Stochastic Sensor Scheduling for Multiple Dynamical Processes over a Shared Channel

Duo Han*, Junfeng Wu†, Yilin Mo*, Lihua Xie*

Abstract—We consider the problem of multiple sensor scheduling for remote state estimation over a shared link. A number of sensors monitor different dynamical processes simultaneously but only one sensor can access the shared channel at each time instant to transmit the data packet to the estimator. We propose a stochastic event-based sensor scheduling framework in which each sensor makes transmission decisions based on both the channel accessibility and the self event-triggering condition. The corresponding optimal estimator is explicitly given. By utilizing the realtime information, the proposed schedule is shown to be a generalization of the time-based ones and outperform the time-based ones in terms of the estimation quality. By formulating an Markov decision process (MDP) problem with average cost criterion, we can find the optimal parameters for the event-based schedule. For practical use, we also design a simple suboptimal schedule to mitigate the computational complexity of solving an MDP problem. We also propose a method to quantify the optimality gap for any suboptimal schedules.

I. INTRODUCTION

State estimation in cyber-physical systems (CPS) has been an active research area for the past decade. The main challenge is that the estimation performance and the communication infrastructure are usually strongly coupled and thus difficult to analyze. Typically, a class of estimation problems for CPS is to optimize the estimation quality subject to some communication constraints [1]–[3]. In this paper we focus on the bandwidth constrained state estimation problem. 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 an optimal sensor schedule deciding which sensor is able to access the channel at each time in order to maximize the overall state estimation quality.

Considering the problem of multiple sensor scheduling for different processes, Savage and La Scala [4] studied the multiple sensors scheduling for a group of scalar Gauss-Markov systems over a finite horizon. They considered the optimality in terms of terminal estimation error covariance.

This research is partially funded by the Republic of Singapore National Research Foundation (NRF) through a grant to the Berkeley Education Alliance for Research in Singapore (BEARS) for the Singapore-Berkeley Building Efficiency and Sustainability in the Tropics (SinBerBEST) Program. BEARS has been established by the University of California, Berkeley as a center for intellectual excellence in research and education in Singapore. Nanyang Technological University, Singapore, 639798. E-mail: dhanaaa, yilmo, elhxie@ntu.edu.sg.

†: ACCESS Linnaeus Center, School of Electrical Engineering, KTH Royal Institute of Technology, Stockholm, Sweden. E-mail: junfengw@kth.se.

The optimal policy is to schedule the transmissions in the end of the horizon in a specified order. However, the terminal error covariance is only suitable for modelling finite horizon problem. To study the infinite horizon scheduling, Shi et al. [5] employed a metric of averaged estimation error. The authors studied two multi-dimensional systems over an infinite horizon and proposed an explicit optimal periodic sensor schedule for this two-sensor case. Han et al. [6] improved the work by designing an online schedule which depends on the importance of the measurement of one sensor. The framework in [5], [6], however, cannot be generalized into the case of three sensors or more, which limits its scope. The optimality condition in [5] benefits from the mutual exclusiveness of two sensors. When the number of sensors is three or more, the complexity increases dramatically and closed-form optimality conditions are difficult to obtain.

We consider the design of an event-based transmission scheduling framework for a large sensor network. Mimicking the CSMA/CA protocol in wireless sensor networks, the sensors make transmission decisions based on both channel accessibility and self event-triggering. 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 behavior. (3) We model an Markov decision process (MDP) problem with average cost criterion to seek the optimal parameters for the class of proposed event-based schedules. (4) For computational simplicity, we also propose a suboptimal parameter design method and analyze the optimality gap.

Notations: Z is the set of integers and \( \mathbb{Z}_+ \) is the set of positive integers. \( S_n^+ \) (\( S_n^{++} \)) is the set of \( n \) by \( n \) positive semi-definite (definite) matrices. For functions \( f \) with appropriate domains, \( f^i(X) := X \), and \( f^i(X) := f(f^{i-1}(X)) \). \( x[i] \) represents the \( i \)th entry of the vector \( x \). For a matrix \( X \), \( X(i,j) \) represents the entry on the \( i \)th row and \( j \)th column of \( X \). For \( x \in \mathbb{R} \), \( |x| \) is the largest integer that is not larger than \( x \) and \( |x| \) is the smallest one that is not less than \( x \). Denote the set or sequence \( \{x_i\}_{i=j}^k = \{x_j, x_{j+1}, \ldots, x_k\}, j \leq k \) and if \( j > k \) then \( \{x_i\}_{i=j}^k = \emptyset \).

