Automatica 127 (2021) 109498

Contents lists available at ScienceDirect

Automatica

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

Multi-hop sensor network scheduling for optimal remote estimation ☆

Takuya Iwaki [a], Junfeng Wu [b,*], Yuchi Wu [c], Henrik Sandberg [a], Karl Henrik Johansson [a]

[a] School of Electrical Engineering and Computer Science, KTH Royal Institute of Technology, 100 44, Stockholm, Sweden
[b] College of Control Science and Engineering, Zhejiang University, 310027, Hangzhou, PR China
[c] Department of Electronic and Computer Engineering, The Hong Kong University of Science and Technology, Hong Kong

A R T I C L E I N F O

Article history:
Received 4 August 2019
Received in revised form 2 October 2020
Accepted 4 January 2021
Available online 17 February 2021

Keywords:
State estimation
Medium access control
Sensor networks
Sensor scheduling
Markov decision process

A B S T R A C T

This paper studies a design problem of how a group of wireless sensors are selected and scheduled to transmit data efficiently over a multi-hop network subject to energy considerations, when the sensors are observing multiple independent discrete-time linear systems. Each time instant, a subset of sensors is selected to transmit their measurements to a remote estimator. We formulate an optimization problem, in which a network schedule is searched to minimize a linear combination of the averaged estimation error and the averaged transmission energy consumption. It is shown that the optimal network schedule forms a tree with root at the gateway node. From this observation, we manage to separate the optimization problem into two subproblems: tree planning and sensor selection. We solve the sensor selection subproblem by a Markov decision process, showing that the optimal solution admits a periodic structure when the transmission cost is sufficiently low. Efficient algorithms are proposed and they are shown to reduce the computational complexity of the original optimization problem. Numerical studies illustrate the effectiveness of the proposed algorithms, and show that they are scalable to large networks.

1. Introduction

Recent development of wireless sensor technology enables control and estimation over multi-hop wireless sensor networks, which is of significant interest for process and automation industries (Lu et al., 2016; Park, Ergen, Fischione, Lu, & Johansson, 2018; Willig, 2008). Wireless sensor networks provide advantages through enhanced and massive sensing, flexible deployment and operation, and more efficient maintenance compared to wired solutions. For instance, wireless sensors can be placed where conventional sensors with cabling cannot be placed such as on mobile robots or rotational machinery. However, since wireless sensors usually have no inexhaustible or reliable energy source, energy limitation affects system performance and lifetime. In this context, energy-aware communication protocols, real-time scheduling algorithms as well as empirical studies for optimizing the performance of wireless sensor networks have been proposed (Chipara, He, Xing, Chen, Wang, Lu, Stankovic, & Abdelzaher, 2006; Hasenfratz, Meier, Moser, Chen, & Thiele, 2010; Sha, Gunatilaka, Wu, & Lu, 2017). In addition, as the number of wireless sensors over an area increases, data packets may be lost due to interference or network congestion. This may lead to poor performance of the overall estimation and control application.

To tackle these problems, sensor scheduling approaches for remote estimation have been investigated by several research groups. In Xu and Hespanha (2005), a communication control scheme is discussed on how to trade off estimation performance and communication cost. A stochastic sensor selection algorithm is proposed in Gupta, Chung, Hassibi, and Murray (2006), where a plant is monitored by multiple sensors but only one of them can access the estimator at every time instance. Optimal estimation with a multiple time-step cost is introduced in Mo, Ambrosino, and Sinopoli (2011), where the authors consider finite time horizon and obtain a suboptimal schedule by formulating a relaxed convex problem. The infinite horizon problem is considered in Mo, Garone, and Sinopoli (2014) and Zhao, Zhang, Hu, Abate, and Tomlin (2014). The authors of Jawaid and Smith (2015) derive conditions for the cost functions to be submodular so that

☆ This work by T. Iwaki, H. Sandberg, and K.H. Johansson was supported in part by the VINNOVA PiiA project "Advancing System Integration in Process Industry," the Knut and Alice Wallenberg Foundation, Sweden, the Swedish Strategic Research Foundation, the Swedish Research Council. The work by J. Wu was supported by NSFC, China No. 62003303 from National Science Foundation of China, and Fundamental Research Funds for the Central Universities, China 2020XZZX003-02. The material in this paper was partially presented at the 56th IEEE Conference on Decision and Control, December 12-15, 2017, Melbourne, Australia. This paper was recommended for publication in revised form by Associate Editor Juan I. Yuz under the direction of Editor Torsten Söderström.

* Corresponding author.
E-mail addresses: takya@kth.se (T. Iwaki), jfwu@zju.edu.cn (J. Wu), ywubj@ust.hk (Y. Wu), hsan@kth.se (H. Sandberg), kallej@kth.se (K.H. Johansson).

https://doi.org/10.1016/j.automatica.2021.109498
0005-1098/

estimation performance can be guaranteed. Schedules designed by greedy algorithms are studied in Zhang, Ayoub, and Sundaram (2017). The minimum mean square error (MMSE) estimation schedule can be obtained in some special cases for example for two sensors in Shi and Zhang (2012), and more sensors in Han, Wu, Zhang, and Shi (2017) and Wu, Ren, Dey, and Shi (2017). While these works provide offline schedules, the authors of Wu, Jia, Johansson, and Shi (2013) offer a deterministic online MMSE schedule by using feedback from the estimator. An MMSE stochastic schedule is proposed in Han et al. (2015). The result of Han et al. (2015) is extended to multiple sensors in Weerakkody, Mo, Sinopoli, Han, and Shi (2016). In a similar setup, remote estimation with variance-based triggering is proposed in Trimpe and D'Andrea (2014), which yields a periodic transmission schedule. In this setup, sensors can directly communicate with the remote estimator through a common bus. The studies above consider the remote estimation under network constraints. Sensor energy consumption and packet dropout are explicitly considered for covariance-based state estimation in Leong, Dey and Quevedo (2017), and LQG control (Leong, Quevedo, Tanaka, Dey & Ahlén, 2017). While all research above consider single-hop networks, i.e., each sensor directly communicates with a gateway or an estimator, multi-hop networks are widely considered for industrial wireless communication, such as WirelessHART (Chen, Nixon, & Mok, 2010), ISA-100 (International Society of Automation, 2009), and Zigbee (ZigBee Alliance, 2006). A co-design framework of multi-hop network scheduling and an optimal controller for a single process are proposed in Demirel, Zou, Soldati, and Johansson (2014). Some recent work considers aspects of the industrial protocols for estimation and control problems (Di Girolamo & D'Innocenzo, 2019; Maass, Nešić, Postoyan, & Dower, 2019a, 2019b). These studies assume that the network nodes are always time-synchronized, since the existing industrial communication protocols have a strict mechanism for such a time-synchronization. In Li, Phillips, and Sanfelice (2018), on the other hand, a distributed estimation problem over asynchronous communications is considered, where the time intervals between consecutive data arrival are bounded but uncertain.

The main contribution of this paper is to provide a framework of how to select and schedule a set of sensors to transmit their measurements efficiently over a time-synchronized multi-hop network. Motivated by an industrial case study at a Swedish paper plant (Agrawal, Ahlén, Olofsson, & Gidlund, 2014; Ahlén et al., 2019), our framework defines the links to be activated to transmit the sensor measurement for optimal remote estimation under sensor energy constraints, when the sensors observe independent discrete-time linear systems. It is important to investigate estimation and control of multiple processes over a shared multi-hop network, in particular, since previous work (Demirel et al., 2014) only deals with a single process. Different from Han et al. (2017), Leong, Dey and Quevedo (2017) and Wu, Ren, Dey, and Shi (2018) and related work, the measurements are not directly sent to the estimator but through some intermediate nodes and a gateway. For the medium access and communication, we consider a periodic superframe structure common to many existing wireless sensor network protocols (Araújo, Mazo, Anta, Tabuada, & Johansson, 2014). A superframe repeated every sampling interval is divided into timeslots. We assume only one point-to-point link is activated at a time. Then, by activating links in a certain order, the measurement data of selected sensors can be efficiently conveyed to the estimator. The link activation is jointly determined with the sensor selection, by considering data aggregation techniques (Heinzelman, Chandrakasan, & Balakrishnan, 2002; Rajagopalan & Varshney, 2006), and constrained by the energy consumption of the sensors. In such a set-up, we first find some structures of the multi-hop network schedule, so that the problem can be decomposed into two subproblems. Then it is shown that this multi-hop network scheduling problem can be solved using a similar approach to the single-hop network (Leong, Dey & Quevedo, 2017) by formulating a Markov Decision Process (MDP). Second, we exploit the MDP formulation to obtain a sufficient condition on the existence of a periodic optimal sensor network schedule. Our condition does not exclude stable plant, which was the case in Han et al. (2017) and our preliminary work (Iwaki, Wu, Wu, Sansberg, & Johansson, 2017). Third, we provide algorithms to realize the periodic optimal schedule. Fourth, to make our approach scalable for larger networks, we present algorithms to obtain suboptimal schedules. The performance of the optimal and suboptimal algorithms is illustrated and evaluated in numerical examples. It is shown that the suboptimal algorithms effectively generate suboptimal schedules with slight performance degradation in small networks and is scalable to large networks.

The remainder of this paper is organized as follows. Section 2 describes the system including wireless network, process, communication, and energy consumption models together with the remote estimator. The problem formulation is also presented. Section 3 presents the main result. Suboptimal schedules are obtained in Section 4. Numerical examples are provided in Section 5. Section 6 concludes the paper.

Notations

The symbols $\mathbb{N}$, $\mathbb{N}_0$, and $\mathbb{R}$ are the sets of integers larger than zero, nonnegative integers, and real numbers, respectively. The set of n by n positive semi-definite (positive definite) real matrices is denoted by $\mathbb{S}_+^n$ ($\mathbb{S}_{++}^n$). For simplicity, we write $X \geq Y$ ($X > Y$), where $X, Y \in \mathbb{S}_+^n$, if $X - Y \in \mathbb{S}_+^n$ ($X - Y \in \mathbb{S}_{++}^n$) and $X \geq 0$ ($X > 0$) if $X \in \mathbb{S}_+^n$ ($X \in \mathbb{S}_{++}^n$). For a matrix A, we use $\lambda_{\max}(A)$ to denote the eigenvalue of A with largest magnitude. For a vector x, we denote its element i as $x[i]$. A vector $\mathbf{1}_N$ denotes a row vector of all ones. The sequence of all vectors x_t, $t = 0, \ldots, k$ is represented by $x_{0:k}$.

2. Problem formulation

In this paper, we discuss an optimal remote estimation problem, where the estimator generates state estimates based on the received information from sensors. The objective is to choose the network scheduler to minimize the estimation error subject to energy considerations. We elaborate on the main components of the system in the following subsections.

2.1. Wireless sensor network

A set of sensors $\mathcal{V}_s \triangleq \{1, 2, \ldots, N\}$ is deployed in an area, monitoring N decoupled discrete-time linear time-invariant (LTI) processes. The sensors are interconnected via a wireless network and they upload measurements through the network to a remote estimator via a gateway. We denote the gateway as node 0, so the whole node set is given by $\mathcal{V} \triangleq \mathcal{V}_s \cup \{0\}$. The network is modeled by a directed graph $\mathcal{G} \triangleq (\mathcal{V}, \mathcal{E})$, where $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ is the set of communication links. The link (i, j) is included in $\mathcal{E}$ if there is a link from node i to node j. For a link $e = (i, j) \in \mathcal{E}$, we introduce the maps to the sending node $v_{\text{out}}(e) = i$ and to the receiving node $v_{\text{in}}(e) = j$. Let $\mathcal{N}_i^{\text{in}}$ and $\mathcal{N}_i^{\text{out}}$ denote the in- and out-neighbors of node i, respectively, i.e.,

$$\mathcal{N}_i^{\text{in}} \triangleq \{j \in \mathcal{V} \mid (j, i) \in \mathcal{E}\},$$
$$\mathcal{N}_i^{\text{out}} \triangleq \{j \in \mathcal{V} \mid (i, j) \in \mathcal{E}\}.$$

Furthermore, we denote $d(e)$ as the distance between nodes i and j. By arranging an order for the links $e_1, \ldots, e_\ell, \ldots, e_{|\mathcal{E}|}$, the

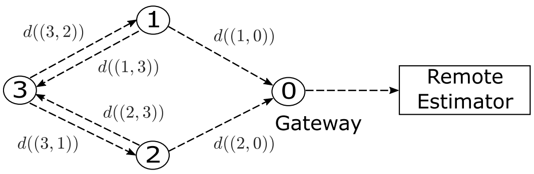


Fig. 1. In a multi-hop wireless sensor network, each sensor transmits its data to a remote estimator through intermediate sensors and a gateway.

node–arc incidence matrix of the graph $\mathcal{G}$ is defined as $G \in \{-1, 0, 1\}^{(N+1)\times|\mathcal{E}|}$, where (i, ℓ)-th element of G is 1 if $v_{\text{out}}(e_\ell) = i$, and -1 if $v_{\text{in}}(e_\ell) = i$, otherwise 0.

