Automatica 123 (2021) 109342

Contents lists available at ScienceDirect

Automatica

journal homepage: www.elsevier.com/locate/automatica

Brief paper

Remote state estimation with usage-dependent Markovian packet losses ☆

Jiazheng Wang [a], Xiaoqiang Ren [b,*], Subhrakanti Dey [c], Ling Shi [a]

[a] Department of Electronic and Computer Engineering, Hong Kong University of Science and Technology, Hong Kong
[b] School of Mechatronic Engineering and Automation, Shanghai University, China
[c] Hamilton Institute, Maynooth University, Ireland

ARTICLE INFO

Article history:
Received 19 September 2019
Received in revised form 20 May 2020
Accepted 6 October 2020
Available online 9 November 2020

Keywords:
Kalman filters
State estimation
Networked control systems
Markov decision processes

ABSTRACT

In this paper, we consider a problem of packet scheduling in the setting of remote estimation with usage-dependent Markovian packet losses. A sensor measures the state of a discrete-time linear process, computes the estimate via a local Kalman filter, and sends the packets to a remote estimator via a network. The link state evolves as a two-state Markov chain, and its state transition depends on the network usage. The aim is to design the scheduling policy which balances the estimation quality and the energy consumption. We identify the problem as a Markov decision process (MDP) and prove the structural properties of the optimal policy. Furthermore, based on the structural properties, we derive the sufficient and necessary condition of the mean square stability of the remote estimator. Simulation examples are provided to illustrate the results.

1. Introduction

With fast developments in the electronic techniques, the number of devices embedded with the computation and wireless communication capabilities has a dramatic increase over the last decade (Gunes et al., 2014). The systems which integrate the physical world with the computation and communication capacities have been known as the Cyber-physical systems (CPSs) (Rajkumar et al., 2010). An important class of CPSs is the networked state estimation systems, like environment monitoring and multi-agent tracking (Lee, 2008), where sensors are deployed to collect data from the physical system and transmit the signals over a network. The design of such networked state estimation system faces several challenges. The uncertainties like packet losses or delays in the data transmission have strong influence on the system performance. As the sensors are usually powered by the batteries, the energy saving mechanism is also crucial as well.

The time-average error covariance of the remote estimate has been widely used as a performance metric to describe the estimation quality. Despite the fact that the time-average error covariance serves as the nice performance metric for stable LTI systems, it can be unbounded for some unstable systems and thus fails to capture the system characteristics. The condition which ensures the stability of remote estimator is needed. Sinopoli et al. (2004) studied the mean-square stability under the assumption that packet dropping processes are i.i.d., and derived a threshold condition that ensures the stability of the remote estimator. Fletcher et al. (2004) modeled the packet dropping processes as a finite state Markov chain, proposed a suboptimal linear estimator, and studied the stability using the linear matrix inequality. Huang and Dey (2007) studied a similar problem, and derived sufficient conditions for the stability of the optimal estimator in the general vector case. The results have been extended by Xu et al. (2018) to more general cases under signal-to-noise ratio constraint, which presented a necessary and sufficient condition for mean-square stability of second-order system.

Motivated by the important role which resource scheduling plays in the online operation, the design problem of optimal scheduling policy has been widely studied in recent years. There are mainly two types of packet scheduling policies in the literature. The first type is offline where the data transmission operation solely depends on the realization of packet dropping processes, e.g. Leong et al. (2017) and Trimpe and D'Andrea (2014). Compared with online polices, the offline policies are much easier to obtain via numerical algorithm and to be implemented. The second type is online and the data transmission also depends on the realization of the system state (or measurement). Threshold-type online policy has been studied in Battistelli et al. (2012), Lipsa and Martins (2011), Ren et al. (2018) and Wu et al. (2013). The online policies usually outperform the offline

☆ The material in this paper was not presented at any conference. This paper was recommended for publication in revised form by Associate Editor Juan C. Aguero under the direction of Editor Torsten Söderström.

* Corresponding author.
E-mail addresses: jwangck@connect.ust.hk (J. Wang), xqren@shu.edu.cn (X. Ren), subhrakanti.dey@angstrom.uu.se (S. Dey), eesling@ust.hk (L. Shi).

https://doi.org/10.1016/j.automatica.2020.109342

policies. However, the optimal design of online policies in general vector state space inevitably involves a partially observable Markov decision process, which is computationally intractable (or PSPACE-complete) (Papadimitriou & Tsitsiklis, 1987). In this work we will focus on the offline polices which in general are more computationally efficient.

Most existing results are established on the assumption that the transition of the link state is independent of the network usage, wherein we use the term *link state* to denote the quality of the network service. However, in many scenarios where the channel access is contention-based (e.g. Avrachenkov et al., 2013; Füßler et al., 2003; Kong et al., 2004), the simultaneous data transmission without coordination between users may easily lead to the congestion, which results in the performance degeneration of the overall system. For example, in the scheme of K (user's) flows sharing a bottleneck router, the probability of successful data transmission relies on the number of packets in the buffer of the router. In this particular example, the number of packets in the buffer indicates the quality of network service, and its transition is influenced by the transmission decision of users. To capture this effect, we model the *link state* as a two-state Markov chain, and the results can be easily generalized to any finite-state Markov chains. We note that the previous results about Markovian packet losses cannot be directly applied, for the link state transition depends on both the link state and the scheduling policy. We extend the results in Huang and Dey (2007) by considering the correlation between network usage and link state transition. We study the optimal transmission scheduling under Markovian packet losses, and then identify the problem as an Markov decision process (MDP). Further, the stability of the remote state estimator is addressed. The main contributions of the paper are summarized as follows,

1 Unlike previous works where the link state transition solely depends on the current link state, we consider the correlation between the link state transition and the network usage, and model the link state transition as a switched two-state Markov chain.
2 We formulate the packet scheduling over the Markovian packet dropping network as an MDP, and prove the structural property of the optimal policy.
3 Based on the above structural result, we provide the necessary and sufficient condition of the stability of the remote estimator under the optimal policy. We are able to recover the results in Huang and Dey (2007) for scalar systems.

The remainder of this paper is organized as follows. In Section 2, the mathematical model of the considered problem is given. In Section 3, the problem is identified as an MDP, and the structural properties of the optimal policy is proved. Based on the structural properties, the necessary and sufficient condition of the stability of the remote estimator is derived. In Section 4, numerical examples are given to demonstrate the results. In Section 5 we state our conclusions. All proofs are reported in Appendix.

Notations. The set of real numbers is denoted as $\mathbb{R}$, the set of (positive) natural number is denoted as $\mathbb{N}$ ($\mathbb{N}^+$). When a matrix X is positive definite (semidefinite), we write $X \succ 0$ ($X \succeq 0$). For a matrix X, $X^\intercal$, TrX and $\rho(X)$ denotes its transpose, trace, and spectral radius, respectively. $\Pr(\cdot)$ ($\Pr(\cdot|\cdot)$) denotes the (conditional) probability, and $\mathbb{E}(\cdot|\cdot)$, ($\mathbb{E}(\cdot|\cdot)$) denotes the (conditional) expectation. $\mathcal{B}(\mathbb{S})$ denotes the collection of bounded function mapping from $\mathbb{S}$ to $\mathbb{R}$. For a given policy π, $\mathbb{E}^\pi$ denotes the expectation under the policy π. For a function h and an integer i, h^i denotes ith function composition of h.

