On an Estimation Oriented Routing Protocol

V. Gupta

Abstract—Consider the problem of estimation of a linear time-invariant process across a communication channel such that the sensor data is delayed by a stochastic time-varying delay that can potentially be infinite. Thus, the data may arrive at the receiver delayed and out of order, or may simply be lost. There are two main contributions of this work. We show that performance of the estimator can not be characterized through a few moments of the delay distribution. Thus, conjectures such as a delay distribution with lesser mean or lesser maximum value always yields better estimation performance are incorrect. For a graph with each edge introducing a random delay to the sensor data, we also provide a routing algorithm that searches for the path that is optimal from an estimation theoretic perspective. A minor contribution of the work is to prove that even if the delay distribution has an infinite support and the delays are correlated across time with a finite memory, the necessary and sufficient condition for stability depends merely on the erasure probability.

I. INTRODUCTION

Networked control systems are now an active area of research (e.g., [1], [4] and the references therein). The performance of such systems is adversely affected by the detrimental effects such as random delays, data loss, data corruption, and so on introduced by the underlying communication network. To combat such performance degradation, two design architectures have emerged:

1) One block design: A compensator block at the output of the communication channel (co-located with the estimator or the controller) is designed to optimally compensate for the imperfections introduced by the communication channel. Such designs have been considered for communication channels that introduce data loss (e.g., [22]) or delay (e.g., [15]).

2) Two block design: Both an encoder at the input of the channel and a decoder at the output of the channel are designed. Such design have been considered for digital noiseless channels (e.g., [17]), analog erasure channels (e.g., [7]), and additive white Gaussian noise channels (e.g., [3]).

In this work, we consider the two block design problem of estimation across communication links that exhibit stochastic data delay. Since we allow the delay distributions to have infinite support, the links may also introduce random data loss. There is significant existing work on the effect of delay and loss on estimation and control. The effect of a constant delay is classically treated through concepts such as phase margin. Using stochastic models such as delays being independent and identically distributed (i.i.d.) from one time step to the next, or delays occurring according to a Markov chain, compensators for stochastically varying delays have been proposed in works such as [20], [15], [11], [24], [26]. Works such as [18], [25] can also be viewed as considering delays in networked control systems. However, most of these works consider delay distributions with finite support, and assume no packet reordering. Similarly control loops with data loss have also been extensively considered.

Within the one block design framework, various approaches to compensate for the data loss to counteract the degradation in performance have been proposed (see, e.g., [5], [20], [14], [23], [2], [12], [22], [6]). The two-block design paradigm has also been considered in works like [8], [7], [10] for channels between sensor and the controller, and [9], [13], [16] for channels between controller and the actuator. Some works [21], [19] have considered estimation across channels that introduce both delays and data loss. Schenato [21] showed that for delay distributions with a finite support and ignoring the possibility of packet reordering, the stability of the estimate error covariance depends only on the erasure probability. Robinson and Kumar [19] showed that if the sensor can transmit a vector of unbounded dimension at every time (the so-called long packet assumption), then i.i.d. delays do not impact the stability of the estimation error covariance.

In this work, we consider delay distributions that can have infinite support (thus introducing packet erasure) and can introduce packet reordering. We begin by presenting the solution to the two block design problem. There are two main contributions of this work:

1) We characterize the effect of delay on the performance and present some counter-examples to the intuitive results that the performance is necessarily improved by lowering the mean or the variance or the maximum of the delay distributions

2) For a graph with each edge introducing a random delay to the sensor data, we provide a Dijkstra-like routing algorithm that identifies the optimal path from an estimation theoretic perspective. In view of the negative result mentioned above, a routing protocol that minimizes average delay or maximum delay may not be optimal for estimation oriented performance.

A minor contribution of the work is to prove that even if the delay distribution has an infinite support and the delays are correlated across time with a finite memory, the necessary and sufficient condition for stability depends merely on the erasure probability.

