s_i^\ell,1\right)\right]\\
&=\zeta\left[\lambda_1 s_0^\ell+(1-\lambda_1)s_{i+2}^\ell\right]+(1-\zeta)W\\
&\quad+\zeta\left[\lambda_2 s_0^\ell+(1-\lambda_2)s_{i+1}^\ell\right]-\zeta\left[\lambda_2 s_0^\ell+(1-\lambda_2)s_{i+2}^\ell\right]\\
&\quad-\zeta\left[\lambda_1 s_0^\ell+(1-\lambda_1)s_{i+1}^\ell\right]-(1-\zeta)W\\
&=-\zeta(\lambda_1-\lambda_2)\left(s_{i+2}^\ell-s_{i+1}^\ell\right)\\
&\le 0.
\end{aligned}$$

This inequality implies the subadditivity of $c\left(s_i^\ell,\gamma_i^\ell\right)$, which verifies condition (3). The superadditivity in condition (4) is shown as follows:

$$\begin{aligned}
&\left[\sum_{j=0}^{\infty}p\left(s_j^\ell \mid s_{i+1}^\ell,1\right)u\left(s_j^\ell\right)+\sum_{j=0}^{\infty}p\left(s_j^\ell \mid s_i^\ell,0\right)u\left(s_j^\ell\right)\right]\\
&\quad-\left[\sum_{j=0}^{\infty}p\left(s_j^\ell \mid s_{i+1}^\ell,0\right)u\left(s_j^\ell\right)+\sum_{j=0}^{\infty}p\left(s_j^\ell \mid s_i^\ell,1\right)u\left(s_j^\ell\right)\right]\\
&=\lambda_1 u\left(s_0^\ell\right)+(1-\lambda_1)u\left(s_{i+2}^\ell\right)+\lambda_2 u\left(s_0^\ell\right)+(1-\lambda_2)u\left(s_{i+1}^\ell\right)\\
&\quad-\left[\lambda_2 u\left(s_0^\ell\right)+(1-\lambda_2)u\left(s_{i+2}^\ell\right)\right]\\
&\quad-\left[\lambda_1 u\left(s_0^\ell\right)+(1-\lambda_1)u\left(s_{i+1}^\ell\right)\right]\\
&=(\lambda_1-\lambda_2)\left[u\left(s_{i+1}^\ell\right)-u\left(s_{i+2}^\ell\right)\right]\\
&\ge 0,
\end{aligned}\tag{36}$$

where the last line holds since $u(\cdot)$ is a nonincreasing function. As all the conditions are verified, this completes the proof. ■

Intuitively, Theorem 4 indicates that, once the price W is designed, the client's optimal policy is to choose to use channel 1 when the trace of the remote estimation error covariance is larger than some threshold.

B. Server's Optimal Policy

In the last subsection, we obtained the threshold structure of each client's optimal strategy once the server's price strategy W is set. We will next show that the threshold is nondecreasing in the price W.

Theorem 5: The threshold of each client's optimal strategy is nondecreasing in W.

Proof: Suppose that when the server's policy is W', client ℓ's optimal threshold structure strategy is $\{\gamma_i\}^\ell$ and it begins to use channel 1 at the state s_j^ℓ. The induced probability distribution of the states is denoted by $\{\pi_i\}^\ell$. Due to the optimality of $\{\gamma_i\}^\ell$,

$$J_C^\ell=\sum_{i=0}^{\infty}\pi_i s_i^\ell+\sum_{i=j}^{\infty}\pi_i W'\le\sum_{i=0}^{\infty}\pi_i^\flat s_i^\ell+\sum_{i=j-1}^{\infty}\pi_i^\flat W', \tag{37}$$

where $\{\pi_i^\flat\}^\ell$ is induced by the policy $\{\gamma_i^\flat\}^\ell$ as illustrated in Fig. 7. $\{\gamma_i^\flat\}^\ell$ means starts to use channel 1 at s_{j-1}^ℓ.