II. PROBLEM SETUP

A. System Model

Denote \( Q := \{1, \ldots, n\} \) as the index set of the processes or sensors and denote \( s_i \) as the \( i \)th sensor for shorthand.
Consider the following $n$ linear time-invariant (LTI) systems:

\[ x_i(k+1) = A_i x_i(k) + \omega_i(k), \quad i \in \mathcal{Q} \quad (1a) \]

\[ y_i(k) = C_i x_i(k) + \nu_i(k), \quad i \in \mathcal{Q} \quad (1b) \]

where \( x_i(k) \in \mathbb{R}^m \) is the state of the \( i \)th process at time \( k \) and \( y_i(k) \in \mathbb{R}^m \) is the measurement obtained by the \( i \)th sensor at time \( k \). 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 \( Q_i > 0, R_i > 0 \), respectively. Assume that \( (A_i, C_i) \) is observable. Furthermore, we assume that \( A_i \) is unstable 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 unstabilized too long.

Each sensor measures its corresponding system state and generates a local estimate first by running a Kalman filter [7]. The information set of the \( i \)th sensor at time \( k \) is given as: \( \mathcal{I}_{i,\text{local}} : = \{ y_i(0), \ldots, y_i(k) \} \), with \( \mathcal{I}_{i,\text{local}}(-1) : = \emptyset \). Let us define the estimate and its covariance to be \( \hat{x}_{i,\text{local}} : = E[x_i(k)|\mathcal{I}_{i,\text{local}}(k)], P_{i,\text{local}}(k) : = E[(x_i(k) - \hat{x}_{i,\text{local}}(k))(x_i(k) - \hat{x}_{i,\text{local}}(k))^\top |\mathcal{I}_{i,\text{local}}(k)] \). The error is thus \( \epsilon_{i,\text{local}} : = x_i(k) - \hat{x}_{i,\text{local}}(k) \). It is well known that the error covariance of the Kalman filter converges exponentially fast to a steady-state constant matrix, i.e., \( P_{i,\text{local}}(k) = \mathcal{P}_i \) for all \( k \), where \( \mathcal{P}_i \) is the solution to the algebraic Riccati equation [7]. We assume that each local Kalman filter has entered the steady state.

The sensors transmit their local estimate to the remote central 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. 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. 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 \( \theta \) is defined as a sequence of \( \gamma_i(k) \), i.e., \( \theta : = \{ \{ \gamma_i(k) \} \}_{k=1}^{\infty} \). Moreover, we define a collection of the information sets \( \mathcal{I}_{i,\theta}(k) \)'s at the estimator side as

\[ \mathcal{I}_{i,\theta}(k) : = \{ \gamma_i(0)\hat{x}_{i,\text{local}}(0), \ldots, \gamma_i(k)\hat{x}_{i,\text{local}}(k) \}, \quad i \in \mathcal{Q}. \]

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

Similar to [5], we use the overall average expected estimation error covariance as a performance metric. For a given schedule \( \theta \in \Theta \) where \( \Theta \) is some feasible set, define a cost function \( J_\theta(\theta) \) over an infinite horizon for estimating the \( i \)th process as follows:

\[ J_\theta(\theta) : = \lim_{T \to \infty} \frac{1}{T} \mbox{Tr} \left( \frac{1}{T} \sum_{k=0}^{T-1} (E[P_i(k)]) \right). \quad (2) \]

We are interested in the following optimization problem:

### Problem 1

\[ \begin{align*}
\text{minimize} & \quad J(\theta) = \sum_{i \in \mathcal{Q}} J_i(\theta), \\
\text{subject to} & \quad \sum_{i \in \mathcal{Q}} \gamma_i(k) = 1, \forall k \in \mathbb{Z}_+. \end{align*} \quad (3) \]

#### B. Event-based Sensor Schedule

Mimicking the protocol of Carrier Sense Multiple Access with Collision Avoidance (CSMA/CA) 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 (AP) and the transmission phase (TP). We assume the transmission time is much larger than the awaiting time. Any sensor listens to the channel carrier for a short period before it sends anything.

If the channel is occupied, then the sensor holds its data; otherwise, it sends the packet to the estimator. The ends of awaiting phase of all sensors are made different in order to avoid collision. In other words, the sensors form a queue with different queuing time in the awaiting phase. Without loss of generality, assume the queue to be \( s_1, s_2, \ldots, s_n \) which means \( s_i \) has higher priority to access the channel than \( s_{i+1} \).

