Delay-Reliability Tradeoffs in Wireless Networked Control Systems

Min Xie and Martin Haenggi

Department of Electrical Engineering
University of Notre Dame, IN 46556 USA
{mxie,mhaenggi@nd.edu}

***Abstract*—**

Networked control systems (NCS) require the data to be communicated timely and reliably. However, random transmission errors incurred by wireless channels make it difficult to achieve these qualities simultaneously. Therefore, a tradeoff between reliability and latency exists in NCSs with wireless channels. Previous work on NCSs usually assumed a fixed transmission delay, which implied that the failed packet will be discarded without retransmission, and thus reliability is reduced. When the channel errors are severe, the NCS cannot afford the resulting packet loss. In this paper, a delay-bounded (DB) packet-dropping strategy is associated with automatic repeat request (ARQ) at the link layer such that the failed packets will be retransmitted unless they are out-of-date. On the one hand, the packet delay is controlled to be below the predetermined delay bound. On the other hand, reliability will be improved because failed packets are given more retransmission opportunities. This paper investigates the tradeoff between packet delay and packet loss rate with DB strategies.

Due to the multihop topology of the NCS, medium access control (MAC) schemes are needed to schedule the transmission of multiple nodes. Moreover, spatial reuse should be taken into account to improve the network throughput. In this paper, two MAC schemes are considered, m-phase time division multiple access (TDMA) and slotted ALOHA. They are compared for different sampling rates and delay bounds. TDMA outperforms ALOHA in terms of both end-to-end (e2e) delay and loss rate when the channel reception probability is above 0.5 and/or traffic is heavy. However, ALOHA shows a self-regulating ability in that the node effective transmit probability depends only on the sampling rate and channel reception probability, but is essentially independent of the ALOHA-dependent parameters. Then, for light traffic, a simple ALOHA with transmit probability 1 is preferred over TDMA in NCSs. The derived relationship between the sampling rate, the e2e delay (or delay

bound), and the packet loss rate is accurate and realistic and can be used in NCSs for more accurate performance analyses.

1 Introduction

Networked sensing and control systems require real-time and reliable data communication. In the wireless environment, due to the random channel errors, it is difficult, if not impossible, to guarantee hard delay bounds with full reliability. Many applications of networked control systems (NCS) are delay-sensitive, but they can tolerate a small amount of data loss. Therefore, it is sufficient to provide NCSs a balanced guarantee between the delay and the loss rate.

Previous work on the QoS in NCS usually assumed that the packet transmission delay is fixed [2, 8, 9, 11]. In the practical wireless environment, this assumption hardly holds since the network-induced delay is random. In order to verify the assumption, in many previous works the packets are allowed only one transmission attempt. In [9,11], the wireless channel is modeled by a Bernoulli process with a success probability p_s. The failed packets are immediately discarded without any retransmission attempts. In this case, the packet delay is a constant equal to one time slot. But the network is rather unreliable. Reliability completely depends on the channel parameter p_s. If $p_s \ll 1$, the packet loss rate $p_L = 1 - p_s$ will be too large to be tolerated.

On the other hand, in the fully reliable network, in order to guarantee 100% reliability, the failed packets will be retransmitted until they are received successfully. Then, the resulting delay is tightly controlled by the channel parameter p_s. Similarly, as $p_s \ll 1$, the delay will be very long and cannot be tolerated by real-time applications. Meanwhile, more energy will be wasted to retransmit packets that are outdated and thus useless for the controller. In addition, from the perspective of network stability, the traffic rate (or the data sampling rate) is constrained to be smaller than the channel reception probability p_s. In an interference-limited network, p_s at the hotspot nodes are very small. Then, only light traffic can be accommodated by such networks.

A feasible solution to balance latency and reliability is to drop a small percentage of packets. A simple strategy is "finite buffer" (FB) [23]. If the buffer is full, some packets will be thrown out. To guarantee a hard delay bound, consider a *bounded delay* (BD) dropping strategy [23], in which the failed packets will be retransmitted until they are received correctly or their delay exceed a delay bound B. The maximum packet delay is guaranteed to be B. The dropping strategy is associated with the node scheduling algorithm to determine which packets are eliminated. For instance, the NCS applications always prefer the newly packets over the old packets. Therefore, using priority scheduling (high priority to new packets) or Last-Come-First-Serve (LCFS) scheduling, the old packets will be dropped to yield the buffer space for new packets.

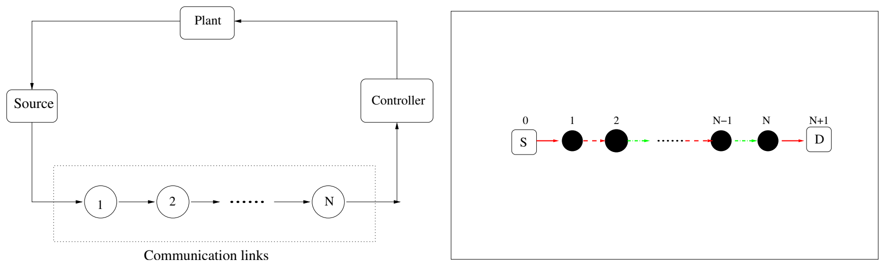
With respect to real-time NCS applications, the BD strategy is employed in this paper. The NCS using the BD scheme is referred to as Delay Bounded (DB)-network. Compared to the fully reliable network, the DB network has several advantages. First, network stability is not an issue. In case that the traffic load is too heavy to be accommodated by the channel, some packets will be discarded, and the network can be finally self-stabilized. Hence, a traffic rate higher than the channel reception probability p_s is allowed in the DB network. Second, less energy is wasted to retransmit packets that will be dropped eventually (refer these packets to as "marked" packets). As a matter of fact, given the multihop topology, the sooner the marked packets are dropped, the better. Third, in interference-limited networks, as the overall traffic load decreases, the channel reception probability p_s will be enhanced to admit more traffic. This feature is particularly helpful for NCS applications. At the cost of discarding outdated packets, more recent packets will be transmitted to the controller.

The disadvantage caused by the BD strategy is *unreliability*, which is measured by the packet loss rate p_L in this paper. However, with coding or network control packets, the transmission errors can be combated and reliability will be improved. Using B and p_L as the maximum delay and reliability and plugging into the controller [9, 11], a more accurate lower bound on the NCS performance can be obtained.

