Delay Minimization with Channel-Adaptive Packetization Policy for Random Data Traffic

Abolfazl Razi*, Ali Abedi† and Anthony Ephremides‡
*ECE Dept., Duke University, Email: abolfazl.razi@duke.edu
†ECE Dept., University of Maine, Email: ali.abedi@maine.edu
‡ECE Dept., University of Maryland, Email: etony@umd.edu

***Abstract*—Packet lengths have crucial impact on the end-to-end delay in a queuing system over noisy time-varying wireless channels that is overlooked in most system designs. In this paper, a delay optimal adaptive *packetization* policy in network layer is proposed that adjusts framing intervals based on the underlying physical layer parameters in order to minimize the end-to-end packet delivery time.**

The arrival stream of fixed-length symbols with exponential inter-arrival time distribution are bundled into packets in an optimal manner and are scheduled for transmission over a noisy wireless channel with Automatic Repeat Request (ARQ) retransmission mechanism. The underlying channel quality in terms of BER is considered known to the scheduler. The proposed solution not only minimizes the end-to-end latency but also prohibits potential queue instability in dynamic channel situations.

***Index Terms*—Packetization policy, joint optimization, channel adaptive scheduling, delay analysis.**

I. Introduction

Packet length significantly affects transmission efficiency in wireless networks with dynamic channel conditions. Longer packet lengths ensure underlying channel coding efficiency by utilizing longer codewords. In a packet-based transmission system, larger packet sizes also reduce the *packetization* overhead for the payload data. This overhead may be due to the packet header (e.g. addressing bits, control bits, and CRC codes), channel setup time, or even channel contention period in wireless networks with opportunistic scheduling. On the other hand, longer packets may increase transmission time by imposing longer packet formation time, since the payload data is not accessible at the destination until a packet is formed at the transmitter and is fully delivered to the destination. Moreover, in a noisy environment with a certain bit error probability, a longer packet size increases the packet error rate, which in turn induces additional retransmission delay to the system [1]–[6].

In simple words, packet length has two opposite effects on the end-to-end latency per symbol. Addressing this essential trade-off and finding the optimum packet length is important for the communication efficiency, which has been mostly overlooked in the former system designs.

We note that the majority of new communication protocols such as IEEE 802.16 and IPv6 allow variable packet lengths. In this work, we exploit this feature and propose a delay optimal packetization policy for a simple single-hop communication system with a First Come First Serve (FCFS) scheduler and an Automatic Repeat Request (ARQ) retransmission mechanism. Although simple, this model highlights and solves the relevant trade-offs and provides insights for more general systems. The end-to-end latency consists of three delay sources including: i) packet formation delay to bundle symbols into transmit packets, ii) waiting time in transmit buffer, and iii) service time including both initial transmission and potential retransmissions.

Recently, several attempts have been made to increase communication efficiency by customizing packet lengths based on the channel quality factors. For instance, the idea of local packet length adaptation is introduced in [7] in order to maximize throughput in WLAN channels. An approximate blocking probability is found for general packet length distributions in [8]. However, the impact of packet length on the end-to-end data delivery time has not been comprehensively studied. Besides, most of the previously reported approaches consider the saturated traffic model [9] and aim at maximizing throughput by reducing the number of dropped packets. The saturated traffic model does not cover random traffic in most real world applications such as web-based applications and Ad-hoc sensor networks. Therefore, we consider a probabilistic traffic model, where a stream of input symbols (e.g. measurements of a data source in a sensor network) are generated according to a Poisson process.

The rest of this paper is organized as follows. In Section II, the system model is defined. In Section III, the end-to-end latency is analyzed and the optimizing packetization criterion is found, while satisfying the system stability conditions. Simulation results are provided in Section IV, followed by concluding remarks in Section V.

II. System Model

A sequence of input symbols $\{x_i\}_{i=0}^{\infty}$ arrive according to a Poisson process with rate λ. The symbols are assumed to have a fixed length of N bits. The arrival symbols are then combined into packets with a constant header size H, scheduled with FCFS discipline in an infinite length queue and transmitted through a wireless channel with bit rate R to the destination, as depicted in Fig. 1.

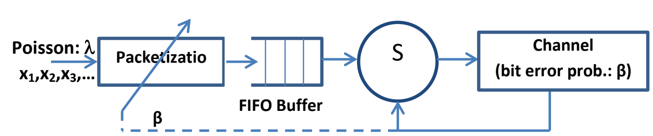

Fig. 1: System Model.