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 channel accessibility and the importance of local data.

We first define two indicators before formally proposing the schedule. Let \( \mu_i(k) \) denote the channel accessibility indicator, i.e.,

\[ \mu_i(k) = \begin{cases} 1, & \text{if } \sum_{j=1}^{i-1} \gamma_j(k) = 0, \\ 0, & \text{otherwise}. \end{cases} \quad (5) \]

Next we define the data importance indicator based on the difference between the MMSE estimate 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 \( \Sigma_i(k) := E[(\epsilon_i(k))^\top \epsilon_i(k) | \mathcal{I}_{i,\theta}(k-1)] \). Furthermore, denote the leave duration of \( s_i \) as \( \tau_i(k) := \min\{k-k_0 : \gamma_i(k_0) = 1, k_0 \leq k \} \).

Various event-triggering criteria for determining the importance of a single measurement have been investigated in [1], [3], [8]. Inspired by [3], [8], we design a stochastic event-triggering rule which can maintain the Gaussianity of the estimation process to check the importance of the data. Define the mapping \( \phi_i(z, \Pi) : \mathbb{R}^{n_1} \times S_+^{n_1} \to [0, 1] \) as

\[ \phi_i(z) = \exp \left( -\frac{1}{2} z^\top \Pi^{-1} z \right). \]
Now we define the data importance indicator \( \eta_i(k) \) as follows:
\[
\eta_i(k) = \begin{cases} 
1, & \text{if } \xi_i(k) > \phi_i(\epsilon_i(k), \alpha_i(k) \Sigma_i(k)), \\
0, & \text{otherwise},
\end{cases}
\]
(6)
where \( \xi_i(k) \sim U[0,1] \) is an i.i.d. auxiliary random variable. The time-varying parameters \( \{\alpha_i(k) \geq 0\}_{i \in Q}^{\infty} \) depend on the historical arrival pattern of all sensors, i.e., strong computation capability and large energy storage. The matrix \( \Sigma_i(k) \) has full rank for any \( i \in Q \) and \( k \in \mathbb{N} \) which we will show later. Now we are ready to propose an event-based schedule \( \theta_+ := \{\{\gamma_i(k)\}_{i \in Q}^{\infty}\}_{k=1}^{\infty} \) and denote \( \Theta_+ \) to be the set of all event-based schedules in (7) subject to all possible \( \{\alpha_i(k)\}_{i \in Q}^{\infty} \), given a specific queue:
\[
\gamma_i(k) = \begin{cases} 
\mu_i(k) \eta_i(k), & \text{if } i \neq n, \\
\mu_i(k), & \text{otherwise}.
\end{cases}
\]
(7)
Note that if \( s_n \) with the last priority detects the channel is idle, it sends the data without checking (6). The condition (6) implies that if the prediction error is small, \( \eta_i(k) \) is very likely to be 0 which prevents unnecessary transmission. The proposed schedule is comprised of two ingredients: compliance via carrier sensing and self-driven motivation. In other words, only when the sensors before \( s_i \) in the queue abandon their rights to transmit and \( s_i \) itself has a strong motivation to reduce the estimation error, \( s_i \) will transmit its data packet. In the subsequent sections, we shall investigate the estimation and communication of the systems under \( \theta_+ \) and optimally design the parameters in a certain sense.

### III. TOWARDS ESTIMATION AND COMMUNICATION

In this section we first investigate what is the MMSE estimator under \( \theta_+ \). To facilitate derivations, we define the following operators for \( X \in \mathbb{R}^{n_i} \) and \( \alpha \in \mathbb{R} \):
\[
h_i(X) := A_i X A_i^\top + Q_i, 
\]
\[(8)\]
\[
g_i(X, \alpha) := \frac{\alpha}{1 + \alpha} (A_i X A_i^\top + h_i(\overline{P}_i) - \overline{P}_i),
\]
\[(9)\]
\[
t_i(X, \alpha) := \frac{1}{1 + \alpha} \overline{P}_i + \frac{\alpha}{1 + \alpha} h_i(X).
\]
(10)
The following lemma on the properties of the innovation and estimate error is useful for proving the main result whose proof can be found in [9]. Let the incremental innovation for \( s_i \) be denoted as \( \hat{\delta}_i(k) := \hat{x}_i,\text{local}(k) - A_i \hat{x}_i,\text{local}(k-1) \).