Fig. 1 illustrates a network $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ with $\mathcal{V} = \{0, 1, 2, 3\}$ and $\mathcal{E} = \{(1, 0), (1, 3), (2, 0), (2, 3), (3, 1), (3, 2)\}$. Assume that the links are arranged in ascending order, G is then given by

$$G = \begin{bmatrix} -1 & 0 & -1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & -1 & 0 \\ 0 & 0 & 1 & 1 & 0 & -1 \\ 0 & -1 & 0 & -1 & 1 & 1 \end{bmatrix}.$$

For sensor 1, the distances to sensor 3 and to gateway 0 are expressed by $d((1, 3)) = d((3, 1))$ and $d((1, 0))$, respectively. The in- and out-neighbors are given by $\mathcal{N}_1^{\text{in}} = \{3\}$ and $\mathcal{N}_1^{\text{out}} = \{0, 3\}$.

2.2. Process model

We consider N discrete-time LTI processes

$$x_{k+1}^{(i)} = A_i x_k^{(i)} + w_k^{(i)}, \quad i \in \mathcal{V}_s, \tag{1}$$

where $x_k^{(i)} \in \mathbb{R}^n$ is the state of process i at time k, $w_k^{(i)} \in \mathbb{R}^n$ is process noise assumed to be Gaussian process with zero-mean independent and identically distributed (i.i.d.) and covariance $W_i \triangleq \mathbb{E}[w_k^{(i)}(w_k^{(i)})^\top] > 0$. The initial state $x_0^{(i)}$, independent of $w_k^{(i)}$, $k \in \mathbb{N}_0$, is also assumed to be Gaussian with mean $\mathbb{E}[x_0^{(i)}]$ and covariance $\Sigma_0^{(i)}$. Without loss of generality, we assume $\mathbb{E}[x_0^{(i)}] = 0$, as nonzero-mean can be translated into zero-mean by the coordinate change $\tilde{x}_k^{(i)} = x_k^{(i)} - \mathbb{E}[x_0^{(i)}]$. We assume that the state $x_k^{(i)}$ can be observed directly by sensor i.

2.3. Communication model and network scheduling

The sensors communicate to the estimator through intermediate sensors and a gateway which define the underlying communication network. Time horizons of the sensors are partitioned into strips of identical time intervals (see Fig. 2). Each time interval is divided into two phases: a sensing phase and a communication phase, where the former is a time period for sensor i to acquire the process state $x_k^{(i)}$ and the latter is a time period for message delivery. The communication phase between time k and $k+1$, which we call superframe at time instance k, is divided into L timeslots. Superframe structures are used in many industrial wireless communication protocols (Chen et al., 2010; International Society of Automation, 2009; ZigBee Alliance, 2006), built upon the IEEE 802.15.4 MAC layer (IEEE 802.15.4, 2006). These MAC schemes are characterized by time-division multiple access (TDMA) protocols. In TDMA protocols, some given frequency channels are shared by the network nodes. At each timeslot, some links are allocated to the channels to transmit their data from a sending node to a receiving node. By repeating this, the data will finally arrive at the gateway. To model the protocols, we make the following natural assumptions for the communication:

Assumption 1. We assume the following properties:

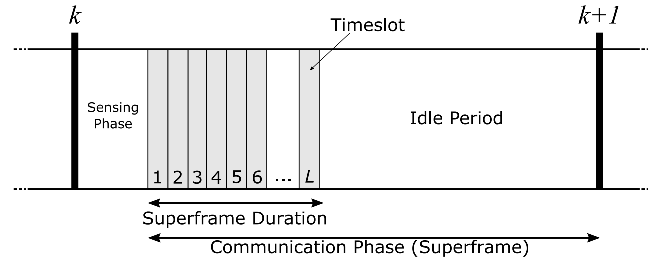


Fig. 2. Each time interval is divided into a sensing and a communication phase. The superframe duration is divided into timeslots. After the duration ends, sensors are in idle period (Araújo et al., 2014).

(*i*) All sensors have the same sampling interval and are perfectly time-synchronized.
(*ii*) Data can be transmitted among the nodes without failure, i.e., no packet dropout occurs.
(*iii*) The number L of timeslots in a single superframe is sufficiently large for accommodating all links in $\mathcal{G}$.

The data packet generated by sensor i is a tuple ('index', 'time', 'value'), where 'index' indicates the sensor index, 'time' the time-stamp when the data is generated, 'value' the measurement value. Thus, formally we can describe the data from sensor i generated at time k as $\left(i, k, x_k^{(i)}\right)$.

In this paper, for the sake of presentation simplicity, we assume that only one frequency channel is available.[1] Thus, at most one link is activated at each timeslot ℓ of superframe k, i.e., a link e is determined by a pair (k, ℓ). To indicate this link, we denote the link activated at timeslot ℓ of superframe k as $e(k, \ell)$. Let us denote $\mathcal{I}^{(i)}(k, \ell)$ as the data set that sensor i holds at timeslot $\ell \in \mathcal{L} \triangleq \{1, \ldots, L\}$ in superframe k, and let $\mathcal{D}(k, \ell) \subseteq \mathcal{I}^{(i)}(k, \ell)$ with $i = v_{\text{out}}(e(k, \ell))$ and $j = v_{\text{in}}(e(k, \ell))$ be the set of data transmitted by sensor i. That is,

$$\mathcal{D}(k, \ell) \triangleq \mathcal{I}^{(j)}(k, \ell+1) \backslash \mathcal{I}^{(j)}(k, \ell).$$

The set of the sensor indices ('index') of which the data is to be transmitted through $e(k, \ell)$ is expressed by

$$\mathcal{S}(k, \ell) \triangleq \left\{ i \in \mathcal{V}_s : (i, k', x_{k'}^{(i)}) \in \mathcal{D}(k, \ell), k' \leq k \right\}.$$

Then, given the initial data set $\mathcal{I}^{(i)}(-1, L)$, $\mathcal{I}^{(i)}(k, \ell)$ can be recursively written as

$$\mathcal{I}^{(i)}(k, \ell) = \begin{cases} \mathcal{I}^{(i)}(k-1, \ell) \oplus \left(i, k, x_k^{(i)}\right), & \text{if } \ell = 1, \\ \mathcal{I}^{(i)}(k, \ell-1) \oplus \mathcal{D}(k, \ell), & \\ \qquad \text{if } \ell \geq 2, \ i = v_{\text{in}}(e(k, \ell-1)), & \\ \mathcal{I}^{(i)}(k, \ell-1), & \\ \qquad \text{if } \ell \geq 2, \ i \neq v_{\text{in}}(e(k, \ell-1)), & \end{cases}$$

where the operation $\mathcal{I} \oplus \mathcal{D}$ is union but only the data packet with large time-stamp ('time') is preserved if $\mathcal{I}$ and $\mathcal{D}$ hold measurements from the same sensor. When L timeslots terminate, the gateway transmits all the measurement $\mathcal{D}_k \triangleq \mathcal{I}^{(0)}(k, L)$ to the estimator. We denote the elapsed time of the data from sensor i in $\mathcal{D}_k$ as $\tau_k^{(i)}$ which can be calculated from the current time and the time-stamp ('time'). Assuming that the data of sensor i in $\mathcal{D}_k$ is generated at time k', and therefore described as $\left(i, k', x_k^{(i)}\right)$, then we write $\tau_k^{(i)} = k - k'$.

The gateway is responsible for coordinating which sensors to be activated at which timeslots and which communication links

[1] Under Assumption 1, the results in this paper can be straightforwardly extended to the case of multiple frequency channels.

to be established. This function is called network scheduling. That is, the gateway decides the network schedule

$$\mathcal{C}_k \triangleq \big(e(k,1), \mathcal{S}(k,1), \ldots, e(k,L), \mathcal{S}(k,L)\big)$$

given the available information after superframe duration k (after timeslot L) denoted $\mathcal{K}_k^{(0)} \triangleq \{\mathcal{C}_{0:k}, \mathcal{D}_{0:k}\}$. The network scheduler chooses the schedule $\mathcal{C}_{k+1}$ for the next superframe according to

$$\mathcal{C}_{k+1} = f_k\big(\mathcal{K}_k^{(0)}\big)$$

where f_k is the map from the set of available information at the gateway to the set of network schedules. We also define $\mathbf{f} \triangleq (f_k)_{k\in\mathbb{N}_0}$ as the network scheduling strategy.

2.4. Energy consumption

The sensors consume a certain amount of energy when they receive data from and transmit data to other sensors. Here we introduce an energy consumption model often employed in wireless communication protocols (Heinzelman et al., 2002; Rajagopalan & Varshney, 2006). The energy consumption for receiving a packet, which contains p bits information, is

$$E_r(p) = E_{\text{elec}}p \tag{2}$$

where the energy coefficient E_{elec} is determined by the electronics, coding and other implementation aspects. The energy consumption for sending p bits information is

$$E_s(p,d) = E_{\text{elec}}p + E_{\text{amp}}d^2p \tag{3}$$

where E_{amp} is the energy coefficient for the amplifier and d is the distance to the receiving sensor or gateway. When transmitting multiple measurements, a sensor can aggregate them into a single packet in order to reduce the transmission overhead. This technology is called packet aggregation (Rajagopalan & Varshney, 2006). Assume that a single measurement from any sensor has c bits. Then the bits of information after aggregation are given by

$$p(q) = c[1 + (q-1)(1-r)] \tag{4}$$

where $q \in \mathbb{N}$ is the number of measurements and $r \in [0,1]$ is the data aggregation rate (Dou, Guo, Cao, & Zhang, 2007). If $r = 1$ the data is aggregated perfectly and the bits after aggregation are independent of the number of measurements, which is, for instance, the case for the LEACH protocol (Heinzelman et al., 2002). If $r = 0$, no packet aggregation is used. Notice that it is difficult to aggregate collected data from different sensors perfectly, but some parts of the data such as header can be removed when aggregating.

Let $q(k,\ell) \triangleq |\mathcal{S}(k,\ell)|$ be the number of measurements transmitted to node $v_{\text{in}}(e(k,\ell))$. Notice that $q(k,\ell)$ is determined by the network schedule $\mathcal{C}_k$, so the total energy consumption for sensor i to receive and send packets in the superframe at time k is given by

$$\begin{aligned} E_k^{(i)}(\mathcal{C}_k) = &\sum_{\ell: v_{\text{in}}(e(k,\ell))=i} E_r\big(p(q(k,\ell))\big) \\ &+ \sum_{\ell: v_{\text{out}}(e(k,\ell))=i} E_s\big(p(q(k,\ell)), d(e(k,\ell))\big). \end{aligned} \tag{5}$$

2.5. Remote estimation

After superframe duration k, the remote estimator computes an estimate

$$\hat{X}_k \triangleq \big(\hat{x}_k^{(1)}, \ldots, \hat{x}_k^{(N)}\big)$$

where $\hat{x}_k^{(i)}$ denotes the estimate of $x_k^{(i)}$. Let $\mathcal{K}_k^{(R)}$ denote the information set at the estimator. Notice that the sensor measurements sent by node 0 and the estimation history are accessible to the remote estimator. In other words, the information available to the remote estimator is

$$\mathcal{K}_k^{(R)} \triangleq \Big\{\hat{X}_{0:k-1}, \mathcal{D}_{0:k}\Big\}.$$

In this paper, as a metric of the estimator performance, we use the mean square error $\mathbb{E}[(\epsilon^{(i)})^\top\epsilon^{(i)}]$ with $\epsilon^{(i)} \triangleq x_k^{(i)} - \hat{x}_k^{(i)}$. Note that the optimal estimate for process i is computed recursively following the modified Kalman filter (Shi, Epstein, & Murray, 2010; Sinopoli et al., 2004) as

$$\begin{aligned} \hat{x}_k^{(i)} &= \mathbb{E}[x_k^{(i)} | \mathcal{K}_k^{(R)}] \\ &= \mathbb{E}[x_k^{(i)} | \mathcal{K}_{k-1}^{(R)}, \hat{X}_{k-1}, \mathcal{D}_k] \\ &= A^{\tau_k^{(i)}} x^{(i)}_{k-\tau_k^{(i)}}, \end{aligned} \tag{6}$$

with initial estimate $\hat{x}_0^{(i)} = 0$. Correspondingly, the error covariance of $x_k^{(i)}$ is denoted as

$$P_k^{(i)} \triangleq \mathbb{E}\Big[(x_k^{(i)} - \hat{x}_k^{(i)})(x_k^{(i)} - \hat{x}_k^{(i)})^\top | \mathcal{K}_k^{(R)}\Big].$$

Note that possible values of the error covariance are included in a set

$$P_k^{(i)} \in \{0, h_i(0), h_i^2(0), \ldots\}, \quad i \in \mathcal{V}_s, \quad \forall k \in \mathbb{N}_0, \tag{7}$$

where $h_i : \mathbb{S}_n^+ \to \mathbb{S}_n^+$ is the operator $h_i(X) = A_iXA_i^\top + W_i$, and $h_i^n(X)$ is the n-hold composition of $h_i(\cdot)$ with $h_i^0(X) = X$, since $P_k^{(i)}$ evolves with $h_i(\cdot)$ from 0 once the estimator receives the measurement (Leong, Dey & Quevedo, 2017). Then the error covariance is computed as $P_k^{(i)} = h_i^{\tau_k^{(i)}}(0)$ and we have

$$\mathbb{E}[(\epsilon^{(i)})^\top\epsilon^{(i)}] = \text{tr}\Big(h_i^{\tau_k^{(i)}}(0)\Big). \tag{8}$$

With this, the estimation error (8) is determined only by $\tau_k^{(i)}$, which is included in $\mathcal{D}_k$.

