[18] A. D. Roy, "Safety first and the holding of assets," *Econometric*, vol. 20, pp. 431–449, 1952.
[19] M. K. Sain, "Control of linear systems according to the minimal variance criterion: A new approach to the disturbance problem," *IEEE Trans. Automat. Control*, vol. AC-11, no. 1, pp. 118–122, Jan. 1966.
[20] M. K. Sain and S. R. Liberty, "Performance measure densities for a class of LQG control systems," *IEEE Trans. Automat. Control*, vol. AC-16, no. 5, pp. 431–439, Oct. 1971.
[21] J. F. Sturm, "Using SeDuMi 1.02, a Matlab toolbox for optimization over symmetric cones," *Optimiz. Methods and Softw.*, vol. 11–12, pp. 625–653, 1999.
[22] H. T. Toivonen and P. M. Mäkilä, "Computer-aided design procedure for multiobjective LQG control problems," *Int. J. Control*, vol. 49, pp. 655–666, 1989.
[23] P. Whittle, *Risk-Sensitive Optimal Control*. Chichester, U.K.: Wiley, 1990.
[24] K. M. Zhou and J. C. Doyle, *Essentials of Robust Control*. Upper Saddle River, NJ: Prentice-Hall, 1998.

Kalman Filtering Over Graphs: Theory and Applications

Ling Shi, *Member, IEEE*

***Abstract*—In this technical note we consider the problem of distributed discrete-time state estimation over sensor networks. Given a graph that represents the sensor communications, we derive the optimal estimation algorithm at each sensor. We further provide a closed-form expression for the steady-state error covariance matrices when the communication graph reduces to a directed tree. We then apply the developed theoretical tools to compare the performance of two sensor trees and convert a random packet-delay model to a random packet-dropping model. Examples are provided throughout the technical note to support the theory.**

***Index Terms*—Kalman filter.**

I. Introduction

Advances in fabrication, modern sensor and communication technologies, and computer architecture have enabled a variety of new networked sensing and control applications. For example, wireless sensor networks form an important class of such applications, which have attracted much attention in the past few years. Sensor networks can be used for environment and habitat monitoring, health care, home and office automation, traffic control, etc. [1]. This area of research brings together researchers from computer science, communication, control, etc. [2].

In many wireless sensor network applications, there is an economic incentive towards using off-the-shelf sensors and standardized communication solutions. A consequence of this is that the individual hardware components might be of relatively low quality and that communication resources are quite limited. Due to the limited communication resources, data packets generated at a particular time may arrive at the sensors at variable times, not necessarily in order, and sometimes not at all. Estimation and control over such resource-constrained networks thus require new design paradigms beyond traditional sampled-data control. For example, consider the problem of state estimation over such a network using a Kalman filter. The Kalman filter is a well-established methodology for model-based fusion of sensor data [3]. In the standard Kalman filter, it is assumed that sensor data are transmitted along perfect communication channels and are available to the estimator instantaneously, and no interaction between communication and control is considered.

Manuscript received March 27, 2009; revised June 05, 2009. First published August 11, 2009; current version published September 04, 2009. Recommended by Associate Editor S. Mascolo.

The author is with the Department of Electronic and Computer Engineering, The Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong (e-mail: eesling@ust.hk).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TAC.2009.2026851

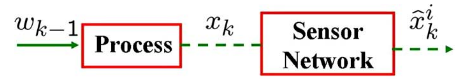

Fig. 1. State estimation using a wireless sensor network.

Kalman filtering under certain information constraints, such as decentralized implementation, has been extensively studied [4]. Implementations for which the computations are distributed among network nodes were considered by Alriksson and Rantzer [5]. Sinopoli *et al.* [6] studied Kalman filtering with intermittent sensor observations, and they showed that there exists a critical packet arrival rate below which the expected value of the estimation error covariance matrix becomes unbounded. The problem of Kalman filtering for systems with delayed measurements is not new and has been studied even before the emergence of networked control [7], [8]. It is well known that discrete-time systems with constant or known time-varying bounded measurement delays may be handled by state augmentation in conjunction with the standard Kalman filtering or by the reorganized innovation approach [9].

This technical note focuses on developing theoretical tools for distributed estimation over sensor networks. The main contributions are summarized as follows.

1) Given an undirected graph $\mathcal{G}$ that represents the sensor communications, we provide an optimal estimation algorithm at each sensor. The algorithm is fully distributed, and can deal with arbitrary data packet drops in the network.
2) When the communication graph reduces to a directed tree, we provide an exact expression on the steady-state error covariance matrices at each sensor.
3) We apply the developed theoretical tools to compare the performance of two sensor trees and convert a random packet-delay model to a random packet-dropping model.