V. Gupta is with the Department of Electrical Engineering, University of Notre Dame. Email: vgupta2@nd.edu. Research supported in part by NSF awards 0846631 and 0834771.
The paper is organized as follows. We begin in the next section by describing the problem framework. Section III provides the main results in the paper. In Section III-A, we provide the solution of the two block problem. In Section III-B, we derive the performance of the estimator in presence of delays and loss. Using this result, in Section III-C we prove that stability only depends on erasure probability even for correlated delays. Section III-D then provides counter-examples to prove that the first few moments are not enough to characterize the estimation performance, and provides a partial ordering of delay probability functions in terms of the estimation performance. Section III-E provides a Dijkstra-like algorithm for estimation oriented routing. Finally, we conclude with some avenues for future work in Section IV.

II. FRAMEWORK DESCRIPTION AND PROBLEM FORMULATION

Consider a process evolving as

\[ x(k + 1) = Ax(k) + w(k), \quad k \geq 0 \tag{1} \]

where \( x(k) \in \mathbb{R}^n \) is the process state and \( w(k) \) is the process noise assumed to be white, Gaussian, zero mean with covariance \( R_w > 0 \). The initial state \( x(0) \) is a zero mean and Gaussian random variable with covariance matrix \( P(0) \). The process state is observed using a sensor that generates measurements, or observations, of the form

\[ y(k) = Cx(k) + v(k), \quad k \geq 0 \tag{2} \]

where \( y(k) \in \mathbb{R}^m \) and the measurement noise \( v(k) \) is also assumed to be white, Gaussian, zero mean with positive definite covariance matrix \( \Sigma_v \). We assume that the pair \((A, C)\) is observable.

The sensor communicates with an estimator across a communication channel. In this paper, we consider the two block design problem, in which the designer specifies an encoder block situated at the input of the channel (collocated with the sensor) and a decoder block at the output of the channel (collocated with the estimator). At every time step \( k \), the sensor / encoder transmits information in the form of a \( n_d \)-dimensional real vector \( s(k) \), where \( n_d \) is a finite number, across the channel to the estimator. The information \( s(k) \) is a function of the measurements \( \{y(t)\}_{t=0}^{\infty} \) and the system matrices. The communication channel introduces a stochastic time-varying delay \( d(k) \) to information \( s(k) \) with packet reordering and erasure (equivalent to infinite packet delay) being possible. We assume the provision of a time-stamp so that the estimator knows the time \( k \) at which the sensor transmitted any vector \( s(k) \) that it receives at time \( k + d(k) \). While, in general, the delay suffered by various transmitted packets can be correlated, or dependent on the time when the transmission occurred; in this paper, we assume that the correlation has a finite memory and is time-invariant. Thus, if the correlation has memory \( t \),

\[
\text{Prob}(d(k) > n|d(k-1) > m_1, d(k-2) > m_2, \cdots, d(k-t) > m_t, d(0) > m_k) \tag{3}
\]

\[
= \text{Prob}(d(k) > n|d(k-1) > m_1, d(k-2) > m_2, \cdots, d(k-t) > m_t), \tag{4}
\]

for all \( k, n \) and \( \{m_i\} \) and this probability is independent of \( k \). Special cases of such a structure include \( t = 0 \), which implies independent and identically distributed delays, and \( t = 1 \), which implies delays according to a Markov chain. Finally, we assume that the sources of randomness \( x(0), \{d(k)\}_{k=0}^{\infty}, \{v(k)\}_{k=0}^{\infty}, \{w(k)\}_{k=0}^{\infty} \) are mutually independent.