2.6. Problem formulation

The problem of interest is to find an optimal network scheduling strategy that minimizes long-term estimation errors penalized by sensor transmission energy usage. We define the cost at time k as

$$C(\mathcal{C}_k, \mathcal{D}_k) \triangleq \sum_{i\in\mathcal{V}_s} \text{tr}\Big(h_i^{\tau_k^{(i)}}(0)\Big) + E(\mathcal{C}_k)$$

where $E(\mathcal{C}_k) \triangleq \sum_{i\in\mathcal{V}_s} \beta_i E_k^{(i)}(\mathcal{C}_k)$ with $\beta_i > 0$. We formulate the following problem:

Problem 1.

$$\min_{\mathbf{f}=(f_0,f_1,\ldots)} \quad J(\mathbf{f}) \triangleq \limsup_{T\to\infty} \frac{1}{T}\sum_{k=0}^{T-1} C(\mathcal{C}_k, \mathcal{D}_k). \tag{9}$$

Remark 1. Problem 1 jointly optimizes a weighted average of the estimation error and sensor energy consumption. Minimization of (9) with given values of β_i corresponds to a minimum-cost schedule with energy consumption constraint given by some $\alpha_i > 0$:

$$\begin{aligned} \min_{\mathbf{f}} \quad & \limsup_{T\to\infty} \frac{1}{T}\sum_{k=0}^{T-1}\sum_{i\in\mathcal{V}_s} \mathbb{E}\Big[(\epsilon^{(i)})^\top\epsilon^{(i)}\Big] \\ \text{s.t.} \quad & \limsup_{T\to\infty} \frac{1}{T}\sum_{k=0}^{T-1} E_k^{(i)}(\mathcal{C}_k) \le \alpha_i, \ i \in \mathcal{V}_s. \end{aligned}$$

For Problem 1 to be well-posed, we make the following assumption.

Assumption 2. The graph $\mathcal{G}$ contains a spanning tree with the root being the gateway node 0.

Assumption 2 guarantees that persistently exciting protocols (Tabbara, Nesic, & Teel, 2007) can be configured over the network $\mathcal{G}$. Therefore, Problem 1 is well-posed as long as $\mathcal{G}$ contains a spanning tree.

3. Structures of network scheduler

In this section, we discuss structural properties of the network scheduler solving Problem 1. First, we show that an optimal schedule requires the network to carry sensor data within a superframe through a tree network formed by a set of activated links, by which the search space for an optimal strategy can be reduced. With this finding, we manage to separate Problem 1 into two subproblems: network routing and sensor selection.

3.1. Necessary conditions for network schedule optimality

We consider all communication links within a single superframe jointly and analyze the resulting graph by treating these links as a whole, where the notion of joint graph arises. Let us define the joint graph for a superframe k under a network scheduling strategy $\mathbf{f}$ in the following way. Denote $\mathcal{E}_k \triangleq (e(k,1), \ldots, e(k,L))$ the sequence of communication links in the superframe k selected from $\mathcal{E}$ of the underlying graph $\mathcal{G}$. Then we call $\mathcal{G}_k \triangleq (\mathcal{V}, \mathcal{E}_k)$ the joint graph of the superframe k. Let us also denote $\mathcal{S}_k \subseteq \mathcal{V}_s$ as the set of sensor indices that the latest data $\left(i, k, x_k^{(i)}\right)$ departs sensor i at the one of timeslots in superframe k, i.e., $i \in \mathcal{S}_k$ if and only if there exists $\ell \in \mathcal{L}$ such that the sending node of $e(k,\ell)$ is i and its data are included in this transmission. That is,

$$\mathcal{S}_k \triangleq \Big\{ i \in \mathcal{V}_s : \exists \ell \in \mathcal{L} \quad \text{s.t. } i = v_{\text{out}}(e(k,\ell)),\ \ i \in \mathcal{S}(k,\ell) \Big\}.$$

For an optimal scheduling strategy $\mathbf{f}^* = (f_0^*, \ldots, f_k^*, \ldots)$, denote the optimal network schedule at time k as $\mathcal{C}_k^* = \big(e^*(k,\ell), \mathcal{S}^*(k,\ell)\big)_{\ell=1}^{L}$, and the optimal set of the communication links and joint graph as $\mathcal{E}_k^*$ and $\mathcal{G}_k^*$, respectively. Furthermore, under a given optimal network schedule $\mathcal{C}_k^*$, we denote the index set $\mathcal{S}_k$ and the data set $\mathcal{D}_k$ as $\mathcal{S}_k^*$ and $\mathcal{D}_k^*$, respectively. Then we have the following lemma.

Lemma 1. *Suppose that Problem 1 has an optimal solution $\mathbf{f}^*$. Then the followings hold:*

(*i*) *If $i \in \mathcal{S}_k^*$, then $\left(i, k, x_k^{(i)}\right) \in \mathcal{D}_k^*$.*
(*ii*) *The joint graph $\mathcal{G}_k^*$ is a tree with node 0 being its unique root.*

Proof. Suppose that there exists $\ell \in \mathcal{L}$ such that $i = v_{\text{out}}(e^*(k,\ell)) \in \mathcal{S}^*(k,\ell)$. Obviously, the data $(i, k, x_k^{(i)})$ arrives at the gateway through a single path without a circle path from sensor i to the gateway. Let the arrival time of the data $(i, k, x_k^{(i)})$ be $k+m$, $m \in \mathbb{N}_0$. We show that $m = 0$ for an optimal network schedule. The proof is by contradiction. Suppose that $m > 0$. Consider a sequence of graphs $\tilde{\mathcal{G}}_{k:k+m} \triangleq \left(\tilde{\mathcal{G}}_k, \ldots, \tilde{\mathcal{G}}_{k+m}\right)$ which is the same as $\mathcal{G}_{k:k+m}^* \triangleq \left(\mathcal{G}_k^*, \ldots, \mathcal{G}_{k+m}^*\right)$ except that links $e \in \mathcal{G}_{k:k+m}^*$ that are used to transmit $x_k^{(i)}$ and the measurements aggregated into $x_k^{(i)}$ are removed, but rescheduled in $\tilde{\mathcal{G}}_{k+m}$ with the latest data $x_{k+m}^{(i)}$. Notice that $\mathcal{G}_{k:k+m}^*$ and $\tilde{\mathcal{G}}_{k:k+m}$ consume the same or smaller amount of energy, but $\tilde{\mathcal{G}}_{k:k+m}$ has smaller estimation error due to the monotonicity of $h_i(\cdot)$ starting from $X = 0$ (Shi & Zhang, 2012). This contradicts the optimality of $\mathcal{G}_{k:k+m}$. Thus, $m = 0$, hence $\left(i, k, x_k^{(i)}\right) \in \mathcal{D}_k$. The second statement is obvious from $m = 0$. The proof is now completed. □

Remark 2. Lemma 1 suggests that data $(i, k, x_k^{(i)})$ will arrive at the remote estimator within superframe k through tree graph $\mathcal{G}_k^*$, if it departs from sensor i in superframe k.

We give another lemma that indicates the order of link activation in an optimal network schedule. Suppose that the network schedule satisfies (*i*) and (*ii*) of Lemma 1. We introduce a partial order to the links in tree graph $\mathcal{G}_k$. That is, for any $e, e' \in \mathcal{E}_k$, we say $e \succeq e'$ if there exists a directed path from $v_{\text{in}}(e)$ to $v_{\text{out}}(e')$. It defines a partial order on $\mathcal{E}_k$ since we can readily show that it is reflexive, antisymmetric, and transitive.

Lemma 2 (*Upstream-first Rule*)**.** *Suppose that Problem 1 has an optimal solution $\mathbf{f}^*$. Then, $\ell_1 \le \ell_2$ if $e^*(k,\ell_1) \succeq e^*(k,\ell_2)$ for $e^*(k,\ell_1), e^*(k,\ell_2) \in \mathcal{E}_k^*$.*

Proof. By letting each sensor i in $\mathcal{G}_k^*$ send $x_k^{(i)}$ following upstream-first order, all measurements sampled and sent within the superframe k reach node 0 free of delays. Otherwise, a part of measurements received by node 0 will arrive with delays. In other words, any strategy $\mathbf{f}$ in this case can never be optimal. □

Remark 3. The upstream-first rule requires each sensor i in $\mathcal{G}_k$ to wait until all the scheduled upstream sensor data arrive. After their arrival, sensor i transmits its data to its downstream neighbor node. It is immaterial in what order of the upstream branches of node i are activated for transmission.

Lemmas 1 and 2 jointly suggest that, to construct a network schedule, it is essential to select which sensors need to transmit data to the remote estimator and to plan communication paths. The sensor selection fully determines the estimation error while the communication paths fully determines the communication cost. To investigate an optimal network scheduling strategy, we only need to focus on the path planning of data communication and the sensor selection. These two steps are separably studied in the sense that given a selected sensor set, we only need to account for the communication cost when we plan the communication paths. Therefore, in the sequel, we will investigate two subproblems: tree planning and sensor selection. The tree planning is studied with respect to sensor energy cost when a subset of sensors is selected. Then the sensor selection is investigated given the optimal communication paths.

3.2. Tree planning subproblem

In the previous subsection, we see that $\mathcal{G}_k^*$ should be always a tree with the unique root node 0 and the links are activated according to the upstream-first rule. In this subsection, we introduce a necessary condition to satisfy the statements (*i*) and (*ii*) of Lemma 1. Imposing this condition to $\mathcal{E}_k$, we formulate an integer linear problem called the tree planning problem, which gives a tree $\mathcal{G}_k$ minimizing the energy consumption $E(\mathcal{C}_k)$.

Let $z_k^{(i)}(e) \in \{0,1\}$ be an index function for any $i \in \mathcal{V}_s$, denoting whether $(i, k, x_k^{(i)})$ is transmitted through link $e \in \mathcal{E}$ at time k. That is, $z_k^{(i)}((j,m)) = 1$ if there exists $\ell \in \mathcal{L}$ such that $e(k,\ell) = (j,m)$ and $i \in \mathcal{S}(k,\ell)$, otherwise 0. To fulfill conditions (*i*) and (*ii*) of Lemma 1, it is necessary to satisfy the following constraints:

(*i*) Each node in $\mathcal{S}_k$ has outgoing flow of its own measurement, i.e., for $i \in \mathcal{S}_k$,

$$\sum_{m \in \mathcal{N}_i^{\text{out}}} z_k^{(i)}((i,m)) - \sum_{m \in \mathcal{N}_i^{\text{in}}} z_k^{(i)}((m,i)) = 1. \tag{10}$$

(*ii*) The gateway has only incoming flow, i.e., for $i \in \mathcal{S}_k$,

$$\sum_{m \in \mathcal{N}_0^{\text{in}}} z_k^{(i)}((m,0)) = 1. \tag{11}$$

(*iii*) Intermediate nodes of a path obey a flow balance, i.e., for $i \in \mathcal{S}_k$ and $j \neq i$,

$$\sum_{m \in \mathcal{N}_j^{\text{out}}} z_k^{(i)}(j,m) - \sum_{m \in \mathcal{N}_j^{\text{in}}} z_k^{(i)}((m,j)) = 0. \tag{12}$$

(*iv*) The nodes that are not in $\mathcal{S}_k$ also obey the flow balance, i.e., for $i \notin \mathcal{S}_k$, the constraint (12) holds.

Let $z_k^{(i)} = [z_k^{(i)}(e_1), \dots, z_k^{(i)}(e_{|\mathcal{E}|})]^\top \in \{0,1\}^{|\mathcal{E}|}$ be the vector of index functions for node i, where links are aligned in an appropriate order, and

$$z_k = [z_k^{(1)\top}, \dots, z_k^{(N)\top}]^\top \in \{0,1\}^{|\mathcal{E}|\cdot N}. \tag{13}$$

Then, using the node–arc incidence matrix G, the constraints (10)–(12) can be written in a compact form as

$$G z_k^{(i)} = b^{(i)}(\mathcal{S}_k), \quad i \in \mathcal{V}_s, \tag{14}$$

where $b^{(i)}(\mathcal{S}_k) \in \mathbb{R}^{N+1}$ is a vector with elements taking one of the values $0, \pm 1$ according to the right terms of (10)–(12).

Example 1. As an example of the flow constraint (14), consider the network shown in Fig. 1. For node 1, we denote

$$z_k^{(1)} = \begin{bmatrix} z_k^{(1)}((1,0)) \\ z_k^{(1)}((1,3)) \\ z_k^{(1)}((2,0)) \\ z_k^{(1)}((2,3)) \\ z_k^{(1)}((3,1)) \\ z_k^{(1)}((3,2)) \end{bmatrix}.$$

Then, by (10)–(12), we obtain (14) when $1 \in \mathcal{S}_k$ as

$$\begin{bmatrix} -1 & 0 & -1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & -1 & 0 \\ 0 & 0 & 1 & 1 & 0 & -1 \\ 0 & -1 & 0 & -1 & 1 & 1 \end{bmatrix} z_k^{(1)} = \begin{bmatrix} -1 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

where the left term matrix corresponds the node–arc incidence matrix G and the right term vector $b^{(1)}(\mathcal{S}_k)$.