The rest of the technical note is organized as follows. In Section II, we give the mathematical models of the considered problems, and provide some preliminary results on Kalman filtering to facilitate the analysis in the remaining sections. In Section III, we present the main result of the technical note. Some concluding remarks and discussions are given in Section IV.

II. Problem Setup

Consider the problem of distributed state estimation over a wireless sensor network (Fig. 1). The process dynamics is described by

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

A wireless sensor network consisting of N sensors $\{S_0, S_1, \ldots, S_{N-1}\}$ is used to measure the state. When S_i takes a measurement of the state in (1), it returns

$$y_k^i = H_i x_k + v_k^i.(i = 0, 1, \ldots, N-1.) \tag{2}$$

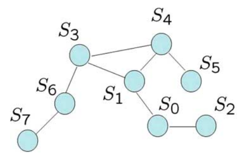

Fig. 2. Graph Example.

In (1) and (2), $x_k \in \mathbf{R}^n$ is the state vector in the real n-dimensional vector space, $y_k^i \in \mathbf{R}^{m_i}$ is the observation vector at S_i, $w_k \in \mathbf{R}^n$ and $v_k^i \in \mathbf{R}^{m_i}$ are zero-mean Gaussian random vectors with $\mathbb{E}[w_k w_j{}'] = \delta_{kj} Q, Q \geq 0$, $\mathbb{E}[v_k^i v_t^{i}{}'] = \delta_{kt}\Pi_i, \Pi_i > 0$, $\mathbb{E}[v_k^i v_t^{j}{}'] = 0\ \forall t,k$ and $i \neq j$, $\mathbb{E}[w_k v_t^{i}{}'] = 0\ \forall i,t,k$, where $\delta_{kj} = 0$ if $k \neq j$ and $\delta_{kj} = 1$ otherwise. We assume that $(A, \sqrt{Q})$ is controllable.

We use an undirected graph $\mathcal{G}$ (e.g., Fig. 2) to represent the sensor communications. The nodes $\mathcal{N}$ of $\mathcal{G}$ correspond to the N sensors $\{S_0, S_1, \ldots, S_{N-1}\}$, and the edges $\mathcal{E}$ of $\mathcal{G}$ correspond to the active communication links between the sensors, e.g., $e_{ij} \in \mathcal{E}$ means S_i communicates with S_j, etc. We assume $\mathcal{G}$ is connected, i.e., there exists a path between any S_i and S_j for $i \neq j$.

At time k, S_i generates the measurement packet y_k^i and sends it with all previously received measurements from its neighbor sensors that are after $k - D$ to all its neighbors, where $D \geq 1$ is a constant. This is reasonable, as when D is sufficiently large, the late-arriving measurement related to the system state in the far past may not contribute much to the improvement of the accuracy of the current estimate. We assume all data packets y_k^i are time-stamped, therefore when a sensor receives a data packet, it will know when the measurement is taken and from which sensor it comes.

Let $\mathcal{B}_{k\,|\,k-l}^i (l = 0, \ldots, D-1)$ be the set of all measurement packets that are taken at time $k - l$ and are available at S_i at time k. For example, consider S_1 in Fig. 2, when $D = 2$, S_1 sends $\{y_k^1, y_{k-1}^0, y_{k-1}^3, y_{k-1}^4\}$ to S_0, S_3, S_4 and receives $\{y_k^0, y_{k-1}^1, y_{k-1}^2\}$ from S_0, $\{y_k^3, y_{k-1}^1, y_{k-1}^4, y_{k-1}^6\}$ from S_3, and $\{y_k^4, y_{k-1}^1, y_{k-1}^3, y_{k-1}^5\}$ from S_4. Therefore

$$\begin{aligned}\mathcal{B}_{k\,|\,k-1}^1 &= \left\{y_{k-1}^0, y_{k-1}^1, y_{k-1}^2, y_{k-1}^3, y_{k-1}^4, y_{k-1}^5, y_{k-1}^6\right\}\\ \mathcal{B}_{k\,|\,k}^1 &= \left\{y_k^0, y_k^1, y_k^3, y_k^4\right\}.\end{aligned}$$

Remark 2.1: Notice that since we only require $\mathcal{G}$ to be connected, $\mathcal{G}$ may contain cycles. This implies that the same measurement packet may arrive at a sensor node multiple times, e.g., y_{k-1}^1 is received twice by S_1 in the previous example. When this happens, the sensor node simply discards any packet that has been received before and the set $\mathcal{B}_{k\,|\,k-l}^i$ only includes distinct measurement packets that are taken at time $k - l$. It is the set $\mathcal{B}_{k\,|\,k-l}^i$ that will be processed by S_i at time k as we shall see in Section III.