The system model for a NCS is outlined in Fig. 1(a), where the source (*e.g.*, sensor) data are transmitted over multiple wireless hops to the controller. Node 1 to N are relays. The data generated at the source node is time-critical, *e.g.*, periodic data used for updating controller output. A loop exists between the source, the controller and plant. The set of communication links is modeled as a tandem queueing network (Fig. 1(b)). Given the multihop topology, multiple nodes in tandem compete for transmission opportunities. Then MAC schemes are needed to schedule the node transmission order to avoid collision and take advantage of spatial reuse. If packets are of fixed length, the objective is to analyze the discrete-time tandem queueing network controlled by wireless MAC schemes.

1.1 Related work

The analysis of multihop MAC schemes in wireless networks essentially involves two issues, the *wireless MAC scheme* and the *tandem queueing system*. Previous work usually analyzed these two issues separately. The analysis on tandem networks focuses on the queueing delay, without consideration of the MAC-dependent access and backoff delay, and the retransmission delay induced by collision. For example, in [14, 15], the delay performance is derived assuming that the node transmits immediately if the server is free, and the transmission is always successful. In [6, 7], real-time queueing theory is proposed to explore real-time Jackson networks with Poisson arrivals and exponential servers for heavy traffic. Poisson arrival and exponential service



(a) Multihop NCS model (b) Regular line network

Fig. 1. System models

possess the memoryless property, which significantly simplifies the analysis from the perspective of queueing theory. However, most traffic models are not as memoryless as Poisson and cause more difficulties in the analysis, particularly when associated with MAC schemes like TDMA. Moreover, the traffic load in NCS is generally not heavy. The assumption of heavy traffic in [6, 7] leads to an underestimation of the network performance.

On the other hand, the analysis of MAC schemes concentrates on only the access and transmission delay while ignoring the queueing delay. For simplicity, it is usually assumed that the network has a single-hop topology, an infinite number of nodes, and a single packet at each node such that no queueing delay is incurred to the packet. More importantly, the traffic flow is generated in an unpractical way that every node always has packets to transmit whenever it is given a transmission opportunity. In [3, 21] study the access delay of Bernoulli and Poisson arrivals. In [18], the queueing delay was claimed to be derived with a special MAC scheme that allows one node to completely transmit all its packets as long as it captures the channel. For a specific node, the so-called queueing delay actually includes only access delay and transmission delay.

In the wireless environment, the analysis of MAC schemes additionally considers the impact of wireless channel characteristics. In [1, 22], a capture model is used to calculate the average packet transmission delay of ALOHA and CSMA for Rayleigh fading channels. [4] proposes optimum scheduling schemes for a line sensory network to minimize the end-to-end (e2e) transmission delay. With respect to the BD scheme, its effect upon a single node is discussed in [5]. In [17], various TDMA schemes combined with the BD scheme are studied in the single-hop scenario. This paper extends to the multihop scenario and studies the e2e performance.

1.2 Our contributions

This paper investigates the tradeoff between the end-to-end (e2) delay, reliability, and the sampling rate in multihop NCS. The main contribution is to *jointly* study the MAC scheme, the dropping strategy and the tandem network. The network performance like delay and packet loss rate is determined by several factors, including traffic, the routing protocol, the channel characteristics, and the MAC scheme. Emphasizing on the MAC scheme, we focus on a regular line network as shown in Fig. 1(b), which disposes of the routing and inter-flow interference problem. The obtained performance provides an upper bound for general two-dimensional networks because 1) the inter-flow interference is zero; 2) networks with equal node distances achieve better performance than those with unequal or random node distances.

Two MAC schemes are studied, m-phase TDMA and slotted ALOHA. In the former, every node is allocated to transmit once in m time slots, and nodes m hops apart can transmit simultaneously. In the latter, the node independently transmits with transmit probability p_m. The wireless channel is characterized by its reception probability p_s. The sampled source data (generated by the source in Fig. 1(a)) is modeled as constant bit rate (CBR). The CBR traffic flow is not only easy to be generated but also more practical than the traffic models used in previous work for MAC schemes, where the traffic load is assumed to be so heavy that the node is always busy transmitting. In practice, this heavy traffic assumption leads to an unstable network. There is no restriction on the node scheduling algorithm. To simplify the analysis, we assume First-Come-First-Serve (FCFS). Other scheduling algorithms like LCFS and priority scheduling can be chosen to better serve the NCS application demand.

We will show later that in the TDMA mode, the CBR arrival results in a non-Markovian queueing system, which substantially complicates the analysis. Despite of the enormous queueing theory literature, it is still difficult to track the transient behavior of non-Markovian systems [19], while the BD strategy is implemented based on the system transient behavior. Even if the source is deterministic and smooth, calculating the delay distribution in a time-dependent BD strategy is still a challenge. In addition, non-Poisson arrival causes correlations between the delays and queue lengths at individual nodes. Closed-form solutions exist only for some special networks, like Jackson networks, which do not include the network in this paper. In one word, accurate analyses are almost impossible in a multihop network with a long path. Therefore, we start the analysis with the first node, then investigate the network performance through simulation results.

Combining MAC with the BD strategy, we compare DB-TDMA and DB-ALOHA in terms of the delay and the packet loss rate. The DB-MAC networks are also compared with their non-dropping counterparts to exhibit the advantage of the BD strategy. Since the traffic intensity is not necessarily as heavy as close to 1, ALOHA possesses a self-regulating property, which does not exist in the heavy traffic assumption. That is, although ALOHA attempts to control the packets transmission with the transmit probability p_m, the node actually transmits with a probability p_t, which is independent of p_m.

This paper is organized as follows. Section 2 describes the system model and presents some results for the non-dropping MAC schemes. Then, TDMA and ALOHA are studied in Section 3. The self-regulating property of ALOHA is specially discussed. Their performance are compared through a set of simulation results in Section 4. The paper is concluded with Section 5.