Suppose that the server increases the price to $W'+\epsilon$, where $\epsilon>0$. We need to prove that the client's cost cannot be smaller if the threshold is less than s_j^ℓ. We compare the average cost induced by $\{\gamma_i\}^\ell$ and $\{\gamma_i^\flat\}^\ell$ under $W'+\epsilon$. First, we show that

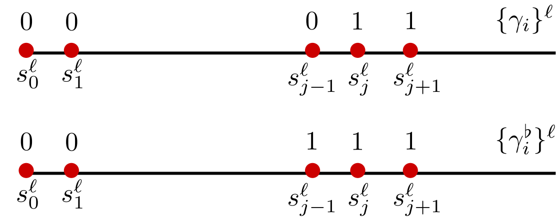


Fig. 7. Illustration of the nondecreasing threshold.

$\sum_{i=j-1}^{\infty}\pi_i^\flat-\sum_{i=j}^{\infty}\pi_i\ge 0$ by direct calculation:

$$\begin{aligned}
&\sum_{i=j-1}^{\infty}\pi_i^\flat-\sum_{i=j}^{\infty}\pi_i\\
&=\frac{\lambda_1\lambda_2}{\lambda_1-(\lambda_1-\lambda_2)(1-\lambda_2)^{j-1}}\frac{(1-\lambda_2)^{j-1}}{\lambda_1}\\
&\quad-\frac{\lambda_1\lambda_2}{\lambda_1-(\lambda_1-\lambda_2)(1-\lambda_2)^{j}}\frac{(1-\lambda_2)^{j}}{\lambda_1}\\
&=\frac{\lambda_1\lambda_2^2(1-\lambda_2)^{j-1}}{\left[\lambda_1-(\lambda_1-\lambda_2)(1-\lambda_2)^{j-1}\right]\left[\lambda_1-(\lambda_1-\lambda_2)(1-\lambda_2)^{j}\right]}\\
&\ge 0.
\end{aligned}\tag{38}$$

And then we show that the client ℓ's average cost induced by the scheduling policy $\{\gamma_i\}^\ell$ is no larger than that induced by $\{\gamma_i^\flat\}^\ell$ under the pricing $W'+\epsilon$:

$$\begin{aligned}
&\sum_{i=0}^{\infty}\pi_i^\flat s_i^\ell+\sum_{i=j-1}^{\infty}\pi_i^\flat\left(W'+\epsilon\right)\\
&=\sum_{i=0}^{\infty}\pi_i^\flat s_i^\ell+\sum_{i=j-1}^{\infty}\pi_i^\flat W'+\sum_{i=j-1}^{\infty}\pi_i^\flat\epsilon\\
&\ge\sum_{i=0}^{\infty}\pi_i s_i^\ell+\sum_{i=j}^{\infty}\pi_i W'+\sum_{i=j-1}^{\infty}\pi_i^\flat\epsilon\\
&\ge\sum_{i=0}^{\infty}\pi_i s_i^\ell+\sum_{i=j}^{\infty}\pi_i W'+\sum_{i=j}^{\infty}\pi_i\epsilon\\
&=\sum_{i=0}^{\infty}\pi_i s_i^\ell+\sum_{i=j}^{\infty}\pi_i\left(W'+\epsilon\right).
\end{aligned}\tag{39}$$

The first inequality holds by virtue of (37) and the second inequality is deduced from (38). For any other threshold which is less than s_j^ℓ, the argument is similar and we omit it here. As a result, when the price W increases, the client cannot pay less if the threshold becomes smaller. Thus the threshold of the client's optimal strategy is nondecreasing in W. ■

For any given W, client ℓ's threshold is denoted as $T_\ell(W)$, which is a nondecreasing function in W. Intuitively, it means that the higher the price is set, the lower the chance of using the high transmission quality channel is. In this asymptotically stable system scenario, once the price W is higher than some value W_ℓ, the client ℓ will never use channel 1. Thus the revenue from this client is 0 for the server. Let $W_{max}=\max\{W_1,W_2,\ldots,W_M\}$.

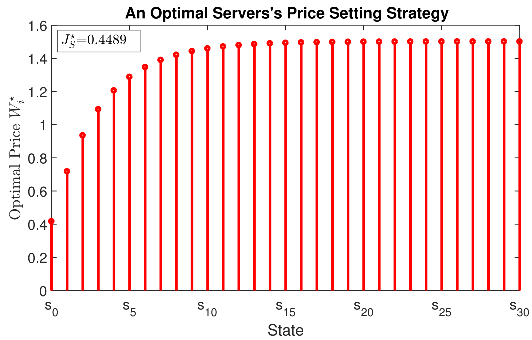


Fig. 8. An optimal price setting strategy.