In this technical note, we are interested in the following problem.

Problem 2.2: Given an undirected graph $\mathcal{G}$ that represents the sensor communications with possible data packet drops, find out the optimal state estimate $\hat{x}_k^i$ computed at each sensor S_i.

Before we provide an optimal estimation algorithm for each sensor i in Section III, we first provide a short summary of Kalman filtering upon which our main result relies.

A. Kalman Filtering Preliminaries

Consider the process in (1) with the following single sensor measurement equation:

$$y_k = C_k x_k + v_k \tag{3}$$

where v_k is zero-mean Gaussian random vectors with $\mathbb{E}[v_k v_j{}'] = \delta_{kj} R_k, R_k > 0$, and $\mathbb{E}[w_k v_j{}'] = 0\ \forall j,k$. Notice that we consider time-varying C_k and R_k here. The Kalman filter in its most general form can assume time-varying A and Q. The special form we look at here suffices for deriving the optimal estimation algorithms in subsequent sections.

Assume a linear estimator receives y_k and computes the optimal state estimate at each time k. Let $\mathbf{Y}_k$ be the set of all measurements received by the estimator at time k. Define

$$\hat{x}_k \triangleq \mathbb{E}[x_k \,|\, \mathbf{Y}_k] \tag{4}$$
$$P_k \triangleq \mathbb{E}[(x_k - \hat{x}_k)(x_k - \hat{x}_k)' \,|\, \mathbf{Y}_k] \tag{5}$$
$$\bar{P} \triangleq \lim_{k\to\infty} P_k, \text{if the limit exists.} \tag{6}$$

It is well known that $\hat{x}_k$ and P_k can be computed as

$$(\hat{x}_k, P_k) = \mathbf{KF}(\hat{x}_{k-1}, P_{k-1}, y_k, C_k, R_k)$$

where $\mathbf{KF}$ denotes the Kalman filter which consists of the following update equations at time k:

$$\hat{x}_{k\,|\,k-1} = A\hat{x}_{k-1}, \tag{7}$$
$$P_{k\,|\,k-1} = AP_{k-1}A' + Q \tag{8}$$
$$K_k = P_{k\,|\,k-1}C_k'[C_k P_{k\,|\,k-1} C_k' + R_k]^{-1} \tag{9}$$
$$\hat{x}_k = A\hat{x}_{k-1} + K_k(y_k - C_k\hat{x}_{k\,|\,k-1}) \tag{10}$$
$$P_k = (I - K_k C_k)P_{k\,|\,k-1}. \tag{11}$$

Let $\mathbb{S}_+^n$ be the set of n by n positive semi-definite matrices. For functions $f_1, f_2 : \mathbb{S}_+^n \to \mathbb{S}_+^n$, define $f_1 \circ f_2$ as $f_1 \circ f_2(X) \triangleq f_1(f_2(X))$. Define the functions $h, \tilde{g}_{[C,R]}, g_{[C,R]} : \mathbb{S}_+^n \to \mathbb{S}_+^n$ as

$$h(X) \triangleq AXA' + Q \tag{12}$$
$$\tilde{g}_{[C,R]}(X) \triangleq X - XC'[CXC' + R]^{-1}CX \tag{13}$$
$$g_{[C,R]}(X) \triangleq h \circ \tilde{g}_{[C,R]}(X). \tag{14}$$

We write $g_{[C,R]}(X)$ and $\tilde{g}_{[C,R]}$ as g_C and $\tilde{g}_C$ when there is no confusion on the underlying parameters R. With some manipulation, it can then be shown that $P_{k\,|\,k-1}$ and P_k from (8) and (11) evolve as

$$P_{k\,|\,k-1} = g_{C_{k-1}}(P_{k-1\,|\,k-2}) \tag{15}$$
$$P_k = \tilde{g}_{C_k}(P_{k\,|\,k-1}). \tag{16}$$

When the parameters C_k and R_k are not time-varying, i.e., $C_k = C$ and $R_k = R$, we have the following result on the steady-state error covariance matrix $\bar{P}$.