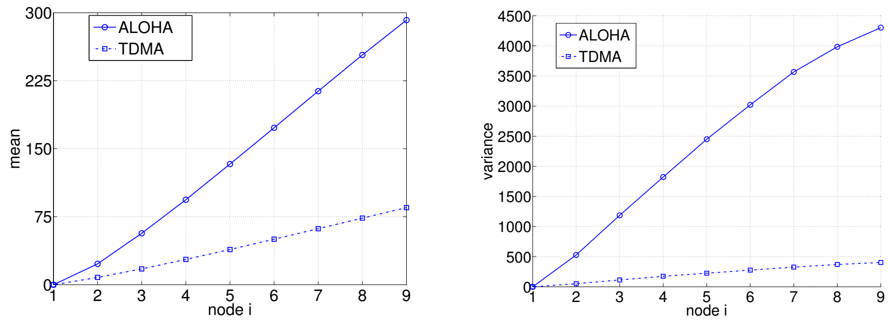
2 System Model

The set of communication links in NCS (Fig. 1(a)) is modeled as a tandem queueing network, which is composed of a source node (node 0), N relay nodes, and the unique destination node (node $N+1$, also referred to as the sink node). A CBR flow of fixed-length packets is generated at the source node. Relay nodes do not generate traffic. Time is divided into time slots. One time slot corresponds to the transmission time of one packet. The sampling rate of CBR traffic is $1/r$, $r \in \mathbb{N}$. Given the channel reception probability p_s, the channel is modeled as a Bernoulli process with parameter p_s. The e2e delay bound is B.

From the perspective of energy efficiency, it is not recommended to drop the packet only when its e2e delay exceeds B, which often happens when the packet reaches the last few nodes to the sink. The longer the route the packet traverses, the more energy is wasted if the packet will be eventually dropped. So, a local BD strategy is preferred. In order to determine how the local BD strategy is implemented, we first review the cumulated delay distribution in non-dropping tandem networks.

In [20], a fully reliable tandem queueing network is studied. CBR Traffic is transformed to correlated and bursty through the error-prone wireless channels. Even with correlation, the e2e delay is approximately linear with the number of nodes with respect to both the delay mean and delay variance, as confirmed by simulation results in Fig. 2. Then, it is reasonable to uniformly allocate B among nodes based on their relative distances to the source node [16], *i.e.*, the local delay bound D_i is set to be $D_i = iB/N = iD$ ($i \in [1, N]$, $D_i \in \mathbb{N}$). Packets are dropped at node i if their cumulated delay exceeds D_i. Intuitively, if a node experiences a delay at node i longer than D_i, then it is highly possible that it has delay longer than B at the final node N. Parameters of interest include the cumulated delay d_i and the packet loss rate p_L^i at node i ($1 \le i \le N$).

As proved in [20], the e2e delay mean of ALOHA is about $p_s/(1-p_s)$ times than that of TDMA. The gap is even larger for the delay variance as shown in Fig. 2(b). However, in the wireless multihop network, TDMA is not feasible to be implemented, and simple MAC schemes like ALOHA are more desirable, even though TDMA substantially outperforms ALOHA in terms of both throughput and delay. The BD strategy is a solution to reduce the performance gap between TDMA and ALOHA.



(a) delay means (b) delay variances

Fig. 2. Comparison of delay performance in the TDMA and ALOHA network with $m = 3, r = 4, p_s = 0.8, N = 8$

Note that in the DB network, the packets are dropped according to their delays. Conventional queueing theory keeps track of the buffer size and cannot capture the packet dropping event [23]. So, we use a delay model [5], in which the system state is denoted by the packet delay such that the delay-dependent packet dropping event can be directly depicted through the system state.

3 Delay Bounded Wireless Multihop Networks

3.1 m-phase TDMA

The m-phase TDMA scheduler takes advantage of spatial reuse so that nodes $i, m+i, 2m+i, \ldots$ $(1 \le i \le m)$ can transmit simultaneously. A transmission can be either a transmission of a new packet or a retransmission of a failed packet. Instead of being divided into time slots, now the time is divided into frames of m time slots. For a node, the beginning of a frame is the beginning of the time slot allocated to this node. The transmission rate is $1/m$, and the transmission is successful with probability p_s. To guarantee system stability, $r > m$. For heavy traffic, we assume $m < r < 2m$. At the frame level, the service time is geometric with p_s. We start with the first node since it determines the traffic pattern of all subsequent nodes.

At the frame level, the interarrival time $1 < r/m < 2$ is not an integer even though it is a constant. As a matter of fact, the number of packet arrivals in one frame jumps between 0 and 1, depending on the arrival pattern of all previous

$r-1$ frames. This dependence makes the resulting queueing system more complicated than the G/M/1 system, where the interarrival times are identical and independent. Hence, a standard Markov chain cannot be established as usual to keep track of the buffer size. Instead, we resort to the delay model that denotes the delay of the Head of Line (HOL) packet as the system state [5].

The system state is the waiting time of the HOL packet in terms of time slots, but the state transitions happen at the frame boundaries. Since the waiting time is the difference between the present time and the packet arrival time, the state value might be negative when the queue is empty and the next packet arrival does not happen. The absolute value of the negative state represents the remaining time till the next packet arrival. With a constant interarrival time r, the transition probability matrix $\mathbf{P} = \{P_{ij}\}$ is:

$$P_{ij} = \begin{cases} p_s & 0 \le i \le D-m, \quad j = i - \Delta, \\ q_s & 0 \le i \le D-m, \quad j = i + m, \\ 1 & D-m < i \le D, \; j = i - \Delta \;\text{ or }\; i < 0, \; j = i + m, \end{cases} \tag{1}$$

where $\Delta := r - m > 0$. At frame t, let the HOL packet be packet k and its waiting time $w_k(t)$. If the transmission is successful, packet k departs at frame t, and the subsequent packet $k+1$ becomes the HOL packet at frame $t+1$. The waiting time of packet $k+1$ at frame t is $w_{k+1}(t) = w_k(t) - r$. It increases by m up to $w_{k+1}(t+1) = w_k(t) - r + m = w_k(t) - \Delta$ at frame $t+1$. Therefore, the system state transit from $w_k(t)$ to $w_k(t) - \Delta$ with probability p_s. If $w_k(t) < \Delta$, packet k is the last packet in the buffer and the buffer becomes empty after its transmission. Then, the system transits to a negative state $i = -(\Delta - w_k(t)) < 0$. For $m < r < 2m$, the server idle time does not exceed one frame. Then, there must be a packet arrival during frame $t+1$. This new packet may arrive in the middle of frame $t+1$ and cannot be transmitted immediately. The waiting time to access the channel is $m + i > 0$. Then the negative state i transits to a positive state $m + i$ with probability 1.