Note that we do not assume that the encoder knows the delays suffered by the previous packets through some mechanism such as an acknowledgement. At every time \( k \), the estimator uses the outputs of the channel till time \( k \), i.e. the set \( \{s(t) : 0 \leq t \leq k, d(t) + t \leq k\} \), to generate an estimate \( \hat{x}_{dec}(k+1) \) of the process state \( x(k+1) \) to minimize the estimation error covariance \( P(k+1) = E[(x(k+1) - \hat{x}_{dec}(k+1))(x(k+1) - \hat{x}_{dec}(k+1))^T] \), where the expectation is taken over the initial condition, and the process and measurement noises. Thus, the estimator is a minimum mean squared error estimator. Due to the stochastic delays and erasures introduced by the channel, \( P(k) \) is a random variable. Thus, we need to characterize various moments of \( P(k) \). We concentrate on the expected value of \( P(k) \), although the techniques we use can be extended to arbitrary moments. If the term \( \lim_{k \to \infty} E[P(k)] \) is bounded, we say that the estimate error is stable. Given the above assumptions, we wish to analyze the stability and performance of the system as a function of the statistics of the delay for the optimal design of the encoder (the vector \( s(k) \)) and the estimator (the estimate \( \hat{x}_{dec}(k) \)). In particular, we wish to determine if the first few moments of the distribution (e.g. the mean and the variance) suffice to determine the error covariance.

Note that the delay \( d(k) \) in the above discussion may arise due to a single link or a network. If the sensor and the estimator are connected via a network of communication links that each introduce stochastic delay (e.g., due to queuing or medium access control), we also consider the problem of routing, i.e., identifying the optimal path in the network to transmit data. Let the network be represented by a graph \( G = (V, E) \) where \( V \) is the set of vertices that includes the sensor node \( s \) and the estimator node \( d \), and \( E \) is the edge set that represents the communication links joining the various nodes. Each edge has an associated delay distribution introduced by the corresponding communication link. We assume that the delays introduced by different links are mutually independent. The routing problem is to identify a path from \( s \) to \( d \) in \( G \) such that the steady state error covariance at the estimator is minimized if all packets transmitted by the sensor to the estimator follow that path.
At every time step \( k \), define a set \( D(k) \) as the set of all time steps such that the vectors \( s(j) \) transmitted at those time steps have been received by the decoder, plus the element -1, i.e.,

\[
D(k) = \{ j : j + d(j) \leq k \} \cup \{-1\}.
\]

We denote the maximal element in the set \( D(k) \) by \( t_s(k) \), i.e., \( t_s(k) = \max_j \) such that \( j \in D(k) \). Thus, \( t_s(k) \) defines the last time such that the vector \( s(t_s(k)) \) has been received at the decoder. Moreover, if no packet has been received, \( t_s(k) = -1 \).

### A. Optimal Encoder and Decoder

Denote the estimate of the state \( x(k) \) given the measurements \( \{ y(j) \}_{j=0}^T \) by \( \hat{x}(k) \), and consider the following encoder and decoder design that is an extension of the design proposed in [7] in the context of channels that introduce erasures but no stochastic delays:

- **Encoder Design:** At every time \( k \), the encoder calculates the estimate \( \hat{x}(k) \) of \( x(k) \) using a Kalman filter. It transmits \( \hat{x}(k) \) using a Kalman filter. The decoder does not receive any packet from the encoder. It updates the estimate at every time step \( k \), the encoder were to transmit at every time step \( k \), the estimate with a better error covariance than an estimate based on all measurements in the set \( D(k) \). Thus, even if the encoder were to transmit at every time step \( k \), all the information it has access to till time \( k \), the error covariance at the decoder is identical to the one that would be formed if the encoder had access to all the measurements in the set \( D(k) \).