Since $E(\mathcal{C}_k)$ is only a function of the network schedule at the current time k, the variable z_k that satisfies the constraints (10)–(12) can be pre-calculated with fixed $\mathcal{S}_k$. That is, at time k, given the set $\mathcal{S}_k$, recalling the energy consumption models (2)–(5) and the definition (13) of z_k, we obtain a tree network graph $\mathcal{G}_k$ by solving the following problem:

Problem 2 (*Tree Planning Subproblem*)**.**

$$E_{\min}(\mathcal{S}_k) \triangleq$$

$$\begin{aligned} \min_{z_k} \quad & E(\mathcal{C}_k) = \sum_{e \in \mathcal{E}} c\eta(e)\Big[(1-r)\sum_{i \in \mathcal{V}_s} z_k^{(i)}(e) + r \max_{i \in \mathcal{V}_s} z_k^{(i)}(e)\Big] \\ \text{s. t.} \quad & G z_k^{(i)} = b^{(i)}(\mathcal{S}), \quad i \in \mathcal{V}_s, \\ & z_k^{(i)}(e) \in \{0,1\}, \quad i \in \mathcal{V}_s, \quad e \in \mathcal{E}, \end{aligned}$$

where

$$\eta(e) \triangleq \begin{cases} \beta_{v_{\text{out}}(e)}\big(E_{\text{elec}} + E_{\text{amp}} d^2(e)\big) + \beta_{v_{\text{in}}(e)} E_{\text{elec}}, \\ \qquad\qquad\qquad\qquad \text{if } v_{\text{in}}(e) \in \mathcal{V}_s; \\ \beta_{v_{\text{out}}(e)}(E_{\text{elec}} + E_{\text{amp}} d^2(e)), \ \text{if } v_{\text{in}}(e) \in \{0\}. \end{cases}$$

Problem 2 is a binary integer problem, which is in general NP-hard. Nevertheless, due to a special algebraic property of the constraints, we manage to find the global minimizer of Problem 2 by solving a relaxed problem. The result is formally presented as follows.

Theorem 1. *A vector $z^* \in \{0,1\}^{N|\mathcal{E}|}$ is a minimizer of Problem 2 if and only if it is a minimizer of the following problem:*

Problem 3.

$$\begin{aligned} \min_{z_k, t} \quad & \sum_{e \in \mathcal{E}} c\eta(e)\Big[(1-r)\sum_{i \in \mathcal{V}_s} z_k^{(i)}(e) + rt(e)\Big] \\ \text{s. t.} \quad & G z_k^{(i)} = b^{(i)}(\mathcal{S}_k), \quad i \in \mathcal{V}_s, && (15) \\ & 0 \le z_k^{(i)}(e) \le 1, \quad i \in \mathcal{V}_s, && (16) \\ & z_k^{(i)}(e) \le t(e), \quad i \in \mathcal{V}_s, \quad e \in \mathcal{E}, && (17) \\ & t(e) \in \{0,1\}, \quad e \in \mathcal{E}. \end{aligned}$$

Proof. See Appendix A. □

Remark 4. In general, binary integer problems can be solved by a branch and bound algorithm. Theorem 1 shows that an optimal solution can be obtained by a relaxed problem without loss of performance. This extremely reduces the number of iterations in the algorithm—the number of possible branches are reduced to $2^{|\mathcal{E}|}$ from $2^{N|\mathcal{E}|}$.

3.3. Sensor selection subproblem

In the previous subsection, we saw that $\mathcal{C}_k$ is determined by solving Problem 3 with given $\mathcal{S}_k$ and applying the upstream-first rule to the resulted graph. We can rewrite the immediate cost using $\tau_k \triangleq [\tau_k^{(1)}, \dots, \tau_k^{(N)}]^\top$ and $\mathcal{S}_k$ as

$$C(\tau_k, \mathcal{S}_k) = \sum_{i \in \mathcal{V}_s} \text{tr}\Big(h_i^{\tau_k^{(i)}}(0)\Big) + E_{\min}(\mathcal{S}_k).$$

Due to the necessary condition (*i*) of Lemma 1, τ_k is determined by $\mathcal{S}_k$. To obtain the network scheduler, we need to find a map from τ_k to $\mathcal{S}_{k+1}$. This problem is called the sensor selection problem formulated as an MDP.

Define the MDP $\mathcal{M} \triangleq (\mathcal{Q}, \mathcal{A}, F(\cdot,\cdot), C(\cdot,\cdot))$ as follows:

(*i*) The state space is given by

$$\mathcal{Q} \triangleq \left\{\tau \in \mathbb{N}_0^N : \tau[i] \in \mathbb{N}_0, \ i \in \mathcal{V}_s\right\}.$$

(*ii*) The action space is given by

$$\mathcal{A} \triangleq \left\{\mathcal{S} : \mathcal{S} \in 2^{\mathcal{V}_s}\right\}.$$

(*iii*) The deterministic transition function from state τ to τ' with action $\mathcal{S} \in \mathcal{A}$ is defined as $F(\tau, \mathcal{S}) = \tau'$ where

$$\tau'[i] = \begin{cases} 0, & \text{if } i \in \mathcal{S}; \\ \tau[i] + 1, & \text{otherwise.} \end{cases}$$

(*iv*). The immediate cost for a transition from τ to τ' with action $\mathcal{S} \in \mathcal{A}$ is given by

$$C(\tau, \mathcal{S}) = \sum_{i \in \mathcal{V}_s} \text{tr}\left(h_i^{\tau'[i]}(0)\right) + E_{\min}(\mathcal{S}).$$

Fig. 3 illustrates the MDP $\mathcal{M}$ for a two-sensor case. A state $[\tau[1], \tau[2]]^\top = [u_1, u_2]^\top$ corresponds to sensors 1 and 2 transmitted u_1 and u_2 time units ago, respectively.

With this set-up, let us introduce a policy $\pi_k : \mathcal{Q} \to \mathcal{A}$ for MDP $\mathcal{M}$ and $\pi \triangleq (\pi_0, \pi_1, \dots)$. We are interested in a policy that minimizes the average cost by choosing the sensor set to be transmitting:

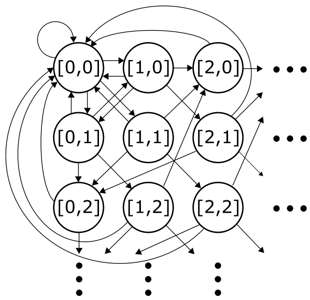


Fig. 3. MDP $\mathcal{M}$ states and their transitions for two sensors. $[u_1, u_2]$ in a circle indicates the MDP state $[\tau[1], \tau[2]]^\top = [u_1, u_2]^\top$. The arrows indicate state transitions.

Problem 4 (*Sensor Selection Problem*)**.**

$$\rho^* \triangleq \min_{\pi \in \Pi} \rho_\pi(\tau_0, \mathcal{S}_0) = \min_{\pi \in \Pi} \lim_{T \to \infty} \frac{1}{T} \sum_{k=0}^{T-1} C(\tau_k, \mathcal{S}_k) \tag{18}$$

where $\mathcal{S}_k = \pi_{k-1}(\tau_{k-1})$ for $k \in \mathbb{N}$ given the initial state and action $(\tau_0, \mathcal{S}_0)$, and Π the set of all possible policies.

Now we will show that $\mathcal{M}$ has an optimal stationary deterministic policy $\pi^* \in \Pi$. The idea is to check the sufficient conditions for the existence of an optimal solution for countable infinite-state MDPs (Sennott, 2009) as discussed in Leong, Dey and Quevedo (2017) and Wu et al. (2018).

Theorem 2. *Consider the MDP $\mathcal{M}$. There exist a constant ρ^* and a relative-value function $H(\cdot)$ satisfying the Bellman equation*

$$\rho^* + H(\tau) = \min_{\mathcal{S} \in \mathcal{A}} \Big\{ C(\tau, \mathcal{S}) + H\big(F(\tau, \mathcal{S})\big) \Big\}. \tag{19}$$

Proof. See Appendix B. □

Theorem 2 shows that there exists a stationary deterministic optimal policy. To find such a policy, next we show that $\mathcal{M}$ can be restricted into a finite-state MDP without loss of performance under the assumption that the estimation error is expensive compared to the communication cost. For the finite-state MDP, we can use classical algorithms such as value iteration. To do this, we make the following assumption.

Assumption 3. Each process $i \in \mathcal{V}_s$ either satisfies:

(*i*) $\lambda_{\max}(A_i) \ge 1$, or
(*ii*) $\lambda_{\max}(A_i) < 1$ and $\mathrm{tr}(X_i) > E_{\min}(\{i\})$, where $X_i \in \mathbb{S}^n_{++}$ is the unique solution to the Lyapunov equation $A_i^\top X_i A_i + W_i - X_i = 0$.

Lemma 3. *Suppose that Assumption 3 holds. Then there exists a constant*

$$\delta_i \triangleq \min_{\kappa} \Big\{ \kappa \in \mathbb{N}_0 : \mathrm{tr}\big(h_i^\kappa(0)\big) > E_{\min}(\{i\}) \Big\}$$

for all $i \in \mathcal{V}_s$.

Proof. It is immediate from the monotonicity of $h_i^n(X)$ along n starting from $X = 0$ (Shi & Zhang, 2012) and Assumption 3. □

Finally, we have the following theorem.

Theorem 3. *Suppose that Assumption 3 holds. Consider MDP $\mathcal{M}$. If $\tau[i] \ge \delta_i$, then $i \in \pi^*(\tau)$.*

Proof. See Appendix C. □

Let us define the finite-state MDP:

$$\mathcal{M}_f \triangleq (\mathcal{Q}_f, \mathcal{A}_f(\cdot), F(\cdot,\cdot), C(\cdot,\cdot))$$

with

$$\mathcal{Q}_f \triangleq \left\{ \tau \in \mathbb{N}_0^N : \tau[i] \le \delta_i,\ i \in \mathcal{V}_s \right\}$$

and $\mathcal{A}_f(\tau) \triangleq \mathcal{A} \backslash \bar{\mathcal{A}}_f(\tau)$ where $\bar{\mathcal{A}}_f(\tau) \triangleq \{\mathcal{S} \in \mathcal{A} : \exists i \in \mathcal{V}_s, \tau[i] = \delta_i, i \notin \mathcal{S}\}$. That is, sensor i is always selected at state τ when $\tau[i] = \delta_i$. In the optimal policy of the MDP $\mathcal{M}$, the state will move into $\mathcal{Q}_f$ in the next transition even if its initial state τ_0 is outside of $\mathcal{Q}_f$. After that, the states never leave $\mathcal{Q}_f$. Thus, the initial cost will be ignorable since its contribution to the average cost is reduced to zero as T tends to infinity (Bertsekas, 2017). Consequently, we can derive the optimal policy of $\mathcal{M}$ by solving the finite state MDP $\mathcal{M}_f$ without loss of performance. We show that the optimal sensor selection is periodic.

Corollary 1. *Suppose that Assumption 3 holds. Then there exists an optimal periodic schedule generated by an optimal policy π^*.*

Proof. Since the MDP $\mathcal{M}_f$ is deterministic, we can fix an arbitrary action as an optimal one at any state in $\mathcal{Q}_f$. Furthermore, since $\mathcal{Q}_f$ is finite, there exists a recurrent state over π^*. Thus, if the system reaches the recurrent state again, the state transition will repeat. Hence the result follows. □

3.4. Two-step value iteration algorithm

Previously, we showed that the set of optimally selected sensors over time is periodic under Assumption 3 and can be obtained by solving a finite-state MDP $\mathcal{M}_f$ with pre-calculated $E_{\min}(\mathcal{S})$. We present a two-step algorithm based on relative value iteration (Puterman, 2005):

Step 1. (Algorithm 1) Calculate an optimal tree network for each candidate set of sensor selection.
Step 2. (Algorithm 2) Calculate an optimal policy of the MDP $\mathcal{M}_f$.

Algorithm 1 Computation of an optimal tree network and energy cost

1: **INPUT:** $\eta(e)$, r
2: **OUTPUT:** $E_{\min}(\mathcal{S})$
3: **for** $\mathcal{S} \in \mathcal{A}$ **do**
4: Compute $E_{\min}(\mathcal{S})$ in Problem 3
5: **end for**

Algorithm 2 has in general high computational complexity. The reasons are twofold: first, the number of states of $\mathcal{M}_f$ depends on δ_i, which increases exponentially by the number of sensors. Second, since we allow to pick any sensor at every time instance, the size of action space is 2^N at every iteration. These issues motivate us to construct suboptimal schedules in the next section.

4. Construction of suboptimal solutions

In this section, we introduce algorithms to compute a suboptimal solution in an efficient way.