Lemma 2.3: Assume $C_k = C, R_k = R$ for all $k \geq 0$. Further assume that (A, C) is observable and $(A, \sqrt{Q})$ is controllable. Then $\bar{P}$ exists and satisfies $\bar{P} = \tilde{g} \circ h(\bar{P})$.

Proof: Standard result from Kalman filtering analysis (e.g., [3]) and the proof is omitted. ■

Now consider the case when data packet can be dropped by the network. Let γ_k be the indicator functor for y_k at time k, i.e., $\gamma_k = 1$ means y_k is received and $\gamma_k = 0$ otherwise. In this case, $(\hat{x}_k, P_k)$ is known to be computed by a modified Kalman filter ($\mathbf{MKF}$) [6]. We write $(\hat{x}_k, P_k)$ in compact form as

$$(\hat{x}_k, P_k) = \mathbf{MKF}(\hat{x}_{k-1}, P_{k-1}, \gamma_k, y_k, C_k, R_k)$$

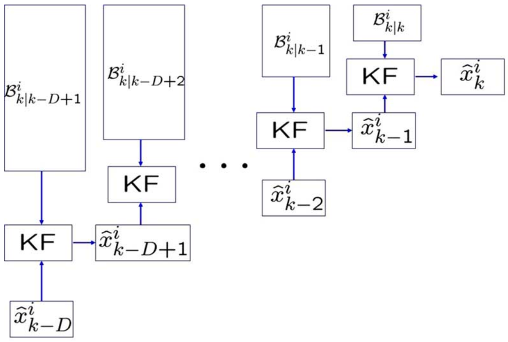

Fig. 3. Kalman filter iterations at time k.

which represents the same set of update equations as in (7)–(9) together with

$$\hat{x}_k = A\hat{x}_{k-1} + \gamma_k K_k(y_k - C_k A\hat{x}_{k-1}) \tag{17}$$
$$P_k = (I - \gamma_k K_k C_k)P_{k\,|\,k-1}. \tag{18}$$

Notice that if $\gamma_k = 1$ for all k, then $\mathbf{MKF}$ simply reduces to the standard Kalman filter.

III. Kalman Filtering Over Graphs

Let the undirected graph $\mathcal{G}$ which represents the sensor communications be given, and consider any $S_i \in \mathcal{G}$. Define $\hat{x}^i_k(\mathcal{G}), P^i_k(\mathcal{G}), \bar{P}^i(\mathcal{G})$ at S_i similar as that in (4)–(6). We write $\hat{x}^i_k(\mathcal{G})$ as $\hat{x}^i_k$, etc., for convenience in this section.

In Section II, we denote $\mathcal{B}^i_{k\,|\,k-l}$ as the set of all measurement packets that are taken at time $k-l$ and are available at S_i at time k. It is easy to verify that

$$\mathcal{B}^i_{k-l|k-l} \subset \mathcal{B}^i_{k\,|\,k-l}, \forall i, k, \text{ and } 0 \leq l \leq D-1.$$

In other words, S_i has more measurements of time $k-l$ at time k than at time $k-l$. Therefore, S_i can obtain a better estimate of x_{k-l} at time k than at time $k-l$. This inspires us to recompute the optimal estimate of the previous states and use them to generate the current estimate. That is the basic idea contained in Theorem 3.1, where we recompute the optimal estimate of $x_{k-D+1}, \ldots, x_{k-1}$ at time k and then make use of the updated estimates to compute the current estimate $\hat{x}^i_k$. Fig. 3 shows the overall estimation scheme at time k.

Theorem 3.1: For S_i in the undirected graph $\mathcal{G}$ with communication depth D, the optimal estimate $\hat{x}^i_k$ and its error covariance matrix P^i_k can be computed from D Kalman filters in sequence as

$$\begin{aligned}
&(\hat{x}^i_{k-D+1}, P^i_{k-D+1})\\
&\quad = \mathbf{KF}\Big(\hat{x}^i_{k-D}, P^i_{k-D}, \mathcal{B}^i_{k\,|\,k-D+1},\\
&\qquad C^i_{k\,|\,k-D+1}, R^i_{k\,|\,k-D+1}\Big)\\
&\quad\vdots\\
&(\hat{x}^i_{k-1}, P^i_{k-1})\\
&\quad = \mathbf{KF}\Big(\hat{x}^i_{k-2}, P^i_{k-2}, \mathcal{B}^i_{k\,|\,k-1}, C^i_{k\,|\,k-1}, R^i_{k\,|\,k-1}\Big)\\
&(\hat{x}_k, P_k)\\
&\quad = \mathbf{KF}\Big(\hat{x}^i_{k-1}, P^i_{k-1}, \mathcal{B}^i_{k\,|\,k}, C^i_{k\,|\,k}, R^i_{k\,|\,k}\Big)
\end{aligned}$$