Packetization over symbols are performed according to the proposed method in Section III to generate transmit packets X_i with length l_i. The resulting packets are transmitted over a single-hop wireless channel, whose effective bit error rate is β. For identically independent distributed (i.i.d) bit error probability and zero error tolerance, a packet is in error if at least one bit is corrupted. This yields the following packet error probability:

$$P_i^e = \sum_{j=1}^{l_i} \binom{l_i}{j} \beta^j (1-\beta)^{l_i - j} = 1 - \alpha^{l_i}, \tag{1}$$

where $\alpha = 1 - \beta$ is bit success probability used for notation convenience in the subsequent equations. The erroneous packets are successfully detected at the destination using CRC codes and are retransmitted using an Automatic Repeat Request (ARQ) scheme with instantaneous feedback channel until they are successfully delivered to the destination. The number of transmissions, denoted by $r \in \{1, 2, 3, ...\}$ is a random variable (r.v.) and depends on the packet length l and bit error probability β with the following Geometric distribution:

$$\begin{aligned} f_{\mathbf{r}}(r) = p(\mathbf{r} = r) &= (1 - P^e)\big(P^e\big)^{r-1} \\ &= \alpha^l (1 - \alpha^l)^{r-1}. \end{aligned} \tag{2}$$

Hence, the expected value of number of retransmissions can be simply found by calculus of power series as:

$$E[r] = \sum_{r=1}^{\infty} r f_{\mathbf{r}}(r) = \sum_{r=1}^{\infty} r \alpha^l (1 - \alpha^l)^{r-1} = \alpha^{-l}. \tag{3}$$

The i^{th} symbol that belongs to the j^{th} packet experiences the end-to-end delay d_i, which includes packet formation delay, waiting time and service time, denoted by f_i, w_j and s_j, respectively. Packet formation delay is the time difference between the symbol arrival epoch and the corresponding packet formation epoch. Thus, may be different for symbols inside a packet. Whereas, w_j and s_j are packet-based delays and are equal for all symbols in packet j and account for waiting and service times for both primary and retransmit periods. The goal is to find the optimal packetization policy that minimizes the expected average delay defined as

$$E[d] = \lim_{t \to \infty} \frac{1}{M(t)} \sum_{i=1}^{M(t)} E[d_i] = E[d_i], \tag{4}$$

where $M(t) = \max\ \{i : t_i < t\}$ is the number of symbols arrived by time t. The last equality is based on the system symmetry and ergodicity of the queue.

III. Packetization Policy

In this section, a time-based packetization policy according to Fig. 2 is proposed. A basic packetization interval T is defined such that the time axis is partitioned into consecutive equal packetization intervals $\big[(i-1)T, iT\big)$. Symbols arrived at each interval (if any) are combined to form a single packet X_i that is scheduled in the queue. Therefore, interval i includes k_i symbols and has the following packet length:

$$\begin{aligned} &l_i = h(k_i)H + k_i N, \\ &p(k_i = k) = e^{-\lambda T}(\lambda T)^k / k!, \end{aligned} \tag{5}$$

where $k_i = M(iT) - M(iT - T) \in \{0, 1, 2, ...\}$ and $h(k_i)$ is an indicator function defined as follows

$$h(k_i) = \begin{cases} 1 & k_i > 0, \\ 0 & k_i = 0. \end{cases} \tag{6}$$

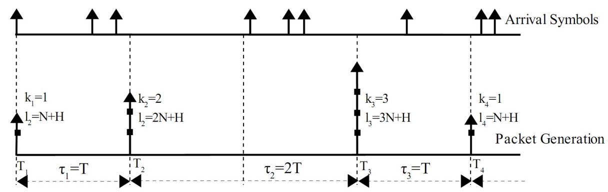

Fig. 2: Time based Packetization of Arrival Symbols.