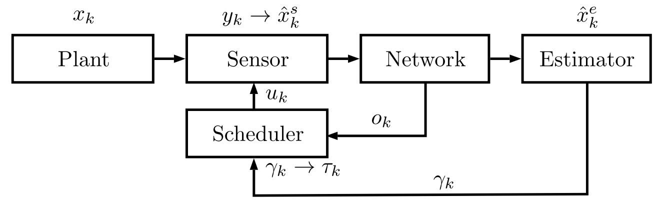


Fig. 1. System model.

2. Problem setup

2.1. System model

We consider the remote state estimation of a linear time-invariant system. The processes dynamics and sensor measurement equation are given as follows:

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

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

where $x_k \in \mathbb{R}^n$ is the state of the process at time k, $y_k \in \mathbb{R}^m$ is the sensor measurement, The noises $w_k \in \mathbb{R}^n$ and $v_k \in \mathbb{R}^m$ are i.i.d. zero-mean white Gaussian random vectors with finite covariance $\Sigma_w \geq 0$, $\Sigma_v \geq 0$, respectively. The initial state x_0 is also a zero-mean Gaussian with covariance $\Sigma_0 \geq 0$, which is independent from w_k, v_k, $k \in N$. We assume that the pair (A, Σ_w) is stabilizable, and the pair (A, H) is detectable.

The sensor is assumed to be smart in the sense that it has sufficient computation capability to run a local Kalman filter to compute the optimal MSE estimate. We denote by $\hat{x}_k^s$ the optimal MSE at the sensor,

$$\hat{x}_k^s \triangleq \mathbb{E}[x_k|y_0, \ldots, y_k]. \tag{3}$$

As the error covariance matrix of a Kalman filter reaches its steady state exponentially fast, without loss of generality, we make the following assumption:

Assumption 1. The error covariance matrix of the local estimate $\hat{x}_k^s$ has reached its steady state, which is denoted by $\overline{P} \triangleq \mathbb{E}[(x_k - \hat{x}_k^s)(x_k - \hat{x}_k^s)^\top]$.

The sensor is equipped with a transmission scheduler (See Fig. 1), which determines whether or not the local estimate $\hat{x}_k^s$ should be sent to the estimator via the network, and the transmission mechanism will be explained in detail in the next part.

2.2. Transmission model and mechanism

We start to introduce the transmission model of the network which delivers the data packet from the sensor to the remote estimator. There exist a group of other network users which also have access to the network, and the load of the network varies with the network usage. The link state at time k is denoted by the random variable o_k: $o_k = 1$ denotes the network is of low load (good state); $o_k = 0$ denotes the network is of high load (bad state). It is assumed that to avoid the persistent congestion, the active queue management, such as the random early detection (RED) method (Floyd & Jacobson, 1993), is adopted in the network, which affects the packet-dropping probability.

The link state evolves as a two-state Markov chain, and an illustration of its transition is shown in Fig. 2. Since the data transmission of the sensor leads to the load change, the transition of the link state depends on both the current link state and the transmission decision of the sensor. Let $u_k \in \{0, 1\}$ denote the

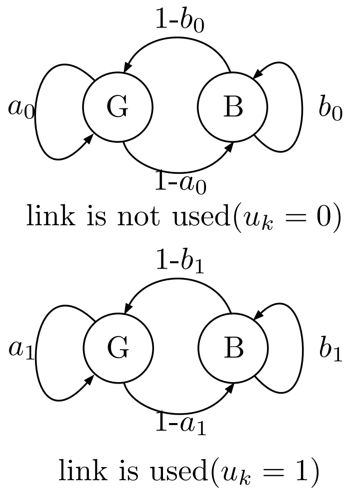


Fig. 2. Illustration of data transmission processes. (G) is for good state or low load, and (B) is for bad state or high load.

communication decision of the sensor at time k: if $u_k = 1$, the data packet containing the local estimate $\hat{x}_k^s$ is sent to the estimator via the network; otherwise it is not sent. We define the following transition probabilities to describe the transition of the link state,

$$\Pr(o_{k+1} = 1 | o_k = 1, u_k = 0) = a_0,$$
$$\Pr(o_{k+1} = 0 | o_k = 0, u_k = 0) = b_0,$$
$$\Pr(o_{k+1} = 1 | o_k = 1, u_k = 1) = a_1,$$
$$\Pr(o_{k+1} = 0 | o_k = 0, u_k = 1) = b_1,$$

where $0 \le a_0, b_0, a_1, b_1 < 1$. Since the data transmission of the sensor leads to higher probability of the high load (or congestion), we further assume that $0 \le a_1 \le a_0 \le 1$, $0 \le b_0 \le b_1 \le 1$.

The uncertainty in the data transmission processes is modeled by a packet-dropping process, which is denoted by a binary random variable γ_k. If $\gamma_k = 1$, the data packet containing the local state estimate $\hat{x}_k^s$ is successfully received by the remote estimator; $\gamma_k = 0$ means that the packet is dropped. The packet-dropping process is mainly due to the fading effects during the wireless communication and the intentional packet rejection caused by RED mechanism in the bottleneck device. γ_k is fedback to the sensor as the acknowledgment (ACK) by the remote estimator. Here we assume that compared with the time scale of the process, the transmission delay related to γ_k can be neglected, which is valid for applications like process control (Cao et al., 2016) and environmental monitoring (Huang et al., 2010). Nevertheless, our result could be extended straightforwardly if the instantaneous link state o_k is not available but with a fixed delay d, say, the sensor only knows o_{k-d} at time k for every k. This can be done by augmenting the state space of the considered MDP problem to be (o_{k-d}, u_{k-1}). We make the following assumption on the data transmission processes and the initial state.

Assumption 2. Given the link state o_k and u_k, the data transmission γ_k at time k is conditionally independent with the data transmission process at other time steps. The probability of successful transmission are given by

$$\Pr(\gamma_k = 1 | o_k, u_k) = \begin{cases} p_0, & \text{if } o_k = 0, u_k = 1, \\ p_1, & \text{if } o_k = 1, u_k = 1, \\ 0, & \text{otherwise,} \end{cases} \quad (4)$$

where $0 \le p_0 \le p_1 \le 1$. To facilitate the analysis, without loss of generality, we assume that the initial packet transmission at $k = 0$ is successful, i.e., $\gamma_0 = 1$.

Since under RED mechanism the probability of packet rejection for high load ($o_k = 0$) is higher that for low load ($o_k = 1$), it is assumed that $p_0 \le p_1$. The link state o_k is broadcast by the network to inform the network users of potential congestion.

2.3. Optimal estimator

The MMSE estimate of the system state at the remote estimator is the conditional mean of the system state, which is denoted by

$$\hat{x}_k^e \triangleq \mathbb{E}(x_k^s | \gamma_0 \hat{x}_0^s, \ldots, \gamma_k^s \hat{x}_k^e, \gamma_0, \ldots, \gamma_k).$$

It is shown in Shi et al. (2011) that the optimal estimate $\hat{x}_k^e$ at the remote estimator and corresponding error covariance P_k^e can be computed as

$$\hat{x}_k^e = \begin{cases} \hat{x}_k^s, & \text{if } \gamma_k = 1, \\ A\hat{x}_{k-1}^e, & \text{otherwise.} \end{cases} \quad (5)$$

$$P_k^e = \begin{cases} \overline{P}, & \text{if } \gamma_k = 1, \\ h(P_{k-1}^e), & \text{otherwise.} \end{cases} \quad (6)$$

where $h(X) \triangleq AXA^\intercal + \Sigma_w$, and h^j denotes jth function composition of h. We denote by τ_k the number of successive packet drops at time k, $\forall k \in \mathbb{N}^+$,

$$\tau_k \triangleq k - 1 - \max\{j | \gamma_j = 1, 0 \le j \le k-1\}. \quad (7)$$

By the basic algebra, from (6) we have

$$P_{k-1}^e = h^{\tau_k}(\overline{P}). \quad (8)$$

2.4. Problem of interests

At time k the scheduler determines the transmission decision u_k based on both the number of successive packet drops τ_k and the link state o_k. An admissible (deterministic) policy π of the transmission scheduler is a sequence of functions $\{\pi_1, \ldots, \pi_k, \ldots\}$ mapping from the number of successive packet drops τ_k and the link state o_k to the action u_k, i.e.,

$$u_k = \pi_k(\tau_k, o_k).$$

We denote by Π the collection of all admissible policies. A policy π is stationary if the policy is time-invariant, i.e. $\pi_k = \pi_1, \forall k \in \mathbb{N}^+$.