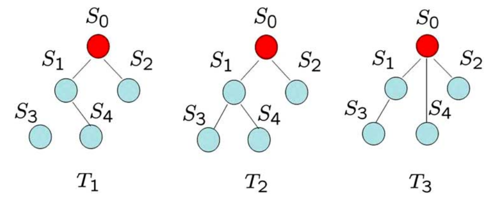

Fig. 4. Three sensor trees.

where $C^i_{k\,|\,k-l}, R^i_{k\,|\,k-l} (l = 0, \ldots, D-1)$ are the joint measurement matrix and measurement noise covariance matrix of the measurement data set $\mathcal{B}^i_{k\,|\,k-l}$. In case $\mathcal{B}^i_{k\,|\,k-l} = \emptyset$, $\mathbf{MKF}$ is used to update $\hat{x}^i_{k-l}$ and P^i_{k-l}.

Proof: We know that the estimate $\hat{x}^i_k$ is generated from the estimate of $\hat{x}^i_{k-1}$ together with all the available measurements at time k through a Kalman filter. Similarly, the estimate $\hat{x}^i_{k-1}$ is generated from the estimate of $\hat{x}^i_{k-2}$ together with all the available measurements for time $k-1$ at time k, etc. This recursion for D steps corresponds to the D Kalman filters stated in the theorem. ■

Remark 3.2: Notice that when implementing the algorithm in Theorem 3.1, each S_i only processes $\mathcal{B}^i_{k\,|\,k-l}, l = 0, \ldots, D-1$. As seen in Section II, $\mathcal{B}^i_{k\,|\,k-l}$ is obtained through communication with its one-hop neighbors, and no complete knowledge of the graph is needed. Therefore the estimation algorithm presented in Theorem 3.1 is fully distributed.

A. Kalman Filtering Over a Tree

In this section, we apply the estimation procedure in Theorem 3.1 to a directed tree that is rooted at S_i with depth D. The joint measurement matrix $C^i_{k\,|\,k-l}$ and noise covariance matrix $R^i_{k\,|\,k-l}$ in this case can be written as C^i_l and R^i_l respectively $(l = 0, \ldots, D-1)$. It is easy to verify that C^i_l and R^i_l satisfy the following:

$$C^i_0 = \Gamma^i_0, C^i_l = [C^i_{l-1}; \Gamma^i_l], \forall l = 1, \ldots, D-1$$

and

$$R^i_0 = \Upsilon^i_0, R^i_l = \mathrm{diag}(R^i_{l-1}, \Upsilon^i_l), \forall l = 1, \ldots, D-1$$

where Γ^i_l and Υ^i_l are the joint measurement matrix and the joint noise covariance matrix of those sensors that are exactly $l+1$ hops away from S_i. Following the optimal estimation algorithm over a graph in Theorem 3.1, we have the following result:

Corollary 3.3: Consider a sensor tree T_i with depth D_i that is rooted at S_i. If (A, C^i_{D-1}) is observable, then the steady-state error covariance matrix $\bar{P}^i$ satisfies

$$\bar{P}^i = \tilde{g}_{C^i_0} \circ g_{C^i_1} \circ \cdots \circ g_{C^i_{D-2}}(P^i_\infty) \tag{19}$$

where P^i_∞ is the unique solution to $g_{C^i_{D-1}}(P^i_\infty) = P^i_\infty$.

Proof: Equation (19) follows directly from (15) and (16). ■

For a given directed tree T with root at S_0, define

$$\mathcal{S}_{l-\text{hop}}(T) \triangleq \{S_i : S_i \text{ is within } l - \text{hops away from } S_0\} \tag{20}$$

for $l = 1, \ldots, D$. For example, in Fig. 4, $\mathcal{S}_{1-\text{hop}}(T_2) = \{S_1, S_2\}$, and $\mathcal{S}_{2-\text{hop}}(T_2) = \{S_1, S_2, S_3, S_4\}$.

Theorem 3.4: For two trees T_1 and T_2, if $\mathcal{S}_{l-\text{hop}}(T_1) \subset \mathcal{S}_{l-\text{hop}}(T_2) \forall l = 1, \ldots, D$, then $\bar{P}(T_2) \leq \bar{P}(T_1)$.