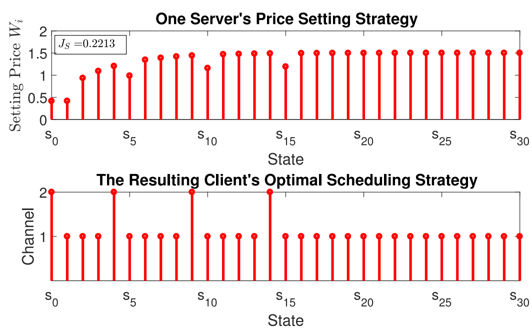


Fig. 9. Reduce the prices at some states.

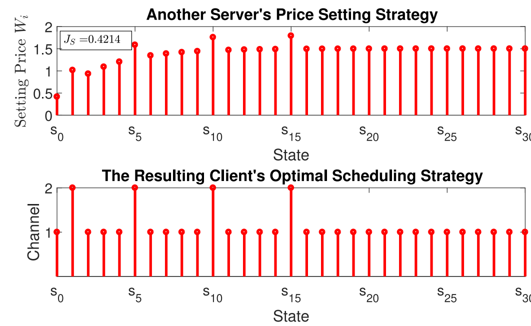


Fig. 10. Improve the prices at some states.

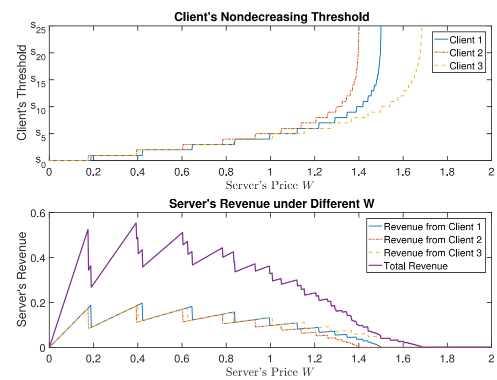


Fig. 11. Client ℓ's threshold and the server's revenue from ℓ.

The server's total revenue is

$$J_S = \sup_{W \in [0, W_{max}]} \sum_{\ell=1}^{M} \frac{\lambda_2 (1-\lambda_2)^{T_\ell(W)}}{\lambda_1 - (\lambda_1 - \lambda_2)(1-\lambda_2)^{T_\ell(W)}} W, \quad (40)$$

where the supremum always exists. In other words, there exists an optimal $W^\star$ such that the server benefits most.

VI. Simulation

In this section, numerical examples are provided to illustrate the scenarios of both the SSSC and SSMC Stackelberg game frameworks.

Consider a single-process case with

$$A = \begin{bmatrix} 0.85 & 0 \\ 0 & 0.9 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 0 \end{bmatrix},$$

$$Q = \begin{bmatrix} 0.3 & 0 \\ 0 & 0.3 \end{bmatrix}, \quad R = 0.3,$$

and communication network channels given by $\lambda_1 = 0.9, \lambda_2 = 0.2$. Let the weight parameter in (8) be $\zeta = 0.5$, which means attaching the same importance to both the estimation error covariance and channel costs. In this scenario, Fig. 8 shows that the critical optimal price setting curve of the server is nondecreasing in the state. The resulting client's optimal strategy is to use channel 1 all the time. The server's optimal revenue under this policy pair $(\{W_i^\star\}, \{\gamma_i^\star\})$ is $J_S^\star = 0.4489$.

Suppose now that the server reduces the prices at states s_1, s_5, s_{10} and s_{15} by 0.3, respectively. The resulting client's optimal strategy is then as shown in Fig. 9. The server's revenue is $J_S = 0.2213$ i.e., is reduced by half. On the contrary, if the server increases the prices at states s_1, s_5, s_{10} and s_{15} by 0.3, respectively, the resulting client's optimal strategy is as shown in Fig. 10. The server's revenue is $J_S = 0.4214$. We notice that the client uses channel 2 at those states due to the expensive prices. These two cases illustrate the optimality of $\{W_i^\star\}$.