In many applications of wireless networks, the data transmission processes are heavily energy-consuming while the size of battery is limited. We assume that the transmission action $u_k = 1$ results in a energy consumption $\mathcal{E} > 0$. We denote by $l : \mathbb{N} \times \{0, 1\} \to \mathbb{R}^+$ the one-stage cost of the system,

$$l(\tau_k, u_k) \triangleq \mathrm{Tr} h^{\tau_k}(\overline{P}) + \lambda \mathcal{E} u_k, \quad (9)$$

where the constant λ can be interpreted as a Lagrange multiplier which balances the estimation quality of the remote estimator against the energy consumption of data transmissions. For an application where the energy consumption is very crucial, a large λ is selected and the sensor is more likely to reduce the frequency of data transmission in order to save the energy consumption. On the contrary, if λ is selected to be quite small, it implies that the transmission energy consumption is minor compared with estimation error.

In this work, we consider the system performance under the average criteria, and the expected average cost over the infinite horizon is given by

$$L(\pi) \triangleq \mathbb{E}^\pi \left[\limsup_{T \to \infty} \frac{1}{T} \sum_{k=1}^{T} l(\tau_k, u_k) \right], \quad (10)$$

where $\mathbb{E}^\pi$ stands for the expected value when using the policy π. We are interested in the following optimization problem.

Problem 1.

$$\min_{\pi\in\Pi} L(\pi).$$

In this work, we focus on both the feasibility of Problem 1 and its optimal solution, which is denoted by π^*. We note that the feasibility of Problem 1 $\big(C(\pi^*) < \infty\big)$ is same with the stability of the remote state estimator $\big(\mathbb{E}^{\pi^*}(\mathrm{Tr}P_k^e) < \infty, \forall k \in \mathbb{N}\big)$. The link state transition has a strong influence on the stability of the remote estimator, and for unstable systems, Problem 1 can be infeasible. In other words, the remote estimator can be unstable under any admissible policies. The following simple example serves as a special case when the stability condition can be derived easily. Suppose that the system parameters are given by

$$A = 2, C = 1, \Sigma_w = 1, \Sigma_v = 1, \mathcal{E} = 0,$$
$$a_0 = 0.5, b_0 = 0.7, p_0 = 0,$$
$$a_1 = 0, b_1 = 1, p_1 = 1.$$

One can easily verify that the optimal policy π^* is to transmit the data packets only if the link state is good, i.e.,

$$\pi^*(\tau_k, o_k) = \begin{cases} 1, & \text{if } o_k = 1, \\ 0, & \text{otherwise.} \end{cases}$$

For this particular case, the system under the optimal policy can be viewed as an identical system sending data packets over an i.i.d. packet dropping network with the probability of successful transmission $p = 1 - b_0 = 0.3$. For $\rho(A)^2(1-p) = 2.8 > 1$, the remote state estimator is unstable under the optimal policy (Theorem 3 in Sinopoli et al., 2004). For the general cases when $p_0 > 0$, it remains difficult to directly obtain the optimal policy as in the above example. The transition of the link state o_k depends on the transmission action u_k, and this coupling results in the difficulty in analyzing the stability of the remote estimator. Before obtaining the optimal policy π^*, it remains formidable to analyze the stability of the remote estimator under the policy π^*.

3. Main results

We first formulate Problem 1 as an MDP problem, and analyze the structural properties of the optimal policy π^*. Based on the threshold structure, we derive the sufficient and necessary condition of the stability of the remote state estimator.

3.1. MDP formulation

In this section, we formulate Problem 1 as an MDP problem with denumerable state and finite action space. We define a tuple $\{\mathbb{S}, \mathbb{U}, \mathbb{P}(\cdot|\cdot,\cdot), c(\cdot,\cdot)\}$ to describe the MDP.

(1) The state space $\mathbb{S} = \mathbb{N} \times \{0, 1\}$: the state $s_k = (\tau_k, o_k)$ is the combination of the holding time of the system τ_k and the link state o_k.
(2) The action space $\mathbb{U} = \{0, 1\}$: the action u_k is the transmission action $u_k \in \{0, 1\}$.
(3) The transition kernel $\mathbb{P}(\cdot|\cdot,\cdot)$: the transition kernel $\mathbb{P}(s_{k+1}|s_k, u_k)$ describes the transition of the state s_k when action u_k is chosen. We have

$$\mathbb{P}(s_{k+1}|s_k, u_k) = \Pr(\tau_{k+1}|\tau_k, o_k, u_k)\Pr(o_{k+1}|o_k, u_k)$$

where

$$\Pr(\tau_{k+1}|\tau_k, o_k, u_k) = \begin{cases} p_j, & \text{if } u_k = 1, \tau_{k+1} = 0, o_k = j, \\ 1 - p_j, & \text{if } u_k = 1, \tau_{k+1} = \tau_k + 1, o_k = j, \\ 1, & \text{if } u_k = 0, \\ 0, & \text{otherwise,} \end{cases}$$

$$\Pr(o_{k+1}|o_k, u_k) = \begin{cases} a_j, & \text{if } u_k = j, o_{k+1} = 1, o_k = 1, \\ 1 - a_j, & \text{if } u_k = j, o_{k+1} = 0, o_k = 1, \\ b_j, & \text{if } u_k = j, o_{k+1} = 0, o_k = 0, \\ 1 - b_j, & \text{if } u_k = j, o_{k+1} = 1, o_k = 0, \\ 0, & \text{otherwise.} \end{cases}$$

(4) The one-stage cost function $c(\cdot,\cdot)$: the one-stage cost $c(s_k, u_k)$ is the immediate cost received when choosing action u_k at state s_k, which is in accordance with the one-stage cost l defined in (9),

$$c(s_k, u_k) \triangleq l(\tau_k, u_k) = \mathrm{Tr}h^{\tau_k}(\overline{P}) + \lambda\mathcal{E}u_k. \tag{11}$$

For an admissible policy π, the expected average cost is

$$\mathcal{C}(\pi) \triangleq \mathbb{E}^{\pi}\Big[\limsup_{T\to\infty} \frac{1}{T}\sum_{k=0}^{T} c(s_k, u_k)\Big]. \tag{12}$$

We formulate Problem 1 as an MDP as follows.

Problem 2.

$$\min_{\pi} \mathcal{C}(\pi).$$

We note that as the above MDP is a reformulation of the original Problem 1, under the same policy π, the expected average cost $\mathcal{C}(\pi)$ is equal to the expected average cost $L(\pi)$ in Problem 1. We denote by π^* the optimal solution of Problem 2. Recall that the stability of the remote estimator is equivalent to the feasibility of Problem 2. The following theorem shows that the optimal policy π^* satisfies the Bellman optimality equality if Problem 2 is feasible.