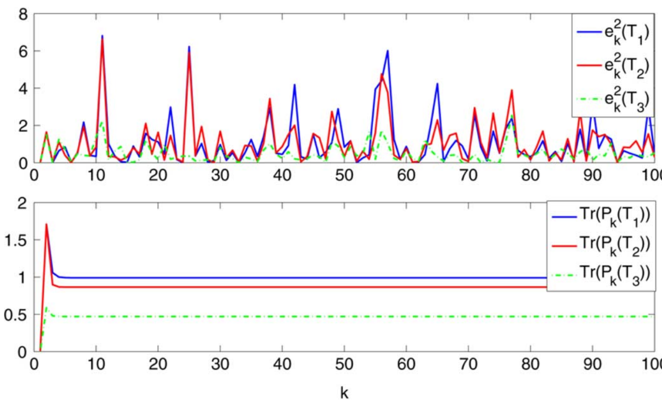

Fig. 5. Performance of the three sensor trees.

Proof: Since $\mathcal{S}_{l-\mathrm{hop}}(T_1) \subset \mathcal{S}_{l-\mathrm{hop}}(T_2) \forall l = 1, \ldots, D$, it is easy to verify that $g_{C_{l-1}(T_1)} \geq g_{C_{l-1}(T_2)}$ and $\tilde{g}_{C_{l-1}(T_1)} \geq \tilde{g}_{C_{l-1}(T_2)}$. Therefore the theorem follows immediately from (19). ■

Corollary 3.5: If $T_1 \subset T_2$, then $\bar{P}(T_2) \leq \bar{P}(T_1)$.

These results provide an easy way to compare the performance of different sensor trees.

Example 3.6: Consider the three sensor trees in Fig. 4. Apparently, $T_1 \subset T_2$, and $\mathcal{S}_{j-\mathrm{hop}}(T_2) \subset \mathcal{S}_{l-\mathrm{hop}}(T_3), l = 1, 2$, therefore from Theorem 3.4 and Corollary 3.5, we immediately obtain

$$\bar{P}(T_3) \leq \bar{P}(T_2) \leq \bar{P}(T_1).$$

This is indeed verified through the simulation in Fig. 5, where the following parameters are used:

$$A = \begin{bmatrix} 1 & 0.1 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} H_1 \\ H_2 \\ H_3 \\ H_4 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{bmatrix}, Q = \begin{bmatrix} 0.5 & 0 \\ 0 & 0.5 \end{bmatrix}$$

and $\Pi_i = 0.5 (i = 1, \ldots, 4)$.

B. From Packet Delay to Packet Drop

Consider the problem of state estimation over a packet-delaying network as seen from Fig. 6. The process dynamics is the same as in (1) and sensor measurement equation is given by

$$y_k = C x_k + v_k. \tag{21}$$

After taking a measurement at time k, the sensor sends y_k to a remote estimator for generating the state estimate. We assume that the measurement data packets from the sensor are to be sent across a packet-delaying network to the estimator. Each y_k is delayed by d_k times, where d_k is a random variable described by a probability mass function f, i.e.,

$$f(j) = \mathbf{Pr}[d_k = j], j = 0, 1, \ldots \tag{22}$$

We assume d_{k_1} and d_{k_2} are independent if $k_1 \neq k_2$, and the estimator discards any data y_k (or $\hat{x}_k^s$) that are delayed by D times or more.

Given the system and the network delay models in (1), and (21)–(22), we are interested in computing $\mathbf{Pr}[P_k \leq M]$, the probability that P_k is bounded by a given matrix $M \in \mathbb{S}_+^n$. The probabilistic metric was proposed in [10] for state estimation over packet-dropping networks.

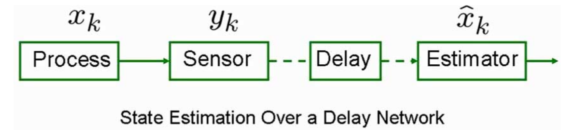

Fig. 6. Estimation over a packet-dropping network.

Let $\gamma_{k-i}^k = 1$ or 0 be the indicator function whether the measurement packet generated at time $k - i$ arrives at the estimator at time k or not. Define $\gamma_{k-i} \triangleq \sum_{j=0}^{i} \gamma_{k-i}^{k-j}$, i.e., γ_{k-i} indicates whether y_{k-i} is received by the estimator at or before k.