In the SSMC scenario (multiple-processes), consider a three clients case with system parameters as follows:

$$A_1 = \begin{bmatrix} 0.85 & 0 \\ 0 & 0.9 \end{bmatrix}, \quad C_1 = \begin{bmatrix} 1 & 0 \end{bmatrix},$$

$$Q_1 = \begin{bmatrix} 0.3 & 0 \\ 0 & 0.3 \end{bmatrix}, \quad R_1 = 0.3,$$

$$A_2 = \begin{bmatrix} 0.85 & 0 \\ 0 & 0.89 \end{bmatrix}, \quad C_2 = \begin{bmatrix} 1 & 0 \end{bmatrix},$$

$$Q_2 = \begin{bmatrix} 0.28 & 0 \\ 0 & 0.28 \end{bmatrix}, \quad R_2 = 0.28,$$

$$A_3 = \begin{bmatrix} 0.85 & 0 & 0 \\ 0 & 0.87 & 0 \\ 0 & 0 & 0.85 \end{bmatrix}, \quad C_3 = \begin{bmatrix} 0 & 1 & 0 \end{bmatrix},$$

$$Q_3 = \begin{bmatrix} 0.27 & 0 & 0 \\ 0 & 0.27 & 0 \\ 0 & 0 & 0.27 \end{bmatrix}, \quad R_3 = 0.27.$$

Let all the weight parameters ζ^ℓ in (34) be 0.5. The relative value iteration algorithm (See Chapter 8.5.5 in [18]) is applied to solve the MDP problem. Fig. 11 portrays the nondecreasing threshold of the client ℓ with respect to the price W given by the server, and the server's revenues from ℓ, $\ell = \{1, 2, 3\}$. In

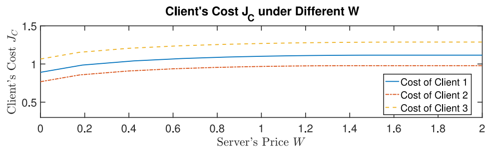


Fig. 12. Client ℓ's cost J_C^ℓ under the optimal selection policy for each W.

this SSMC scenario, we can directly find from Fig. 11 that the optimal pricing strategy is $W^\star = 0.3925$, and the maximum total revenue is $J_S^\star = 0.5541$. Fig. 12 displays the client ℓ's cost J_C^ℓ under its optimal selection policy for each W given by the server. As W increases, the client ℓ's minimum cost J_C^ℓ converges to $J_C^{\ell\star}$, which can be calculated using Equation (21) in the SSSC scenario. It can be seen that under some price W, its resulting minimum J_C^ℓ can be smaller than $J_C^{\ell\star}$. This implies that Assumption 3 on the server's "constant" pricing strategy breaks the absolute monopoly and brings more vitality to clients.

VII. Conclusion and Future Work

In this work, we presented a Stackelberg game-theoretic framework for communication channel pricing and selection. The server aims to reap the most revenue, while the clients aim to minimize the remote estimation error covariance and the communication cost. In both the SSSC and SSMC scenarios, we proved the existence of an optimal stationary and deterministic strategy for each client. Our results further showed that the optimal policy pair of the server and clients features special structures (under suitable assumptions).

There are several directions from which this work can progress. First, if the processes of some clients are unstable (i.e., $\rho(A_\ell) > 1$), then it may be possible that the server's maximum revenue will become extremely large if the successful channel transmission probabilities are quite small. Second, the single server is a monopoly player and thus clients have little freedom. One may consider scenarios with multiple servers. Finally, imperfect information Stackelberg games can be explored when clients cannot receive ACKs reliably.

Appendix A
Proof of Lemma 1

Proof: Lemma A.1 in [19] shows that the inequality (3) holds. Since the trace of a positive semi-definite matrix is always non-negative, it is trivial to verify that the inequality (4) holds. ■