Theorem 1. *If Problem 2 is feasible, then for the optimal policy π^*, there exist a constant $\mu^* \geq 0$, a function $q : \mathbb{S} \to \mathbb{R}$ and a stationary and deterministic policy $\pi^* \in \Pi$ such that (μ^*, q, π^*) satisfies the average reward optimality equation (ACOE); that is,*

$$\mu^* + q(s) = \min_{u\in\mathbb{U}}[c(s,a) + \sum_{s'\in\mathbb{S}} q(s')\mathbb{P}(s'|s,u)] = c(s,\pi^*(s)) + \sum_{s'\in\mathbb{S}} q(s')\mathbb{P}(s'|s,\pi^*(s)), \tag{13}$$

for all $s \in \mathbb{S}$. Moreover, π^* is the optimal policy and μ^* is the corresponding optimal averaged cost.

3.2. Structural properties of optimal policies

Before deriving the condition ensuring the stability of the remote estimator, we need some preliminary results about the structure of the optimal policy π^*. In this section, we assume that the optimal policy π^* exists and $\mathcal{C}(\pi^*) < \infty$.

Theorem 2. *The optimal policy π^* has a threshold structure, i.e., there exist two nonnegative integers $\tau_g, \tau_b \in \mathbb{N} \cup \{\infty\}$, such that for any state $s_k = (\tau_k, o_k) \in \mathbb{S}$,*

$$\pi^*(s_k) = \begin{cases} 1, & \text{if } \tau \geq \tau_g, o_k = 1, \\ 1, & \text{if } \tau \geq \tau_b, o_k = 0, \\ 0, & \text{otherwise,} \end{cases} \tag{14}$$

where τ_g and τ_b denote the thresholds of data transmission for the link state $o_k = 1$ and $o_k = 0$, respectively. Furthermore, when other parameters are kept fixed.

The above theorem shows that the optimal policy π^* processes a threshold structure. Intuitively, for a fixed link state o, larger the holding time τ is, more reward the system can gain through the data transmission.

3.3. Stability of remote estimator

Based on the structural properties of the optimal policy π^*, we derive the sufficient and necessary condition of the mean square stability of the remote estimator under π^*.

Theorem 3. *The remote estimator is mean square stable under the optimal policy π^* iff*

$$\rho(A)^2 \cdot \min\left[\rho(B_1), \rho(B_2)\right] < 1, \tag{15}$$

where

$$B_1 \triangleq \begin{bmatrix} a_1(1-p_1) & 1-b_0 \\ (1-a_1)(1-p_1) & b_0, \end{bmatrix} \tag{16}$$

$$B_2 \triangleq \begin{bmatrix} a_1(1-p_1) & (1-b_1)(1-p_0) \\ (1-a_1)(1-p_1) & b_1(1-p_0) \end{bmatrix}. \tag{17}$$

Remark 3.1. We are able to recover Theorem 8 in Huang and Dey (2007) for scalar systems by directly setting $a_0 = a_1, b_0 = b_1$. The results can be generalized to the network with multiple link states. Similarly, the set of $\{B_j\}$ matrices can be constructed by the permutation, and the stability condition can be checked by computing the spectral radius of each element.

Corollary 1. *If the optimal policy π^* exists ($\mathcal{C}(\pi^*) < \infty$) and the inequality $\rho(A)^2\rho(B_2) \geq 1$ holds, then the threshold parameter τ_b in (14) is ∞.*

The above result can be interpreted that in the case when the congestion plays the crucial role $\left(\rho(A)^2\rho(B_2) \geq 1\right)$, whenever there exists a congestion (link state is $o = 0$), the sensor ceases to transmit the packet ($u = 0$) and transmission restarts until the congestion vanishes (link state is $o = 1$).

4. Simulation

To illustrate our results, we provide several numerical examples in this section. We first consider a two-dimensional system with the following parameters,

$$A = \begin{bmatrix} 1.2 & 1 \\ 0 & 1, \end{bmatrix} C = [1, 1],\ \Sigma_w = \begin{bmatrix} 1 & 0 \\ 0 & 0.5, \end{bmatrix} \Sigma_v = 1,$$

$$a_0 = 0.8, b_0 = 0.1, p_0 = 0.1, \tag{18}$$

$$a_1 = 0.6, b_1 = 0.8, p_1 = 0.8.$$

Theorem 3 shows that the value of the energy consumption $\mathcal{E}$ is irrelevant to the stability of estimator under the optimal policy π^*. The parameters in (15) can be computed as

$$\rho(A)^2\rho(B_1) = 0.5451 < 1,$$

$$\rho(A)^2\rho(B_2) = 1.0701 > 1.$$

By Theorem 3 the remote estimator is mean square stable under π^*.

We use the MDP toolbox (Chadès et al., 2009) to obtain the optimal policy π^* with different energy consumption $\lambda\mathcal{E} \geq 0$, and the results are shown in Table 1. We observe that the threshold τ_g is monotonically non-decreasing w.r.t. the energy consumption $\lambda\mathcal{E}$, and another threshold τ_b remains to be ∞ regardless of the energy consumption $\lambda\mathcal{E}$. The results are in accordance with Corollary 1.

Table 1
Numerical experiments with various energy consumption $\lambda\mathcal{E} \geq 0$ for $p_0 = 0.1$.

$\lambda\mathcal{E}$	Value		
	$p_0 = 0.1$		
	τ_g	τ_b	$C(\pi^*)$
0	0	∞	5.0314
10	1	∞	11.0828
50	2	∞	27.2118
100	3	∞	42.9965
200	4	∞	70.3328
600	5	∞	155.2087

Table 2
Numerical experiments with various energy consumption $\lambda\mathcal{E} \geq 0$ for $p_0 = 0.5$ and $p_0 = 0.8$.

$\lambda\mathcal{E}$	Value					
	$p_0 = 0.5$			$p_0 = 0.8$		
	τ_g	τ_b	$C(\pi^*)$	τ_g	τ_b	$C(\pi^*)$
0	0	0	4.7264	0	0	2.4025
10	1	3	10.8257	1	1	9.0335
50	2	5	26.9770	3	3	24.5553
100	3	6	42.6988	3	3	39.2614
200	4	7	69.9775	4	4	64.6199
600	5	9	154.9411	5	5	148.7552

We continue to consider the system with different value of p_0 while other parameters remain fixed. For the case when $p_0 = 0.5$, the parameters in (15) can be computed as

$$\rho(A)^2\rho(B_1) = 0.5451 < 1,$$

$$\rho(A)^2\rho(B_2) = 0.6136 < 1.$$

For the case when $p_0 = 0.8$, the parameters in (15) can be computed as

$$\rho(A)^2\rho(B_1) = 0.5451 < 1,$$

$$\rho(A)^2\rho(B_2) = 0.2880 < 1.$$

Similarly, we compute the optimal policy π^* with different values of the energy consumption $\lambda\mathcal{E}$, and the results are shown in Table 2. We observe that for the case when p_0 is either 0.5 or 0.8, the threshold τ_b is always finite, and is monotonically increasing with respect to the energy consumption $\lambda\mathcal{E}$. The optimal policy π^* is also threshold-type, and what is different from the first case is that the sensor will transmit the data packet regardless the link state o_k if the holding time τ_k is large enough.

To illustrate the performance of the remote estimator, we then consider a stable system with the following parameters,

$$A = \begin{bmatrix} 0.8 & 1 \\ 0 & 0.8, \end{bmatrix} \lambda\mathcal{E} = 10,$$

and the other parameters are the same with that in (18). One realization of the first element of the local estimate $\hat{x}_k^s$ and that of the remote estimate $\hat{x}_k^e$ are shown in Fig. 3, and we observe that both local and remote estimators provide good estimates of the physical plant.

