Title	On stochastic sensor network scheduling for multiple processes
Author(s)	Han, Duo; Wu, Junfeng; Mo, Yilin; Xie, Lihua
Citation	Han, D., Wu, J., Mo, Y., & Xie, L. On stochastic sensor network scheduling for multiple processes. IEEE Transactions on Automatic Control, 62(12), 6633-6640. doi:10.1109/TAC.2017.2717193
Date	2016
URL	http://hdl.handle.net/10220/44021
Rights	© 2016 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works. The published version is available at: [http://dx.doi.org/10.1109/TAC.2017.2717193].

On Stochastic Sensor Network Scheduling for Multiple Processes

Duo Han[*], Junfeng Wu[†], Yilin Mo[*1], Lihua Xie[*]

Abstract—We consider the problem of multiple sensor scheduling for remote state estimation of multiple process over a shared link. In this problem, a set of sensors monitor mutually independent dynamical systems in parallel but only one sensor can access the shared channel at each time to transmit the data packet to the estimator. We propose a stochastic event-based sensor scheduling in which each sensor makes transmission decisions based on both channel accessibility and distributed event-triggering conditions. The corresponding minimum mean squared error (MMSE) estimator is explicitly given. Considering information patterns accessed by sensor schedulers, time-based ones can be treated as a special case of the proposed one. By ultilizing realtime information, the proposed schedule outperforms the time-based ones in terms of the estimation quality. Resorting to solving an Markov decision process (MDP) problem with average cost criterion, we can find optimal parameters for the proposed schedule. As for practical use, a greedy algorithm is devised for parameter design, which has rather low computational complexity. We also provide a method to quantify the performance gap between the schedule optimized via MDP and any other schedules.

I. Introduction

Sensor scheduling is crucial for remote state estimation in cyber-physical systems (CPS). Typically, the main task for sensor scheduling is to improve estimation quality subject to communication constraints [1]–[6]. In this paper we focus on the problem of bandwidth constrained sensor scheduling for state estimation. To be specific, the distributed sensors in charge of different monitoring tasks are sharing a common channel for data transmission. At each time instant, only one sensor can access the communication channel. We consider a sensor schedule deciding which sensor is able to access the channel at each time in order to optimize the overall state estimation quality.

The problem of single sensor scheduling subject to limited transmission rate has been studied in [1], [3], [5], [7]. Generally speaking, if a sensor uses only prior knowledge of systems, an optimal policy for transmission is very likely to transmit the data packets periodically [3], [7], [8]. This kind

of policies are referred to as time-based schedules. When the sensor makes transmission decisions based on realtime innovations, which generally has better estimation performance then periodic ones if properly designed but induces higher computational complexity, the transmission is likely to be random [1], [5], [6]. This kind of policies are referred as event-based schedules. Weerakkody et al. [9] extended [5] by considering multiple sensors monitoring one process, but without any constraint on channel accessibility.

Compared with the single system case, not enough research efforts have been put in the case of different sensors monitoring different systems, which is widely encountered in practice. A simple motivating example is the underground petroleum storage using WirelessHART technology in [10]. Underground salt caverns are often used for crude oil storage. Brine and crude oil flowing in both directions are measured by sensors and reported to the control center through a gateway device. The control center aims to maintain a certain pressure inside the caverns. Obviously, each sensor competes with others for the gateway access to achieve their own goal. Therefore, a schedule for optimizing an objective function taking the benefits of all sensors into account is desirable. A few preliminary works have considered this case. Savage and La Scala [11] studied multiple sensor scheduling for a set of scalar Gauss-Markov systems over a finite horizon. They considered the optimality in terms of terminal estimation error covariance. An optimal policy is to schedule all the transmissions in the end of the horizon. However, the terminal error covariance metric is only suitable for finite horizon scheduling problems. To study an infinite horizon scheduling, Shi et al. [12] adopted a metric of averaged estimation error, studied two multi-dimensional systems over an infinite horizon and proposed an explicit optimal periodic sensor schedule. The conclusion in [12] benefits from the mutual exclusiveness of two sensors. When the number of sensors is beyond three, closed-form optimal scheduling is formidable to obtain. Another similar problem where one sensor is scheduled to monitor multiple processes was studied in [13]. The problem is same with the multiple sensor scheduling problem in nature. The authors proposed algorithms to search schedules such that the error covariance of each system is bounded by some constant matrix. Unlike the single-sensor case, multi-sensor event-based scheduling is not fully investigated to push the limit of performance. A preliminary work by Han et al. [14] improved [12] by designing an event-based schedule which depends on the importance of a sensor's measurements.

In this note we consider an event-based sensor scheduling design for a set of sensors. At each time every sensor makes

a transmission decision based on both channel accessibility by sensing the carrier and the importance of its own data by checking some criteria. Only when the channel is accessible and the data is considered as being sufficiently important, the sensor will transmit the data packet. The sensor scheduling studied in this note is restricted to a kind of stochastic event-based sensor scheduling since it can maintain Gaussianity of estimation process and bypass the nonlinear problem induced by the event-triggering mechanism, e.g., [1], [6]. Compared with time-based schedules, the proposed event-based schedule dynamically assigns the communication resource according to the needs of sensors. From operating principle point of view, the time- and event-based sensor schedules are analogous to time division multiple access (TDMA) and carrier sense multiple access with collision avoidance (CSMA/CA) medium access control (MAC) in communication networks, respectively. The main contributions are summarized as follows: (1) We first propose an event-based scheduling infrastructure with network cooperation and self event-triggering mechanism. Any time-based schedules can be treated as a special case of the framework. (2) Based on the underlying schedule, we derive the minimum mean squared error (MMSE) estimator and analyze the communication behaviour of each sensor. (3) We model an Markov decision process (MDP) problem with average cost criterion to seek the optimal parameters for the class of proposed stochastic schedules. (4) For computational simplicity, we also propose a greedy parameter design method. Moreover, we are able to analytically quantify the performance any schedule by showing a lower bound of the optimal cost.

Notations: $\mathbb{Z}_+$ is the set of positive integers. $\mathbb{S}^n_+$ ($\mathbb{S}^n_{++}$) is the set of n by n positive semi-definite (definite) matrices. $\mathrm{Tr}[\cdot]$ denotes the trace of a matrix. $X^{1/2}$ denotes the square root of $X \in \mathbb{S}^n_+$. For functions f with appropriate domains, $f^0(X) := X$, and $f^t(X) := f(f^{t-1}(X))$. $x[i]$ represents the ith entry of the vector x. For a matrix X, $X(i,j)$ represents the entry on the ith row and jth column of X. For $x \in \mathbb{R}$, $\lfloor x \rfloor$ is the largest integer that is not larger than x and $\lceil x \rceil$ is the smallest one that is not less than x. Denote the set or sequence $\{x_i\}_{i=j}^k = \{x_j, x_{j+1}, \ldots, x_k\}$, $j \le k$ and if $j > k$ then $\{x_i\}_{i=j}^k = \emptyset$.