If the transmission is failed and $w_k(t) \le D - m$, the HOL packet remains at the buffer and will be retransmitted after one frame. Its delay increases by m up to $w_k(t+1) = w_k(t) + m \le D$ with probability q_s. If $w_k(t) > D - m$, this HOL packet k will experience a delay greater than D after one frame and be discarded (maybe in the middle of the frame). Then, at the beginning of frame $t+1$, packet $k+1$ becomes the HOL packet with a delay $w_{k+1}(t+1) = w_k(t) - \Delta$. Recall that if the transmission is successful, the positive state $w_k(t)$ transits to state $w_k(t) - \Delta$, as well. In other words, if $w_k(t) > D - m$, the system state always transits to $w_k(t) - \Delta$ with probability 1, regardless of whether the transmission is successful or failed.

The steady-state probability distribution$\{\pi_i\}$ can be obtained either iteratively or by using mathematical tools to solve $\pi = \pi\mathbf{P}$. For a critical case $\Delta = 1$, $\{\pi_i\}$ is derived in terms of π_0 as follows ($q_s \triangleq 1 - p_s$):

$$\pi_i = \begin{cases} \frac{\pi_0}{p_s^i}\left(1 + q_s \sum_{k=1}^{K_i}(-q_s p_s^m)^k g(k)\right) & i \le D-m \\ q_s \sum_{j=i-m}^{D-m} \pi_j & i > D-m \end{cases} \tag{2}$$

where

$$g(k) = \binom{i-km+1}{k} - \binom{i-km}{k} p_s, \ K_i = \left\lfloor \frac{i+1}{m+1} \right\rfloor$$

If the HOL packet is transmitted successfully, its delay at the first node is $w_k(t)$ plus one time slot for transmission. Therefore, the delay distribution $\{d_i\}$ is completely determined by the probabilities of non-negative states,

$$d_i = \frac{\pi_{i-1}}{\sum_{j\ge 0} \pi_j}. \tag{3}$$

In the m-phase TDMA network, the node transmits at the frame boundaries. A packet may be dropped in the middle of a frame if it experiences a delay greater than $D-m$ at the beginning of this frame and fails to be transmitted. The packet dropping probability is

$$p_L^{(1)} = \frac{q_s \sum_{i=D-m+1}^{D} \pi_i}{\lambda}, \quad \lambda = \frac{m}{r} \tag{4}$$

By observing the balance equations, we obtain

$$\pi_i = \begin{cases} q_s \sum_{j=1}^{i} \pi_j + \pi_0 & 1 \le i < m \\ q_s \sum_{j=i-m}^{k} \pi_j & i \ge m, \end{cases} \tag{5}$$

where $k = \min\{i, D-m\}$. (5) holds for the delay distribution $\{d_i\}$ as well. Apparently, d_i is jointly determined by $l = \min\{i, m, B-i\}$ consecutive states below i, which results in a backward iteration. If $D-m > m$, the probability mass function (pmf) of the first node delay is composed of three sections, $[1, m]$, $[m+1, D-m+1]$, and $[D-m+2, D+1]$. If $D < 2m$, the pmf will be simpler. Since a smaller D causes a higher dropping probability, a general condition for the delay constraint is $D > 2m$ to ensure that every packet has at least one transmission opportunity. Because both m and D are positive integers, the smallest possible value of D is $2m+1$.

Note that as $D \to \infty$, [20] has shown that the output of the first node is a correlated on-off process. This correlation exists even if $D < \infty$. Then, the following relay nodes are fed with bursty and correlated traffic, which makes it difficult to analyze the resulting network. So, for $D < \infty$, the network performance is investigated through simulation results.

The e2e delay is the sum of all local delays. The delay mean μ and the dropping probability p_L can be upper bounded as follows:

$$\mu = \sum_{i=1}^{N} d_i \leq N\mu_1 \tag{6}$$

$$p_L = 1 - \prod_{i=1}^{N}(1 - p_L^{(i)}) \leq 1 - (1 - p_L^{(1)})^N. \tag{7}$$

The tightness of these upper bounds depends on $p_L^{(1)}$. The fewer packets are dropped, the closer $p_L^{(i)}$ to $p_L^{(1)}$, and the tighter is the bound.

3.2 Slotted ALOHA

In slotted ALOHA, each node independently transmits with probability p_m. Note that p_m represents the node transmission opportunity. The node actually transmits only if it is given a transmission opportunity and it has packets to transmit, which depends on its buffer occupancy. Traffic and the channel model are the same as in TDMA. Again, we start with the first node, which is observed at the time slot level. Given the success probability p_s, a packet departs the node if and only if the node is scheduled to transmit and the transmission is successful, with a probability $a = p_s p_m$. Otherwise, the packet is retained in the buffer or discarded. The service time is geometric with a. The system state is the waiting time of the HOL packet. The state transition probabilities are

$$P_{ij} = \begin{cases} 1 - a & i \in [0, D),\ j = i + 1 \\ a & i \in [0, D),\ j = i - r + 1 \\ 1 & i < 0, \quad\ \ j = i + 1 \quad \text{or } i = D,\ j = D - r + 1. \end{cases} \tag{8}$$

At slot t, assume that the HOL packet is packet k with delay $w_k(t)$. If $w_k(t) < D$ and the packet successfully departs the node with probability a, the next packet becomes the HOL packet at slot $t+1$ with delay $w_{k+1}(t+1) = w_{k+1}(t) + 1 = w_k(t) - r + 1$. Otherwise, if the packet fails to depart the node with probability $(1 - a)$, it remains as the HOL packet with its delay increased by one. If $w_k(t) = D$, either the packet is transmitted successfully or not, it has to be deleted from the buffer since its delay exceeds the bound D at slot $t+1$. In this case, the system state transits from D to $D - r + 1$ with probability 1. The negative states indicate an empty buffer and the system is waiting for the next new packet arrival.