Appendix B
Proof of Theorem 1

Proof: When the state space $\mathbb{S}$ of an MDP is countable, instead of finite, an optimal stationary and deterministic policy may not exist and sometimes, the average cost may not be bounded (See Example 7.1.3 and Example 7.1.4 in [11]). Some conditions are needed to guarantee the existence of an optimal stationary and deterministic policy in the countable state space case. One way to verify the existence is to show that the (SEN) assumptions in [11] hold. As a consquence, according to Theorem 7.5.6. in [11], the average cost optimality equation (ACOE) holds, which here refers to Equation (15) in our problem. Hence, we need to check the three conditions in the (SEN) assumptions which are stated in [11] as follows. Define the discounted value function f_α under the optimal policy $\{\gamma_k\}$ with the discount parameter α as:

$$f_\alpha(s_i) \triangleq \inf_{\{\gamma_k\}} \sum_{k=1}^{\infty} \alpha^k \mathbb{E}\left[c_k\left(\mathrm{Tr}\{P_{k-1}\}, \gamma_k\right) \mid \mathrm{Tr}\{P_0\} = s_i\right].$$

Let s_0 be a distinguished state. Define a function $\Delta f_\alpha(s_i) \triangleq f_\alpha(s_i) - f_\alpha(s_0)$.

(SEN1). The quantity $(1-\alpha) f_\alpha(s_0)$ is bounded for all $\alpha \in (0,1)$.

(SEN2). There exists a non-negative finite function $\mathcal{M}$ such that $\Delta f_\alpha(s_i) \le \mathcal{M}(s_i)$ for all $s_i \in \mathbb{S}$ and $\alpha \in (0,1)$.

(SEN3). There exists a non-negative finite constant $\mathcal{L}$ such that $-\mathcal{L} \le \Delta f_\alpha(s_i)$ for all $s_i \in \mathbb{S}$ and $\alpha \in (0,1)$.

We verify these three conditions one by one. First, we consider a simple "always using channel 2" policy, i.e. $\gamma_k = 0$ for all $k \ge 1$. Then the resulting one-stage cost function is

$$\begin{aligned} c_k\left(\mathrm{Tr}\left\{P_{k-1}\right\}, 0\right) &= \zeta \mathrm{Tr}\left\{\mathbb{E}\left[P_k\right]\right\} \\ &= \zeta\left[\lambda_2 \mathrm{Tr}\{\overline{P}\} + (1-\lambda_2)\,\mathrm{Tr}\{h\left(P_{k-1}\right)\}\right]. \end{aligned}$$

Since that our problem considers an asymptotically stable system, $\mathrm{Tr}\{h\left(P_k\right)\}$ is always upper bounded by some positive constant. And therefore, there exists a positive constant $\mathcal{C}$ such that

$$\lim_{N\to\infty} \frac{1}{N} \sum_{k=1}^{N} \mathbb{E}\left[c_k\left(\mathrm{Tr}\{P_{k-1}\}, 0\right)\right] = \mathcal{C} < \infty$$

for any initial value $\mathrm{Tr}\{P_0\}$. Notice that the discounted cost is less than the average cost and hence we have

$$\sum_{k=1}^{\infty} \alpha^k \mathbb{E}\left[c_k\left(\mathrm{Tr}\{P_{k-1}\}, 0\right)\right] < \mathcal{C} < \infty$$

which implies the condition (SEN1).

Second, for any initial value s_i, the following inequality always holds:

$$0 < f_\alpha(s_i) \le \sum_{k=1}^{\infty} \alpha^k \mathbb{E}\left[c_k\left(\mathrm{Tr}\{P_{k-1}\}, 0\right)\right] < \mathcal{C} < \infty.$$

It is trivial to see that $\Delta f_\alpha(s_i)$ has both upper and lower bound for all $s_i \in \mathbb{S}$ and $\alpha \in (0,1)$. The conditions (SEN2) and (SEN3) are then verified. Once the three conditions are checked, we can conclude that there exists an optimal stationary and deterministic sensor selection policy which satisfies the ACOE. ■

Appendix C
Proof of Lemma 2

Proof: Assume that $s_{-1} = 0$ and let k be arbitrary. Then

$$\sum_{j=0}^{\infty} m_j s_j = \sum_{j=0}^{\infty} \left[m_j \sum_{i=0}^{j} \left(s_i - s_{i-1}\right)\right]$$

$$
\begin{aligned}
&= \sum_{i=0}^{\infty}\left[(s_i - s_{i-1})\sum_{j=i}^{\infty} m_j\right] \\
&= \sum_{i=1}^{\infty}\left[(s_i - s_{i-1})\sum_{j=i}^{\infty} m_j\right] + s_0\sum_{j=0}^{\infty} m_j \\
&\geq \sum_{i=1}^{\infty}\left[(s_i - s_{i-1})\sum_{j=i}^{\infty} n_j\right] + s_0\sum_{j=0}^{\infty} n_j \\
&= \sum_{j=0}^{\infty} n_j s_j. \qquad (41)
\end{aligned}
$$