II. PROBLEM SETUP

A. System Model

Consider the following n mutually independent linear time-invariant (LTI) systems, which are monitored by n sensors respectively:

$$x_i(k+1) = A_i x_i(k) + \omega_i(k), \tag{1a}$$
$$y_i(k) = C_i x_i(k) + \nu_i(k), \ i \in \mathcal{S} \tag{1b}$$

where $\mathcal{S} := \{1, \ldots, n\}$ is the index set of the n processes or sensors, $x_i(k) \in \mathbb{R}^{n_i}$ is the state of the ith process at time k and $y_i(k) \in \mathbb{R}^{m_i}$ is the measurement obtained by the ith sensor at time k. For shorthand denote s_i as the ith sensor. The system noise $\omega_i(k)$'s, the measurement noise $\nu_i(k)$'s and the initial system state $x_i(0)$ are mutually independent zero-mean Gaussian random variables with covariances $Q_i > 0$, $R_i > 0$, and $\Pi_i \ge 0$, respectively. Assume that (A_i, C_i) is

detectable. Furthermore, we assume that A_i is unstable like [12] for two reasons: (1) unstable systems bring out stability issues rather than stable ones do; (2) most process estimation tasks will become unpredictable if left unattended too long.

Each sensor measures its corresponding system state and generates a local estimate first. More specifically, at each time s_i locally runs a Kalman filter to compute the minimum mean square error (MMSE) estimate of $x_i(k)$, *i.e.*, $\hat{x}_{i,\mathrm{local}}(k) := \mathbb{E}[x_i(k)|\{y_i(l)\}_{l=0}^k]$. The corresponding estimation error and error covariance are

$$\epsilon_{i,\mathrm{local}}(k) := x_i(k) - \hat{x}_{i,\mathrm{local}}(k), \tag{2}$$
$$P_{i,\mathrm{local}}(k) := \mathbb{E}\left[\epsilon_{i,\mathrm{local}}(k)(\epsilon_{i,\mathrm{local}}(k))^\top | \{y_i(l)\}_{l=0}^k\right]. \tag{3}$$

The quantities $\hat{x}_{i,\mathrm{local}}(k)$ and $P_{i,\mathrm{local}}(k)$ can be obtained through a standard Kalman filter [15]. It is well known that $P_{i,\mathrm{local}}(k)$ converges exponentially fast to $\overline{P}_i \in \mathbb{S}^{n_i}_+$ which is the solution of a discrete algebraic Riccati equation [15]. Since we consider a problem over the infinite horizon in the sequel, we ignore the transient behaviour and make a standing assumption that $P_{i,\mathrm{local}}(k) = \overline{P}_i, \forall k \in \mathbb{Z}_+$.

For the purpose of state estimation, all sensors transmit their own local estimates to the remote estimator over a shared channel. In this work we consider a bandwidth-limited sensor network which allows one sensor to access the channel each time instant. In other words, *only a single sensor can send its estimate over the shared channel at each time instant while the others still keep their local copies*.

B. Scheduling Problem

The problem of interest is *how to efficiently schedule the transmission of sensors at each time in terms of some performance metric*.

We first define some mathematical notations. Denote the transmission indicator for s_i at time k as a binary variable $\gamma_i(k)$, *i.e.*, $\gamma_i(k) = 1$ means s_i sends data and vice versa. A sensor schedule θ is defined as a sequence of $\gamma_i(k)$, *i.e.*,

$$\theta := \{\gamma_i(k)\}_{i \in \mathcal{S}, k \in \mathbb{Z}_+}.$$

Moreover, we define a collection of the information sets $\mathcal{I}_{i,\theta}(k)$'s for $i \in \mathcal{S}$ at the estimator side as

$$\mathcal{I}_{i,\theta}(k) := \{\gamma_i(0)\hat{x}_{i,\mathrm{local}}(0), \ldots, \gamma_i(k)\hat{x}_{i,\mathrm{local}}(k)\}. \tag{4}$$

Due to the mutual independency among the systems, the estimator computes the estimate of $x_i(k)$ as $\hat{x}_i(k) := \mathbb{E}[x_i(k)|\mathcal{I}_{i,\theta}(k)]$ with the corresponding error covariance $P_i(k) := \mathbb{E}[(x_i(k) - \hat{x}_i(k))(x_i(k) - \hat{x}_i(k))^\top | \mathcal{I}_{i,\theta}(k)]$. Note that $\gamma_i(k), \hat{x}_i(k)$ and $P_i(k)$ are functions of θ though we do not explicitly show that in the notations.

Similar to [12], we use the overall average expected estimation error covariance as a performance metric, *i.e.*,

$$J(\theta) := \limsup_{T \to \infty} \frac{1}{T} \sum_{i \in \mathcal{S}} \mathrm{Tr}\left[\sum_{k=0}^{T-1} (\mathbb{E}[P_i(k)])\right]. \tag{5}$$

Fig. 1: Illustration of the awaiting phase and transmission phase for three sensors between time k and $k + 1$.

Let $J_i(\theta)$ denote the individual expected estimation error covariance correspondingly. The problem of interest is formally stated as:

$$\underset{\theta}{minimize} \quad J(\theta) \quad subject\ to \quad \sum_{i \in \mathcal{S}} \gamma_i(k) = 1. \quad (6)$$

Time-based schedules refer to that θ is a function of time only. In that case, $\gamma_i(k)$ is independent of $\hat{x}_{i,\text{local}}(k)$. The optimal scheduling solution turns out to be a periodic TDMA-like schedule like [12]. The main advantage of this type of schedule is its simplicity. From (4), however, we can see that when $\gamma_i(k) = 0$, $\mathcal{I}_{i,\theta}(k)$ contains no information on $\hat{x}_{i,\text{local}}(k)$ and thus the estimator gains nothing about the system state from s_i at time k when $\gamma_i(k) = 0$.

To outperform the time-based schedule, we study a type of event-based schedule meaning that $\gamma_i(k)$ is a function of the real-time estimates or measurements which contain information on the underlying system state. Both the sensor and the estimator know how $\gamma_i(k)$ at each time is determined based on some triggering rules. Hence, even with $\gamma_i(k) = 0$, the estimator can still infer some information about the system. Therefore, the event-triggering mechanism improves the estimation performance by providing extra information.

C. Stochastic Event-based Sensor Schedule

Mimicking the protocol of Carrier Sense Multiple Access with Collision Avoidance (CSMA/CA) widely used in wireless sensor networks, we propose a distributed event-based sensor schedule to enhance the overall estimation performance compared with the time-based schedules. Now we introduce the infrastructure. There are two phases for each sensor in each transmission frame: the awaiting phase and the transmission phase. Since the packet transmission time is often much larger than the uncertainties in communication networks such as the propagation delays, it is safe to assume the transmission time is much larger than the awaiting time during each epoch. Any sensor listens to the channel carrier for a short period before it sends anything. If the channel is occupied during the awaiting phase, then the sensor holds its data; otherwise, it sends the packet to the estimator in the transmission phase. The ends of awaiting phase of all sensors are made different in order to avoid collision (see Fig. 1). In other words, the sensors form a queue $q(k)$ with different queueing time in the awaiting phase. For example, the queue $(s_1, \ldots, s_n)$ means s_i has higher priority to access the channel than s_{i+1}. The queue $q(k)$ can be time varying or invariant and denote $\mathcal{Q}$ to be the set of possible priority queues, *i.e.*, all permutations of $s_1, \ldots, s_n$.

The idea of event-based scheduling behind is as follows. If the data of s_i contains little innovative information, which can be checked by some criteria introduced later, then s_i will be unlikely to transmit the data and s_{i+1} in the queue can take the chance to use the channel. The queue implemented on the top of carrier sensing is useful here to resolve the conflict when more than one sensor want to transmit. In other words, the transmission decision of s_i depends on the importance of local data and the channel accessibility.