5. Conclusion

In this paper, we study a problem of packet scheduling for remote state estimation with Markovian packet losses. We identify the problem as an MDP, and show that the optimal policy processes a structural property. Based on the structural property, we derive the necessary and sufficient condition of the stability of the remote estimator. Consequently the stability of the estimator can be efficiently checked.

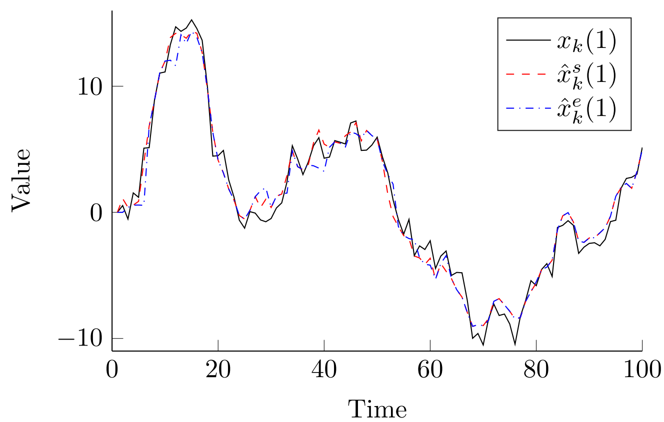


Fig. 3. The first element of true state x_k, local estimate $\hat{x}_k^s$, and remote estimate $\hat{x}_k^e$.

Acknowledgments

The work by J. Wang and L. Shi is supported by a Hong Kong RGC General Research Fund 16208517. The work of X. Ren is supported in part by National Key R&D Program of China (No. 2018AAA0102800, No. 2018AAA0102804), in part by the Natural Science Foundation of Shanghai Municipality, China, under Grant 20ZR1420200, and in part by Shanghai Key Laboratory of Power Station Automation Technology, China. The work of S. Dey was supported by the Swedish Research Council (VR) grant Dnr 621-2013-5395.

Appendix A. Proof of Theorem 1

The theorem is proved separately for stable and unstable systems. For the stable systems ($\rho(A) < 1$), we use Theorem 8.4.5 in Puterman (2014) to prove Theorem 1. We start by showing some preliminary properties of $h(\cdot)$.

Lemma 1 (*Lemma A.2. Shi et al., 2010*)**.** *The function $h(\cdot)$ is monotonically nondecreasing on $\mathbb{N}$, i.e.,*

$$h^{\tau}(\overline{P}) \preceq h^{\tau'}(\overline{P}), \forall \tau \leq \tau'.$$

Lemma 2 (*Lemma 1 Wang et al., 2020*)**.** *For stable system ($\rho(A) < 1$), the following limit of function composition exists, i.e.,*

$$\lim_{j\to\infty} h^j(\overline{P}) = Y,$$

where Y is the solution to the Lyapunov function $AYA^\intercal + \Sigma_w = Y$.

We continue to show that for any stable system, the model can be approximated by an MDP with a finite state space. By Lemma 2, the error covariance matrix P_k^e converges to its upper bound Y as the holding time τ_k goes to infinity.

For any $\epsilon > 0$, the sets $T_{\epsilon,0}$ and $T_{\epsilon,1}$ are defined as

$$T_{\epsilon,0} \triangleq \{s_k = (\tau, o_k) | \mathrm{Tr}Y - \mathrm{Tr}h^j(\overline{P}) \leq \epsilon, o_k = 0\},$$
$$T_{\epsilon,1} \triangleq \{s_k = (\tau, o_k) | \mathrm{Tr}Y - \mathrm{Tr}h^j(\overline{P}) \leq \epsilon, o_k = 1\}.$$

We truncate the countable state space $\mathbb{S}$ to be finite by augmenting $T_{\epsilon,0}$ and $T_{\epsilon,1}$ as two single states with the following cost function,

$$c(T_{\epsilon,1}, u) = c(T_{\epsilon,0}, u) = \mathrm{Tr}Y + \lambda \mathcal{E} u.$$

It is easy to see that the difference between the optimal value of the original MDP problem and that of the truncated one is bounded by ϵ. Then Problem 2 can be approximated to an arbitrary degree of closeness by the truncated version by setting the value of ϵ. By Theorem 8.4.5 in Puterman (2014), for the truncated MDP there exists a stationary average optimal policy π^* satisfies ACOE (13). It follows from Theorem 16.3 in Altman (1999) that the optimal policy π, which satisfies ACOE, is a limit of optimal stationary policies for the truncated MDP (as ϵ tends to 0).

For unstable systems, we prove the theorem by verifying the conditions (C1)(C2) of Theorem 2 in Cavazos-Cadena and Sennott (1992). The condition (C1) trivially holds since we assume the existence of the policy π' that $\mathcal{C}(\pi') < \infty$. For any unstable system, the function $h(\cdot)$ is unbounded and nondecreasing on $\mathbb{N}$. Thus for any $M > 0$, the set $\{s | c(s, a) < M\}$ is finite, which implies the condition (C2). By Theorem 2 in Cavazos-Cadena and Sennott (1992) we can show the existence of the stationary average optimal policy π^*. The proof is thus completed.

Appendix B. Proof of Theorem 2

To prove the theorem, it suffices to show that if there exists a state $s = (\tau, o)$ such that $\pi^*(s) = 1$, then for any $s' = (\tau', o) \in \mathbb{S}$, $\tau' \geq \tau$, we have

$$\pi^*(s') = 1. \tag{B.1}$$

Before proving (B.1),

Definition 1. A measurable function $v : \mathbb{S} \to \mathbb{R}$ is monotonically nondecreasing on $\mathbb{S}$ if for any $\tau \leq \tau'$,

$$v(\tau, 0) \leq v(\tau', 0),$$
$$v(\tau, 1) \leq v(\tau', 1).$$

Definition 2. A measurable function $v : \mathbb{S} \to \mathbb{R}$ is submodular on $\mathbb{S}$ if the function $f(\tau) = v(\tau, 1) - v(\tau, 0)$ is monotonically nondecreasing on $\mathbb{N}$.

Lemma 3. *If the optimal policy π^* exists, then the value function $q(\cdot)$ is monotonic and submodular on $\mathbb{S}$.*

Proof. We start to prove that the value function can be obtained by the value iteration if π^* exists. We use the vanishing discount approach in Hernández-Lerma and Lasserre (2012) to analyze the value function $q(\cdot)$. For any policy π and the initial state s, the discounted average reward with the discounted factor $0 < \alpha < 1$ is defined as

$$C_\alpha(\pi, s) \triangleq (1-\alpha) E_s^\theta \left[\sum_{k=1}^{\infty} \alpha^k c(s_k, u_k) \right]. \tag{B.2}$$

The optimal discounted cost $C_\alpha^*(s)$ and the discounted relative value function $Q_\alpha(s)$ are defined as follows,

$$C_\alpha^*(s) \triangleq \min_\pi C_\alpha(\theta, s), \tag{B.3}$$

$$Q_\alpha(s) \triangleq C_\alpha^*(s) - C_\alpha^*(s^0), \tag{B.4}$$

where $s^0 = (0, 0) \in \mathbb{S}$. The relative value function $q(\cdot)$ is a limit of a sequence of the discounted value function $Q_\alpha(s)$ with a discount factor α approaching 1 (Hernández-Lerma & Lasserre, 2012), i.e.,

$$q(s) = \lim_{\alpha\to 1} Q_\alpha(s), \forall s \in \mathbb{S}. \tag{B.5}$$

The value iteration can be viewed as a dynamic programming (DP) operator operating on the function space $\mathcal{B}(\mathbb{S})$. We then introduce the DP operator $\mathcal{T}_\alpha$ related to the discounted cost model.