- **Decoder Design:** The decoder calculates the estimate \( \hat{x}(k) \) of \( x(k) \) using the following recursive filter with the initial value \( \hat{x}_{\text{dec}}(0) = 0 \).

\[
\hat{x}_{\text{dec}}(k+1) = A\hat{x}_{\text{dec}}(k).
\]

- Of the (possibly multiple) packets received by the decoder, the packet with the maximal time-stamp has time stamp \( m \) (hence contains the vector \( \hat{x}(m) \)), and \( m \) is larger than the time stamp of any packet that the decoder has received till time \( k - 1 \). It sets

\[
\hat{x}_{\text{dec}}(k+1) = A^{k-m+1}\hat{x}(m).
\]

- Of the (possibly multiple) packets received by the decoder, the packet with the maximal time-stamp has time stamp \( m \), and \( m \) is lesser than the time stamp of at least one packet that the decoder has received till time \( k - 1 \). It ignores the packets and updates

\[
\hat{x}_{\text{dec}}(k+1) = A\hat{x}_{\text{dec}}(k).
\]

**Proposition 3.1:** (Optimality of the Encoder-Decoder Design.) Consider the problem formulation as stated in Section II. The encoder-decoder design given above leads to the estimate with the minimum error covariance at the decoder among all causal designs, at every time step.

**Proof:** The proof follows from two observations:

1. The above encoder decoder design leads to the estimate

\[
\hat{x}_{\text{dec}}(k+1) = \hat{x}(k+1)\{ y(j) \}_{j=0}^{t_s(k)},
\]

at every time \( k \). Thus, the estimate at the decoder is identical to the one that would be formed if the decoder had access to all the measurements in the set \( D(k) \).

2. No causal encoder decoder design can lead to an estimate with a better error covariance than an estimate based on all measurements in the set \( D(k) \). Thus, even if the encoder were to transmit at every time step \( k \) all the information it has access to till time \( k \), the error covariance cannot be lesser than that achieved by the encoder decoder design proposed above.

Note that the encoder-decoder design is optimal even though a vector with a constant dimension is recursively calculated and transmitted at every time step. A design that transmitted all measurements till time \( k \) would, on the other hand, involve increasing amount of data transmission. Moreover, the design provides the optimal estimate for an arbitrary realization of the delay process, irrespective of whether the delays are i.i.d. or correlated across time. The design also results in the optimal estimate at every time step for any realization of the delay process.

### B. Performance

Define by \( M(k+1) \) the error covariance of the estimate \( \hat{x}(k+1) \), if the error covariance of the estimate of the state \( x(k+1) \), if no packet has been received, the estimate is the error covariance of the estimate of state \( x(k+1) \), if no further measurement was received. Finally define

\[
f_m(S) = f(f(\cdots f(S) \cdots)),
\]

with the convention that \( f_0(S) = S \). The expected error covariance for estimate of state \( x(k+1) \) can now be calculated by conditioning it on the value of \( t_s(k) \). Since

\[
E[P(k+1)] = \sum_{m=1}^{k} \text{Prob}(t_s(k) = m)E[(x(k+1) - \hat{x}_{\text{dec}}(k+1)^T t_s(k) = m]
\]

\[
= \sum_{m=1}^{k} \text{Prob}(t_s(k) = m)M(m+1).
\]

If \( t_s(k) = m \geq 0 \), then the information transmitted at time \( m \) was delayed by no more than \( k - m \) time steps, and all information after that time suffered enough delay not to be received till time \( k \). If \( t_s(k) = -1 \), all the packets were delayed enough not to be received till time \( k \). Thus,

\[
\text{Prob}(t_s(k) = m) = \text{Prob}(d(k) > 0, \cdots, d(m+1) > k - m - 1, d(m) \leq k - m, m \geq 0)
\]

\[
\text{Prob}(t_s(k) = -1) = \text{Prob}(d(k) > 0, d(k-1) > 1, \cdots, d(0) > k).
\]
Given equations (5)-(6), the expected error covariance can be evaluated. In fact, the same technique can be used to obtain arbitrary moments of \( P(k + 1) \) at any time step. Thus, the entire probability mass function of the error covariance can be characterized. The result can also be used to constrain the delay guarantees that a communication network must provide for an acceptable level of error covariance performance.