Let us first define two binary indicators before formally proposing the schedule: the data importance indicator $\eta_i(k)$ and the channel accessibility indicator $\mu_i(k)$. Various event-triggering criteria for determining the importance of a single measurement have been investigated in [1], [5], [9]. In the spirit of [5], [9], we design a stochastic event-triggering rule to check the importance of the data. The dominant feature of the stochastic event-triggering mechanism is to maintain the Gaussianity of the estimation process and bypass the nonlinear problem such as [1]. We define $\eta_i(k)$ based on *the difference between the MMSE estimate under θ at local sensors and the prediction at the estimator side*, *i.e.*,

$$\epsilon_i(k) := \hat{x}_{i,\text{local}}(k) - A_i \hat{x}_i(k-1),$$

and the corresponding covariance as

$$\Sigma_i(k) := \mathbb{E}\left[(\epsilon_i(k))^\top \epsilon_i(k) | \mathcal{I}_{i,\theta}(k-1) \right]. \quad (7)$$

Define the mapping $\phi_i(z, \Pi) : \mathbb{R}^{n_i} \times \mathbb{S}_+^{n_i} \mapsto [0, 1]$ as

$$\phi_i(z, \Pi) := \exp\left(-\frac{1}{2} z^\top \Pi^{-1} z \right)^1.$$

Now we define the data importance indicator $\eta_i(k)$ as follows:

$$\eta_i(k) := \begin{cases} 1, & \text{if } \xi_i(k) > \phi_i(\epsilon_i(k), \alpha_i(k)\Sigma_i(k)), \\ 0, & \text{otherwise.} \end{cases} \quad (8)$$

where $\xi_i(k) \sim U[0,1]$ is an i.i.d. auxiliary random variable, and $\alpha_i(k)$ is a tunable parameter which reflects the importance of the data packet. For smaller $\alpha_i(k)$, $\eta_i(k)$ is more likely to be 1. Namely, for a more important sensor, we would like to pick a smaller $\alpha_i(k)$ to ensure it can transmit more data.

The tuning parameter $\{\alpha_i(k) \geq 0\}_{i \in \mathcal{S}, k \in \mathbb{Z}_+}$ and the queue parameter $\{q(k)\}_{k \in \mathbb{Z}_+}$, depending on the historical arrival pattern of all sensors, are designed and broadcasted by the central estimator. The estimator usually has advantageous resources compared with sensors, *i.e.*, stronger computation capability and larger energy storage. The design procedure will be discussed in Section IV.

Let $\mu_i(k)$ denote the channel accessibility indicator [2], *i.e.*,

$$\mu_i(k) := \begin{cases} 1, & \text{if } \sum_{j=1}^{i-1} \gamma_j(k) = 0 \\ 0, & \text{otherwise} \end{cases}. \quad (9)$$

Now we are ready to propose an event-based schedule $\theta_+ := \{\gamma_i(k)\}_{i \in \mathcal{S}, k \in \mathbb{Z}_+}$ and denote Θ_+ to be the set of

---

[1]We abuse the notation X^{-1} to represent the inverse of $X \in \mathbb{S}_{++}^{n_i}$ or the Moore-Penrose pseudo-inverse of a singular $X \in \mathbb{S}_+^{n_i}$. Similarly, we use $\det(X)$ to denote the determinant of a square matrix X or its pseudo-determinant if X is singular.

[2]For concise presentation, by relabelling the sensors we assume the queue $q(k) = (s_1, \ldots, s_n)$ by default. More discussion on the design of the optimal queue is shown in Section IV.

all event-based schedules in (10) subject to all possible $\{\alpha_i(k)\}_{i\in\mathcal{S},k\in\mathbb{Z}_+}$ and $\{q(k)\}_{k\in\mathbb{Z}_+}$:

$$\gamma_i(k) := \begin{cases} \mu_i(k)\eta_i(k), & \text{if } i \neq n \\ \mu_i(k), & \text{otherwise} \end{cases}. \qquad (10)$$

Notice that $\gamma_i(k)$ depends on $\epsilon_i(k)$ and $\Sigma_i(k)$, and $\epsilon_i(k+1)$ and $\Sigma_i(k+1)$ depends on $\gamma_i(k)$. Thus θ_+ cannot be determined offline. Also note that if s_n with the last priority detects that the channel is idle, it sends the data without checking (8). The condition (8) implies that if the prediction error is small, $\eta_i(k)$ is very likely to be 0 which prevents unecessary transmission.

In the subsequent sections, we analyze the estimation performance using such a schedule and design the parameters $\{\alpha_i(k) \geq 0\}_{i\in\mathcal{S},k\in\mathbb{Z}_+}$ and the queue parameters $\{q(k)\}_{k\in\mathbb{Z}_+}$ optimally by formulating optimization problems.

III. OPTIMAL FILTERING

After proposing the stochastic event-based schedule θ_+, we are keen in deriving the MMSE estimator under θ_+ and quantifying the estimation performance of θ_+. Moreover, from the viewpoint of communication, we present the transmission probability of each sensor.

To facilitate derivations, we define the following operators for $X \in \mathbb{S}_+^{n_i}$ and $\alpha \in \mathbb{R}$:

$$h_i(X) := A_i X A_i^\top + Q_i, \qquad (11)$$

$$g_i(X, \alpha) := \frac{\alpha}{1+\alpha}(A_i X A_i^\top + h_i(\overline{P}_i) - \overline{P}_i), \qquad (12)$$

$$t_i(X, \alpha) := \frac{1}{1+\alpha}\overline{P}_i + \frac{\alpha}{1+\alpha}h_i(X). \qquad (13)$$

The following lemma on the properties of the innovation and estimation error is useful for proving the main result whose partial proof can be found in [16]. Let the incremental innovation for s_i be denoted as

$$\delta_i(k) := \hat{x}_{i,\text{local}}(k) - A_i\hat{x}_{i,\text{local}}(k-1). \qquad (14)$$

Lemma 1 *Given $\epsilon_{i,\text{local}}(k), \delta_i(k)$ defined in (2) and (14), the following statistical properties hold:*

(i) $\delta_i(k)$ is zero mean Gaussian distributed, i.e., $\delta_i(k) \sim \mathcal{N}(0, h_i(\overline{P}_i) - \overline{P}_i)$.
(ii) $\mathbb{E}\left[\delta_i(k)(\delta_i(j))^\top\right] = 0$ for any $k \neq j$.
(iii) $\mathbb{E}\left[\epsilon_{i,\text{local}}(k)(\delta_i(k_0))^\top\right] = 0$ for any $k_0 \leq k$.
(iv) $\mathbb{E}\left[\hat{x}_{i,\text{local}}(k)(\epsilon_{i,\text{local}}(k_1))^\top\right] = 0$ for any $k_1 \geq k$ and $\mathbb{E}\left[\hat{x}_{i,\text{local}}(k)(\delta_i(k_2))^\top\right] = 0$ for any $k_2 > k$.

The following lemma on Bayes' inference is presented before showing the main result.

Lemma 2 *Suppose $z \in \mathbb{R}^{n_i}$ is a Gaussian random variable with $z \sim \mathcal{N}(0, Z)$ with respect to the Lebesgue measure on $\mathcal{Z} := \{Z^{1/2}x : x \in \mathbb{R}^{n_i}\}$ and ξ is uniformly distributed over $[0, 1]$. The following statements hold:*

(i) The occurring probability of the following event is $\Pr\{\xi \leq \phi_i(z, \Pi)\} = \det(I + Z\Pi^{-1})^{-1/2}$.
(ii) The conditional pdf of z is $f(z|\xi \leq \phi_i(z, \Pi)) \sim \mathcal{N}\left(0, (Z^{-1} + \Pi^{-1})^{-1}\right)$.