4.1. Reduced MDP schedule

The first algorithm solves an approximate MDP by restricting the size of the state and action spaces. To do this, we introduce some sets of sensors and assume that the sensors in the same set are always scheduled to transmit together. The reduced MDP (R-MDP), $\tilde{\mathcal{M}} \triangleq (\tilde{\mathcal{Q}}, \mathcal{A}_f(\cdot), F(\cdot,\cdot), \tilde{C}(\cdot,\cdot))$ is obtained as follows:

Algorithm 2 Computation of an optimal schedule

1: **INPUT:** $\mathcal{M}_f$, $E_{\min}(\mathcal{S})$, v^0, $\epsilon > 0$, $\bar{\tau}$
2: **OUTPUT:** $\pi^*(\tau)$
3: $v^0 \leftarrow v^0 - v^0(\bar{\tau}) \cdot \mathbf{1}_{|\mathcal{Q}_f|}$ and $k = 0$
4: **for** $\tau \in \mathcal{Q}_f$ **do**
5: Compute

$$v^{k+1}(\tau) = \min_{\mathcal{S} \in \mathcal{A}_f} \left\{ C(\tau, \mathcal{S}) + v^k(F(\tau, \mathcal{S})) \right\} \tag{20}$$

6: **end for**
7: $v^{k+1} \leftarrow v^{k+1} - v^{k+1}(\bar{\tau}) \cdot \mathbf{1}_{|\mathcal{Q}_f|}$
8: **if** $\max(v^{k+1}(\tau) - v^k(\tau)) - \min(v^{k+1}(\tau) - v^k(\tau)) \leq \epsilon$ **then**
9: Go to Step 13
10: **else**
11: $k \leftarrow k + 1$ and return to Step 4
12: **end if**
13: For each $\tau \in \mathcal{Q}_f$, set

$$\pi^*(\tau) = \arg\min_{\mathcal{S} \in \mathcal{A}_f} \left\{ C(\tau, \mathcal{S}) + v^k(F(\tau, \mathcal{S})) \right\}$$

(*i*) Split $\mathcal{V}_s$ into M disjoint subsets $\tilde{\mathcal{V}} \triangleq \{\mathcal{V}_1, \ldots, \mathcal{V}_M\}$.
(*ii*) Define the bounds $\tilde{\delta}_j \triangleq \min\{\delta_i : i \in \mathcal{V}_j\}, j = 1, \ldots, M$.
(*iii*) Define the state space

$$\tilde{Q} \triangleq \{\tau \in \mathbb{N}_0^M : \tau[j] = 0, \ldots, \delta_j, j = 1, \ldots, M\}$$

and the action space $\mathcal{A}_f(\tau)$.
(*iv*). Define the cost

$$\tilde{C}(\tau, \mathcal{S}) \triangleq \sum_{i=1}^{N} \mathrm{tr}\left(h_i^{\tau[j(i)]}(0)\right) + E_{\min}(\mathcal{S})$$

where $j(i)$ indicates the subset j in $\tilde{\mathcal{V}}$ to which sensor i belongs.

We compute the R-MDP schedule by calling Algorithms 1 and 1 with $\mathcal{M}$ replaced by $\tilde{\mathcal{M}}$.

4.2. Fixed-period algorithm

The idea of our second algorithm is to fix the transmission period of each sensor obtained by solving smaller MDPs. Then the whole schedule is obtained by combining all such schedules. The procedure is given by the fixed period algorithm (FPA) in Algorithm 3.

Let us denote the sensor selection obtained by Algorithm 3 as $\mathcal{S}_{\mathrm{FPA},k}$. For this algorithm, we have the following result.

Proposition 1. *Suppose that the data aggregation rate $r = 0$. Then the schedule obtained by Algorithm 3 is optimal, i.e.,*

$$\lim_{T \to \infty} \frac{1}{T} \sum_{k=0}^{T-1} C(\tau_k, \mathcal{S}_{FPA,k}) = \rho^*.$$

Proof. We have $E(\mathcal{C}_k) = \sum_{e \in \mathcal{E}} \sum_{i \in \mathcal{V}_s} \eta(e) z_k^{(i)}(e)$, which means that the energy consumption $E(\mathcal{C}_k)$ is a linear combination of $z_k^{(i)}(e)$. That is, we have

$$E_{\min}(\mathcal{S}_{\mathrm{FPA},k}) = \sum_{i \in \mathcal{V}_s} \sigma_i(\mathcal{S}_{\mathrm{FPA},k}^{(i)}) E_{\min}(\{i\})$$

where $\sigma_i(\mathcal{S}) = 1$ if $i \in \mathcal{S}$, otherwise 0. Then we have

Algorithm 3 Fixed Period Algorithm

INPUT: $\eta(e)$
OUTPUT: $\{\mathcal{S}_{\mathrm{FPA},k}\}_{k=0}^{D}$
for $i \in \mathcal{V}_s$ **do**
Compute $E_{\min}(\{i\})$
Set $\mathcal{M}_i = (\mathcal{Q}_i, \mathcal{A}_i, F(\cdot, \cdot), C(\cdot, \cdot))$ with $\mathcal{Q}_i = \{\tau_i \in \mathbb{N}_0 : \tau_i = 0, \ldots, \delta_i\}$ and $\mathcal{A}_i = \{\emptyset, i\}$
Solve $\mathcal{M}_i$ and compute a period D_i
Set

$$\mathcal{S}_{\mathrm{FPA},k}^{(i)} = \begin{cases} \{i\}, & \text{if } k \equiv 0 \mod D_i \\ \emptyset, & \text{if } k \not\equiv 0 \mod D_i \end{cases}$$

end for
Compute D, the least common multiple of D_i, $i = 1, \ldots, N$
for $k = 0, 1, \ldots, D$ **do**
Set $\mathcal{S}_{\mathrm{FPA},k} = \bigcup_{i \in \mathcal{V}_s} \mathcal{S}_{\mathrm{FPA},k}^{(i)}$
Compute $E_{\min}(\mathcal{S}_{\mathrm{FPA},k})$
end for

$$\begin{aligned} &C(\tau, \mathcal{S}_{\mathrm{FPA},k}) \\ &= \sum_{i \in \mathcal{V}_s} \mathrm{tr}\left(h_i^{\tau[i]}(0)\right) + E_{\min}(\mathcal{S}_{\mathrm{FPA},k}) \\ &= \sum_{i \in \mathcal{V}_s} \left[\mathrm{tr}\left(h_i^{\tau[i]}(0)\right) + \sigma_i(\mathcal{S}_{\mathrm{FPA},k}^{(i)}) E_{\min}(\{i\}) \right]. \end{aligned}$$

Thus, minimization of

$$\lim_{T \to \infty} \frac{1}{T} \sum_{k=0}^{T-1} \left[\mathrm{tr}\left(h_i^{\tau[i]}(0)\right) + \sigma_i(\mathcal{S}_{\mathrm{FPA},k}^{(i)}) E_{\min}(\{i\}) \right]$$

for each $i \in \mathcal{V}_s$ yields the minimum cost. This completes the proof. □

5. Numerical examples

In this section, we present three numerical examples to illustrate our results in this paper. In the first example, we evaluate the performance of the R-MDP and FPA schedules by comparing them with the optimal one for the small network depicted in Fig. 1. In the second example, we provide a larger network, and show that the two suboptimal algorithms can obtain the schedules efficiently even if the size of the original MDP is too large to efficiently compute the optimal schedule. The third example shows that these suboptimal schedules are scalable to networks consisting of a hundred nodes.

5.1. Optimal and suboptimal schedules for a small network ($N = 3$)

To see the performances of the proposed algorithms, we consider the small network depicted in Fig. 1. The system parameters of the three plants are

$$A_1 = \begin{bmatrix} 1.3 & 1.2 \\ 0 & 1.4 \end{bmatrix}, \ A_2 = \begin{bmatrix} 1.5 & 0.8 \\ 0 & 1.2 \end{bmatrix}, \ A_3 = \begin{bmatrix} 3.5 & 2.0 \\ 0 & 3.1 \end{bmatrix},$$

with $W_i = 0.1 I_2$, for $i = 1, 2, 3$ where I_2 is the 2×2 identity matrix. For communication parameters, we assume that $E_{\mathrm{elec}} = E_{\mathrm{amp}} = 1$, $c = 1$, $\beta_i = 1$ for $i = 1, 2, 3$, $d((1, 0)) = d((2, 0)) = d((1, 3)) = d((2, 3)) = 1$, and $r = 0.5$. The action set consists of every possible subset of sensors selected to transmit accompanied by all possible routes as shown in Table 1.

Table 1
All possible sensor selections and their routes to the gateway for the network in Fig. 1.

Action index	Sensor selection	Path
0	∅	–
1	1	$1 \to 0$
2	1	$1 \to 3 \to 2 \to 0$
3	2	$2 \to 0$
4	2	$2 \to 3 \to 1 \to 0$
5	3	$3 \to 1 \to 0$
6	3	$3 \to 2 \to 0$
7	1, 2	$1 \to 0, 2 \to 0$
8	1, 2	$1 \to 3 \to 2 \to 0$
9	1, 2	$2 \to 3 \to 1 \to 0$
10	2, 3	$3 \to 2 \to 0$
11	2, 3	$3 \to 1 \to 0, 2 \to 0$
12	2, 3	$2 \to 3 \to 1 \to 0$
13	3, 1	$3 \to 1 \to 0$
14	3, 1	$3 \to 2 \to 0, 1 \to 0$
15	3, 1	$1 \to 3 \to 2 \to 0$
16	1, 2, 3	$3 \to 1 \to 0, 2 \to 0$
17	1, 2, 3	$3 \to 2 \to 0, 1 \to 0$
18	1, 2, 3	$1 \to 3 \to 2 \to 0$
19	1, 2, 3	$2 \to 3 \to 1 \to 0$

Table 2
The optimal paths for each sensor selection.

Sensor selection	Optimal path	Energy cost
∅	–	0
1	$1 \to 0$	1
2	$2 \to 0$	1
3	$3 \to 1 \to 0$	3
1, 2	$1 \to 0, 2 \to 0$	2
2, 3	$3 \to 2 \to 0$	3.5
1, 3	$3 \to 1 \to 0$	3.5
1, 2, 3	$3 \to 1 \to 0, 2 \to 0$	4.5

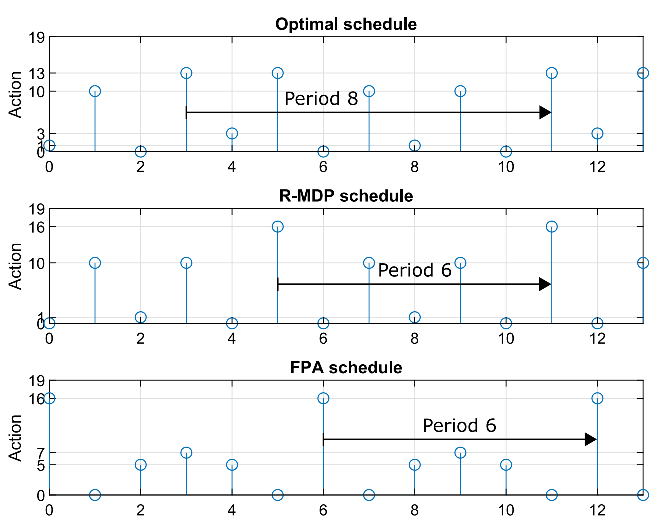


Fig. 4. Three schedules obtained by the proposed optimal and suboptimal Algorithms. Top: Optimal schedule, middle: R-MDP schedule, bottom: FPA schedule.

Optimal schedule

First, we derive the optimal schedule. Algorithm 1 gives the optimal paths and their energy costs, see Table 2. By the value of $E_{\min}(\{i\})$, $i = 1, 2, 3$, we obtain the bounds of the MDP state space as $\delta_1 = 3$, $\delta_2 = 4$, $\delta_3 = 3$ with Theorem 3. Then we can find the optimal schedule by Algorithm 1. The result is shown in Fig. 4 (top). The period of the optimal schedule is 8 in which actions 0, 1, 3, 10, and 13 from Table 1 are taken.

Table 3
Averaged costs and the sizes of MDP.

	Averaged cost	$\lvert\mathcal{Q}\rvert$	$\lvert\mathcal{A}\rvert$
Optimal	4.09	80	8
R-MDP	4.17	16	4
FPA	4.35	≤5	≤2

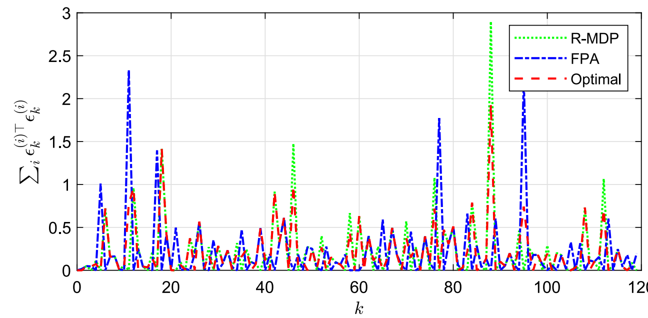


Fig. 5. Estimation performance comparison of the three schedules: optimal (red), R-MDP (green), and FPA (blue). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

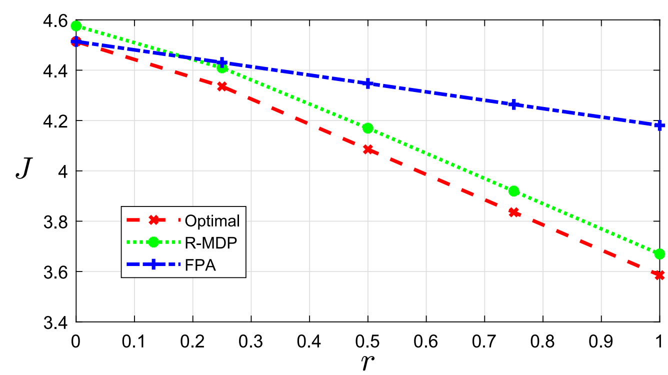


Fig. 6. The averaged cost J of the optimal, R-MDP and FPA schedules with respect to the data aggregation rate r for the network with $N = 3$.