**Lemma 1** The following 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) \( \mathbb{E} [\delta_i(k) \delta_i(j)^\top] = 0 \) for any \( k \neq j \).
- (iii) \( \mathbb{E} [\epsilon_i,\text{local}(k) \delta_i(k_0)^\top] = 0 \) for any \( k_0 \leq k \).
- (iv) \( \mathbb{E} [\hat{x}_i,\text{local}(k_1) \epsilon_i,\text{local}(k_1)] = 0 \) for any \( k_1 \geq k \) and \( \mathbb{E} [\hat{x}_i,\text{local}(k) \delta_i(k_2)^\top] = 0 \) for any \( k_2 > k \).

For notational simplification, denote
\[
\hat{\gamma}_i(k) = 1 - \gamma_i(k), \hat{\mu}_i(k) = 1 - \mu_i(k).
\]
(11)

**Theorem 1** Under the proposed schedule \( \theta_+ \), the MMSE state estimate for each system is given by
\[
\hat{x}_i(k) = \begin{cases} 
\hat{x}_i,\text{local}(k), & \text{if } \hat{\gamma}_i(k) = 1 \\
A_i \hat{x}_i(k-1), & \text{if } \hat{\gamma}_i(k) = 0
\end{cases},
\]
and the corresponding estimation error covariance is
\[
P_i(k) = \begin{cases} 
\overline{P}_i, & \text{if } \hat{\gamma}_i(k) = 1 \\
h_i(P_i(k-1)), & \text{if } \hat{\mu}_i(k) = 1 \\
t_i(P_i(k-1), \alpha_i(k)), & \text{otherwise}
\end{cases}.
\]
(12)
Moreover, \( \Sigma_i(k) \) in (6) is given by
\[
\Sigma_i(k) = h_i(P_i(k-1)) - \overline{P}_i.
\]
(13)
The following lemma whose proof is omitted here is used for the derivation of main results.

**Lemma 2** Suppose \( z \in \mathbb{R}^n \) is a Gaussian random variable, i.e., \( z \sim \mathcal{N}(0, Z) \) and \( \xi \) is uniformly distributed over \([0,1]\).
The following statements hold:

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

Now we are ready to prove Theorem 1.

**Proof:** When \( \hat{\gamma}_i(k) = 1 \), from [7] we know that
\[
\mathbb{E} [x_i(k) | \hat{x}_i,\text{local}(k), I_i,\theta(k-1)] = \hat{x}_i,\text{local}(k) \text{ and } P_i(k) = \overline{P}_i.
\]
When \( \hat{\mu}_i(k) = 0 \), \( \hat{\gamma}_i(k) \) must be 0 which implies the channel is occupied by other sensors, the estimator can only do prediction on the estimate of \( x_i(k) \), i.e.,
\[
\hat{x}_i(k) = A_i \hat{x}_i(k-1), \quad P_i(k) = h_i(P_i(k-1)).
\]
(15)
The two cases above are easy to analyze. Next we consider the remaining case, i.e., when \( \hat{\mu}_i(k) \hat{\gamma}_i(k) = 1 \).

The following equation is easy to verify and useful for subsequent derivations.
\[
x_i(k) = A_i \hat{x}_i(k-1) + \epsilon_i,\text{local}(k) + \epsilon_i(k).
\]
(16)

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 \( \mu_i(k+1) \gamma_i(k+1) = 1 \) occurs at time \( k + 1 \). Moreover, the following recursive equation always holds:
\[
\epsilon_i(k) = A_i \epsilon_i(k+1) + \delta_i(k).
\]
(17)
Also from (15) and (16) we have
\[
x_i(k+1) = A_i^{\tau_i(k+1)} \hat{x}_i,\text{local}(k - \tau_i(k + 1)) + \epsilon_i,\text{local}(k+1) + \epsilon_i(+1) = A_i^{\tau_i(k+1)} \hat{x}_i,\text{local}(k - \tau_i(k + 1)) + \epsilon_i,\text{local}(k+1) + \sum_{j=0}^{\tau_i(k)} A_i^{j} \delta_i(k+1 - j),
\]
where the last two terms on the RHS are mutually independent from Lemma 1.(iii). Due to Lemma 1(i) and the fact [7] that \( f(\epsilon_i,\text{local}(k + 1) | I_i,\theta(k + 1)) \sim \mathcal{N}(0, \overline{P}_i) \), we have
\[
\mathbb{E} [x_i(k+1) | I_i,\theta(k+1)] = A_i \hat{x}_i(k).
\]
Therefore, from (16) we know that for \( j \in [k - \tau_i(k), k] \) the equality holds:
\[
P_i(k) = \overline{P}_i + \Lambda_i(k).
\]
(18)
where \( \Lambda_i(k) := \mathbb{E} \left[ (\varepsilon_i(k))^\top \varepsilon_i(k) | \mathcal{I}_t, \theta_t(k) \right] \). Hence we have \( \Lambda_i(k) = P_i(k) - \bar{P} \). Then together with (17) we can conclude that