Proof: First we have

$$\Pr\{\xi \leq \phi_i(z, \Pi)\} = \mathbb{E}\left[\exp\left(-\frac{1}{2}z^\top\Pi^{-1}z\right)\right]$$

$$= \int_{\mathcal{Z}} \frac{(2\pi)^{-r/2}}{\det(Z)^{1/2}}\exp\left(-\frac{1}{2}z^\top Z z\right)\exp\left(-\frac{1}{2}z^\top\Pi^{-1}z\right)dz$$

$$= \det(I + Z\Pi^{-1})^{-\frac{1}{2}}, \qquad (15)$$

where $r = \text{rank}(Z)$. Statement (ii) is a direct result of the Bayes' theorem. ∎

For notational simplification, denote $\tilde{\gamma}_i(k) := 1 - \gamma_i(k)$, $\tilde{\mu}_i(k) := 1 - \mu_i(k)$. Furthermore, denote the leave duration of s_i as $\tau_i(k) := \min\{k - k_0 : \gamma_i(k_0) = 1, k_0 \leq k\}$. Now we are ready to present the main result.

Theorem 1 *Under the proposed schedule θ_+ given in (10), the MMSE state estimate for each process is given by*

$$\hat{x}_i(k) = \begin{cases} \hat{x}_{i,\text{local}}(k), & \text{if } \gamma_i(k) = 1 \\ A_i\hat{x}_i(k-1), & \text{if } \gamma_i(k) = 0 \end{cases}, \qquad (16)$$

and the corresponding estimation error covariance is

$$P_i(k) = \begin{cases} \overline{P}_i, & \text{if } \gamma_i(k) = 1 \\ h_i(P_i(k-1)), & \text{if } \mu_i(k) = 0 \\ t_i(P_i(k-1), \alpha_i(k)), & \text{otherwise} \end{cases}. \qquad (17)$$

Moreover, $\Sigma_i(k)$ in (7) is given by

$$\Sigma_i(k) = h_i(P_i(k-1)) - \overline{P}_i. \qquad (18)$$

Proof: When $\gamma_i(k) = 1$, from [15] we know that $\mathbb{E}[x_i(k)|\hat{x}_{i,\text{local}}(k), \mathcal{I}_{i,\theta}(k-1)] = \hat{x}_{i,\text{local}}(k)$, and $P_i(k) = \overline{P}_i$.

When $\mu_i(k) = 0$ ($\gamma_i(k)$ must be 0) which implies the channel is occupied by another sensor, the estimator can only do prediction on the estimate of $x_i(k)$, *i.e.*

$$\hat{x}_i(k) = A_i\hat{x}_i(k-1), \quad P_i(k) = h_i(P_i(k-1)). \qquad (19)$$

The two cases above are easy to analyze. Next we consider the remaining case, *i.e.*, when $\mu_i(k)\tilde{\gamma}_i(k) = 1$. The following equation is easy to verify and useful for subsequent derivations.

$$x_i(k) = A_i\hat{x}_i(k-1) + \epsilon_{i,\text{local}}(k) + \epsilon_i(k). \qquad (20)$$

Without loss of generality, assume that $\gamma_i(k) = 1$ or $\mu_i(j) = 0$ for any $j \in [k-\tau_i(k), k]$, and $\mu_i(k+1)\tilde{\gamma}_i(k+1) = 1$ occurs at time $k + 1$. Moreover, the following recursive equation always holds:

$$\epsilon_i(k) = A_i\epsilon_i(k-1) + \delta_i(k). \qquad (21)$$

Also from (19) and (20) we have $x_i(k + 1) = A_i^{\tau_i(k+1)}\hat{x}_{i,\text{local}}(k - \tau_i(k + 1)) + \epsilon_{i,\text{local}}(k + 1) + \epsilon_i(k + 1) = A_i^{\tau_i(k+1)}\hat{x}_{i,\text{local}}(k - \tau_i(k + 1)) + \epsilon_{i,\text{local}}(k + 1) + \sum_{j=0}^{\tau_i(k)} A_i^j\delta_i(k + 1 - j)$, where the last two terms on the RHS are mutually independent in view of Lemma 1(iii). Due to Lemma 1(i) and the fact [15] that $f(\epsilon_{i,\text{local}}(k + 1)|\mathcal{I}_{i,\theta}(k + 1)) = f(\epsilon_{i,\text{local}}(k + 1)) \sim \mathcal{N}(0, \overline{P}_i)$, we have $\mathbb{E}[x_i(k + 1)|\mathcal{I}_{i,\theta}(k + 1)] = A_i\hat{x}_i(k)$. Therefore, from (20) we know that for $j \in [k - \tau_i(k), k + 1]$ the equality holds:

$$P_i(k) = \overline{P}_i + \Lambda_i(k), \qquad (22)$$

where $\Lambda_i(k) := \mathbb{E}\left[(\epsilon_i(k))^\top \epsilon_i(k) | \mathcal{I}_{i,\theta}(k)\right]$. Hence we have $\Lambda_i(k) = P_i(k) - \overline{P}_i$. Then together with (21) we can conclude that

$$f(\epsilon_i(k+1)|\mathcal{I}_{i,\theta}(k)) \sim \mathcal{N}(0, A_i(P_i(k) - \overline{P}_i)A_i^\top + h_i(\overline{P}_i) - \overline{P}_i)$$

which proves (18). Then from Lemma 2, we have

$$f(\epsilon_i(k+1)|\mathcal{I}_{i,\theta}(k+1)) \sim \mathcal{N}(0, g_i(P_i(k) - \overline{P}_i, \alpha_i(k))).$$

Thus, from (22) we have

$$\begin{aligned} f(x_i(k+1)|\mathcal{I}_{i,\theta}(k+1)) \\ \sim \mathcal{N}(A_i\hat{x}_i(k), \overline{P}_i + g_i(P_i(k) - \overline{P}_i, \alpha_i(k))). \end{aligned} \quad (23)$$

Notice that $\hat{x}_i(k+1)$ is actually $A_i^{\tau_i(k+1)}\hat{x}_{i,\text{local}}(k - \tau_i(k+1))$ which is the same with the predicted estimate but with a smaller error covariance. Consequently, no matter what $\mu_i(k+2)$ and $\gamma_i(k+2)$ are at time $k+2$, the mutual independence of the three terms in (20) and the recursion in (21) hold. Therefore, the conditional pdf of $\epsilon_i(k+2)$ can be computed, which is zero mean Gaussian distributed. Thus $\hat{x}_i(k+2)$ is still a predicted estimate, *i.e.*, $A_i\hat{x}_i(k+1)$, with the corresponding error covariance dependent on $\mu_i(k+2)$ and $\gamma_i(k+2)$. Recursively, we can verify (16) and (17) which completes the proof. ∎

The result in (17) shows that if s_i is idle due to other competitors, *i.e.*, $\mu_i(k) = 0$, then the estimator simply runs a prediction. If s_i decides to hold its packet even if it has the access to the channel, *i.e.*, $\mu_i(k)\tilde{\gamma}_i(k) = 1$, the estimator updates the estimate using the information encoded by (8). Using the following lemma, we can see the estimation performance under $\mu_i(k)\tilde{\gamma}_i(k) = 1$ is lower bounded by that under $\gamma_i(k) = 1$ and upper bounded by that under $\mu_i(k) = 0$.