The delay distribution $\{d_i\}$ is calculated based on $\{\pi_i\}$ and (3) like in TDMA. Since the packet dropping possibly occurs at the time slot boundaries, the packet dropping probability is

$$p_L^{(1)} = \frac{(1 - a)\pi_D}{\lambda} = r(1 - p_s p_m)\pi_D. \tag{9}$$

Rewriting the balance equations, we obtain

$$\pi_i = \begin{cases} a\sum_{j=k}^{i+r-1}\pi_j & i < D-r+1 \\ (1-a)\pi_{i-1} & D-r+1 \leq i \leq D, \end{cases} \tag{10}$$

where $k = \max\{0, i\}$. (10) holds for $\{d_i\}$, as well. Different from the TDMA network, (10) exhibits that d_i is jointly determined by $r-1$ consecutive states above i, which results in a forward iteration. The pmf essentially consists of three sections, $[1, D-2r+2]$, $[D-2r+3, D-r+1]$, and $[D-r+2, D+1]$. A general setting is $D > 2(r-1)$, so that the pmf contains all three sections.

Note that the node transmits with probability p_m when its buffer is non-empty. In other words, the effective transmit probability is $p_t = P_B p_m$, where P_B is the node busy probability. In previous work, the traffic load is so heavy that $P_B = 1$. Then the effective transmit probability p_t is identical to the transmit probability p_m so that the performance of ALOHA networks can be optimized by manipulating p_m. However, if the traffic load is light and $P_B \ll 1$[1], which is highly possible in ALOHA networks, simply optimizing p_m does not necessarily lead to optimization of the network performance.

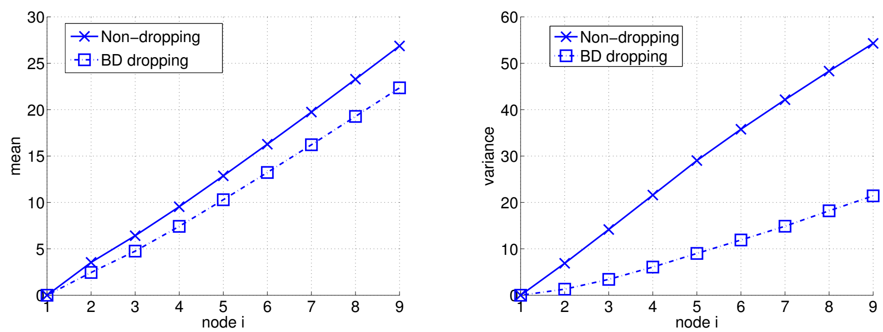
As a matter of fact, based on queueing theory, the busy probability of node i is $P_B^i = \lambda_i/(p_s p_m)$, where λ_i is the arrival rate to node i. As the delay bound D goes to infinity, it is easy to show that $\lambda_i = 1/r$ for all i and thus $P_B = 1/(rp_s p_m)$. Naturally, the effective transmit probability $p_t = 1/(rp_s)$ depends only on the traffic rate $1/r$ and the channel reception probability p_s, and is completely independent of p_m. Since the network performance is essentially determined by the effective transmit probability p_t, this observation implies that the ALOHA parameter p_m does not contribute to the change of the network performance. In the DB-ALOHA network, due to packet dropping, the arrival rate λ_i to node i is a function of the loss rate $p_L^{(i)}$, $\lambda_i = \lambda_{i-1}(1-p_L^{(i)}$. The packet loss rate $p_L^{(i)}$ depends on the delay bound D, the service rate $a = 1/(p_s p_m)$ and the arrival rate λ_{i-1}. Through intuition, p_t is not completely independent of p_m. However, in most cases the loss rate is required to be relatively small. Then $P_B^i \approx 1/(rp_s p_m)$, and it is reasonable to say that p_t is independent of p_m.

From the analysis in [20] we can see that the longer delay of ALOHA is mostly caused by the longer access delay, which is proportional to $1/p_m$. When the traffic load is light, a large p_m will lead to a small access delay and significantly improve the e2e delay. In this sense, to optimize the delay performance of ALOHA, p_m should be chosen based on the traffic load, which has been ignored in previous work because of the heavy traffic assumption. For example, in [10], the throughput is maximized by $p_m = 1/N$ without consideration that the node does not need contend for transmission opportunities when its buffer is empty.

[1] P_B is essentially a function of p_m.

4 Simulation Results

A set of extensive simulation results are provided to expose the performance of the DB-TDMA and DB-ALOHA network. First of all, the fully reliable (non-dropping) TDMA network is compared with the DB-TDMA network in Fig. 3. We set $m = 3$ and $p_s = 0.8$ for both networks. For the DB network, we additionally set the interarrival time $r = 4$ and $D = 4$, which results in an e2e dropping probability $p_L \approx 0.20$. It implies that 20% packets will be discarded and the throughput is $(1 - 0.2)/r = 1/5$. In the non-dropping network, all generated packets will be successfully delivered to the sink and the throughput is exactly the traffic rate $1/r$. Accordingly, for the non-dropping network, we set the interarrival time $r = 5$ such that both networks are compared under the identical throughput.



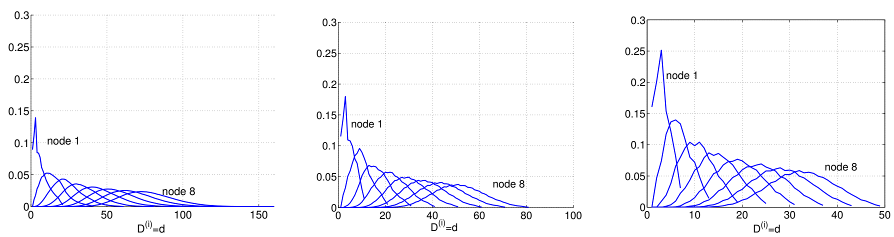
(a) Delay means (b) Delay variances

Fig. 3. Comparison of the non-dropping and DB-TDMA network with $m = 3, N = 9, p_s = 0.8$

With the BD strategy, the e2e delay decreases. A substantial improvement is particularly reflected on the delay variance that is reduced by 60%. On the one hand, simply reducing the traffic load at the first node does not improve the network performance significantly. On the other hand, although introducing redundant packets does increase the traffic load, it enhances the delay performance. In addition, the lost packets can be compensated for by the redundant packets, which ensures reliability. In this sense, the BD strategy is very helpful to achieve a good balance between latency and reliability.