\[
    f(\varepsilon_i(k+1)|\mathcal{I}_t, \theta_t(k)) \sim \mathcal{N}(0, A_i(P_i(k) - \bar{P})A_i^\top + h_i(P_i(k) - \bar{P})_i)
\]

which proves (14). From Lemma 2 and (19), we have

\[
    f(\varepsilon_i(k+1)|\mathcal{I}_t, \theta_t(k+1)) \sim \mathcal{N}(0, g_i(P_i(k) - \bar{P}, \alpha_i(k))).
\]

Thus, from (18) and (20) we have

\[
    f(x_i(k+1)|\mathcal{I}_t, \theta_t(k+1)) \sim \mathcal{N}(A_i \hat{x}_i(k), \bar{P}_i + g_i(P_i(k) - \bar{P}, \alpha_i(k))). \tag{21}
\]

Notice that \( \hat{x}_i(k+1) \) is actually \( A_i^{(k+1)} \hat{x}_i, local(k-\tau_i(k+1)) \) which is same with the predicted estimate but with a smaller error covariance. Consequently, no matter what \( u_i(k+2) \) and \( \gamma_i(k+2) \) are at time \( k+2 \), the mutual independence between the three terms in (16) and the recursion in (17) hold. Therefore, the conditional pdf of \( \varepsilon_i(k+2) \) can be computed in a similar fashion as (19) and (20), which is zero mean Gaussian distributed. Thus \( x_i(k+2) \) is still a predicted estimate, i.e., \( x_i \), \( \hat{x}_i, local(k+1) \), with a corresponding error covariance dependent on \( u_i(k+2) \) and \( \gamma_i(k+2) \). Recursively, we can verify (12) and (13) which completes the proof.

The result in (13) shows that if \( s_i \) is idle due to other competitors, i.e., \( u_i(k) = 0 \), 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., \( u_i(k) \gamma_i(k) = 1 \), the estimator updates the estimate using the information encoded by (6).

From [10], for any realization of \( P_i(k-1) \) we know that \( P_i \leq t_i(P_i(k-1), \gamma_i(k)) \leq h_i(P_i(k-1)) \). 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.

The probability of kinds of events during transmission can be computed as follows. For concise notation, denote

\[
    \beta_i(k) := (\alpha_i(k)/(1 + \alpha_i(k)))^{n/2}, \quad \tilde{\beta}_i(k) := 1 - \beta_i(k).
\]

**Theorem 2** The probability mass function of \( \eta_i(k), \mu_i(k) \) and \( \gamma_i(k) \) is given respectively:

(i) \( \Pr \{ \eta_i(k) = 0 \} = \beta_i(k) \).

(ii) \( \Pr \{ \mu_i(k) = 1 \} = \begin{cases} 1, & \text{if } i = 1 \\
 \prod_{j=1}^{i-1} \beta_j(k), & \text{if } i > 1 \end{cases} \).

(iii) \( \Pr \{ \gamma_i(k) = 1 \} = \begin{cases} \prod_{j=1}^{n-1} \beta_j(k), & \text{if } i = n \\
 \tilde{\beta}_i(k) \prod_{j=1}^{i-1} \beta_j(k), & \text{if } i < n \end{cases} \).

The proof is omitted here for space saving.

**IV. PARAMETER DESIGN FOR OPTIMAL TRADEOFF**

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 \( \Theta_+ \) from which \( \theta_+ \) is at least as well as any time-based schedule \( \theta_i \) does, i.e., \( \{ \theta_+ \in \Theta_+ : J(\theta_+) \leq J(\theta_i), \forall \theta_i \neq \emptyset \} \).