Lemma 3 *The following statements hold for any $i \in \mathcal{S}$:*
(i). For any $l_1, l_2 \in \mathbb{Z}_+$ with $l_1 < l_2$, $h_i^{l_1}(\overline{P}_i) \leq h_i^{l_2}(\overline{P}_i)$.
(ii). For any $l \in \mathbb{Z}_+$, $\mathrm{Tr}\left[\overline{P}_i\right] < \mathrm{Tr}\left[h_i(\overline{P}_i)\right] \leq \cdots \leq \mathrm{Tr}\left[h_i^l(\overline{P}_i)\right]$.
(iii). For any X which is a convex combination of $\{h_i^j(\overline{P}_i)\}_{j=0}^\infty$ and any $\alpha \geq 0$, $X \leq t_i(X, \alpha)$ and $\mathrm{Tr}\left[t_i^{l_1}(X, \alpha)\right] \leq \mathrm{Tr}\left[t_i^{l_2}(X, \alpha)\right]$, for $l_1 < l_2$.

Proof: From [17, Lemma A.1], we can prove that $\overline{P}_i \leq h_i(\overline{P}_i)$. Since $h_i(\cdot)$ is affine, we can take $h_i(\cdot)$ on both sides and draw the conclusion (i). Next we show the strictiveness in (ii), *i.e.*, $\mathrm{Tr}\left[\overline{P}_i\right] < \mathrm{Tr}\left[h_i(\overline{P}_i)\right]$. Assuming $\overline{P}_i = h_i(\overline{P}_i)$, then we can find that $Q = 0$ which contradicts the assumption $Q > 0$, or $C_i = 0$ which contradicts the fact (A_i, C_i) is detectable. Thus $\overline{P}_i \neq h_i(\overline{P}_i)$. From (i) we immediately have $\mathrm{Tr}\left[\overline{P}_i\right] \leq \mathrm{Tr}\left[h_i(\overline{P}_i)\right] \leq \cdots \leq \mathrm{Tr}\left[h_i^l(\overline{P}_i)\right]$. Moreover, we know that

$$\left\{ \begin{array}{l} \overline{P}_i \neq h_i(\overline{P}_i) \\ \overline{P}_i \leq h_i(\overline{P}_i), \overline{P}_i \in \mathbb{S}_+^{n_i} \end{array} \right. \Rightarrow \mathrm{Tr}\left[\overline{P}_i\right] \neq \mathrm{Tr}\left[h_i(\overline{P}_i)\right].$$

The last statement is a direct result of (iii). ∎

From Lemma 3(iii), for any realization of $P_i(k-1)$ we know that $\overline{P}_i \leq t_i(P_i(k-1), \alpha_i(k)) \leq h_i(P_i(k-1))$. In other words, compared with the time-based schedules, even if

s_i does not transmit anything, the estimator is still likely to obtain an estimate better than a pure prediction.

Denote as $r_i(k)$ the rank of $\Sigma_i(k)$ which can be computed offline from (18). The probability of kinds of events during transmission can be computed as follows. For concise notation, denote

$$\widehat{\alpha}_i(k) := \alpha_i(k)/(1 + \alpha_i(k)) \quad (24)$$
$$\beta_i(k) := \widehat{\alpha}_i(k)^{r_i(k)/2}, \quad \tilde{\beta}_i(k) := 1 - \beta_i(k). \quad (25)$$

Theorem 2 *The following statements hold true:*

$$\Pr\{\eta_i(k) = 1\} = \tilde{\beta}_i(k),$$

$$\Pr\{\mu_i(k) = 1 | \Xi_i(k)\} = \left\{ \begin{array}{ll} 1, & \text{if } i = 1 \\ \prod_{j=1}^{i-1} \beta_j(k), & \text{if } i > 1 \end{array} \right.,$$

$$\Pr\{\gamma_i(k) = 1 | \Xi_i(k)\} = \left\{ \begin{array}{ll} \prod_{j=1}^{n-1} \beta_j(k), & \text{if } i = n \\ \tilde{\beta}_i(k)\prod_{j=1}^{i-1} \beta_j(k), & \text{if } i < n \end{array} \right.,$$

where $\Xi_i(k) := \{\eta_{i-1}(k), \dots, \eta_1(k)\}$.

Proof: From Lemma 2(i), we have $\Pr\{\xi_i(k) < \phi_i(\epsilon_i(k), \alpha_i(k)\Sigma_i(k))\} = [\alpha_i(k)/(1 + \alpha_i(k))]^{r_i(k)/2}$. For (ii), it is easy to see that $\Pr\{\mu_1(k) = 1 | \Xi_i(k)\} = 0$ for s_1. From (9) and (10), we know that $\Pr\{\mu_i(k) = 1 | \Xi_i(k)\} = 1 - \sum_{j=1}^{i-1}\Pr\{\gamma_j(k) = 1 | \Xi_i(k)\}$ and $\Pr\{\gamma_i(k) = 1 | \Xi_i(k)\} = (1 - \Pr\{\eta_i(k) = 0\} | \Xi_i(k))\Pr\{\mu_i(k) = 1 | \Xi_i(k)\}$. After some calculation, omitted for longevity, we obtain (ii) and (iii). ∎

Remark 1 *Since $\Sigma_i(k)$ in (18) can be written as $\sum_{j\geq 1}\kappa_j h_i^j(\overline{P}_i) - \overline{P}_i$ where $\sum_{j\geq 1}\kappa_j = 1$, we can guarantee $\Sigma_i(k)$ to be full rank as long as $h_i^j(\overline{P}_i) - \overline{P}_i > 0, \forall j \geq 1$ holds. A sufficient condition is that A_i has full rank.*

Since we have known how the error covariance of each process is updated under different conditions of $\{\gamma_i(k), \mu_i(k), \eta_i(k)\}$ and the probability of each outcome in Theorem 2, we can next investigate how to optimally set the event-triggers for each sensor.

IV. PARAMETER DESIGN

In this section, we aim to design the optimal parameters for the class of stochastic event-based schedules based on the results in Section III. It turns out that finding the optimal schedule is computationally challenging. Thus, we propose a greedy algorithm to achieve suboptimal schedules. Moreover, we analytically quantify the performance gap between the optimal and suboptimal schedules.

We first show that the proposed schedule performs at least as good as any time-based schedule.

Theorem 3 *There exists a nonempty subset of Θ_+ from which θ_+ is at least as good as any time-based schedule $\theta_\ddagger$ does, i.e., $\{\theta_+ \in \Theta_+ : J(\theta_+) \leq J(\theta_\ddagger), \forall \theta_\ddagger\} \neq \emptyset$.*

The proof is straightforward since any time-based schedule is a special case of θ_+ in (10). For example, if s_i is deterministically scheduled to transmit at time k, then one can set

$\alpha_i(k) = 0$ and $\alpha_j(k) = \infty$ for $j \neq i$. Thus by optimizing the parameters in the schedule θ_+, one can always find a schedule in Θ_+ performs as well as any time-based schedule can.

A. Optimal Parameter Design by Solving an MDP Problem

Now we are keen on the optimal $\theta_+ \in \Theta_+$ of minimizing (5). By formulating an Markov decision process (MDP) problem with average cost criterion, we can design the optimal time-varying parameters $\{\alpha_i(k)\}_{i\in\mathcal{S},k\in\mathbb{Z}_+}$ and the queue $\{q(k)\}_{k\in\mathbb{Z}_+}$.