Based on the proposed packetization policy, if no symbol arrived at a framing interval, no packet is formed and symbol collection is continued until the end of the next packetization interval. Therefore, packet inter-arrival times can be extended to multiple packetization intervals. Probability of zero symbol arrival in an interval of length T is $P_0 = p(k_i = 0) = e^{-\lambda T}$. Thus, the packet inter-interval time, τ_i is Geometrically distributed with success parameter $1 - P_0$ and the following moments:

$$\begin{aligned} E(\tau_i) &= \frac{1}{1 - P_0} T, \\ E(\tau_i^2) &= \frac{1 + P_0}{(1 - P_0)^2} T^2. \end{aligned} \tag{7}$$

To obtain the probability mass function (pmf) of packet length l_i, we simply combine (5) and (6) that yields:

$$p(l) = \begin{cases} e^{-\lambda T}(\lambda T)^k / k! & l = H + kN, k \neq 0, \\ e^{-\lambda T} & l = 0, k = 0. \end{cases} \tag{8}$$

The service time of a packet of length l accounts for sending l bits over the channel, which might be repeated for r times due to the ARQ retransmission mechanism. Noting the fixed transmission rate R bit/sec, retransmission probability in (2) and packet length pmf in (8), the following pmf is derived for service time, s:

$$p\big(\mathbf{s} = \frac{r}{R}.h(k)H + kN\big) = \alpha^{H+kN}\big[1 - \alpha^{H+kN}\big]^{r-1} e^{-\lambda T}\frac{(\lambda T)^k}{k!},$$
$$\text{for } r = 1, 2, 3, ...,\ k = 0, 1, 2, ... \ . \tag{9}$$

The countable discrete support set of s is $\mathcal{S} = \{0\} \cup \{r(H + kN)/R :\ r, k = 1, 2, 3, ...\}$, modeling empty intervals with zero-length packets. We note that the pmf of service time for ARQ retransmission over noisy channels with Permission arrivals is not a monotonically decreasing function in contrast to the error-free channels.

To calculate the first and second moments of service time that is required to calculate the waiting time, we first state the following proposition:

Proposition 3.1: If k is a Poisson random variable (r.v.) with mean μ, then $E_{\mathbf{k}}[\frac{k^n}{\zeta^k}] = e^{-\mu(1-1/\zeta)} \sum_{i=1}^{n} S_2(n, i)(\mu/\zeta)^i$ for $n = 1, 2, 3...$, where $S_2(n, i)$ is the Stirling numbers of the second kind which counts the number of ways to partition a set of n elements into i nonempty subsets [10].

Proof: See Appendix. ■

Corollary: For $n = 1, 2$ as two important special cases, we recall that $S_2(n, 0) = 0$ and $S_2(1, 1) = S_2(2, 1) = S_2(2, 2) = 1$. Therefore,

$$E_{\mathbf{k}}[\frac{k}{\zeta^k}] = \frac{\mu}{\zeta} e^{-\mu(1-1/\zeta)}, \tag{10}$$

$$E_{\mathbf{k}}[\frac{k^2}{\zeta^k}] = \frac{\mu}{\zeta}(1 + \frac{\mu}{\zeta}) e^{-\mu(1-1/\zeta)}. \tag{11}$$

Lemma 3.2: If k is a Poisson r.v. with parameter μ and $x(k)$ is defined as $x(k) \triangleq \frac{h^m(k)}{\zeta^k}$ for $m = 1, 2, 3, ...$, then $E_{\mathbf{k}}[x] = e^{-\mu}(e^{\mu/\zeta} - 1)$.

The proof follows from definition of the expected value. We notice that $h^m(k) = h(k)$. Thus,

$$x(k) = \begin{cases} \frac{1}{\zeta^k} & k \neq 0 \\ 0 & k = 0 \end{cases}$$

$$\Rightarrow E_{\mathbf{k}}[x] = \sum_{k=0}^{\infty} x(k) p_{\mathbf{k}}(k) = \sum_{k=0}^{\infty} \frac{1}{\zeta^k} p_{\mathbf{k}}(k) - p_{\mathbf{k}}(k = 0) \frac{1}{\zeta^0}$$
$$= E_{\mathbf{k}}[\frac{1}{\zeta^k}] - e^{-\mu} = e^{-\mu(1-1/\zeta)} - e^{-\mu} = e^{-\mu}(e^{\mu/\zeta} - 1). \tag{12}$$

Lemma 3.3: If k is a Poisson r.v. with parameter μ and $x(k)$ is defined as $x(k) \triangleq \frac{h(k)^m k^n}{\zeta^k}$, where $m, n \in \{1, 2, 3, ...\}$, then $E_{\mathbf{k}}[x] = e^{-\mu(1-1/\zeta)} \sum_{i=1}^{n} \binom{n}{i} (\mu/\zeta)^i$.

The proof immediately follows from Proposition 3.1 and noting the fact that $x(k) = \frac{h(k)^m k^n}{\zeta^k} = \frac{k^n}{\zeta^k}$ for any positive integer m.