Due to the boundedness of sequence $\{s_j\}$ and (18), the limits $\sum_{j=0}^{\infty} m_j s_j$ and $\sum_{j=0}^{\infty} n_j s_j$ always exist. Thus the inequality (41) makes sense. ■

References

[1] R. A. Gupta and M.-Y. Chow, "Networked control system: Overview and research trends," *IEEE Trans. Ind. Electron.*, vol. 57, no. 7, pp. 2527–2535, Jul. 2010.

[2] W. Zhang, M. S. Branicky, and S. M. Phillips, "Stability of networked control systems," *IEEE Control Syst.*, vol. 21, no. 1, pp. 84–99, Feb. 2001.

[3] R. M. Murray, *Control in an Information Rich World: Report of the Panel on Future Directions in Control, Dynamics, and Systems*. Philadelphia, PA, USA: SIAM, 2003.

[4] G. C. Walsh, H. Ye, and L. G. Bushnell, "Stability analysis of networked control systems," *IEEE Trans. Control Syst. Technol.*, vol. 10, no. 3, pp. 438–446, May 2002.

[5] J. N. Al-Karaki and A. E. Kamal, "Routing techniques in wireless sensor networks: A survey," *IEEE Wireless Commun.*, vol. 11, no. 6, pp. 6–28, Dec. 2004.

[6] W. Sun, L. Guijarro, V. Pla, and J. R. Vidal, "Joint resource and price competition in wireless sensor network-based service provision," in *Proc. IEEE Consum. Commun. Netw. Conf.*, 2017, pp. 254–259.

[7] N. C. Luong, P. Wang, D. Niyato, Y.-C. Liang, Z. Han, and F. Hou, "Applications of economic and pricing models for resource management in 5G wireless networks: A survey," *IEEE Commun. Surveys Tut.*, to be published, doi: 10.1109/COMST.2018.2870996.

[8] B. Gu, M. Dong, C. Zhang, Z. Liu, and Y. Tanaka, "Real-time pricing for on-demand bandwidth reservation in SDN-enabled networks," in *Proc. IEEE Consum. Commun. Netw. Conf.*, 2017, pp. 696–699.

[9] S. Yun, J. Lee, S. S. Newaz, and J. K. Choi, "Energy efficient pricing scheme for multi-homing in heterogeneous wireless access networks: A game theoretic model and its analysis," in *Proc. IEEE Wireless Commun. Netw. Conf.*, 2015, pp. 1672–1677.

[10] Y. Ni, A. S. Leong, D. E. Quevedo, and L. Shi, "Game-theoretic pricing and selection with fading channels," in *Proc. IEEE 11th Asian Control Conf.*, 2017, pp. 934–939.

[11] L. I. Sennott, *Stochastic Dynamic Programming and the Control of Queueing Systems*. Hoboken, NJ, USA: Wiley, 2009, vol. 504.

[12] B. D. Anderson and J. B. Moore, *Optimal Filtering*. North Chelmsford, MA, USA: Courier Corp., 2012.

[13] L. Schenato, "Optimal estimation in networked control systems subject to random delay and packet drop," *IEEE Trans. Autom. Control*, vol. 53, no. 5, pp. 1311–1317, Jun. 2008.

[14] M. Simaan and J. B. Cruz, "On the Stackelberg strategy in nonzero-sum games," *J. Optim. Theory Appl.*, vol. 11, no. 5, pp. 533–555, 1973.

[15] D. Fudenberg and J. Tirole, *Game Theory*. Cambridge, MA, USA: MIT Press, 1991, vol. 726.

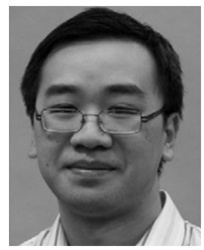
[16] M. J. Osborne and A. Rubinstein, *A Course in Game Theory*. Cambridge, MA, USA: MIT Press, 1994.

[17] M. B. Lieberman and D. B. Montgomery, "First-mover advantages," *Strategic Manage. J.*, vol. 9, no. S1, pp. 41–58, 1988.