C. Stability

We can now calculate the stability conditions. Due to Proposition 3.1, these conditions are necessary for stability with any encoder and decoder design. For the optimal encoder and decoder design presented in Section III-A, these conditions are both necessary and sufficient. The proof is omitted for space constraints.

Proposition 3.2 (Stability Conditions): Consider the problem formulation as stated in Section II. A necessary condition for stability with any causal encoder decoder design is that

\[
\limsup_{k \to \infty} (1 - \Pr(d(0) \leq k|d(t) > k - t, \ldots, d(1) > k - 1)) \rho(A)^2 < 1. \tag{7}
\]

Conversely, the encoder-decoder design specified in Section III-A leads to stability if (7) is satisfied.

This result implies that only the erasure probability can impact stability even if the delay distribution has infinite support. In particular, for a delay distribution that is independent and identically distributed, the necessary and sufficient condition for stability is

\[
\limsup_{k \to \infty} (1 - F(k)) \rho(A)^2 < 1,
\]

where \( F(k) \) is the cumulative distribution function (cdf) of the delay. For this special case our result agrees with the one in [19]. For a Markovian delay distribution, the necessary and sufficient condition for stability is

\[
q_{kk} \rho(A)^2 < 1,
\]

where

\[
q_{kk} = \limsup_{k \to \infty} \Pr(d(0) > k|d(1) > k - 1)).
\]

D. Characterization of Delay Distributions

In practice, it may be desirable to obtain a proxy of the delay distribution for its effect on the estimation error covariance. Thus, e.g., it may be supposed that a reasonable goal of a network protocol for optimizing the control performance is to minimize the expected delay or the maximum delay induced. In this section, we provide some counter-examples to show that such statements do not hold universally. For the rest of the paper, we assume that delays occur in an independent and identically distributed manner. For any value of the delay \( m \geq 0 \), the probability of delay being equal to \( m \) time steps is denoted by \( p(m) \).

Consider a scalar process evolving as

\[
x(k + 1) = x(k) + w(k), \tag{8}
\]

with noise covariance \( R_w = 1 \) and the covariance of the initial state \( P(0) = 1 \). The process is observed by a sensor

\[
y(k) = x(k) + v(k),
\]

with noise covariance \( \Sigma_v = 1 \). If all the measurements are available to the estimator without delay, the steady state error covariance \( M^* = 1.9522 \). Let the delays probability mass functions be

\[
p_1(m) : \Pr(d(m) = 5) = 1
\]

\[
p_2(m) : \Pr(d(m) = m) = \begin{cases} \frac{1}{2}, & m = 0 \\ \frac{1}{3}, & m = 1 \\ \frac{1}{3}, & m = 6. \end{cases}
\]

Both the probability mass functions \( p_1(m) \) and \( p_2(m) \) have mean \( = 5 \). However, \( p_1(m) \) has a smaller variance and a smaller maximum value for delay. The steady state expected error covariance for the two functions is given by \( E[P(k)] = 23.88 \) for \( p_1(m) \) and \( E[P(k)] = 17.34 \) for \( p_2(m) \). Thus the probability mass function with the higher variance (and the higher maximum value) has the lower cost. Instead, if the process evolves as

\[
x(k + 1) = 2x(k) + w(k),
\]

the probability mass function with the higher variance (i.e. \( p_2(m) \)) has the higher cost. Thus, there is no general relation between the variance or the maximum value of the delay and the cost achieved even if the means of two probability mass functions for the delays are the same.

Similarly, it is not necessarily true that a probability mass function of the delay with a higher mean leads to a higher expected error covariance. For instance, consider the same process as in (8) and independent and identically distributed delays, but with the probability mass functions

\[
p_1(m) : \Pr(d(m) = 2) = 1
\]

\[
p_2(m) : \Pr(d(m) = \begin{cases} \frac{1}{3}, & m = 0 \\ \frac{1}{3}, & m = 1 \\ \frac{1}{3}, & m = 6. \end{cases}
\]

The probability mass functions \( p_1(m) \) and \( p_2(m) \) have means 2 and 2.33 respectively. The steady state expected error covariance for the two functions is given by \( E[P(k)] = 6.49 \) for \( p_1(m) \) and \( E[P(k)] = 4.31 \) for \( p_2(m) \). Thus, lesser mean of the delay probability mass function does not necessarily lead to better performance.