Proposition 3.4: Service time s has the following first and second order moments:

$$E[s] = \frac{N e^{-\lambda T}}{(1 - e^{-\lambda T}) R \alpha^H} \Big[(\eta + \lambda T \alpha^{-N}) e^{\lambda T \alpha^{-N}} - \eta \Big], \tag{13}$$

$$\begin{aligned} E[s^2] = &\frac{N^2 e^{-\lambda T}}{(1 - e^{-\lambda T}) R^2 \alpha^H} \Big[\big(2\lambda T \alpha^{-N(2+\eta)} \\ &+ 2\lambda^2 T^2 \alpha^{-N(4+\eta)} + 4\eta\lambda T \alpha^{-N(2+\eta)} \\ &+ 2\eta^2 \alpha^{-H} \big) e^{\lambda T \alpha^{-2N}} \\ &- \big(\lambda T \alpha^{-N} + \lambda^2 T^2 \alpha^{-2N} + 2\eta\lambda T \alpha^{-N} + \eta^2 \big) e^{\lambda T \alpha^{-N}} \\ &- \big(2\eta^2 \alpha^{-H} - \eta^2 \big) \Big], \end{aligned} \tag{14}$$

where $\eta = H/N$ is defined as the ratio of header size to symbol size.

Proof: See Appendix. ■

Remark: There are two extreme cases. When packetization interval is chosen very small, $(\lambda T \to 0)$, then using approximation $(e^{-\lambda T} \approx 1 - \lambda T \approx 1)$ results in $E(s) \approx \frac{(H+N)}{R \alpha^{H+N}}$. In this case, the impact of H on the service time is as large as N, since most packets include only one symbol and their lengths are $N + H$. On the other hand, for extremely large packetization interval $(\lambda T \to \infty)$ and finite N and H, we have $N\lambda T \gg H$. After some mathematical manipulation, we have $E[s] \approx \frac{N\lambda T}{R\alpha^N} e^{-\lambda T(1-\alpha^{-N})}$. In this case, each packet includes a large number of symbols and the impact of header size on the service time is negligible. For error free channel, we have $\alpha = 1 - \beta = 1$, hence service time is reduced to $\frac{N\lambda T}{R}$, which grows linearly with T. This is shown in simulation results presented in Section IV.

A. Delay Optimal Packetization Policy

For the n^{th} packet, the arrival, service and waiting times are presented by T_n, s_n, and w_n, respectively. The inter-arrival time is also denoted by $\tau_n = T_{n+1} - T_n$, as depicted in Fig. 2. The waiting time for the next packet w_{n+1} is a non-negative value and can not exceed the difference between the previous packets inter-arrival time (τ_n) and sojourn time $(w_n + s_n)$, hence we have the following relation, which is known as Lindley's equation [11]:

$$w_{n+1} = [w_n + s_n - \tau_n]^+, \tag{15}$$

with initial condition $w_0 = 0$, where $(x)^+ = \max(0, x)$. We note that s_n depends on the number of arrivals during the n^{th} interval, while τ_n is multiples of the basic packetization time T and is defined only based on the probability of the first arrival after interval n. Thanks to the memoryless property of exponential distribution, they are independent and hence:

$$\begin{aligned} &p(w_{n+1} = j, \tau_n \leq t | w_0, w_1, ..., w_n, \tau_0, \tau_1, ...\tau_{n-1}) \\ &= p(w_{n+1} = j, T_{n+1} - T_n \leq t | w_n). \end{aligned} \tag{16}$$

This means that the process $\{w, \tau\} = \{w_n, \tau_n\}$ forms a renewal Markov Process. To find the transition kernel for the embedded Markov chain $\{w_n\}$, we note:

$$\begin{aligned} &p(w_{i+1} = y | w_i = x, \tau_i = \tau) \\ &= f_{\mathbf{s}}(y - x + \tau) U(y) \\ &+ \sum_{\substack{s \in \mathcal{S} \\ 0 \leq s \leq \max(0, \tau - x)}} f_{\mathbf{s}}(s) \delta(y), \end{aligned} \tag{17}$$

where $U(.)$ and $\delta(.)$ are step and Dirac impulse functions and $f_{\mathbf{s}}(s)$ is pmf of s_i defined in (9). The first term in (17) accounts for the case where $w_i + s - \tau_i > 0$. Hence, $w_{i+1} = w_i + s - \tau_i$ is a positive value occurs for $s = y - x + \tau_i$. The second term is corresponding to the case where $x < \tau_i$ and $0 \leq s_i \leq |x - \tau_i| = \tau_i - x$. In this case, packet $n + 1$ is arrived after completion of the previous packet service and thus $w_{i+1} = \max(0, w_i + s - \tau_i)$ is mapped to 0.