Definition 3. Given a measurable function $v : \mathbb{S} \to \mathbb{R}$ and a discounted factor $\alpha \in (0, 1)$, the operator $\mathcal{T}_\alpha$ is defined by

$$\mathcal{T}_\alpha v(s) \triangleq \min_{u\in\mathbb{A}}[c(s, u) + \alpha \sum_{\bar{s}\in\mathbb{S}} v(\bar{s})\mathbb{P}(\bar{s}|s, u)], s \in \mathbb{S}. \tag{B.6}$$

It is known that the value function q can be obtained by the value iteration algorithm (Bertsekas, 1995).

To prove that q is monotonic and submodular on $\mathbb{S}$, it suffices to show that for any $v \in B(\mathbb{S})$, the monotonicity and submodularity of $v : \mathbb{S} \to \mathbb{R}$ are preserved by $\mathcal{T}_\alpha$. We start to show the monotonicity. Suppose that v is monotonic on $\mathbb{S}$, then the following inequalities hold for any $\tau \le \tau', u \in \mathbb{U}$,

$$c((\tau, 0), u) + \alpha \sum_{\bar{s}\in\mathbb{S}} v(\bar{s})\mathbb{P}(\bar{s}|(\tau, 0), u)$$
$$\le c((\tau', 0), u) + \alpha \sum_{\bar{s}\in\mathbb{S}} v(\bar{s})\mathbb{P}(\bar{s}|(\tau', 0), u),$$
$$c((\tau, 1), u) + \alpha \sum_{\bar{s}\in\mathbb{S}} v(\bar{s})\mathbb{P}(\bar{s}|(\tau, 1), u)$$
$$\le c((\tau', 1), u) + \alpha \sum_{\bar{s}\in\mathbb{S}} v(\bar{s})\mathbb{P}(\bar{s}|(\tau', 1), u),$$

which imply that the monotonicity of $v(\cdot)$ propagates through $\mathcal{T}_\alpha$. Similarly, to prove that q is submodular, it suffices to show that for any $\tau \in \mathbb{N}$ the following inequality holds,

$$q(\tau, 1) - q(\tau, 0) \le q(\tau + 1, 1) - q(\tau + 1, 0). \tag{B.7}$$

We will continue to use mathematical induction to prove (B.7). To be more specific, suppose that for any $\tau \in \mathbb{N}$, $v(\tau, 1) + v(\tau + 1, 0) \le v(\tau+1, 1)+v(\tau, 0)$, we intend to show that for any $\tau \in \mathbb{N}$,

$$\mathcal{T}_\alpha v(\tau, 1) + \mathcal{T}_\alpha v(\tau + 1, 0) \le \mathcal{T}_\alpha v(\tau + 1, 1) + \mathcal{T}_\alpha v(\tau, 0). \tag{B.8}$$

Notice that the form of the term $\mathcal{T}_\alpha v(\tau, 1)$ depends on the operator $\mathcal{T}_\alpha$ defined in (B.6), and we use the case analysis to prove (B.7). Recall the assumption that $0 \le a_1 \le a_0 \le 1, 0 \le b_0 \le b_1 \le 1$. Suppose that the following equalities hold,

$$\mathcal{T}_\alpha v(\tau, 1) = c((\tau, 0), 1) + \alpha \sum_{\bar{s}\in\mathbb{S}} v(\bar{s})\mathbb{P}(\bar{s}|(\tau, 0), 1), \tag{B.9}$$

$$\mathcal{T}_\alpha v(\tau, 0) = c((\tau, 1), 1) + \alpha \sum_{\bar{s}\in\mathbb{S}} v(\bar{s})\mathbb{P}(\bar{s}|(\tau, 1), 1), \tag{B.10}$$

and then (B.8) follows from the basic algebra. For other cases, (B.8) can be obtained in a similar way, and the details are omitted here.

For the monotonicity and the submodularity propagate through the DP operator $\mathcal{T}_\alpha$ for any $\alpha \in (0, 1)$, one can initialize $Q_\alpha(\cdot)$ with $Q_\alpha(s) = 0, \forall s \in \mathbb{S}$, and thus by the mathematical induction, one obtains that $Q_\alpha(\cdot)$ is monotonic and submodular on $\mathbb{S}$ for any $\alpha \in (0, 1)$. Recall that $q(\cdot)$ is the limit of the sequence of $Q_\alpha(\cdot)$ as α goes to 1(B.5), which implies that $q(\cdot)$ is monotonic and submodular on $\mathbb{S}$.

We are now ready to prove the theorem. Suppose that there exists a certain state $s = (\tau, o)$ satisfying $\pi^*(s) = 1$, and then the ACOE equation (13) implies that

$$c(s, 1) + \sum_{\bar{s}\in\mathbb{S}} q(\bar{s})\mathbb{P}(\bar{s}|s, 1) \le c(s, 0) + \sum_{\bar{s}\in\mathbb{S}} q(\bar{s})\mathbb{P}(\bar{s}|s, 0).$$

As a consequence of the above lemma, for any $s' = (\tau', o), \tau' \ge \tau$, we have

$$c(s', 1) + \sum_{\bar{s}\in\mathbb{S}} q(\bar{s})\mathbb{P}(\bar{s}|s', 1) \le c(s', 0) + \sum_{\bar{s}\in\mathbb{S}} q(\bar{s})\mathbb{P}(\bar{s}|s', 0),$$

which implies that for any $s' = (\tau', o), \tau' \ge \tau$, the equality $\pi^*(s') = 1$ holds. Thus the proof is completed.

Appendix C. Proof of Theorem 3

In Theorem 2, We have shown that the optimal policy π^* has a threshold structure which can be characterized by two parameters $\tau_g, \tau_b \in \mathbb{N} \cup \{\infty\}$. The structure of the optimal policy π^* depends on the finiteness of the parameters τ_g, τ_b, which can be divided into the following four cases:

1) $\tau_g < \infty, \tau_b = \infty$.2) $\tau_g < \infty, \tau_b < \infty$.
3) $\tau_g = \infty, \tau_b < \infty$.4) $\tau_g = \infty, \tau_b = \infty$.

For any state $s_k = (\tau_k, o_k) \in \mathbb{S}$, we define the following four policies $\pi_1, \pi_2, \pi_3, \pi_4$ as

$$\pi_1(s_k) = \begin{cases} 0, & \text{if } o_k = 0, \\ 1, & \text{otherwise}, \end{cases} \tag{C.1}$$

$$\pi_2(s_k) = 1, \tag{C.2}$$

$$\pi_3(s_k) = \begin{cases} 0, & \text{if } o_k = 1, \\ 1, & \text{otherwise}, \end{cases} \tag{C.3}$$

$$\pi_4(s_k) = 0, \tag{C.4}$$

which corresponds to the above four cases, respectively. The following lemma builds relation between the stability of the remote estimator under π^* and the stability of the estimator under π_1, π_2, π_3 or π_4.