4.1 m-phase TDMA network

Throughout this section, the m-phase TDMA network is assumed to have $m = 3, r = 4, N = 8$. Compared to the pmf of the non-dropping TDMA network, the pmf of the cumulated delays from node 0 to node i ($1 \leq i \leq N$) is truncated based on B/D (Fig. 4) . For $D > 2m$, the pmf of the cumulated delay is scaled. In Fig. 5, the e2e delay mean and variance are shown for $p_s = 0.8, D = 10$. The delay mean approximately linearly increases with the number of nodes. In comparison with the non-dropping TDMA network ($D = \infty$), the delay mean is reduced by 40% and the delay variance by 75%. The resulting e2e packet loss rate $p_L = 0.0414$ (as listed in Table 4.1) is acceptable when the aforementioned packet-level coding scheme is applied to combat the network unreliability.



(a) $D = 20$ (b) $D = 10$ (c) $D = 6$

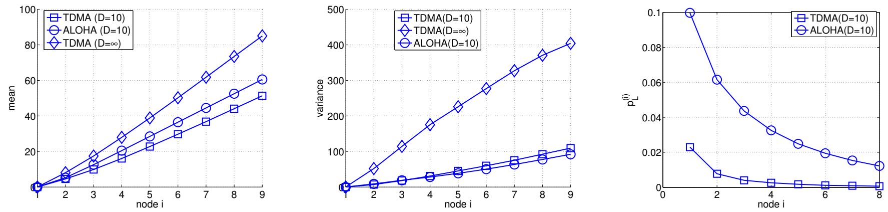
Fig. 4. pmf of the cumulated delay in the TDMA network with $m = 3, r = 4, p_r = 0.8, N = 8$

p_s	D				
	1	5	10	15	20
0.7	0.9615	0.3742	0.2183	0.1678	0.1431
0.75	0.9331	0.2621	0.1229	0.0803	0.0589
0.8	0.8880	0.1550	0.0414	0.0142	0.0058

Table 1. E2e dropping probabilities of the TDMA network with $m = 3, r = 4, N = 8$

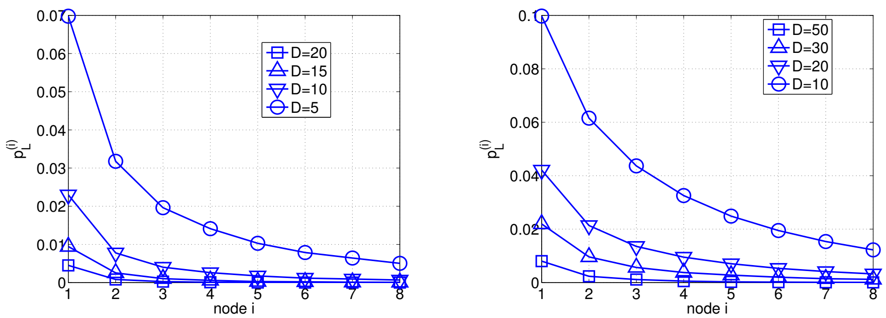
The dropping probability is to be traded off against the delay. A smaller D results in a smaller delay and a higher local dropping probability $p_L^{(i)}$ at node i, which is shown in Fig. 6(a). As D increases, $p_L^{(i)}$ decreases more slowly. It implies that the major packet loss occurs at the first few nodes of the chain. This property is desirable since the downstream nodes do not need to spend energy to transmit packets that will finally be discarded. For large D, the per-node dropping probability asymptotically converges to zero.

Fig. 7 demonstrates the effects of D. Both the delay mean and the delay variance are nearly linear in D, particularly when D is small. Unsurprisingly,



(a) Delay means (b) Delay variances (c) Dropping probability

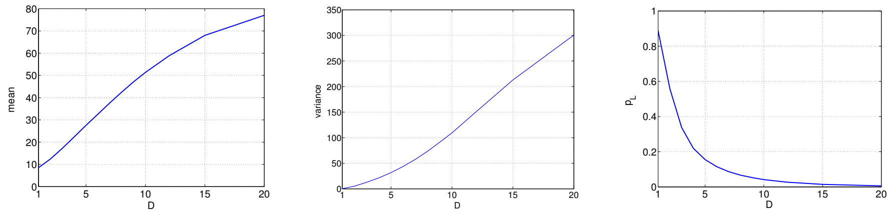
Fig. 5. Performance comparison of the system with $m = 3, r = 4, N = 8, p_s = 0.8, D = 10$



(a) TDMA (b) ALOHA

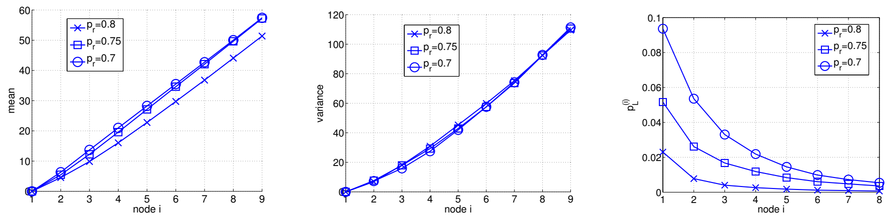
Fig. 6. Packet dropping performance for the system with $m = 3, r = 4, N = 9, p_s = 0.8$

the e2e dropping probability p_L asymptotically decreases with D. For large D, say $D > 10$, the decrease of p_L becomes very slow. Thus, simply increasing D does not help to improve reliability, but does harm the delay performance. There may exist an optimal D to achieve the best balance between the delay and the packet loss. Unlike in fully reliable networks [20], the delay performance is not severely deteriorated by the drop of p_s (Fig. 8). Moreover, as long as D is sufficiently large, even if the traditional stability condition does not hold, the resulting e2e dropping probability is so moderate that both the data latency and reliability are guaranteed. For instance, considering the critical case $p_s = m/r = 0.75$, for $D \geq 20$, the packet loss $p_L \leq 0.05$ is negligible. For small D like $D = 10$ and $p_L \leq 0.13$, it is not difficult to introduce redundant packets for reliability.