Noting that τ_i follows Geometric distribution, (17) yields the following transition kernel:

$$\begin{aligned} &P_i(x, y) \triangleq p(w_{i+1} = y | w_i = x) \\ &\sum_{k=1}^{\infty} p(w_{i+1} = y | w_i = x, \tau_i = kT) p(\tau_i = kT) \\ &= \sum_{k=1}^{\infty} (1 - e^{-\lambda T}) e^{-k\lambda T} \Big[f_{\mathbf{s}}(y - x + kT) U(y) \qquad (18) \\ &+ \sum_{\substack{s \in \mathcal{S} \\ 0 \leq s \leq \max(0, kT - x)}} \delta(y) f_{\mathbf{s}}(s) \Big], \end{aligned} \tag{19}$$

B. Stability Condition

The transition kernel in (18) states the evolution of queuing delays until it reaches the stationary state. It has been shown in [12], [13] that for such a system, the stability is granted if the embedded Markov process complies the sufficient condition of ergodicity $E[s_n - \tau_n] < 0$. Therefore, T must be chosen such that $E(s) < E(\tau_i)$ is satisfied. Using (7,13,36), we have the following stability conditions:

$$\frac{1}{R\alpha^H}\Big[(H + N\lambda T\alpha^{-N})e^{-\lambda T(1-\alpha^{-N})} - He^{-\lambda T}\Big] \leq T, \tag{20}$$

which provides a lower bound on T value. Equation (20) does not admit a neat closed form but can be solved numerically or evaluated approximately. For error free channel $\alpha = 1$, this simplifies to:

$$R \geq \frac{(H + N\lambda T) - He^{-\lambda T}}{T}. \tag{21}$$

To find the necessary conditions on channel rate R, we consider two extreme cases of short and long packetization intervals. If $T \to \infty$, (21) yields: $R \geq N\lambda$, which is the rate capable of handling long packets with negligible overhead bits. The other extreme case occurs when $T \to 0$, where each symbol forms a packet of length $H + N$ upon arrival. In this case we have $e^{-\lambda T} \approx 1 - \lambda T$ and (21) converts to $R \geq (N + H)\lambda$. If $R \in [N\lambda, (N + H)\lambda]$. There exists a subset of T such that the queue is stable and we can obtain arbitrary near optimal throughput of $\rho = \frac{E[s]}{E[\tau]} \to 1^-$.

C. Expected Waiting Time

Noting that the Markov process $\{w_n, \tau_n\}$ is regenerative, if we set $v_0 = 0$ and $v_j = \sum_{i=n-j+1}^{n} (s_i - \tau_i)$, by recurrence, we can rewrite (15) as:

$$w_{n+1} = \max(v_0, v_1, ..., v_n). \tag{22}$$

For a stable queue, we have $\lim_{n\to\infty} E(v_n) = nE(s-\tau) < \infty$. Hence, w_n tends to a r.v. $w = \sup_i v_i$, as n approaches infinity. Therefore, (15) yields:

$$E(w) = E[(w + s - \tau)^+], \tag{23}$$

$$\sigma_w^2 = \sigma^2_{(w+s-\tau)^+}. \tag{24}$$

Using the methodology proposed in [14], which is known as Kingman formula and noting the interarrival time distribution in (7), we obtain the following approximation formula for the waiting time:

$$\begin{aligned} E(w) &\approx \frac{\rho E(s)(\sigma_s^2 + \sigma_\tau^2)}{2(1-\rho)} \\ &= \frac{\big[E(s)\big]^2(\sigma_s^2 + \sigma_\tau^2)}{2\big(T - E(s)\big)}. \end{aligned} \tag{25}$$

Substituting (7, 13, 14) in (25) provides a closed form equation for waiting time. The simulation results in Section IV confirm the accuracy of the above calculations.

D. Packet Formation Delay

To account for the packet formation delay, we note the symmetry and the memoryless property of Poisson arrivals. This follows that the expected value of the average of the time span from a symbol arrival time to the end of the corresponding interval over all arrival symbols during that interval is half the interval length $E[\sum f_i] = T/2$. This is in consistence with the known fact that the Poisson process can be reversed is time [11]. This property allows us to interchangeably calculate the expected distance from either sides of the interval, which in turn implies:

$$\begin{aligned} &E[\sum t_i] = E[\sum f_i] = kT - E[\sum t_i] \\ &\Rightarrow E[\sum t_i] = kT/2 \Rightarrow E[t_i] = T/2, \end{aligned} \tag{26}$$

for an interval with k arrivals.