**Proof:** Any time-based schedule is a special case of \( \theta_+ \) in (7). 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 \( \theta_+ \), one can always find a schedule in \( \Theta_+ \) performs as well as any time-based schedule can. \( \blacksquare \)

The next question is how to determine the optimal \( \{ \alpha_i(k) \}_{i=1}^{\infty} \) to achieve the minimum average cost over the space \( \Theta_+ \), i.e., solving the following optimization problem:

**Problem 2**

\[
    \minimize_{\theta \in \Theta_+} J(\theta), \quad \text{subject to } \sum_{i \in Q} \gamma_i(k) = 1. \tag{22}
\]

By formulating an Markov decision process (MDP) problem with average cost criterion, we can design the optimal parameters.

**A. Optimal Parameter Design by Solving an MDP Problem**

First define the state space \( V \) to be the set of all possible \( \{ P_i(k) \}_{k \in \mathbb{Z}_+} \subset \{ C_1 \times \cdots \times C_n \} \) for any \( k \) and any \( \{ \{ \alpha_i(k) \}_{i \in Q} \}_{k=1}^{\infty} \), where \( C_i \) is the set of all convex combination of \( h_i^j(P_i^k), \forall j, \text{ i.e., } \sum_j \lambda_j h_i^j(P_i^k), \sum_j \lambda_j = 1 \). The compact action space \( B \) is defined as \( \{ v \in \mathbb{R}_+ : 0 \leq v \leq 1, \forall i \in Q \} \), which is identical for each state in \( V \). The action represents the set of \( \beta_i(k)^4 \) broadcast to the sensors by the estimator. Denote the set \( \mathbb{K} := \{ (v, b) : v \in V, b \in B \} \) which is a Borel space. The transition law \( Q(\cdot|v, b) \) is a stochastic kernel on \( V \) given \( \mathbb{K} \), 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 \( \Omega := (V, B, Q(\cdot|v, b), c(v, b)) \). A decision at time \( k \) is a mapping \( d(k) : V \rightarrow B \) and the set of admissible decisions is denoted as \( D \). A policy \( \zeta \) for \( \Omega \) is a sequence of decision rules \( d(1), d(2), \ldots \). The policy \( \zeta \) is said to be stationary if \( d(2) = d \) for all time \( k \). We define the average expected cost of the policy \( \zeta \) per unit time by

\[
    g_\zeta(v) = \limsup_{T \to \infty} \frac{1}{T} \mathbb{E}_{\zeta, v} \left[ \sum_{k=0}^{T-1} r(v(k), b(k)) \right], \tag{23}
\]

where \( v \in V \) is the initial state, and the expectation is taken based on the stochastic process \( \{ v(k), b(k) \} \) uniquely determined by the policy \( \zeta \) and \( v \). The target for an MDP problem with average cost criterion is to search a optimal stationary policy \( \zeta^* \) for \( \Omega \) such that the average cost is minimized, i.e., \( g_\zeta^*(v) \leq g_\zeta(v), \forall v \in V \).

It is easy to see that finding the optimal policy \( \zeta^* \) for \( \Omega \) is equivalent to seeking the optimal solution over \( \Theta_+ \) to Problem 1. Numerous literature have studied the optimality.

\(^1\) In practice, the estimator would transmit \( \alpha_i(k) \) instead of \( \beta_i(k) \) for the implementation of event-triggering in (6). The computation is simple since \( \alpha_i(k) \) is one-to-one correspondence to \( \beta_i(k) \).
conditions for an optimal policy of an MDP problem with average cost criterion in Borel spaces such as [11].

B. Suboptimal Schedules

Due to the computational complexity of solving an MDP, we propose suboptimal schedules for practical use and provide a general lower bound on the optimal cost to quantify the optimality gap between any suboptimal schedules and the optimal one.

Receding Horizon (RH) Algorithm: The RH algorithm, resembling receding horizon control, computes the optimal scheduling decisions over a finite window of size \( \tau \in \mathbb{Z}_+ \) and keeps only the next-step scheduling decision. At the next time instant, it computes a new stretch of scheduling decisions by sliding the window one step forward. We call the special case of RH algorithm with the window size of 1 as Maximum Error First (MEF) Algorithm which is the most greedy one.

C. Optimality gap

Notice that the performance gap between the suboptimal schedules and the optimal \( \theta_\star \) is upperbounded by the gap between the suboptimal performance and a lower bound for the optimal performance. In this section we aim to find a tight lower bound of the optimal performance by using MDP method and thus we can quantify the optimality gap of any suboptimal schedule.