First define the state space $\mathcal{V}$ to be the set of all possible $(P_1(k),\ldots,P_n(k))$, where $P_i(k)$ belongs to $\mathcal{C}_i$ which is the set of any convex combination of $h_i^j(\overline{P}_i),\forall j$, i.e., $\mathcal{C}_i := \{\sum_j \lambda_j h_i^j(\overline{P}_i):\sum_j \lambda_j = 1\}$. The action is a tuple of $b := (u,q) \in \mathcal{U} \times \mathcal{Q}$ where $\mathcal{U} := \{u \in \mathbb{R}^n : 0 \leq u[i] \leq 1, \forall i \in \mathcal{S}\}$. The compact action space is denoted as $\mathcal{B}$, which is identical for each state in $\mathcal{V}$. Denote the set $\mathcal{K} := \{(v,b) : v \in \mathcal{V}, b \in \mathcal{B}\}$ which is a Borel space. The transition law $Q(\cdot|v,b)$ with $(v,b) \in \mathcal{K}$ is a stochastic kernel on $\mathcal{V}$ given $\mathcal{K}$, which can be obtained from Theorem 1 and Theorem 2. The cost function $c(v,b)$ is thus defined as the sum of the trace of each matrix in v. We thus model an MDP denoted as $\Omega := (\mathcal{V}, \mathcal{B}, Q(\cdot|\cdot,\cdot), c(\cdot,\cdot))$. A decision at time k is a mapping $d(k) : \mathcal{V} \mapsto \mathcal{B}$. A policy ζ for Ω is a sequence of decision rules. We define the *average expected cost* of the policy ζ per unit time by

$$g_\zeta(v) = \limsup_{T\to\infty} \frac{1}{T}\mathbb{E}_{\zeta,v}\left[\sum_{k=0}^{T-1} r(v(k),b(k))\right], \quad (26)$$

where $v \in \mathcal{V}$ is the initial state, and the expectation is taken based on the stochastic process $\{v(k),b(k)\}$ uniquely determined by the policy ζ and v. The target for an MDP problem with average cost criterion is to search an optimal stationary policy ζ^* for Ω such that the average cost is minimized, i.e., $g_{\zeta^*}(v) \leq g_\zeta(v)$, $\forall v \in \mathcal{V}$. It is easy to see that finding the optimal policy ζ^* for Ω is equivalent to searching the optimal $\theta_+ \in \Theta_+$ of minimizing (5).

Numerous literature have studied the optimality conditions for a policy of an MDP problem with average cost criterion in Borel spaces such as [18]. Though some researchers have attempted to solve the MDP problem in Borel spaces like [19], obtaining the optimal cost and the optimal policy in a computationally efficient fashion, however, is generally chanllenging. The technique of discretizing the state space and rebuilding the stochastic kernel from the original model is a popular approach to approximately solve the MDP with Borel spaces. Unfortunately, for solving the discrete MDP, the computational burden is high. For example, the classical policy iteration algorithm for solving an MDP requires the computational effort of $|\mathcal{V}_d||\mathcal{B}_d|^2 + \frac{1}{3}|\mathcal{V}_d|$ multiplications/iterations [20, Chapter 8], where $|\mathcal{V}_d|$ represents the cardinality of the discretized state set and $|\mathcal{B}_d|$ the cardinality of the discretized action set. The number of discretized state space is growing at least exponentially with the number of the sensors in order to keep the same grid size.

B. Greedy Algorithm

The optimal stochastic schedule is formidable to obtain in practice. Instead, we propose a suboptimal greedy schedule which minimizes the next-step trace of total expected error covariance. At each time k, the estimator broadcasts the priority queue $q(k)$ and the event-triggering parameters $\{\alpha_i(k)\}_{i\in\mathcal{S}}$ to all sensors which minimizes the one-step cost function:

$$J_{greedy}(k) = \text{Tr}\left[\sum_{i\in\mathcal{S}} \mathbb{E}\left[P_i(k)\right]\right]. \quad (27)$$

Unlike searching the optimal priority queue for the optimal schedule via enumeration, we find a simple rule for the optimal priority queue for the greedy schedule if $r_i(k) = r_j(k), \forall i \neq j$. An example is that the system matrices of several identical vehicles have full rank, i.e., $\text{rank}(A_i) = \text{rank}(A_j) = n_i = n_j, \forall i,j \in \mathcal{S}$.

Theorem 4 *If $r_i(k) = r_j(k)$ for any $i,j \in \mathcal{S}$, the queue $q(k) = \{s_{\varphi_1},\ldots,s_{\varphi_m},\ldots,s_{\varphi_n}\}$, where $\varphi_m \in \mathcal{S}$, is optimal for the greedy schedule if and only if for each $m < n$ the following inequality holds:*

$$\text{Tr}\left[h_{\varphi_m}(P_{\varphi_m}(k-1)) - \overline{P}_{\varphi_m}\right]$$
$$\geq \text{Tr}\left[h_{\varphi_{m+1}}(P_{\varphi_{m+1}}(k-1)) - \overline{P}_{\varphi_{m+1}}\right]. \quad (28)$$

The proof is omitted for longevity. Please refer to [21].

After determining the optimal $q(k)$, we can optimize $J_{greedy}(k)$ in (27) in terms of $\{\alpha_i(k)\}_{i\in\mathcal{S}}$. Based on Theorem 1 and 2, we can explicitly obtain a cost function which turns out to be a signomial of $\{\alpha_i(k)\}_{i\in\mathcal{S}}$. The signomial programming (SP) problem is widely studied in communication society [22]. For example, we exploit an efficient algorithm for searching the global optimum proposed in [23]. The computational time is affected by the number of variables and number of negative monomials (see [23, Table 4]). For our case, if we consider 4-sensor schedule, the algorithm will converge in around 30 iterations[3].

C. Optimality Gap

Notice that the performance gap between the suboptimal schedules and the optimal $\theta_+ \in \Theta_+$ is upperbounded by the gap between the suboptimal performance and a lower bound for the optimal performance. Next we discuss the upper bound of such performance gap to quantify the performance of any suboptimal stochastic schedule.

First we construct an artificial schedule whose performance is better than the optimal $\theta_+ \in \Theta_+$. We relax the original constraint $\sum_i \gamma_i(k) = 1$ by requiring the sum of the rates of all sensors to be less than 1. In other words, we allow that channel can be used by multiple sensors but their sum communication rate must be less than 1. Denote $\Theta_\ddagger$ as the set of all schedules in (10) with forcing $\mu_i(k)$ to be 1 for $i \leq n-1$ and all k. In other words, any absence of $\hat{x}_{i,\text{local}}(k)$ implies (8) for $i \leq n-1$. Moreover, we also implement the

---

[3]Each iteration involves solving one or two convex optimization problems with the same set of variables.

event-triggering mechanism for s_n. By solving the following optimization problem:

Problem 1

$$\underset{\theta \in \Theta_{\ddagger}}{minimize} \quad J(\theta), \quad subject\ to \quad \sum_{i \in \mathcal{S}} \mathbb{E}\left[\gamma_i(k)\right] \leq 1, \quad (29)$$

it is not hard to see the solution to Problem 1 is better than the optimal $\theta_+ \in \Theta_+$ since any absence of each sensor now conveys some information to the estimator. While for $\theta_+ \in \Theta_+$ the information is lost for the sensors whose priority is behind the designated one. Also note that the queue is useless for $\theta \in \Theta_{\ddagger}$.