E. Optimum Packetization Interval Criterion

In stationary situation, the expected value of the end-to-end delay, ($d = w + s + f$) is:

$$\begin{aligned} E[d] &= \lim_{t\to\infty} \frac{1}{M(t)} \sum_{i=1}^{M(t)} E[d_i] = E[d_i] \\ &= E[w_i] + E[s_i] + E[f_i] \\ &= \frac{\big[E[s]\big]^2(\sigma_s + \sigma_\tau)}{2\big(T - E[s]\big)} + E[s] + T, \end{aligned} \tag{27}$$

where moments of service time, s are functions of (N, H, β, T) as defined in (13,14). The convexity of (27) with respect to T can be easily verified by checking positivity of the second derivative, which is straightforward bot involving many terms. It was also confirmed by numerical evaluation as depicted in Fig. 5. Equating derivative of (27) with respect to T to zero ($\frac{\partial E[d]}{\partial T} = 0$) provides the optimum packetization interval T^*, which minimizes the end to end latency $E[d]$.

IV. Simulation results

In this section, simulation results are provided to confirm the accuracy of the derived delay optimal packetization criterion. For each simulation, we performed Monte-Carlo simulation for 100000 packets.

Fig. 3 presents the perfect match between the simulation results and the analytically derived coefficient of variation for service time $K_s = \frac{\sigma_s}{E[s]}$, which verifies the accuracy of equations (13), (14) and (36). The simulation parameters are arbitrarily set to $\lambda = 10, N = 16, H = 40$. It is shown that K_s increases as packetization time T moves away from zero. This is due to the fact that as T approach zero $T \to 0$, each packet tends to include only one symbol and hence presents a fixed length of $N + H$. Thus, the service time which is proportional to the packet length presents a low variation and $K_s = \sigma_s/E[s] \to 0$. For error free channel and moderate T values $\lambda T \approx 1$, the packet length presents more unpredictability. Furthermore, when T grows to infinity, due to the accumulation of Poisson arrivals over long interval, the packet lengths tend to $\lambda TN + H$ with small variations. Therefore, $K_s \to 0$ for error free channels, PER $= 0$. However, for an erroneous channel, larger packet lengths

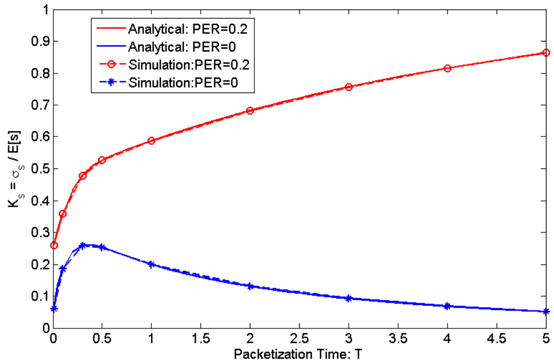

Fig. 3: Coefficient of Variation for Service Time vs Packetization Time, Comparison between Simulations and Analysis $(N = 16, H = 40, \lambda = 10)$.

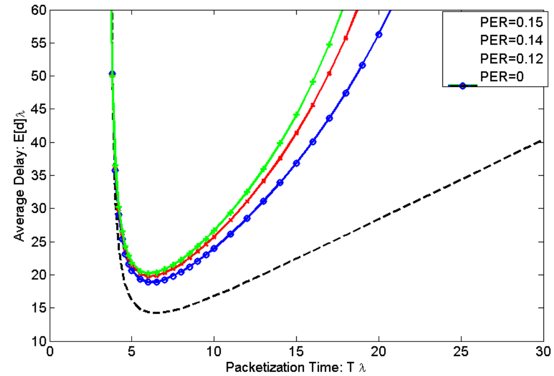

Fig. 5: Expected Delay vs Packetization Time $(N = 16, H = 30, \lambda = 10)$.

experience higher packet drop rates and large variations on service time due to the unpredictable retransmission number. This phenomena is dominant and service time shows higher variations for error-prone channels.