[18] M. L. Puterman, *Markov Decision Processes: Discrete Stochastic Dynamic Programming*, 1st ed. New York, NY, USA: Wiley, 1994.

[19] L. Shi, M. Epstein, and R. M. Murray, "Kalman filtering over a packet-dropping network: A probabilistic perspective," *IEEE Trans. Autom. Control*, vol. 55, no. 3, pp. 594–604, Mar. 2010.

Yuqing Ni received the B.Eng. degree (Hons.) from the College of Control Science and Engineering, Zhejiang University, Hangzhou, China, in 2016. She is currently working toward the Ph.D. degree at the Department of Electronic and Computer Engineering, Hong Kong University of Science and Technology, Kowloon, Hong Kong. From April 2019 to June 2019, she was a Visiting Student with the Department of Automatic, KTH Royal Institute of Technology, Stockholm, Sweden. Her research interests include security and privacy in cyber-physical system, networked state estimation, and wireless sensor networks.



Alex S. Leong was born in Macau in 1980. He received the B.S. degree in mathematics and the B.E. degree in electrical engineering, both in 2003, and the Ph.D. degree in electrical engineering, in 2008, all from the University of Melbourne, Melbourne, Vic., Australia. He was with the Department of Electrical and Electronic Engineering, University of Melbourne from 2008 to 2015. He is currently a Research Associate with Paderborn University, Paderborn, Germany. His research interests include networked control systems, signal processing for sensor networks, and statistical signal processing. He was the recipient of the L.R.East Medal from Engineers Australia in 2003, an Australian Postdoctoral Fellowship from the Australian Research Council in 2009, and the Discovery Early Career Researcher Award from the Australian Research Council in 2012.

Daniel E. Quevedo (S'97–M'05–SM'14) received the Ingeniero Civil Electrónico and M.Sc. degrees from Universidad Técnica Federico Santa María, Valparaso, Chile, both in 2000, and the Ph.D. degree from the University of Newcastle, Newcastle, N.S.W., Australia, in 2005.

He is the Head of the Chair of Automatic Control (*Regelungs- und Automatisierungstechnik*) with Paderborn University, Paderborn, Germany. His research interests are in control of networked systems and of power converters.

Prof. Quevedo is an Associate Editor for the IEEE CONTROL SYSTEMS MAGAZINE, an Editor for the *International Journal of Robust and Nonlinear Control*, and a Past Chair for the IEEE Control Systems Society: *Technical Committee on Networks and Communication Systems*. He was the recipient of the IEEE Conference on Decision and Control Best Student Paper Award in 2003 and was also a finalist in 2002. In 2009, he was the recipient of a five-year research fellowship from the Australian Research Council. He was the co-recipient of the 2018 IEEE Transactions on Automatic Control George S. Axelby Outstanding Paper Award. He was supported by a Full Scholarship from the Alumni Association during his time at the Universidad Técnica Federico Santa María and was the recipient of several university-wide prizes upon graduating.

Ling Shi received the B.S. degree in electrical and electronic engineering from the Hong Kong University of Science and Technology, Kowloon, Hong Kong, in 2002, and the Ph.D. degree in control and dynamical systems from the California Institute of Technology, Pasadena, CA, USA, in 2008. He is currently an Associate Professor with the Department of Electronic and Computer Engineering, and the Associate Director with the Robotics Institute, both at the Hong Kong University of Science and Technology. His research interests include cyber-physical systems security, networked control systems, sensor scheduling, event-based state estimation, and exoskeleton robots. He was an Editorial Board Member for The European Control Conference during 2013–2016. He was a Subject Editor for the *International Journal of Robust and Nonlinear Control* (2015–2017). He has been an Associate Editor for the IEEE TRANSACTIONS ON CONTROL OF NETWORK SYSTEMSsince July 2016, and an Associate Editor for the IEEE CONTROL SYSTEMS LETTERSsince Feb 2017. He was also an Associate Editor for a special issue on secure control of cyber physical systems in the IEEE TRANSACTIONS ON CONTROL OF NETWORK SYSTEMS in 2015–2017. He was the General Chair of the 23rd International Symposium on Mathematical Theory of Networks and Systems (MTNS 2018).