Now $\alpha_i(k)$ solely depends on $\tau_i(k)$ and so does $\beta_i(k)$. From (17) we know that $J_i(\theta)$ is determined by the underlying stochastic process $\{\tau_i(k)\}_{k=1}^{\infty}$. With a little abuse of notations, we use β_i^j to denote the probability of $\Pr\{\eta_i(k) = 0\}$ for sensor i with $\tau_i(k) = j$ and α_i^j accordingly for $i \leq n$.

We can quantify the performance of a suboptimal schedule by using the following result.

Proposition 1 *Suppose $\theta_* \in \Theta_+$ to be an optimal solution of minimizing (5). For any schedule $\theta_\perp \in \Theta_+$ (including any time-based schedules), define the optimality gap to be $\Delta(\theta_\perp) := J(\theta_\perp) - J(\theta_*)$. The following inequality holds:*

$$\Delta(\theta_\perp) \leq J(\theta_\#) - J(\theta_*).$$

The schedule $\theta_\#$ is characterized by $\{\beta_i^\ell\}_{i \in \mathcal{S}, \ell \in \mathbb{Z}_+}$ which is the solution to the following problem:

$$\underset{\pi_i^0, \beta_i^\ell \in \mathbb{R}}{minimize} \quad \sum_{i \in \mathcal{S}} \pi_i^0 \mathrm{Tr}\left[P_i^0\right] + \sum_{j=0}^{\infty} \pi_i^0 (\prod_{\ell=0}^{j} \beta_i^\ell) \mathrm{Tr}\left[P_i^{j+1}\right],$$

$$subject\ to \quad \pi_i^0 + \pi_i^0 \sum_{j=0}^{\infty} \prod_{\ell=0}^{j} \beta_i^\ell \geq 1,\ i \in \mathcal{S},$$

$$\sum_{i \in \mathcal{S}} \pi_i^0 \leq 1,$$

where

$$P_i^j := \prod_{u=0}^{j-1} (\beta_i^u)^{2/r_i(u)} h_i^j(\overline{P}_i) + \sum_{\ell=0}^{j-1} \left(1 - (\beta_i^{j-1-\ell})^{2/r_i(j-1-\ell)}\right)$$

$$\times \left(\prod_{u=0}^{\ell-1} (\beta_i^{j-1-u})^{2/r_i(j-1-\ell)}\right) h_i^\ell(\overline{P}_i), \quad \forall i \in \mathcal{S}. \quad (30)$$

Proof: In Problem 1, the transmission of each sensor is not affected by others and we can first consider such a subproblem for one sensor:

$$\underset{\theta \in \Theta_{\ddagger}}{minimize} \quad J_i(\theta), \quad subject\ to \quad \mathbb{E}\left[\gamma_i(k)\right] \leq \varrho_i. \quad (31)$$

The stochastic process $\{\tau_i(k)\}_{k=1}^{\infty}$ is a Markov chain with countable state space $\mathcal{M} := \{0, 1, 2, \ldots\}$. The transition matrix T_i is given by

$$T_i(j_2, j_1) = \begin{cases} 1 - \beta_i^{j_1}, & \text{if } j_2 = 0 \\ \beta_i^{j_1}, & \text{if } j_2 = j_1 + 1 \\ 0, & \text{otherwise} \end{cases},$$

where $T_i(j_2, j_1) := \Pr\{\tau_i(k) = j_2 | \tau_i(k) = j_1\}$. Though we can design β_i^j freely, we have to choose the set of $\{\beta_i^j\}_{j=0}^{\infty}$ to guarantee that the state $\tau_i(k) = 0$ is *recurrent*. Otherwise, the Markov chain $\{\tau_i(k)\}$ is transient and $J_i(\theta)$ will go unbounded due to Lemma 3(iii) and the fact that $\lim_{j \to \infty} \mathrm{Tr}\left[h_i^j(\overline{P}_i)\right] \to \infty$. Therefore, the stationary distribution $\pi_i = [\pi_i^0, \pi_i^1, \pi_i^2, \ldots]$ must exist [24]. Denote the corresponding estimation error covariance for $\tau_i(k)$ as $P_i^{\tau_i(k)}$. From Theorem 1 and 2(i) we have that $P_i^0 = \overline{P}_i, P_i^j = t(P_i^{j-1}, \alpha_i^{j-1})$ and thus P_i^j in (30). Now the cost function becomes $J_i(\theta) = \varrho_i \mathrm{Tr}\left[P_i^0\right] + \sum_{j=0}^{\ell_{max}-1} \varrho_i (\prod_{\ell=0}^{j} \beta_i^\ell) \mathrm{Tr}\left[P_i^{j+1}\right]$.

Note that the implicit equality constraint for Problem 1 is $\sum_j \pi_i^j = 1$. Since the cost function $J_i(\theta)$ is a decreasing function of π_i^0, we can transform it into the inequality constraint and the optimal solution must be reached only when the equality holds. Therefore, by solving the optimazition problem in the proposition we can obtain a lower bound for the optimal cost and thus the optimality gap bound. ∎

Note that the optimization problem is also a signomial programming problem whose global optimum can be numerically solved [23].

Remark 2 *For a small network, we can enumerate the optimal or suboptimal schedules for different orders and choose the best order. However, it is formidable to do this for a large network. One heuristic technique is to assign a time-invariant priority queue in parameter design according to the optimal π_i^0 in Proposition 1, i.e., letting the queue to be $\{s_{\phi_1}, \ldots, s_{\phi_\ell}, \ldots, s_{\phi_n}\}$, where $\phi_\ell \in \mathcal{S}$, with $\pi_{\phi_\ell}^0 \leq \pi_{\phi_{\ell+1}}^0$ since a larger $\pi_{\phi_\ell}^0$ usually implies that error covariance of s_{ϕ_ℓ} may grow faster than others.*

V. EXAMPLE

In this section, we use a simple example to show the superiority of the proposed event-based schedule and compare event-based schedules with different parameter designs. To compare with the optimal time-based schedule in [12] which is only applicable to the two-sensor case, we also conduct simulations over two processes with the parameters:

$$A_1 = \begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix}, C_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}^\top, Q_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, R_1 = 1 \text{ and}$$

$$A_2 = \begin{bmatrix} 1.1 & 1 \\ 0 & 1 \end{bmatrix}, C_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}^\top, Q_2 = \begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix}, R_2 = 1.$$

We compare the performance among the following schedules:

- $\theta_{offline}$: the optimal time-based schedule [12] is $\{s_2, s_1, s_1, s_2, s_1, s_1, \ldots\}$ with the period 3.
- θ_{greedy}: the suboptimal schedule via the Greedy algorithm.
- θ_{MDP}: the approximate optimal event-based schedule via MDP approach. By discretizing the continuous state and action space and rebuilding the transition law, we approximate the MDP problem by a finite state discrete MDP problem and solve it by value iteration. The number of state grids and action grids is 1000 and 100, respectively.

Schedules and LB	$\theta_{offline}$	θ_{greedy}	θ_{MDP}	LB
$J(\cdot)$	92.64	52.05	55.23	48.21

Table I: Performance comparison among different schedules. The lower bound of the optimal J is also listed as LB.
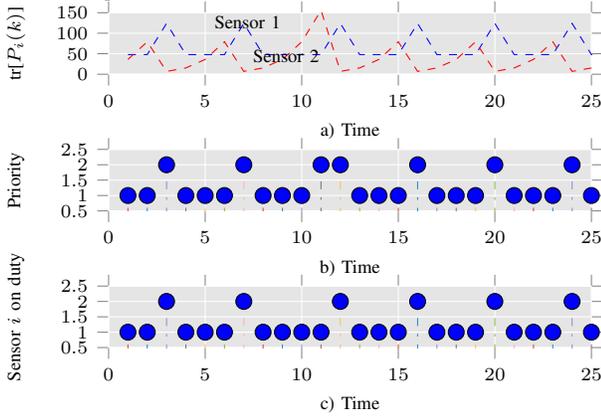
Fig. 2: Realization of the event-based schedule θ_{greedy}: a) $\mathrm{Tr}\,[P_i(k)]$ of each sensor, b) y-axis denotes the sensor index with the higher priority, c) y-axis denotes the on-duty sensor index.