We now present a partial characterization of probability mass functions of delay in terms of their effect on the estimation performance. For i.i.d. delays, the error covariance is given by

\[
E[P(k + 1)] = \sum_{m=0}^{k} \left( \prod_{j=0}^{k-m-1} \Pr(d(k) > j) \right) \Pr(d(k) \leq k - m) f_{k-m}(M(m + 1)) + \prod_{j=0}^{k} \Pr(d(k) > j) f_{k+1}(P(0)). \tag{9}
\]
Define a random variable \( r \in \{-1, \ldots, m\} \), called the modified delay, with the probability mass function
\[
\text{Prob}(r = -1) = \prod_{j=0}^{k} \text{Prob}(d(k) > j),
\]
(10)
\[
\text{Prob}(r = m) = \text{Prob}(d(k) \leq k - m) \prod_{j=0}^{k-m-1} \text{Prob}(d(k) > j), \quad \forall m = 0, \ldots, k.
\]

It can be verified that equation (10) defines a valid probability mass function. Using this auxiliary function, the performance can be rewritten as
\[
E[P(k+1)] = \sum_{m=-1}^{k} \text{Prob}(r = m)f_{k-m}(M(m+1)).
\]
Thus, the expected error covariance is the expected value of \( f_{k-m}(M(m+1)) \). However, the expectation is taken not with respect to the delay probability mass function \( p(m) \), but with respect to the probability mass function of the auxiliary random variable \( r \). The following result provides the desired characterization.

**Proposition 3.3:** Consider two modified delay distributions \( \{\text{Prob}(r_1 = m)\} \) and \( \{\text{Prob}(r_2 = m)\} \) corresponding to delay distributions \( p_1(m) \) and \( p_2(m) \) respectively. If \( \{\text{Prob}(r_1 = m)\} \) (first order) stochastically dominates \( \{\text{Prob}(r_2 = m)\} \), then the expected steady state estimation error covariance achieved under \( p_1(m) \) is no more than that achieved under \( p_2(m) \).

**E. Routing**

For the routing problem, we consider the cost function as steady state expected estimation error covariance. Although the cost over a route cannot be written as the sum of costs due to each edge in the route, the following result similar to the principle of optimality holds.

**Proposition 3.4:** Consider a network with the sensor located at node \( s \) and the estimator at node \( d \), and let \( n \) be any node on the optimal route \( S \). Let \( S_1 \) denote the portion of \( S \) from \( s \) to \( n \). If the estimator is located at node \( n \), let \( S_2 \) be the optimal route from \( s \) to \( n \). Then the expected estimation costs achieved by an estimator at node \( n \) by using \( S_1 \) or \( S_2 \) are identical.

**Proof:** Denote by \( \{\text{Prob}(r_1 = m)\} \) the modified delay distribution for an arbitrary route \( S_t \) from \( s \) to \( n \) and by \( \{\text{Prob}(r_2 = m)\} \) the modified delay distribution for the portion \( S_c \) of the route \( S \) from \( s \) to \( d \). At any time \( k \), let the error covariance achieved by an estimator situated at \( n \) when packets are being transmitted from \( s \) using route \( S_t \) be given by \( P_1(k) \). Let the packets from node \( n \) be transmitted to \( d \) over the route \( S_c \). Since the estimate at \( d \) is obtained by time updating the estimate at \( n \), the cost for an estimator located at node \( d \) that uses these packets is
\[
E[P_2(k)|\{P_1(k)\}] = \sum_{m=-1}^{k} \text{Prob}(r_2 = m)f_{k-m}(P_1(m+1)).
\]