Lemma 4. *The remote estimator is stable under the optimal policy π^* iff the remote estimator under one of the policies $\{\pi_1, \pi_2, \pi_3, \pi_4\}$ is stable, i.e., $\mathcal{C}(\pi^*) < \infty$ iff $\min\{\mathcal{C}(\pi_1), \mathcal{C}(\pi_2), \mathcal{C}(\pi_3), \mathcal{C}(\pi_4)\} < \infty$.*

Proof. As a consequence of Theorem 2, for sufficiently large τ, the structure of π^* is the same with one particular policy of $\pi_1, \pi_2, \pi_3, \pi_4$. We note that the stability of the estimator is not affected by the actions taken on the finite subset of the state space $\mathbb{S}$, thus the stability of the estimator under π^* is directly related to the threshold structure in Theorem 2.

We continue to recall the definition of the matrices B_1, B_2 and then define the matrices B_3, B_4 as follows,

$$B_1 \triangleq \begin{bmatrix} a_1(1-p_1) & 1-b_0 \\ (1-a_1)(1-p_1) & b_0, \end{bmatrix}$$
$$B_2 \triangleq \begin{bmatrix} a_1(1-p_1) & (1-b_1)(1-p_0) \\ (1-a_1)(1-p_1) & b_1(1-p_0) \end{bmatrix},$$
$$B_3 \triangleq \begin{bmatrix} a_0 & (1-b_1)(1-p_0) \\ 1-a_0 & b_1(1-p_0) \end{bmatrix},$$
$$B_4 \triangleq \begin{bmatrix} a_0 & 1-b_0 \\ 1-a_0 & b_0 \end{bmatrix}.$$

We remark that the matrices B_1, B_2, B_3, B_4 correspond to the above four policies $\tau_1, \tau_2, \tau_3, \tau_4$, respectively. The following lemma shows the relation between the stability of π_1 and the spectral radius of the matrix B_1.

Lemma 5. *The estimator under π_1 is stable iff $\rho(A)^2 \cdot \rho(B_1) < 1$.*

Proof. The indicator function of the link state $o \in \{0, 1\}$ is defined as

$$\mathbb{1}_{\{o\}}(w) \triangleq \begin{cases} 1, & \text{if } w = o, \\ 0, & \text{otherwise}. \end{cases}$$

We note that under the policy π_1, the action of the sensor depends solely on the link state. For the link state $o \in \{1, 0\}$, let $P_{k,o} \triangleq \mathbb{E}\big[\mathrm{Var}(x_k - \hat{x}_k^e)\mathbb{1}_{\{o\}}(o_k)\big]$ be the expected error variance of $\hat{x}_k^e$ when $o_k = o$. Then one obtains that for $k \geq 1, o \in \{1, 0\}$, under the policy π_1, there holds

$$P_{k+1,o} = \sum_{j\in\{1,0\}} (1-p_j)\Pr(o_{k+1} = o|o_k = j)AP_{k,j}A + \Sigma_w.$$

Vectorizing the above equation for each m yields

$$\begin{bmatrix} P_{k+1,1} \\ P_{k+1,0} \end{bmatrix} = B_1^\intercal \oplus (A \oplus A) \begin{bmatrix} P_{k,1} \\ P_{k,0} \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \end{bmatrix}^\intercal \oplus \Sigma_w,$$

where $\oplus$ denotes the Kronecker product. It is easy to see that the estimator under π_1 (π_2) is stable iff

$$\rho(B_1^\intercal \oplus (A \oplus A)) = \rho(B_1) \cdot \rho^2(A) < 1.$$

Thus the proof is finished.

Similarly, we can show that the estimator under π_i ($i \in \{1, 2, 3, 4\}$) is stable iff $\rho(A)^2 \cdot \rho(B_i) < 1$. As a consequence of the above lemma, we obtain that $\mathcal{C}(\pi^*) < \infty$ iff one of $\pi_1, \pi_2, \pi_3, \pi_4$ is stable, i.e.,

$$\rho(A)^2 \cdot \min\big[\rho(B_1), \rho(B_2), \rho(B_3), \rho(B_4)\big] < 1. \tag{C.5}$$

The spectral radius of any positive matrix is upper bounded by its maximal column sum, and is lower bounded by its minimal column sum. Hence, we have

$$\begin{aligned} 1 - p_1 &\leq \rho(B_1) \leq 1, \\ 1 - p_1 &\leq \rho(B_2) \leq 1 - p_0, \\ 1 - p_0 &\leq \rho(B_3) \leq 1, \\ &\rho(B_4) = 1, \end{aligned}$$

which implies $\rho(B_1) \leq \rho(B_4)$ and $\rho(B_2) \leq \rho(B_3)$. Then it follows that the inequality (C.5) can be rewritten as (15). The proof is thus completed.

Appendix D. Proof of Corollary 1

We prove the result by the contradiction. Suppose that the threshold parameter τ_b in (14) is finite. Because the stability of the estimator is not affected by the actions taken on the finite subset of the state space $\mathbb{S}$, by Lemma 4 the estimator should be stable under one of the polices $\{\pi_2, \pi_3, \pi_4\}$. By Theorem 3, we have

$$\rho(A)^2\rho(B_2) < 1,$$

which contradicts the inequality $\rho(A)^2\rho(B_2) \geq 1$. Thus $\tau_b = \infty$ and the proof is completed.

References

Altman, Eitan (1999). *Constrained Markov decision processes (vol. 7)*. CRC Press.

Avrachenkov, Konstantin, Ayesta, Urtzi, Doncel, Josu, & Jacko, Peter (2013). Congestion control of TCP flows in internet routers by means of index policy. *Computer Networks*, *57*(17), 3463–3478.

Battistelli, Giorgio, Benavoli, Alessio, & Chisci, Luigi (2012). Data-driven communication for state estimation with sensor networks. *Automatica*, *48*(5), 926–935.

Bertsekas, Dimitri P. (1995). *Dynamic programming and optimal control (vol. 1) (no. 2)*. Athena Scientific Belmont, MA.

Cao, Zhixing, Lu, Jingyi, Zhang, Ridong, & Gao, Furong (2016). Iterative learning Kalman filter for repetitive processes. *Journal of Process Control*, *46*, 92–104.

Cavazos-Cadena, Rolando, & Sennott, Linn I. (1992). Comparing recent assumptions for the existence of average optimal stationary policies. *Operations Research Letters*, *11*(1), 33–37.

Chadès, Iadine, Cros, Marie-Josée, Garcia, Frédéric, & Sabbadin, Régis (2009). *Markov decision processes (MDP) toolbox (vol. 42)*. URL http://www.inra.fr/mia/T/MDPtoolbox.

Fletcher, Alyson K., Rangan, Sundeep, & Goyal, Vivek K. (2004). Estimation from lossy sensor data: Jump linear modeling and Kalman filtering. In *Information processing in sensor networks, 2004. third international symposium on* (pp. 251–258). IEEE.

Floyd, Sally, & Jacobson, Van (1993). Random early detection gateways for congestion avoidance. *IEEE/ACM Transactions on Networking*, (4), 397–413.

Füßler, Holger, Widmer, Jörg, Käsemann, Michael, Mauve, Martin, & Hartenstein, Hannes (2003). Contention-based forwarding for mobile ad hoc networks. *Ad Hoc Networks*, *1*(4), 351–369.

Gunes, Volkan, Peter, Steffen, Givargis, Tony, & Vahid, Frank (2014). A survey on concepts, applications, and challenges in cyber-physical systems. *KSII Transactions on Internet and Information Systems*, *8*(12).

Hernández-Lerma, Onésimo, & Lasserre, Jean B. (2012). *Discrete-time Markov control processes: Basic optimality criteria (vol. 30)*. Springer Science & Business Media.