(a) Delay means (b) Delay variances (c) p_L

Fig. 7. The impact of D in the m-phase TDMA network with $m = 3, r = 4, N = 8, p_r = 0.8$



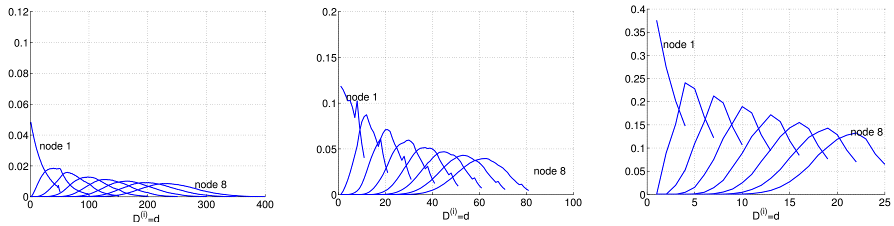
(a) Delay means (b) Delay variances (c) $p_L^{(i)}$

Fig. 8. The impact of p_s in the m-phase TDMA network with $m = 3, r = 4, N = 9, D = 10$

4.2 Slotted ALOHA

This section discusses the performance of the DB-ALOHA. To compare with TDMA, we set $p_m = 1/m$ and $m = 3, r = 4, N = 8$. Different from the TDMA system, the pmf tail is both truncated and twisted by applying the BD scheme (Fig. 9). But the delay central moments and the dropping probability behave similarly as in TDMA. Specifically, the delay mean and variance linearly increase with the number of nodes (Fig. 5(a) andFig. 5(b)), and the per-node dropping probability $p_L^{(i)}$ diminishes with the node index i and the first node experiences the maximum packet loss (Fig. 5(c) and Fig. 6(b)). The e2e dropping probability p_L is listed in Table 2.

Like the TDMA network, both the delay mean and variance are approximately linear with D (Fig. 10(a) and Fig. 10(b)). The dropping probability p_L (Fig. 10(c)) sharply decreases with small D. However, for D sufficiently large, say $D \geq 30$, the decreasing speed is decelerated and p_L eventually converges to zero. Apparently, a larger D is needed for p_L to reach zero. The impact of p_s is displayed in Fig. 11. Again, the drop of p_s causes a very small



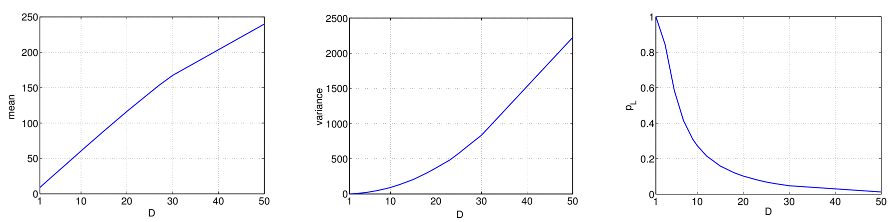
(a) $D = 50$ (b) $D = 10$ (c) $D = 3$

Fig. 9. pmf of the cumulated delay in the ALOHA network with $m = 3, r = 4, p_r = 0.8, N = 8$

p_s	D						
	1	10	20	30	40	50	100
0.70	0.9999	0.4243	0.2505	0.1911	0.1599	0.1423	0.1048
0.75	0.9999	0.3481	0.1784	0.1164	0.0863	0.0692	0.0341
0.80	0.9998	0.2731	0.1020	0.0477	0.0238	0.0127	0.0000

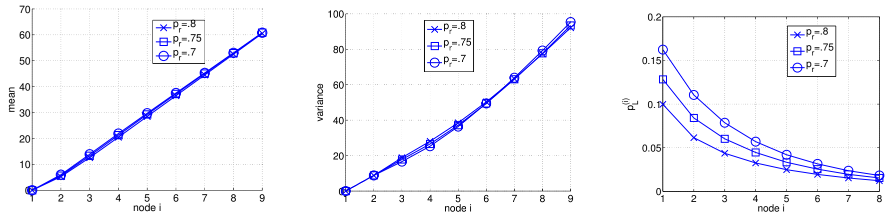
Table 2. E2e dropping probabilities of the ALOHA network

difference in the delay mean and variance, but results in an increase of the per-node packet dropping rate. Moreover, the per-node dropping probability asymptotically converges to zero.



(a) Delay means (b) Delay variances (c) p_L

Fig. 10. The impact of D in the ALOHA network with $m = 3, r = 4, N = 8, p_r = 0.8$



(a) Delay means (b) Delay variances (c) $p_L^{(i)}$

Fig. 11. The impact of p_s in the ALOHA network with $m = 3, r = 4, N = 8, D = 10$

4.3 Comparison

In the fully reliable network, the TDMA network substantially outperforms the ALOHA network in terms of the delay (Fig. 2). In the DB network, the performance gap between TDMA and ALOHA becomes fairly small. As shown in Fig. 5, the difference of the delay mean between the TDMA and ALOHA network decreases from 300% (reliable) to 20% (DB); while the delay variance difference changes from 750% (reliable) to 10% (DB). The main performance degradation caused by the random access is the dropping probability. For ALOHA, the local dropping probability $p_L^{(i)}$ at node i is almost five times than that of TDMA. Moreover, $p_L^{(i)}$ of ALOHA converges to zero more slowly than TDMA. However, if packet-level coding is applied to introduce redundant packets to compensate for the dropped packets, ALOHA is is a feasible MAC scheme that achieves a good delay performance. As a tradeoff, when the gap in p_L is reduced, the gap in the delay moments will be increased.

The pair (B, p_L) can be used in the controller design to optimize the NCS performance as shown in [10, 13]. With nonzero packet loss p_L and the maximum packet delay B, the controller system can be written as a Markovian Jump Linear System (MJLS). Optimizing the MJLS system can optimize the NCS network with MAC schemes.