The cost is conditioned on the values \( \{P_1(k)\} \) or in turn the modified delays \( \{r_1\} \). We can obtain the total cost by using the law of iterated expectation as
\[
E[P_2(k)] = E_{\{r_1\}}[\sum_{m=-1}^{k} \text{Prob}(r_2 = m)f_{k-m}(P_1(m+1))]
\]
\[
= \sum_{m=-1}^{k} \text{Prob}(r_2 = m)f_{k-m}(E_{\{r_1\}}[P_1(m+1)]),
\]
where the last equality follows since \( f_{k-m}(\cdot) \) is an affine operator. Since the operate \( f_{k-m}(X) \) is an increasing operator in the argument \( X \), this equation implies that to minimize the cost \( E[P_2(k)] \), the route \( S_t \) should be chosen so that the cost \( E_{\{r_1\}}[P_1(m+1)] \) is minimized. Since \( S \) is an optimal route, this implies that the cost \( E_{\{r_1\}}[P_1(m+1)] \) when \( S_t = S_1 \) is minimum over all possible routes from \( s \) to \( n \). But, by definition, \( S_2 \) is an optimal route from \( s \) to \( n \), hence the cost when \( S_t = S_1 \) or \( S_2 \) is also minimum over all possible routes from \( s \) to \( n \). Thus the costs whether \( S_1 \) or \( S_2 \) is chosen are equal.

The principle of optimality yields an algorithm based on dynamic programming (or equivalent shortest path algorithms such as Dijkstra’s algorithm) as follows.

1. Assign to every node a cost. The initial cost is \( M^* \) for node \( s \) and \( \infty \) for all other nodes.
2. Mark all nodes as unvisited. Set the node \( s \) as the current node.
3. For the current node, consider all its unvisited neighbors and calculate their cost as follows.
   - If the delay distribution from the node \( s \) to the current node is given by \( p_d \) and the distribution from the current node to the neighbor node \( n \) is given by \( s_d \), then the delay distribution from the node \( s \) to \( n \) is given by the convolution \( t_d = p_d \otimes s_d \).
   - Using the distribution \( t_d \), calculate the modified delay distribution from (10).
   - Calculate the cost for \( n \) using \( t_d \) and (11). If this cost is less than the previously recorded cost for \( n \), overwrite the cost.
4. When all the neighbors of the current node have been considered, mark it as visited. A visited node will not be checked ever again; its cost recorded now is final and minimal.
5. If all the one-hop neighbors of the destination node \( d \) have been marked as visited, then go to step 6; else set the unvisited node with the smallest cost as the next current node and continue from step 3.
6. The optimal path consists of all nodes chosen as current nodes in succession.

**IV. CONCLUSIONS AND FUTURE DIRECTIONS**

Control performance is dependent on both the throughput and the delay characteristics of the communication network. We study the problem of estimation and control of a linear time-invariant process across a communication channel that
introduces both a stochastic delay and erasure for the sensor data at every time step. The delay is time-varying and can cause re-ordering of the transmitted data packets. There are two main contributions of this work. We show that performance of the estimator can not be characterized through a few moments of the delay distribution. We also provide a routing algorithm that searches for the path in a network that is optimal from an estimation theoretic perspective. Finally, we prove that even if the delay distribution has an infinite support and the delays are correlated across time with a finite memory, the necessary and sufficient condition for stability depends merely on the erasure probability.

The work is only a first step towards systematic co-design of communication and control protocols. The characterization of the effect of delay on performance is not yet fully satisfactory. We still lack simple rules of thumb to compare any two delay distributions in terms of their effect on estimation performance. Delays in communication protocols are often caused by mechanisms such as acknowledgements and retransmissions. For real-time control, it would be interesting to see if such mechanisms are useful.

REFERENCES