Fig. 4 presents the impact of packetization interval length (T) on various delay types. Solid lines (marked A) in this figure represent analytically derived delays, while dashed lines (marked S) represent the Monte-Carlo simulation results. The packet formation delay, which is shown with blue curve is proportionally related to T as was intuitively expected and justified in Section III-D. Waiting time is also derived analytically as a function of T (solid green curve), which is well confirmed by the simulation results. As it is shown, waiting time is a convex function of T, and so is the average end-to-end delay (red curve). Convexity of the average end-to-end delay with respect to T guarantees the uniqueness of the optimum packetization interval length for a given set of system parameters.

Fig. 5 demonstrates the behavior of the expected average delay, $E[d]$ derived in (27) with varying packetization interval T for different channel qualities in terms of PER. It is shown that for small T, the high overhead cost may cause the average input rate (in terms of bit per second) exceed the channel rate and hence the queue becomes unstable. On the other hand, for error-free channels, when T grows to infinity, the packet inter-arrival times tend to be less than service time and therefore the dominant delay term is packet formation delay in (26). Hence, the expected delay grows linearly with packetization interval length. For noisy channels with non-zero PER, longer packetization intervals result in a larger average packet lengths. Consequently, PER grows exponentially with T and imposes larger service time and hence waiting time on the queue that ultimately makes the queue unstable with extremely long waiting time. Therefore, the growth of end-to-end delay with T is more crucial. This shows the importance of the given analysis to find an optimal packetization interval, which is not only a function of input traffic properties, but also depends on the underlying channel quality. Therefore, ignoring physical layer parameters such as the utilized channel error-rate in the packet formation at higher protocol layers might cause the whole transmission system to fail.

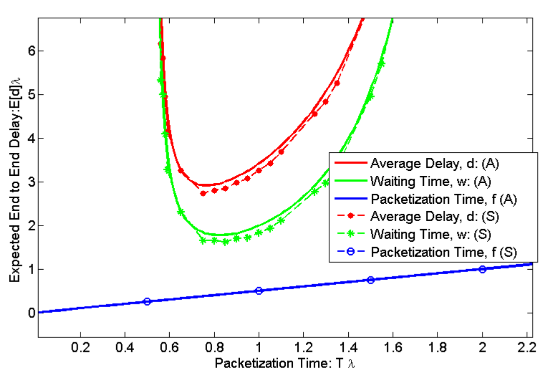

Fig. 4: Expected Waiting Time and Service Time for Different Packet Error Rates $(N = 16, H = 30, \lambda = 10)$.

V. Concluding Remarks

In this paper, the impact of packetization interval length on the end to end latency is investigated to provide an optimal policy to encapsulate Poisson arrival symbols into transmit packets. It was noticed that the Poisson arrival symbols yield three distributions for packet arrivals including: Bernoulli, Geometric, and Deterministic as T departs from zero to infinity.

It was also shown that relatively small packetization interval may cause huge delays due to the low header efficiency and instability of the queuing system. On the other hand larger packetization interval increase delay by means of higher packet formation delay and higher retransmission numbers. This suggests that packet formation policy design should not ignore underlying physical layer parameters such as channel rate and bit error probability. This fundamental study for a single hop communication can be extended further for more complicated system setups in Adhoc networks considering more physical layer parameters.

Appendix

Proof of proposition 3.1 First, we notice that:

$$E_{\mathbf{k}(\mu)}[\frac{k^n}{\zeta^k}] = \sum_{k=0}^{\infty} \frac{k^n}{\zeta^k} e^{-\mu} \frac{\mu^k}{k!} = e^{-\mu(1-1/\zeta)} \sum_{k=1}^{\infty} k^n e^{-\mu/\zeta} \frac{(\mu/\zeta)^k}{k!}$$
$$= e^{-\mu(1-1/\zeta)} E_{\mathbf{k}(\mu/\zeta)}[k^n]. \quad (28)$$

It was shown in [15] that k^n can be represented in the form of falling factorials as

$$k^n = \sum_{i=0}^{n} S_2(n,i) k_{(i)}, \tag{29}$$

where $k_{(i)} = k(k-1)...(k-n+1)$ is the falling factorial. Combining (28) and (29) results in:

$$E_{\mathbf{k}(\mu)}[\frac{k^n}{\zeta^k}] = e^{-\mu(1-1/\zeta)} \sum_{i=0}^{n} S_2(n,i) E_{\mathbf{k}(\mu/\zeta)}[k_{(i)}]. \tag{30}$$

$E_{\mathbf{k}}[k_{(i)}]$ can be easily found by factorial moment generation function for a Poisson distribution as follows:

$$\begin{aligned} E_{\mathbf{k}(\mu/\zeta)}[k_{(i)}] &= \frac{d^i}{dt^i} E_{\mathbf{k}(\mu/\zeta)}[t^k] \mid_{t=0} = \frac{d^i}{dt^i} e^{\mu/\zeta(t-1)} \mid_{t=0} \\ &= \mu^i e^{\mu/\zeta(t-1)} \mid_{t=0} = (\mu/\zeta)^i \end{aligned} \tag{31}$$

Substituting (31) in (30) completes the proof.