The recursive Kalman filtering technique from Theorem 3.1 dealing with delayed measurement provides a promising way to bridge the gap between packet drop analysis and packet delay analysis. The basic ideas is as follows. Since y_{k-i} may arrive at time k, we can improve the estimation quality by recalculating $\hat{x}_{k-i}$ utilizing the new available measurement y_{k-i}. Once $\hat{x}_{k-i}$ is updated, we can update $\hat{x}_{k-i+1}$ in a similar fashion. The following proposition summarizes the estimation process.

Proposition 3.7: Let $y_{k-i}, i \in [0, D-1]$ be the oldest measurement received by the estimator at time k. Then $\hat{x}_k$ is computed by $i+1$ $\mathbf{MKF}$s as

$$\begin{aligned}(\hat{x}_{k-i}, P_{k-i}) &= \mathbf{MKF}(\hat{x}_{k-i-1}, P_{k-i-1}, 1, y_{k-i}) \\ (\hat{x}_{k-i+1}, P_{k-i+1}) &= \mathbf{MKF}(\hat{x}_{k-i}, P_{k-i}, \gamma_{k-i+1}, y_{k-i+1}) \\ &\vdots \\ (\hat{x}_{k-1}, P_{k-1}) &= \mathbf{MKF}(\hat{x}_{k-2}, P_{k-2}, \gamma_{k-1}, y_{k-1}) \\ (\hat{x}_k, P_k) &= \mathbf{MKF}(\hat{x}_{k-1}, P_{k-1}, \gamma_k, y_k).\end{aligned}$$

Proof: Similar to the Proof of Theorem 3.1. ■

Define $\hat{\gamma}_i(D)$ as

$$\hat{\gamma}_i(D) \triangleq \begin{cases} \sum_{j=0}^{i} f(j), & \text{if } 0 \leq i < D, \\ \sum_{j=0}^{D-1} f(j), & \text{if } i \geq D. \end{cases}$$

Then it is easy to verify that

$$\mathbf{Pr}[\gamma_{k-i} = 1] = \hat{\gamma}_i(D). \tag{23}$$

Notice that now $\mathbf{Pr}[\gamma_{k-i} = 1]$ becomes a constant, thus given a stochastic description of the packet delays in (22), we can convert the packet delay model into a packet drop model. Similar to [11], we are then able to obtain similar bounds on $\mathbf{Pr}[P_k \leq M]$ using the corresponding new packet arrival rate $\hat{\gamma}_i(D)$.

IV. Discussions

In this technical note, we consider the problem of distributed estimation over sensor networks. We derive the optimal estimation algorithm at each sensor when the sensor communications are represented by an undirected graph. When the communication graph reduces to a directed tree, we also provide an exact expression on the steady-state error covariance matrices at each sensor. We show in Section IV how the previously developed theoretical tools can be applied to compare the performance of two sensor trees and convert a random packet-delay model to a random packet-dropping model.

There are many interesting directions that can be pursued along the line of this work. For example, if the sensors communicate with their neighbors using their state estimate instead of measurement data, how should the optimal estimation algorithm be in this case? What are the tradeoffs using the state estimate communication and using the measurement communication? Given a desired performance metric, for example, it is required that $\mathbf{Pr}[P_k^i \leq M] \geq 1 - \epsilon_i, \forall i$ for a given $0 < \epsilon_i \leq 1$, how should we determine the minimum number D? This is interesting as D determines the computational load at each sensor (i.e., running a chain of D Kalman filters at each time). If centralized control and coordination is allowed, what is the optimal communication graph for the sensors so that $\max_i P_k^i$ is minimized? Those problems will be pursued in the future.