R-MDP schedule

Next, we formulate an R-MDP by setting $\mathcal{V}_1 = \{1\}$ and $\mathcal{V}_2 = \{2, 3\}$. Then we have $\delta_1 = 3$ and $\delta_2 = 3$. The obtained schedule is shown in Fig. 4 (middle). It has period 6 with actions 0, 1, 10, and 16. We can see that sensors 2 and 3 are always selected together.

FPA schedule

We derive an FPA schedule by Algorithm 3. Now we have $E_{\min}(\{1\}) = E_{\min}(\{2\}) = 2$, and $E_{\min}(\{3\}) = 5$, with which we formulate MDP $\mathcal{M}_i$ for $i = 1, 2, 3$. Then we obtain the fixed activation period for each sensor: $D_1 = 3$, $D_2 = 3$, and $D_3 = 2$, which yields the period 6 schedule as shown in Fig. 4 (bottom).

Performance evaluation

The averaged cost and the sizes of the MDPs for each schedule are summarized in Table 3. We can see that the R-MDP and FPA schedules obtain similar performances compared to the optimal one even though the sizes of the MDPs are considerably reduced. The estimation performance $\sum_{i\in\mathcal{V}_s} \epsilon_k^{(i)\top}\epsilon_k^{(i)}$ is plotted in Fig. 5. We can conclude that the proposed suboptimal schedules obtain well-performing schedules.

We show the averaged cost of these three schedules with respect to the data aggregation rate r in Fig. 6. It confirms

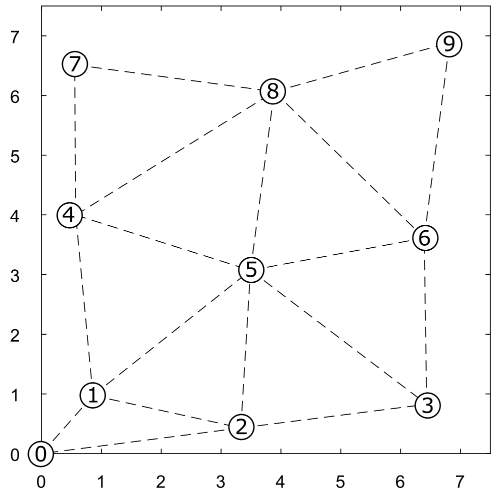


Fig. 7. A sensor network with $N = 9$ sensors.

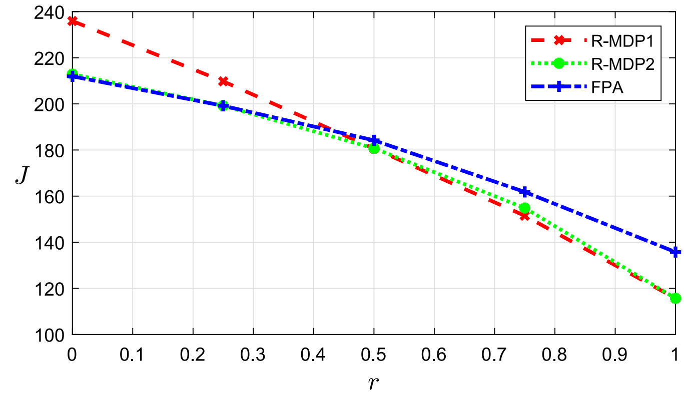


Fig. 8. The averaged cost of the two reduced MDP schedules and the FPA schedule with respect to the data aggregation rate r for the network with $N = 9$ sensors.

Table 4
Averaged costs, periods, sizes of MDP.

	Averaged cost	Period	$\lvert\mathcal{Q}\rvert$	$\lvert\mathcal{A}\rvert$
R-MDP1	180.63	12	400	16
R-MDP2	180.76	12	840	16
FPA	184.20	60	≤ 6	≤ 2

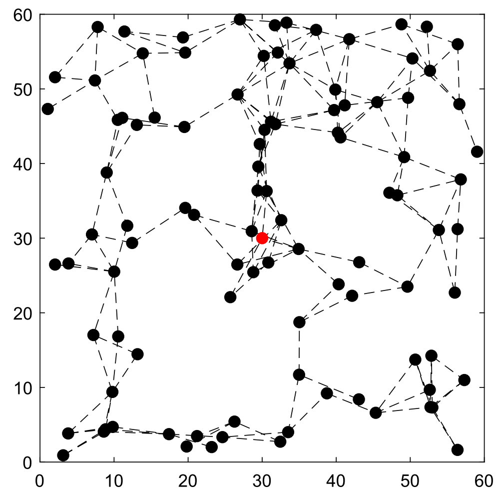


Fig. 9. A sensor network with $N = 99$ sensors. Black dots represent sensors and the red one the gateway. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Proposition 1, i.e., the averaged cost of the FPA schedule is optimal when $r = 0$ and this averaged cost is the upper bound of the FPA averaged cost for any r. The difference of the averaged costs of the optimal and FPA schedules increases with increasing r, since the optimal schedule receives benefit of the data aggregation. In Fig. 4, the FPA schedule takes action 5 once in a period, i.e., sensor 3 is selected alone. However, this is not effective in terms of the energy cost since the data cannot be aggregated even though it passes through sensor 1. In the optimal schedule in Fig. 4, sensor 3 is always selected together with sensor 1 (action 13) or with sensor 2 (action 10). The R-MDP schedule results in larger costs for any r. However, the cost is close to that of the FPA schedule if $r = 1$, since the R-MDP still tries to take advantage of the data aggregation.

5.2. Suboptimal schedules for a larger network ($N = 9$)

To see the performances of the proposed suboptimal scheduling algorithms in a more realistic situation, we consider the network shown in Fig. 7. The network consists of $N = 9$ sensors distributed over a square field and a gateway at the origin. The sensors can communicate with the other sensors when the distances are shorter than $d_{\max} = 4$. The plants are given by

$$A_1 = \begin{bmatrix} 2.3 & 1.2 \\ 0 & 1.9 \end{bmatrix},\ A_2 = \begin{bmatrix} 2 & 0 \\ 0 & 1.6 \end{bmatrix},\ A_3 = \begin{bmatrix} 3 & 2.4 \\ 2.2 & 3.5 \end{bmatrix},$$
$$A_4 = \begin{bmatrix} 1.4 & 0.2 \\ 0.5 & 1.5 \end{bmatrix},\ A_5 = \begin{bmatrix} 2.3 & 0.5 \\ 0.2 & 1.4 \end{bmatrix},\ A_6 = \begin{bmatrix} 2.2 & 0 \\ 0 & 2 \end{bmatrix},$$
$$A_7 = \begin{bmatrix} 2.5 & 0.2 \\ 1.2 & 2.2 \end{bmatrix},\ A_8 = \begin{bmatrix} 2.1 & 1.2 \\ 0 & 1.5 \end{bmatrix},\ A_9 = \begin{bmatrix} 3.5 & 3.6 \\ 2.3 & 3.5 \end{bmatrix},$$

with $W_i = 0.1I_2$, for $i = 1, \ldots, 9$, $E_{\text{elec}} = E_{\text{amp}} = 1$, $c = 4$, $\beta_i = 1$ for $i = 1, \ldots, 9$, and $r = 0.5$. The bounds of the MDP states are obtained as $\delta_1 = 4$, $\delta_2 = 6$, $\delta_3 = 3$, $\delta_4 = 6$ $\delta_5 = 5$, $\delta_6 = 6$, $\delta_7 = 4$, $\delta_8 = 5$, and $\delta_9 = 3$. This means that the original MDP problem is computationally expensive to solve as the size of its state space is of the order of $\prod_{i=1}^{N} \delta_i \sim 10^6$. The averaged cost and size of the MDPs are summarized in Table 4. R-MDP1 uses R-MDP algorithm with grouping sensors based on their locations in order to take advantage of the data aggregation, i.e., we include sensors placed in the near distance into the same set. We take $\mathcal{V}_1 = \{1, 2\}, \mathcal{V}_2 = \{3, 5\}, \mathcal{V}_3 = \{4, 7\}$, and $\mathcal{V}_4 = \{6, 8, 9\}$. For R-MDP2, we make sensor sets based on the bound δ_i to avoid too many or too few transmissions with respect to the divergence speed of each error covariance, i.e., sensors with close bounds are included in the same sets. We use $\mathcal{V}_1 = \{1, 7\}, \mathcal{V}_2 = \{3, 9\}, \mathcal{V}_3 = \{5, 8\}$, and $\mathcal{V}_4 = \{2, 4, 6\}$. The FPA schedule is obtained from small MDPs $\mathcal{M}_i$, $i = 1, \ldots, 9$. The obtained FPA schedule generated by $D_1 = 4, D_2 = 6, D_3 = 3, D_4 = 6, D_5 = 5, D_6 = 6, D_7 = 4, D_8 = 5$, and $D_9 = 3$ has period 60.

Fig. 8 shows the averaged costs of the three schedules with respect to r. As in Proposition 1, the FPA schedule is optimal when $r = 0$. Thus, it has a near optimal performance if r is small. The performance further degrades compared to the R-MDPs when r is large. Both approaches for the R-MDP schedules reduce cost when r is large. The R-MDP2 has a comparatively better performance regardless of the value of r. It implies that a way group sensors influence the performance.

5.3. Suboptimal schedules for a large network ($N = 99$)

To see the scalability of the proposed suboptimal scheduling algorithms, we consider the larger network shown in Fig. 9. The network consists of $N = 99$ sensors distributed over a square field (black dots) and a gateway at the center of the field (red dot).

Table 5
Averaged costs, periods, sizes of MDP.

	Averaged cost	Period	$\lvert\mathcal{Q}\rvert$	$\lvert\mathcal{A}\rvert$
R-MDP	15300.9	5	252	8
FPA	15498.6	2520	≤ 10	≤ 2

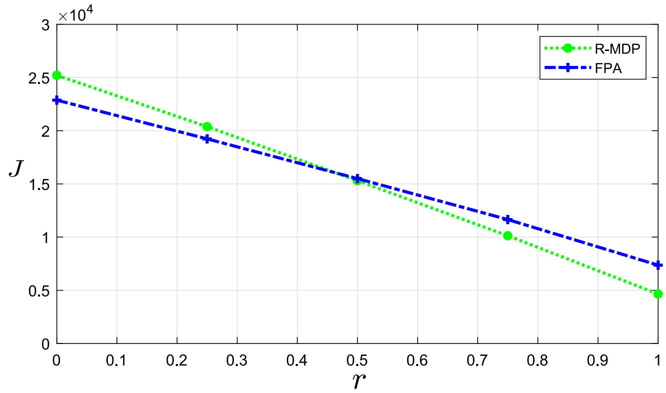


Fig. 10. The averaged cost of the two reduced MDP and the FPA schedules with respect to the data aggregation rate r for the network with $N = 99$.

For the R-MDP algorithm, we divide sensors into three subgroups, so that the number of states is reduced to 252 when $r = 0.5$ (Table 5). The FPA schedule can also be obtained by solving 99 small MDPs where the maximum number of the states is 10. The period of the FPA schedule is 2520. However, the period may further increase in some cases since it is derived by taking the least common multiple among D_i, $i \in \mathcal{V}_s$. In this case, it is difficult to obtain the actual performance J since we need to recompute $E_{\min}(\mathcal{S}_{\text{FPA},k})$ in step 12 of Algorithm 3. Fig. 10 shows the averaged costs of the two schedules with respect to r. Similar to previous examples, the FPA algorithm results in better schedules when r is small while the R-MDP does better when r is large.

6. Conclusion

In this paper, we proposed a co-design framework of multi-hop network scheduling for remote estimation. We formulated an optimization problem minimizing an infinite-time averaged estimation error covariance with sensor energy consumption. We showed that the problem can be divided into two subproblems by exploiting necessary conditions for network scheduling optimality. An existence condition for a periodic optimal solution was derived. To reduce the computational complexity, we proposed two alternative algorithms to obtain suboptimal schedules. It was demonstrated how the proposed algorithms are effective in numerical examples.

There are several possible directions to extend this work. Further analysis of constructing suboptimal schedules will be interesting. Especially, how to group the sensors to obtain an effective suboptimal schedule is still an open problem. Considering delay and packet dropouts induced by wireless communication will be important to include in future studies. Furthermore, introducing this framework to the existing industrial wireless communication protocols, e.g., WirelessHART or ISA-100, is to be considered, and then taking into account more practical constraints, such as the number of timeslots in a superframe.

Appendix A. Proof of Theorem 1

To prove Theorem 1, we need the following definition and supporting lemmas.

Definition 1 (*Total Unimodular Matrix Schrijver, 2003*)**.** A square integer matrix is unimodular if it has determinant $+1$ or -1. A matrix is totally unimodular if every square non-singular submatrix of it is unimodular.

Lemma 4. *Let X be a totally unimodular matrix. Then the following matrices are also totally unimodular:*

(i) $\text{diag}(X, \ldots, X)$,

(ii) $\begin{bmatrix} X \\ I \end{bmatrix}$ *and* $\begin{bmatrix} X \\ -I \end{bmatrix}$,

(iii) $\begin{bmatrix} X & -X \end{bmatrix}$ *and* $\begin{bmatrix} X \\ -X \end{bmatrix}$.