5 Conclusions

This paper aims to provide a more accurate measurement to optimize NCS networks with MAC schemes. Previous work usually assumed the wireless channels as a constant-delay, which is not practical. We derive the e2e delay and packet loss probability of DB multihop wireless networks for two MAC schemes, m-phase TDMA and probabilistic slotted ALOHA. These parameters can be used by the controller to evaluate the performance. The e2e delay mean and variance are approximately linear with the number of nodes. The local dropping probabilities $p_L^{(i)}$ asymptotically converge to zero. A moderate delay bound B is sufficient to guarantee a small packet loss and thus achieve a good balance between reliability and latency.

Compared to fully reliable networks, the e2e delay of DB networks becomes less sensitive to the channel reception probability p_s. This improvement is desirable since the network performance is not expected to rapid fluctuate with p_s, which basically cannot be controlled. Besides, with the BD strategy, the delay performance gap between TDMA and ALOHA is reduced. Due to the implementation complexity and overhead, TDMA is less favored than ALOHA. But ALOHA has poor delay performance. With the reduced performance gap, ALOHA becomes more practical.

We also show that ALOHA possesses the self-regulating property. For light traffic, unlike previous work [10], we find that a large transmit probability p_m does not degenerate the network performance as for heavy traffic. Previously, $p_m = 1/N$ (N is the number of nodes) is thought to optimize the throughput. As the network enlarges, p_m becomes very small and the resulting delay is long. From the perspective of the delay, a large p_m is preferred, particularly when it does not decrease the throughput substantially.

This paper considers a FCFS node scheduling for analysis tractability. Other scheduling algorithms that favor the newly arriving packets like LCFS and priority scheduling, may be more desirable for NCS applications. In the future work we will discuss how NCSs perform when associating with the BD strategy and other scheduling algorithms.

6 Acknowledgment

The authors would like to thank the support of the Center for Applied Mathematics (CAM) Fellowship of the University of Notre Dame and the partial support of NSF (grant ECS03-29766).

References

1. J. C. Arnbak and W. V. Blitterswijk (1984) Capacity of Slotted ALOHA in Rayleigh-Fading Channels. In: IEEE Journal on Selected Areas in Communications, Vol. 5: 261–269, Feburary.
2. Michael S. Branicky, Stephen M. Phillips, Wei Zhang (2000) Stability of Networked Control Systems: Explicit Analysis of Delay, In: IEEE American Control Conference.
3. M. J. Ferguson (1977) An Approximate Analysis of Delay for Fixed and Variable Length Packets in an Unslotted ALOHA Channel. In: IEEE Transactions on Communications. Vol. 25: 644–654, July.
4. C. Florens and R. McEliece (2003) Packet Distribution Algorithms for Sensor Networks. In IEEE INFOCOM 2003.

5. K. K. Lee and S. T. Chanson (2004) Packet Loss Probability for Bursty Wireless Real-Time Traffic Through Delay Model. In: IEEE Transactions on Vehicular Technology, correspondence. Vol. 53: 929– 938, May .
6. J. P. Lehoczky (1996) Real-Time Queueing Theory. In: 17th IEEE Real-Time Systems Symposium, 186–195, December.
7. J. P. Lehoczky (1996) Real-Time Queueing Network Theory. In: 18th IEEE Real-Time Systems Symposium, 58–67, December.
8. Feng-Li Lian, James Moyne, Dawn Tilbury (2001) Time Delay Modeling and Sample Time Selection for Networked COntrol Systems. In: Proceedings of ASME-DSC, 2001 International Mechanical Engineering Congress and Exposition. November, New York.
9. Xiangheng Liu, Andrea Goldsmith (2003) Wireless Communication Tradeoffs in Distributed Control, In: IEEE Conference on Decision and Control.
10. Xiangheng Liu, Andrea Goldsmith (2004) Wireless Medium Access Control in Networked Control Systems, In: IEEE American Control Conference.
11. Xiangheng Liu, Andrea Goldsmith (2005) Cross-layer Design of Distributed Control over Wireless Networks. Systems and Control: Foundations and Applications, (Ed. T. Basar), Birkhauser. (In Press)
12. M. G. Luby, M. Mitzenmacher, M. A. Shokrollahi, D. A. Spielman (2001) Efficient Erasure Correcting Codes, In: IEEE Transactions on Information Theory. Vol. 47: 569–584.
13. J. Nilsson (1998) Real-Time Control Systems with Delays. Ph.D Thesis, Lund Institute of Technology.
14. D. Park, H. G. Perros (1992) Approximate Analysis of Discrete-time Tandem Queueing networks with Customer Loss. In: IEEE GLOBECOM, Vol. 3: 1503 –1507, December.
15. H. Perros, Y. Rhee (1998) An Approximate Analysis of An Open Tandem Queueing Network with Population Constraint and Constant Service Times. In: IIE Transactions, Vol. 30: 973–979, October.
16. R. R. Pillai(2001), A Distributed Overload Control Algorithm for Delay-Bounded Call Setup. In: IEEE/ACM Transactions on Networking, Vol. 9:780–789, December.
17. C. Santivanez and I. Stavrakakis (1999) Study of Various TDMA Schemes for Wireless Networks in the Presence of Deadlines and Overhead. In: IEEE Journal on Selected Areas in Communications, Vol. 17 :1284–1304, July.
18. A. B. Behroozi-Toosi, R. R. Rao (1992) Delay Upper Bounds for a Finite User Random-Access System with Bursty Arrivals. In: IEEE Transactions on Communications. Vol. 40: 59–596, March.
19. K. R. Trivedi(2001) Probability and Statistics with Reliability, Queueing and Computer Science Applications. Wiley, John and Sons. ISBN 0471333417.
20. M. Xie and M. Haenggi (2005), Delay Performance of Different MAC Schemes for Multihop Wireless Networks. IEEE GLOBECOM'05.
21. Y. Yang and T.-S. P. Yum (2003) Delay Distributions of Slotted ALOHA and CSMA. In: IEEE Transactions on Communications, Vol. 51: 1846–1857, November.
22. M. Zorzi and S. Pupolin (1994) Slotted ALOHA for High-Capacity Voice Cellular Communications. In: IEEE Transactions on Vehicular Technology, Vol. 43. :1011–1021, November.

23. Michele Zorzi (2002) Data-Link Packet Dropping Models for Wireless Local Communications, In: IEEE Transactions on Vehicular Technology. Vol. 51, No. 4, July: 710–719.