We relax the constraint in (4) by requiring the sum of the rates of all sensors to be 1. In other words, we allow the channel can be used by multiple sensors but their sum communication rate must be (or less than) 1. Denote \( \Theta_\tau \) as the set of all schedules in (7) with forcing \( \mu_i(k) \) to be 1 for \( i \leq n - 1 \) and all \( k \). In other words, any absence of \( \bar{x}_{i,local}(k) \) implies (6) for \( i \leq n - 1 \). Note that there is no event-triggering mechanism for \( s_n \) and its scheduling strategy is time-based and periodic according to [12]. Then we have the following optimization problem:

**Problem 3**

\[
\text{minimize} \quad J(\theta), \quad \text{subject to} \quad \sum_{i \in Q} \mathbb{E}[\gamma_i(k)] = 1. \tag{24}
\]

Now \( \alpha_i(k) \) solely depends on \( \tau_i(k) \) and so does \( \beta_i(k) \). From (13) 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 \( \beta_i^j \) to denote the probability of \( \Pr\{\tau_i(k) = 0\} \) for sensor \( i \) with \( \tau_i(k) = j \) and \( \alpha_i^j \) accordingly for \( i \leq n - 1 \) and from the optimal time-based scheduling theory [5], [12] \( \beta_n^0 \) is the rate constraint, \( \beta_i^j = 0, j > [r_i^{-1}] \) and \( \beta_n^0 = 1 - [r_n^{-1}] \times r_n, j = [r_n^{-1}] \) where \( r_n \) denotes the rate constraint, i.e., \( \mathbb{E}[\gamma_n(k)] \leq r_n \). Therefore, we minimize \( J_i(\theta) \) in (26) with respect to \( \{\beta_i^j\}_{i=0}^{\infty} \).

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

**Proposition 1** Suppose \( \theta_\star \in \Theta_\tau \) to be the optimal solution to Problem 2. For any schedule \( \theta_\perp \in \Theta_\perp \) (including any time-based schedules), define the optimality gap to be \( \Delta(\theta_\perp) := J(\theta_\perp) - J(\theta_\star) \). The following inequality holds:

\[
\Delta(\theta_\perp) \leq J(\theta_\#) - J(\theta_\star),
\]

where \( \theta_\# \) is characterized by \( \{\beta_i^j\}_{i=0}^{\infty} \) which is the solution to the following problem:

\[
\begin{align*}
\text{minimize} \quad & \sum_{i \in Q} \pi_0^i \text{Tr}[P_i^0] + \sum_{j=0}^{\infty} \pi_0^i \prod_{l=0}^{j} \beta_i^l \text{Tr}[P_i^{j+1}], \\
\text{subject to} \quad & \pi_0^i + \sum_{j=0}^{\infty} \prod_{l=0}^{j} \beta_i^l \geq 1, \quad i \in Q, \\
& \pi_0^i \leq 1, \quad i \in Q, \\
& \text{where } P_i^0 := h_i^n(\mathcal{P}_n) \text{ and }
\end{align*}
\]

\[
P_i^j := \prod_{n=0}^{j-2} (\beta_i^2/n_i) h_i^n(\mathcal{P}_i) + \sum_{l=0}^{j-1} (1 - (\beta_i^{j-1-l}/2n_i)) \times \left( \prod_{n=0}^{l-1} (\beta_i^{l-1-u}/2n_i) \right) h_i^n(\mathcal{P}_i), \quad \forall i \leq n - 1. \tag{25}
\]

**Proof:** It is easy to see the optimal cost in Problem 3 is a lower bound of the optimal cost in Problem 1 since the constraint in (4) is stronger. In Problem 3, the transmission of each sensor is not affected by others and thus we can first consider such a subproblem for one sensor:

\[
\text{minimize} \quad J_i(\theta), \quad \text{subject to} \quad \mathbb{E}[\gamma_i(k)] \leq r_i. \tag{26}
\]