Huang, Jyh-How, Chen, Ying-Yu, Huang, Yu-Te, Lin, Po-Yen, Chen, Yi-Chao, Lin, Yi-Fu, Yen, Shih-Ching, Huang, Polly, & Chen, Ling-Jyh (2010). Rapid prototyping for wildlife and ecological monitoring. *IEEE Systems Journal*, *4*(2), 198–209.

Huang, Minyi, & Dey, Subhrakanti (2007). Stability of Kalman filtering with Markovian packet losses. *Automatica*, *43*(4), 598–607.

Kong, Zhen-ning, Tsang, Danny HK, Bensaou, Brahim, & Gao, Deyun (2004). Performance analysis of IEEE 802.11 e contention-based channel access. *IEEE Journal on Selected Areas in Communications*, *22*(10), 2095–2106.

Lee, Edward A. (2008). Cyber physical systems: Design challenges. In *Object Oriented real-time distributed computing, 2008 11th IEEE international symposium on* (pp. 363–369). IEEE.

Leong, Alex S., Dey, Subhrakanti, & Quevedo, Daniel E. (2017). Sensor scheduling in variance based event triggered estimation with packet drops. *IEEE Transactions on Automatic Control*, *62*(4), 1880–1895.

Lipsa, Gabriel M., & Martins, Nuno C. (2011). Remote state estimation with communication costs for first-order LTI systems. *IEEE Transactions on Automatic Control*, *56*(9), 2013–2025.

Papadimitriou, Christos H., & Tsitsiklis, John N. (1987). The complexity of Markov decision processes. *Mathematics of Operations Research*, *12*(3), 441–450.

Puterman, Martin L. (2014). *Markov decision processes: discrete stochastic dynamic programming*. John Wiley & Sons.

Rajkumar, Ragunathan Raj, Lee, Insup, Sha, Lui, & Stankovic, John (2010). Cyber-physical systems: the next computing revolution. In *Proceedings of the 47th design automation conference* (pp. 731–736). ACM.

Ren, Xiaoqiang, Wu, Junfeng, Johansson, Karl Henrik, Shi, Guodong, & Shi, Ling (2018). Infinite horizon optimal transmission power control for remote state estimation over fading channels. *IEEE Transactions on Automatic Control*, *63*(1), 85–100.

Shi, Ling, Cheng, Peng, & Chen, Jiming (2011). Sensor data scheduling for optimal state estimation with communication energy constraint. *Automatica*, *47*(8), 1693–1698.

Shi, Ling, Epstein, Michael, & Murray, Richard M. (2010). Kalman Filtering over a packet-dropping network: A probabilistic perspective. *IEEE Transactions on Automatic Control*, *55*(3), 594–604.

Sinopoli, Bruno, Schenato, Luca, Franceschetti, Massimo, Poolla, Kameshwar, Jordan, Michael I, & Sastry, Shankar S (2004). Kalman Filtering with intermittent observations. *IEEE Transactions on Automatic Control*, *49*(9), 1453–1464.

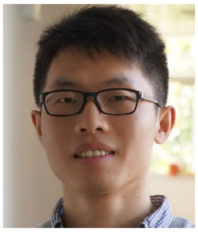
Trimpe, Sebastian, & D'Andrea, Raffaello (2014). Event-based state estimation with variance-based triggering. *IEEE Transactions on Automatic Control*, *59*(12), 3266–3281.

Wang, Jiazheng, Ren, Xiaoqiang, Mo, Yilin, & Shi, Ling (2020). Whittle index policy for dynamic multi-channel allocation in remote state estimation. *IEEE Transactions on Automatic Control*, *65*(2), 591–603.

Wu, Junfeng, Jia, Qing-Shan, Johansson, Karl Henrik, & Shi, Ling (2013). Event-based sensor data scheduling: Trade-off between communication rate and estimation quality. *IEEE Transactions on Automatic Control*, *58*(4), 1041–1046.

Xu, Liang, Xie, Lihua, & Xiao, Nan (2018). Mean square stabilization over Gaussian finite-state Markov channels. *IEEE Transactions on Control of Network Systems*, *5*(4), 1830–1840.

Jiazheng Wang received the B.Eng. degree from the Department of Automatic Control, Zhejiang University, Hangzhou, China, in 2015. He is currently a Ph.D. candidate in the Department of Electronic & Computer Engineering, Hong Kong University of Science and Technology. His research interests include resource allocation in wireless systems, networked estimation, and cyberphysical system security.

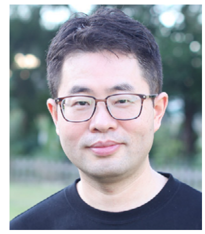


Xiaoqiang Ren is a professor at the School of Mechatronic Engineering and Automation, Shanghai University, China. He received the B.E. degree in Automation from Zhejiang University, Hangzhou, China, in 2012 and the Ph.D. degree in control and dynamic systems from Hong Kong University of Science and Technology in 2016. Prior to his current position, he was a postdoctoral researcher in the Hong Kong University of Science and Technology in 2016, Nanyang Technological University from 2016 to 2018, and KTH Royal Institute of Technology from 2018 to 2019. His research interests include security of cyber-physical systems, sequential decision, and networked estimation and control.

Subhrakanti Dey received the Bachelor in Technology and Master in Technology degrees from the Department of Electronics and Electrical Communication Engineering, Indian Institute of Technology, Kharagpur, in 1991 and 1993, respectively, and the Ph.D. degree from the Department of Systems Engineering, Research School of Information Sciences and Engineering, Australian National University, Canberra, in 1996.

He is currently a Professor with the Hamilton Institute, National University of Ireland, Maynooth, Ireland. Prior to this, he was a Professor with the Dept. of Engineering Sciences in Uppsala University, Sweden (2013–2017), Professor with the Department of Electrical and Electronic Engineering, University of Melbourne, Parkville, Australia, from 2000 until early 2013, and a Professor of Telecommunications at University of South Australia during 2017–2018. From September 1995 to September 1997, and September 1998 to February 2000, he was a Postdoctoral Research Fellow with the Department of Systems Engineering, Australian National University. From September 1997 to September 1998, he was a Postdoctoral Research Associate with the Institute for Systems Research, University of Maryland, College Park. His current research interests include wireless communications and networks, signal processing for sensor networks, networked control systems, and molecular communication systems.

Professor Dey currently serves on the Editorial Board of IEEE Control Systems Letters, IEEE Transactions on Control of Network Systems, and IEEE Transactions on Wireless Communications. He was also an Associate Editor for IEEE Transactions on Signal Processing, (2007–2010, 2014–2018), IEEE Transactions on Automatic Control, (2004–2007) and Elsevier Systems and Control Letters (2003–2013).



Ling Shi received the B.S. degree in electrical and electronic engineering from Hong Kong University of Science and Technology, Kowloon, Hong Kong, in 2002 and the Ph.D. degree in Control and Dynamical Systems from California Institute of Technology, Pasadena, CA, USA, in 2008. He is currently a Professor in the Department of Electronic and Computer Engineering, and the associate director of 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 is a senior member of IEEE. He served as an editorial board member for the European Control Conference 2013–2016. He was a subject editor for International Journal of Robust and Nonlinear Control (2015–2017). He has been serving as an associate editor for IEEE Transactions on Control of Network Systems from July 2016, and an associate editor for IEEE Control Systems Letters from Feb 2017. He also served as 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 served as the General Chair of the 23rd International Symposium on Mathematical Theory of Networks and Systems (MTNS 2018). He is a member of the Young Scientists Class 2020 of the World Economic Forum (WEF).