We also compute the lower bound for the optimal event-based schedule according to Proposition 1.

We summarize the results in Table I. For performance of each schedule, we repeatedly run 500 experiments and take their average. The estimation performance of $\theta_{offline}$ is the worst among all. The two event-based schedules reduce the cost function $J(\theta_{offline})$ by 43.8% and 40.4%, respectively. The approximate solution to the MDP does not outperform the solutions given by the greedy algorithm in this case due to the discretization approximation. Moreover, the gap between $J(\theta_{greedy})$ and LB is small, implying that the suboptimal schedule via the greedy algorithm is a good choice with low computational complexity compared with the MDP-based optimal schedule. The computational time for θ_{greedy} at each time and the lower bound are 3.3s and 25.1s, respectively[4].

We also plot the realization of θ_{greedy} to illustrate the superiority of the event-based mechanism over the time-based one. At time $k = 11$, Sensor 2 has the first priority of using the channel in Fig. 2(b). This reflects the validity of Theorem 4 from the fact $\mathrm{Tr}\,[h_2(P_2(k))] - \overline{P}_2 > \mathrm{Tr}\,[h_1(P_1(k))] - \overline{P}_1$ seen from Fig. 2(a). In Fig. 2(c), however, Sensor 1 is scheduled to transmit its data because the event-triggering mechanism of Sensor 2 classifies its data as unimportant and leaves the channel access to Sensor 1.

VI. CONCLUSION

We have studied a stochastic sensor scheduling framework for sensor networks monitoring different processes. The sensors make transmission decisions based on both channel accessibility and self-triggering events depending on the real-time innovations. The proposed schedule which dynamically allocates the limited communication bandwidth to sensors is shown to be more efficient than the time-based schedules in terms of average estimation performance. We have also discussed the optimal event-based schedule design through an MDP approach. Since solving an MDP problem in Borel spaces is generally computationally difficult, we have also proposed a suboptimal schedule via the greedy algorithm and analyzed the optimality gap.

REFERENCES

[1] J. Wu, Q.-S. Jia, K. H. Johansson, and L. Shi, "Event-based sensor data scheduling: Trade-off between communication rate and estimation quality," *IEEE Trans. Autom. Control*, vol. 58, no. 4, pp. 1041–1046, 2013.

[2] S. Liu, M. Fardad, P. K. Varshney, and E. Masazade, "Optimal periodic sensor scheduling in networks of dynamical systems," *IEEE Trans. Signal Process.*, vol. 62, no. 12, pp. 3055–3068, 2014.

[3] L. Zhao, W. Zhang, J. Hu, A. Abate, and C. J. Tomlin, "On the optimal solutions of the infinite-horizon linear sensor scheduling problem," *IEEE Trans. Autom. Control*, vol. 59, no. 10, pp. 2825–2830, 2014.

[4] Y. Mo, E. Garone, and B. Sinopoli, "On infinite-horizon sensor scheduling," *Systems & control letters*, vol. 67, pp. 65–70, 2014.

[5] D. Han, Y. Mo, J. Wu, S. Weerakkody, B. Sinopoli, and L. Shi, "Stochastic event-triggered sensor schedule for remote state estimation," *IEEE Trans. Autom. Control*, vol. 60, no. 10, pp. 2661–2675, 2015.

[6] D. Shi, R. J. Elliott, and T. Chen, "Event-based state estimation of discrete-state hidden Markov models," *Automatica*, vol. 65, pp. 12–26, 2016.

[7] C. Yang and L. Shi, "Deterministic sensor data scheduling under limited communication resource," *IEEE Trans. Signal Process.*, vol. 59, no. 10, pp. 5050–5056, 2011.

[8] D. Shi and T. Chen, "Optimal periodic scheduling of sensor networks: A branch and bound approach," *Systems & Control Letters*, vol. 62, no. 9, pp. 732–738, 2013.

[9] S. Weerakkody, Y. Mo, B. Sinopoli, D. Han, and L. Shi, "Multi-sensor scheduling for state estimation with event-based, stochastic triggers," *IEEE Trans. Autom. Control*, DOI: 10.1109/TAC.2015.2505066, 2015.

[10] J. Song, S. Han, A. K. Mok, D. Chen, M. Lucas, and M. Nixon, "WirelessHART: Applying wireless technology in real-time industrial process control," in *Proc. IEEE Real-Time and Embedded Technology and Applications Symp.*, 2008, pp. 377–386.

[11] C. O. Savage and B. F. La Scala, "Optimal scheduling of scalar Gauss-Markov systems with a terminal cost function," *IEEE Trans. Autom. Control*, vol. 54, no. 5, pp. 1100–1105, 2009.

[12] L. Shi and H. Zhang, "Scheduling two Gauss–Markov systems: An optimal solution for remote state estimation under bandwidth constraint," *IEEE Trans. Signal Process.*, vol. 60, no. 4, pp. 2038–2042, 2012.

[13] Z. Lin and C. Wang, "Scheduling parallel kalman filters for multiple processes," *Automatica*, vol. 49, no. 1, pp. 9–16, 2013.

[14] D. Han, H. Zhang, and L. Shi, "An event-based scheduling solution for remote state estimation of two LTI systems under bandwidth constraint," in *Proc. Amer. Control Conf.*, 2013, pp. 3314–3319.

[15] B. Anderson and J. Moore, *Optimal Filtering*. New Jersey: Prentice Hall, 1979.

[16] J. Wu, K. H. Johansson, and L. Shi, "A stochastic online sensor scheduler for remote state estimation with time-out condition," *IEEE Trans. Autom. Control*, vol. 59, no. 11, pp. 3110–3116, 2014.

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

[18] O. Hernández-Lerma and J. B. Lasserre, *Further topics on discrete-time Markov control processes*. Springer Science & Business Media, 2012, vol. 42.

[19] Q. Zhu and X. Guo, "Value iteration for average cost Markov decision processes in Borel spaces," *Applied Mathematics Research eXpress*, vol. 2005, no. 2, pp. 61–76, 2005.

[20] M. L. Puterman, *Markov decision processes: discrete stochastic dynamic programming*. John Wiley & Sons, 2014.

[21] D. Han, J. Wu, Y. Mo, and L. Xie, "On stochastic sensor network scheduling for multiple processes," 2016. [Online]. Available: https://arxiv.org/abs/1611.08222

[22] M. Chiang, *Geometric programming for communication systems*. Now Publishers Inc, 2005.

[23] C. D. Maranas and C. A. Floudas, "Global optimization in generalized geometric programming," *Computers & Chemical Engineering*, vol. 21, no. 4, pp. 351–369, 1997.

[24] S. P. Meyn and R. L. Tweedie, *Markov chains and stochastic stability*. Springer Science & Business Media, 2012.

---

[4]All simulations are conducted on MacBook with a processor of 1.3GHz Intel Core i5 and a memory of 4GB DDR3.