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 \( T_i(j_2, j_1) := \Pr\{\tau_i(k) = j_2 | \tau_i(k-1) = j_1\} \). Though we can design \( \beta_i^j \) freely, we have to choose the set of \( \{\beta_i^j\}_{i=0}^{\infty} \) 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 from [10] and the fact that \( \lim_{j \to \infty} \text{Tr}[h_i^j(P_i)] \to \infty \). Therefore, the stationary distribution \( \pi_i = [\pi_0^i, \pi_1^i, \pi_2^i, \ldots] \) must exist [13]. Denote the corresponding estimation error covariance for \( \tau_i(k) \) as \( P_i^r(k) \). From Theorem 1 and 2(ii) we have that \( P_i^0 = \mathcal{P}_i, P_i^j = t(P_i^{j-1}, \alpha_i^{j-1}) \) and thus \( P_i^j \) in (25). Now the cost function becomes \( J_i(\theta) = r_i \text{Tr}[P_i^0] + \sum_{j=0}^{l_{max}-1} r_i \left( \prod_{l=0}^{j} \beta_i^l \right) \text{Tr}[P_i^{j+1}] \).

Note that the implicit equality constraints for Problem 3 are \( \sum_i \pi_i^0 = 1 \) and \( \sum_i \pi_i^1 = 1 \). Since the cost function \( J_i(\theta) \) is a decreasing function of \( \pi_i^0 \), we can transform it into the inequality constraints and the optimal solution must be reached only when the equalities hold. Therefore, by solving the optimization problem in the proposition we can obtain
TABLE I: Performance comparison among different schedules. The lower bound of the optimal $J$ is also listed as LB.

<table>
<thead>
<tr>
<th>Schedules and LB</th>
<th>$\theta_1$</th>
<th>$\theta_2$</th>
<th>$\theta_3$</th>
<th>$\theta_4$</th>
<th>LB</th>
</tr>
</thead>
<tbody>
<tr>
<td>$J(\cdot)$</td>
<td>13.18</td>
<td>8.53</td>
<td>8.23</td>
<td>8.58</td>
<td>7.51</td>
</tr>
</tbody>
</table>

Remark 1 The queue order is assumed to be given in the previous analysis. For a small network, we can enumerate the optimal or suboptimal schedules for different orders and choose the best one. However, it is formidable for a large network. One heuristic technique is to assign the priority to the queue to be $s_1, \ldots, s_i, s_{i+1}, \ldots$ with $\pi_i^0 \leq \pi_{i+1}^0$ since a larger $\pi_i^0$ usually implies the error covariance of $s_i$ may grow faster than others.

V. EXAMPLES

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 [5] which is only applicable to the two-sensor case, we also conduct simulations over two processes with the parameters: $A_1 = 5, C = 1, Q = 1, R = 1$ and $A_1 = 1.1, C = 1, Q = 3, R = 1$, respectively.

We compare the performance of the following four schedules:

- $\theta_1$: the optimal time-based schedule [5] is $\{s_2, s_1, s_1, s_1, s_2, s_1, s_1, \ldots\}$ with the period 3.
- $\theta_2$: the suboptimal event-based schedule via MEF.
- $\theta_3$: the suboptimal event-based schedule via RH algorithm with a window size of 3.
- $\theta_4$: the approximate optimal event-based schedule using the MDP approach through discretization. The number of state grids and action grids is 1000 and 100.

We also compute the lower bound for the optimal event-based schedule according to Proposition 1. We test the maximum $l$ with 2, 3, 4, 5 and the corresponding lower bound is 7.81, 7.55, 7.51, 7.51. So we set the maximum $l$ to be 5. Moreover, we obtain that $\pi_i^0 = 0.68 > \pi_2^3 = 0.31$. Thus we choose the queue to be $s_1, s_2$.

We summarize the results in Table I. The estimation performance of $\theta_1$ is the worst among all. The three event-based schedules reduce the cost function $J(\theta_1)$ by 35.2%, 37.5%, 34.9%, respectively. The approximate solution to the MDP does not outperform the solutions given by MEF or RH algorithms in this case due to the discretization procedure. Moreover, the gap between $J(\theta_2) (J(\theta_3))$ and LB is small, which implies these suboptimal schedules are practically good choices with low computational complexity compared with the MDP-based optimal schedule.

VI. CONCLUSION

We have studied an event-based sensor scheduling framework for sensor networks monitoring different processes. The sensors make transmission decisions based on both channel accessibility and self-event-triggering depending on the realtime innovations. The proposed schedule has been shown to provide a better tradeoff between estimation and communication than the time-based schedule. We have also investigated the optimal event-based schedule through an MDP approach. Since solving an MDP problem in Borel spaces is generally computationally difficult, we have also proposed a suboptimal schedule via RH algorithms and analyzed the optimality gap.

REFERENCES