Lemma 5 (Papadimitriou & Steiglitz, 1998)**.** *If A is totally unimodular, then all the vertices of the polyhedron $\{x : Ax \leq b\}$ are integers for any integer vector b.*

Proof of Theorem 1. First, we transform Problem 2 into an integer linear problem by introducing $t(e) = \max_{i \in \mathcal{V}_s} z_k^{(i)}(e)$ and the constraint $z_k^{(i)}(e) \leq t(e)$ for $i \in \mathcal{V}_s$ and $e \in \mathcal{E}$. We show that relaxing $0 \leq z_k^{(i)}(e) \leq 1$ still obtains a binary integer solution.

The constraints (15)–(17) can be written in a compact form

$$\{z_k : \mathbf{G} z_k \leq b\}$$

where $b = [b(\mathcal{S}_k)^\top, \ldots, -b(\mathcal{S}_k)^\top, 1, \ldots, 1, 0, \ldots, 0]^\top$ with $b(\mathcal{S}_k)^\top = [b^{(1)}(\mathcal{S}_k)^\top, \ldots, b^{(N)}(\mathcal{S}_k)^\top]^\top$ and

$$\mathbf{G} \triangleq \begin{bmatrix} \text{diag}(G, \ldots, G) \\ -\text{diag}(G, \ldots, G) \\ I \\ -I \end{bmatrix}.$$

The matrix G is the node–arc incidence matrix of $\mathcal{G}$, therefore it is totally unimodular. Then by Lemma 4, $\mathbf{G}$ is totally unimodular. Fixing $t(e)$ to 0 or 1 for all $e \in \mathcal{E}$, z_k obtains the integer solution if the corresponding linear problem is feasible (Lemma 5). Thus, a minimizer of Problem 2 is equal to that of Problem 3. This completes the proof. □

Appendix B. Proof of Theorem 2

To prove Theorem 2, we define a standard policy. Consider a Markov chain with countable infinite state space $\mathcal{Q}$. Let us denote $p^n_{q_1,q_2}$ by the probability that the state that is currently at q_1 will be q_2 for the first time exactly after $n \geq 1$ transitions. That is,

$$p^n_{q_1,q_2} = \Pr(\tau_k \neq q_2, k = 1, \ldots, n-1, \tau_n = q_2 | \tau_0 = q_1).$$

The expected first passage time (Sennott, 2009) $t_{\tau,z}$ is denoted as

$$t_{q_1,q_2} = \sum_{n=1}^{\infty} n p^n_{q_1,q_2},$$

and the corresponding averaged total cost, called the expected first passage cost (Sennott, 2009) is denoted by c_{q_1,q_2}.

Definition 2 (*Sennott, 2009*)**.** A randomized stationary policy π is a standard policy if there exists a state $z \in \mathcal{Q}$ such that the expected first passage time $t_{\tau,z}$ from τ to z satisfies $t_{\tau,z} < \infty$ for all $\tau \in \mathcal{Q}$, and the expected first passage cost $c_{\tau,z}$ from τ to z satisfies $c_{\tau,z} < \infty$ for all $\tau \in \mathcal{Q}$.

Lemma 6 (Sennott, 2009, Corollary 7.5.10)**.** *Assume that the following conditions hold:*

(i) There exists a standard policy π such that the positive recurrent class induced by π is equal to $\mathcal{Q}$.

(ii) Given $U > 0$, the set $\mathcal{Q}_U = \{\tau : C(\tau, \mathcal{S}) \leq U$ for some $\mathcal{S}\}$ is finite.

Then there exists a solution to the Bellman equation (19) for the average cost problem with countable infinite state space $\mathcal{Q}$.

Proof of Theorem 2. The proof follows (Leong, Dey & Quevedo, 2017) by considering a randomized policy π such that at any states sensor i transmits its measurement with probability θ_i and $\theta \triangleq \prod_{i\in\mathcal{V}_s} \theta_i$ satisfies $1 - 1/\lambda_{\max}^2(A_i) < \theta < 1$ for all $i \in \mathcal{V}_s$. Let $z = [0, \dots, 0] \in \mathbb{N}_0^N$, then at any states it comes back to z with probability θ. Thus,

$$t_{\tau,z} = \theta + 2(1-\theta)\theta + 3(1-\theta)^2\theta + \cdots = \frac{1}{\theta} < \infty.$$

Notice that $c_{\tau_1,z} \leq c_{\tau_2,z}$ if $\tau_1[i] \leq \tau_2[i], \forall i$ and $E_{\min}(\mathcal{S}_1 \cup \mathcal{S}_2) \leq E_{\min}(\mathcal{S}_1) + E_{\min}(\mathcal{S}_2)$, the expected average cost is

$$\begin{aligned}
c_{\tau,z} &\leq \sum_{i\in\mathcal{V}_s} \operatorname{tr}\left(h_i^{\tau[i]}(0)\right) \\
&\quad + (1-\theta)\left[c_{\tau+\mathbf{1}_N,z} + \sum_{i\in\mathcal{V}_s} E_{\min}(\{i\})\right] \\
&\quad + \theta \sum_{i\in\mathcal{V}_s} E_{\min}(\{i\}) \\
&= \sum_{i\in\mathcal{V}_s} \operatorname{tr}\left(h_i^{\tau[i]}(0)\right) + \sum_{i\in\mathcal{V}_s} E_{\min}(\{i\}) + (1-\theta)c_{\tau+\mathbf{1}_N,z} \\
&= \sum_{n=0}^{\infty}(1-\theta)^n\left[\sum_{i\in\mathcal{V}_s} \operatorname{tr}\left(h_i^{\tau[i]+n}(0)\right) + \sum_{i\in\mathcal{V}_s} E_{\min}(\{i\})\right] \\
&= \sum_{n=0}^{\infty}(1-\theta)^n \sum_{i\in\mathcal{V}_s} \operatorname{tr}\left(h_i^{\tau[i]+n}(0)\right) + \frac{1}{\theta}\sum_{i\in\mathcal{V}_s} E_{\min}(\{i\}) \\
&< \infty
\end{aligned}$$

The boundedness of the last inequality holds from the assumption $1 - 1/\lambda_{\max}^2(A_i) < \theta < 1$ (Schenato, 2008; Schenato, Sinopoli, Franceschetti, Poolla, & Sastry, 2007; Xu & Hespanha, 2005). Hence, π is a standard policy.

Next, we show that the positive recurrent class is equal to $\mathcal{Q}$. Consider an arbitrary state $\tau \in \mathcal{Q}$. This state is reachable from state z after $\tau_{\max} \triangleq \max_i\{\tau[i]\}$ by letting sensor i transmit its data for the first $\tau_{\max} - \tau[i]$ transitions and not transmit $\tau[i]$ transitions after that. Let us denote the probability of this realization as θ'. Then, for any state τ, one can return to this state with probability equal to or higher than $\theta'' \triangleq \theta\cdot\theta'$ after $\tau_{\max}$ transitions. Thus, the probability that one returns to the state τ is

$$\theta'' + (1-\theta'')\theta'' + (1-\theta'')^2\theta'' + \cdots = 1,$$

which shows that the recurrent class is equal to $\mathcal{Q}$, hence the first condition is verified.

The second condition is verified as $C(\tau, \mathcal{S})$ is monotonically increasing in τ. □

Appendix C. Proof of Theorem 3

To prove Theorem 3, we first present the following lemma.

Lemma 7. *Suppose that Assumption 3 holds. Consider MDP $\mathcal{M}$. Then, for all $i \in \mathcal{V}_s$, there exists a time instance $k \in \mathbb{N}_0$ such that $i \in \pi^*(\tau_k)$.*

Proof. It is obvious since no transmission policy is never optimal if the process is unstable or the process is stable but the transmission cost is lower than the steady-state estimation error. □

Next, we introduce a partial order over the state space $\mathcal{Q}$. For the states $\tau, \tau' \in \mathcal{Q}$, we say $\tau \succ \tau'$ if $\tau[i] > \tau'[i]$ for all $i \in \mathcal{V}_s$.

Proof of Theorem 3. Proof is given by contradiction. Suppose that π^* is an optimal policy with $i \notin \pi^*(\tau)$ where $\tau[i] \geq \delta_i$. Consider a policy π' such that $i \in \pi'(\tau)$, but all the other sensors are selected as same as $\pi^*(\tau)$. Let $\tau_1^* \triangleq F(\tau, \pi^*(\tau))$ and $\tau_1' \triangleq F(\tau, \pi'(\tau))$ be the next state of τ according to the policy π^* and π', respectively. Obviously, $\tau_1^* \succ \tau_1'$ since $\tau_1^*[j] = \tau_1'[j]$ for $j \neq i$, and $\tau_1^*[i] = \tau^*[i] + 1$, $\tau_1'[i] = 0$. Define $V(\tau, \mathcal{S}) \triangleq C(\tau, \mathcal{S}) + H(F(\tau, \mathcal{S}))$, then by the Bellman principle, the optimal policy π^* needs to satisfy $V(\tau, \pi^*(\tau)) - V(\tau, \pi'(\tau)) < 0$. We will show that π^* contradicts this principle. Now, we have

$$\begin{aligned}
&V(\tau, \pi^*(\tau)) - V(\tau, \pi'(\tau)) \\
&= C(\tau, \pi^*(\tau)) + H(\tau_1^*) - C(\tau, \pi'(\tau)) - H(\tau_1') \\
&= \sum_{j\in\mathcal{V}_s} \operatorname{tr}\left(h_j^{\tau_1^*[j]}(0)\right) - \sum_{j\in\mathcal{V}_s} \operatorname{tr}\left(h_j^{\tau_1'[j]}(0)\right) \\
&\quad + E_{\min}(\pi^*(\tau)) - E_{\min}(\pi'(\tau)) + H(\tau_1^*) - H(\tau_1') \\
&= \operatorname{tr}\left(h_i^{\tau[i]+1}(0)\right) + E_{\min}(\pi^*(\tau)) - E_{\min}(\pi^*(\tau) \cup \{i\}) \\
&\quad + H(\tau_1^*) - H(\tau_1') \\
&\geq \operatorname{tr}\left(h_i^{\delta_i}(0)\right) - E_{\min}(\{i\}) + H(\tau_1^*) - H(\tau_1') \\
&> H(\tau_1^*) - H(\tau_1'). \qquad \text{(C.1)}
\end{aligned}$$

The first inequality holds due to the subadditivity of $E_{\min}(\cdot)$. Let $\mathcal{S}_1^* \triangleq \pi^*(\tau_1^*)$ be the action given by the optimal policy at state τ_1^*. Consider a policy π_1' such that $\pi_1'(\tau_1') = \mathcal{S}_1^*$. Also let $\tau_2^* \triangleq F(\tau_1, \pi^*(\tau_1^*))$ and $\tau_2' \triangleq F(\tau_1', \pi_1'(\tau_1'))$ be the next state of τ_1^* according to π^* and that of π_1' according to π_1', respectively. The inequality (C.1) continues

$$\begin{aligned}
&H(\tau_1^*) - H(\tau_1') \\
&= C(\tau_1^*, \pi^*(\tau_1^*)) + H(F(\tau_1^*, \pi^*(\tau_1^*))) \\
&\quad - C(\tau_1', \pi^*(\tau_1')) - H(F(\tau_1', \pi^*(\tau_1'))) \\
&\geq C(\tau_1^*, \pi^*(\tau_1^*)) + H(F(\tau_1^*, \pi^*(\tau_1^*))) \\
&\quad - C(\tau_1', \pi_1'(\tau_1')) - H(F(\tau_1', \pi_1'(\tau_1'))) \\
&\geq C(\tau_1^*, \mathcal{S}_1^*) - C(\tau_1', \mathcal{S}_1^*) + H(\tau_2^*) - H(\tau_2') \\
&= \operatorname{tr}\left(h_i^{\tau_2^*[i]}(0)\right) - \operatorname{tr}\left(h_i^{\tau_2'[i]}(0)\right) + H(\tau_2^*) - H(\tau_2') \\
&> H(\tau_2^*) - H(\tau_2'). \qquad \text{(C.2)}
\end{aligned}$$

If $i \in \mathcal{S}_1^*$, then we have $\tau_2^*[i] = \tau_2'[i] = 0$, i.e., $\tau_2^* = \tau_2'$. Then we obtain $V(\tau, \pi^*(\tau)) - V(\tau, \pi'(\tau)) > 0$. If $i \notin \mathcal{S}_1^*$, repeating (C.2), we have

$$H(\tau_1^*) - H(\tau_1') > \cdots > H(\tau_n^*) - H(\tau_n') > \cdots.$$

By Lemma 7, we have a time instance k such that $i \in \mathcal{S}_k^*$. Thus, we have $V(\tau, \pi^*(\tau)) - V(\tau, \pi'(\tau)) > 0$ and this contradicts the optimality of π^*. □

References

Agrawal, P., Ahlén, A., Olofsson, T., & Gidlund, M. (2014). Long term channel characterization for energy efficient transmission in industrial environments.. *IEEE Transactions on Communications*, *62*(8), 3004–3014.

Ahlén, A., Åkerberg, J., Eriksson, M., Isaksson, A. J., Iwaki, T., Johansson, K. H., et al. (2019). Toward wireless control in industrial process automation: A case study at a paper mill. *IEEE Control Systems Magazine*, *39*(5), 36–57.