REFERENCES

[1] N. P. M. , *Sensor Networks and Configuration*. New York: Springer, 2007.
[2] D. Culler, D. Estrin, and M. Srivastava, "Overview of wireless sensor networks," *IEEE Computer, Special Issue in Sensor Networks*, vol. 37, no. 8, pp. 41–49, Aug. 2004.
[3] B. Anderson and J. Moore*, Optimal Filtering*. Englewood Cliffs, NJ: Prentice Hall, 1990.
[4] D. Siljak*, Large-Scale Dynamic Systems: Stability and Structure*. New York: North-Holland, 1978.
[5] P. Alriksson and A. Rantzer, "Distributed kalman filtering using weighted averaging," in *Proc. 17th Int. Symp. Math. Theory Networks Syst.*, Kyoto, Japan, 2006, CD ROM.
[6] B. Sinopoli, L. Schenato, M. Franceschetti, K. Poolla, M. Jordan, and S. Sastry, "Kalman filtering with intermittent observations," *IEEE Trans. Automat. Control*, vol. 49, no. 9, pp. 1453–1464, Sep. 2004.
[7] A. Ray, L. W. Liou, and J. Shen, "State estimation using randomly delayed measurements," *J. Dyn. Syst., Meas. Control*, vol. 115, pp. 19–26, 1993.
[8] E. Yaz and A. Ray, "Linear unbiased state estimation for random models with sensor delay," in *Proc. IEEE Conf. Decision Control*, Dec. 1996, pp. 47–52.
[9] H. Zhang, L. Xie, D. Zhang, and Y. Soh, "A re-organized innovation approach to linear estimation," *IEEE Trans. Automat. Control*, vol. 49, no. 10, pp. 1810–1814, Oct. 2004.
[10] L. Shi, M. Epstein, A. Tiwari, and R. M. Murray, "Estimation with information loss: Asymptotic analysis and error bounds," in *Proc. IEEE Conf. Decision Control*, 2005, pp. 1215–1221.
[11] L. Shi, M. Epstein, and R. M. Murray, "Kalman filtering over a packet dropping network: A probabilistic approach," in *Proc. 10th Int. Conf. Control, Automat., Robot. Vision*, Hanoi, Vietnam, Dec. 2008.

On Structural Properties of the Lyapunov Matrix Equation for Optimal Diagonal Solutions

Sandip Roy and Ali Saberi

***Abstract*—We revisit the classical problem of finding a positive diagonal solution P to the Lyapunov equation $A^TP + PA < 0$ that minimizes the Lyapunov exponent (the maximum eigenvalue of $A^TP + PA$) for $A \in R^{n \times n}$, with the aim of identifying structural properties of the Lyapunov matrix equation at the optimum. Using eigenvalue sensitivity notions together with optimization machinery, we are able to obtain an explicit characterization of the minimum Lyapunov exponent that provides such structural insight.**

***Index Terms*—Lyapunov exponent.**

I. INTRODUCTION

The construction of positive diagonal solutions P to the Lyapunov equation $A^TP + PA < 0$, $A \in \mathcal{R}^{n \times n}$, has been of interest to the linear algebra and control systems communities (e.g., [1], [2])[1]. Researchers in these fields have been motivated by a range of applications, including analysis of economic systems over ranges of market speed adjustment rates [3], characterization of singularly-perturbed systems [4], analysis of interconnected systems [6], and design of neural networks [7], among others. Progress has been made in several fronts in this body of research:

1) Algorithms to check for the existence of a diagonal solution have been developed [2], [5]. More generally, Boyd has shown that the existence of structured solutions to the Lyapunov equation—including diagonal ones—can be found by solving a convex optimization [8].
2) Explicit necessary and sufficient conditions for the existence of a diagonal solution have been obtained in several special cases, including for $A \in R^{3 \times 3}$ (as developed by [9] in the linear algebra literature and revisited by [7] in the controls community), for block-triangular and normal A (e.g., [9]–[11]), and for A in the class of M matrices (see, e.g., [12]). More generally, the the existence of a diagonal solution has been equivalenced with other linear algebraic conditions (e.g., [12], [13]), leading to the explicit necessary and sufficient conditions as well as more general sufficient conditions.
3) The methodologies have been extended (often trivially) to minimize the *Lyapunov exponent*—the largest eigenvalue of $A^TP + PA$, which serves as measure of performance/robustness in several applications—with respect to P (see e.g., [2], [5]).

$\square$

In our recent studies of high-performance controller design for *communicating-agent networks* (e.g., teams of autonomous robots or networks of sensors), we also have encountered the problem of finding a diagonal solution that minimizes the Lyapunov exponent, among other

Manuscript received April 06, 2008; revised November 07, 2008. First published August 18, 2009; current version published September 04, 2009. This work was supported in part by the National Science Foundation under Grants ECS-0528882 and ECS-0725589, by the National Aeronautics and Space Administration under Grant NNA06CN26A, and by the Office of Naval Research under Grant N000140310848. Recommended by Associate Editor M. Fujita.

The authors are with Washington State University, Pullman, WA 99164 USA (e-mail: sroy@eecs.wsu.edu; saberi@eecs.wsu.edu).

Digital Object Identifier 10.1109/TAC.2009.2026852

[1]A matrix that admits such a diagonal solution is termed *diagonally Lyapunov stable* or alternatively *Volterra Lyapunov stable* in the literature.