Araújo, J., Mazo, M., Anta, A., Tabuada, P., & Johansson, K. H. (2014). System architectures, protocols and algorithms for aperiodic wireless control systems. *IEEE Transactions on Industrial Informatics*, *10*(1), 175–184.

Bertsekas, D. P. (2017). *Dynamic programming and optimal control (vol. I)*. Athena Scientific.

Chen, D., Nixon, M., & Mok, A. (2010). *WirelessHART: Real-time mesh network for industrial automation*. Springer.

Chipara, O., He, Z., Xing, G., Chen, Q., Wang, X., Lu, C., Stankovic, J., & Abdelzaher, T. (2006). Real-time power-aware routing in sensor networks. In *Proc. IEEE int. workshop on quality of service* (pp. 83–92).

Demirel, B., Zou, Z., Soldati, P., & Johansson, M. (2014). Modular design of jointly optimal controllers and forwarding policies for wireless control. *IEEE Transactions on Automatic Control*, *59*(12), 3252–3265.

Di Girolamo, G. D., & D'Innocenzo, A. (2019). Codesign of controller, routing and scheduling in wirelessHART networked control systems. *International Journal of Robust and Nonlinear Control*.

Dou, J., Guo, Z., Cao, J., & Zhang, G. (2007). Data aggregation rate assumption for wireless sensors. In *Proc. IEEE int. workshop on radio-frequency integration technology* (pp. 278–281).

Gupta, V., Chung, T. H., Hassibi, B., & Murray, R. M. (2006). On a stochastic sensor selection algorithm with applications in sensor scheduling and sensor coverage. *Automatica*, *42*(2), 251–260.

Han, D., Mo, Y., Wu, J., Weerakkody, S., Sinopoli, B., & Shi, L. (2015). Stochastic event-triggered sensor schedule for remote state estimation. *IEEE Transactions on Automatic Control*, *60*(10), 2661–2675.

Han, D., Wu, J., Zhang, H., & Shi, L. (2017). Optimal sensor scheduling for multiple linear dynamical systems. *Automatica*, *75*, 260–270.

Hasenfratz, D., Meier, A., Moser, C., Chen, J., & Thiele, L. (2010). Analysis, comparison, and optimization of routing protocols for energy harvesting wireless sensor networks. In *Proc. IEEE int. conf. on sensor networks, ubiquitous, trustworthy computing* (pp. 19–26).

Heinzelman, W. B., Chandrakasan, A. P., & Balakrishnan, H. (2002). An application-specific protocol architecture for wireless microsensor networks. *IEEE Transactions on Wireless Communications*, *1*(4), 660–670.

IEEE 802. 15. 4 (2006). IEEE 802.15.4 standard: wireless medium access control (MAC) and pysical layer (PHY) specification for low-rate wireless personal area networks (WPANs). http://www.ieee802.org/15/pub/TG4.html.

International Society of Automation (2009). Wireless systems for industrial automation: Process control and related applications, ISA-100.11a-2009.

Iwaki, T., Wu, Y., Wu, J., Sansberg, H., & Johansson, K. H. (2017). Wireless sensor network scheduling for remote estimation under energy constraints. In *Proc. IEEE conf. on decision and control* (pp. 3362–3367).

Jawaid, S. T., & Smith, S. L. (2015). Submodularity and greedy algorithms in sensor scheduling for linear dynamical systems. *Automatica*, *61*, 282–288.

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

Leong, A. S., Quevedo, D. E., Tanaka, T., Dey, S., & Ahlén, A. (2017). Event-based transmission scheduling and LQG control over a packet dropping link. In *Proc. IFAC world congress* (pp. 8945–8950).

Li, Y., Phillips, S., & Sanfelice, R. G. (2018). Robust distributed estimation for linear systems under intermittent information. *IEEE Transactions on Automatic Control*, *63*(4), 973–988.

Lu, C., Saifullah, A., Li, B., Sha, M., Gonzalez, H., Gunatilaka, D., et al. (2016). Real-time wireless sensor-actuator networks for industrial cyber-physical systems. *Proceedings of the IEEE*, *104*(5), 1013–1024.

Maass, A. I., Nešić, D., Postoyan, R., & Dower, P. M. (2019a). $\mathcal{L}_p$ Stability of networked control systems implemented on WirelessHART. *Automatica*, *109*, Article 108514.

Maass, A. I., Nešić, D., Postoyan, R., & Dower, P. M. (2019b). Observer design for non-linear networked control systems with persistently exciting protocols. *IEEE Transactions on Automatic Control*.

Mo, Y., Ambrosino, R., & Sinopoli, B. (2011). Sensor selection strategies for state estimation in energy constrained wireless sensor networks. *Automatica*, *47*(7), 1330–1338.

Mo, Y., Garone, E., & Sinopoli, B. (2014). On infinite-horizon sensor scheduling. *Systems & Control Letters*, *67*, 65–70.

Papadimitriou, C. H., & Steiglitz, K. (1998). *Combinatorial optimization: algorithms and complexity*. Dover.

Park, P., Ergen, S. C., Fischione, C., Lu, C., & Johansson, K. H. (2018). Wireless network design for control systems: A survey. *IEEE Communications Surveys and Tutorials*, *20*(2), 978–1013.

Puterman, M. L. (2005). *Markov decision processes: discrete stochastic dynamic programming*. John Wiley & Sons.

Rajagopalan, R., & Varshney, P. K. (2006). Data aggregation techniques in sensor networks: A survey. *IEEE Communications Surveys and Tutorials*, *8*(4), 48–63.

Schenato, L. (2008). Optimal estimation in networked control systems subject to random delay and packet drop. *IEEE Transactions on Automatic Control*, *53*(5), 1311.

Schenato, L., Sinopoli, B., Franceschetti, M., Poolla, K., & Sastry, S. S. (2007). Foundations of control and estimation over lossy networks. *Proceedings of the IEEE*, *95*(1), 163–187.

Schrijver, A. (2003). *Combinatorial optimization: polyhedra and efficiency (vol. 24)*. Springer Science & Business Media.

Sennott, L. I. (2009). *Stochastic dynamic programming and the control of queueing systems*. John Wiley & Sons.

Sha, M., Gunatilaka, D., Wu, C., & Lu, C. (2017). Empirical study and enhancements of industrial wireless sensor-actuator network protocols. *IEEE Internet of Things Journal*, *4*(3), 696–704.

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

Shi, L., & Zhang, H. (2012). Scheduling two Gauss–Markov systems: An optimal solution for remote state estimation under bandwidth constraint. *IEEE Transactions on Signal Processing*, *60*(4), 2038–2042.

Sinopoli, B., Schenato, L., Franceschetti, M., Poolla, K., Jordan, M. I., & Sastry, S. S. (2004). Kalman filtering with intermittent observations. *IEEE Transactions on Automatic Control*, *49*(9), 1453–1464.

Tabbara, M., Nesic, D., & Teel, A. R. (2007). Stability of wireless and wireline networked control systems. *IEEE Transactions on Automatic Control*, *52*(9), 1615–1630.

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

Weerakkody, S., Mo, Y., Sinopoli, B., Han, D., & Shi, L. (2016). Multi-sensor scheduling for state estimation with event-based, stochastic triggers. *IEEE Transactions on Automatic Control*, *61*(9), 2695–2701.

Willig, A. (2008). Recent and emerging topics in wireless industrial communications: A selection. *IEEE Transactions on Industrial Informatics*, *4*(2), 102–124.

Wu, J., Jia, Q., Johansson, K. H., & Shi, L. (2013). Event-based sensor data scheduling: Trade-off between communication rate and estimation quality. *IEEE Transactions on Automatic Control*, *58*(4), 1041–1046.

Wu, S., Ren, X., Dey, S., & Shi, L. (2017). Optimal scheduling of multiple sensors with packet length constraint. In *Proc. IFAC world congress* (pp. 14430–14435).

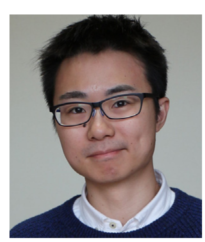
Wu, S., Ren, X., Dey, S., & Shi, L. (2018). Optimal scheduling of multiple sensors over shared channels with packet transmission constraint. *Automatica*, *96*, 22–31.

Xu, Y., & Hespanha, J. P. (2005). Estimation under uncontrolled and controlled communications in networked control systems. In *Proc. IEEE conf. on decision and control and european control conf.* (pp. 842–847).

Zhang, H., Ayoub, R., & Sundaram, S. (2017). Sensor selection for kalman filtering of linear dynamical systems: Complexity, limitations and greedy algorithms. *Automatica*, *78*, 202–210.

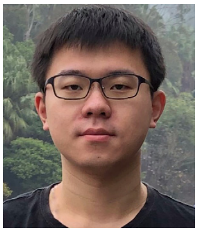
Zhao, L., Zhang, W., Hu, J., Abate, A., & Tomlin, C. J. (2014). On the optimal solutions of the infinite-horizon linear sensor scheduling problem. *IEEE Transactions on Automatic Control*, *59*(10), 2825–2830.

ZigBee Alliance (2006). Zigbee specification. https://www.zigbee.org/download/standards-zigbee-specification/.

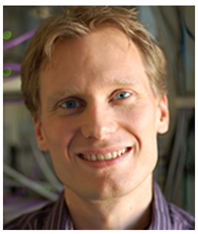


Takuya Iwaki received the B.E. and M.E degrees from Tokyo Institute of Technology, Tokyo, Japan, in 2009 and 2012, respectively. He was a control systems engineer in the oil and gas industry from 2012 to 2016. He is currently a doctoral student at the School of Electrical Engineering and Computer Science, KTH Royal Institute of Technology, Stockholm, Sweden. His main research interests include control and estimation over wireless communication and their application to industrial process control systems.

Junfeng Wu received the B.Eng. degree from the Department of Automatic Control, Zhejiang University, Hangzhou, China, and the Ph.D. degree in electrical and computer engineering from Hong Kong University of Science and Technology, Hong Kong, in 2009, and 2013, respectively. From September to December 2013, he was a Research Associate in the Department of Electronic and Computer Engineering, Hong Kong University of Science and Technology. From January 2014 to June 2017, he was a Postdoctoral Researcher in the ACCESS (Autonomic Complex Communication nEtworks, Signals and Systems) Linnaeus Center, School of Electrical Engineering, KTH Royal Institute of Technology, Stockholm, Sweden. He is currently with the College of Control Science and Engineering, Zhejiang University, Hangzhou, China. His research interests include networked control systems, state estimation, and wireless sensor networks, multiagent systems. He received the Guan Zhao-Zhi Best Paper Award at the 34th Chinese Control Conference in 2015.



Yuchi Wu received the B.Eng. degree from School of Electronic and Information Engineering, Beihang University, Beijing, China, in 2015. Later, he received the M.Phil. degree in 2017 and the Ph.D. degree in 2020, both from Department of Electronic and Computer Engineering, Hong Kong University of Science and Technology, Kowloon, Hong Kong. From March 2020 to May 2020, he was a visiting student in the School of Mathematics and Statistics, Carleton University, Ottawa, ON, Canada. His research interests include mean-field games, large-scale systems and state estimation in wireless sensor networks.



Henrik Sandberg is Professor at the Division of Decision and Control Systems, KTH Royal Institute of Technology, Stockholm, Sweden. He received the M.Sc. degree in engineering physics and the Ph.D. degree in automatic control from Lund University, Lund, Sweden, in 1999 and 2004, respectively. From 2005 to 2007, he was a Postdoctoral Scholar at the California Institute of Technology, Pasadena, USA. In 2013, he was a visiting scholar at the Laboratory for Information and Decision Systems (LIDS) at MIT, Cambridge, USA. He has also held visiting appointments at the Australian National University and the University of Melbourne, Australia. His current research interests include security of cyber–physical systems, power systems, model reduction, and fundamental limitations in control. Dr. Sandberg was a recipient of the Best Student Paper Award from the IEEE Conference on Decision and Control in 2004, an Ingvar Carlsson Award from the Swedish Foundation for Strategic Research in 2007, and a Consolidator Grant from the Swedish Research Council in 2016. He has served on the editorial boards of IEEE Transactions on Automatic Control and the IFAC Journal Automatica.

Karl Henrik Johansson is Professor with the School of Electrical Engineering and Computer Science at KTH Royal Institute of Technology in Sweden and Director of Digital Futures. He received M.Sc. and Ph.D. degrees from Lund University. He has held visiting positions at UC Berkeley, Caltech, NTU, HKUST Institute of Advanced Studies, and NTNU. His research interests are in networked control systems and cyber–physical systems with applications in transportation, energy, and automation networks. He is a member of the Swedish Research Council's Scientific Council for Natural Sciences and Engineering Sciences. He has served on the IEEE Control Systems Society Board of Governors, the IFAC Executive Board, and is currently Vice-President of the European Control Association. He has received several best paper awards and other distinctions from IEEE, IFAC, and ACM. He has been awarded Distinguished Professor with the Swedish Research Council and Wallenberg Scholar with the Knut and Alice Wallenberg Foundation. He has received the Future Research Leader Award from the Swedish Foundation for Strategic Research and the triennial Young Author Prize from IFAC. He is Fellow of the IEEE and the Royal Swedish Academy of Engineering Sciences, and he is IEEE Control Systems Society Distinguished Lecturer.