Proof of proposition 3.4:

For derivation simplicity, we first consider an auxiliary service time $\tilde{s}$ by allowing zero-symbol packet generation, such that an imaginary dummy packet of length 0 is generated for zero symbol accumulation during a packetization interval. Then, we modify the obtained equations by excluding zero-symbol packets. The real and auxiliary service times are related as follows:

$$\begin{cases} k \neq 0 \Rightarrow s = \tilde{s} & \text{with prob.} P_0 = 1 - e^{-\lambda T} \\ k = 0 \Rightarrow \text{s does not exist} & \text{with prob.} 1 - P_0 = e^{-\lambda T} \end{cases} \tag{32}$$

We note that the retransmission parameter, r has Geometric distribution with success parameter $(1-\beta)^l = \alpha^{H+kN}$ dependent on the packet length. Hence, its first and second order moments are $E_{\mathbf{r}}[r] = \alpha^{-(H+kN)}$ and $E_{\mathbf{r}}[r^2] = \frac{2-\alpha^{H+kN}}{\alpha^{2(H+kN)}}$, which are functions of k. Therefore, using Proposition 3.1, Lemma 3.2 and Lemma 3.3, the moments of service time can be calculated as :

$$\begin{aligned} E[\tilde{s}] &= E_{\mathbf{k}}\big[E_{\mathbf{r}}[\tilde{s}]\big] = E_{\mathbf{k}}\big[\frac{h(k)H + kN}{R} E_{\mathbf{r}}[r]\big] \\ &= E_{\mathbf{k}}\Big[\frac{h(k)H + kN}{R\alpha^{H+kN}}\Big] \\ &= \frac{H}{R\alpha^H} E_{\mathbf{k}}\Big[\frac{h(k)}{(\alpha^N)^k}\Big] + \frac{N}{R\alpha^H} E_{\mathbf{k}}\Big[\frac{k}{(\alpha^N)^k}\Big] \\ &= \frac{H}{R\alpha^H}\big(e^{-\lambda T}(e^{\lambda T \alpha^{-N}} - 1)\big) + \frac{\lambda T N}{R\alpha^H \alpha^N} e^{-\lambda T(1-\alpha^{-N})}\Big] \\ &= \frac{N e^{-\lambda T}}{R\alpha^H}\Big[(\eta + \lambda T \alpha^{-N}) e^{\lambda T \alpha^{-N}} - \eta\Big], \end{aligned} \tag{33}$$

where $\eta = H/N$ is the ratio of header size to symbol size. Similarly we have:

$$\begin{aligned} E[\tilde{s}^2] &= E_{\mathbf{k}}[E_{\mathbf{r}}(\tilde{s}^2)] = E_{\mathbf{k}}[\frac{(h(k)H + kN)^2}{R^2} E(r^2)] \\ &= E_{\mathbf{k}}\Big[\frac{(h(k)H + kN)^2 (2 - \alpha^{H+kN})}{R^2 \alpha^{2(H+kN)}}\Big] \\ &= \frac{2H^2}{R^2\alpha^{2H}} E_{\mathbf{k}}\Big[\frac{h^2(k)}{(\alpha^{2N})^k}\Big] + \frac{2N^2}{R^2\alpha^{2H}} E_{\mathbf{k}}\Big[\frac{k^2}{(\alpha^{2N})^k}\Big] \\ &+ \frac{4NH}{R^2\alpha^{2H}} E_{\mathbf{k}}\Big[\frac{kh(k)}{(\alpha^{2N})^k}\Big] - \frac{H^2}{R^2\alpha^H} E_{\mathbf{k}}\Big[\frac{h^2(k)}{(\alpha^N)^k}\Big] \\ &- \frac{N^2}{R^2\alpha^H} E_{\mathbf{k}}\Big[\frac{k^2}{(\alpha^N)^k}\Big] - \frac{2NH}{R^2\alpha^H} E_{\mathbf{k}}\Big[\frac{kh(k)}{(\alpha^N)^k}\Big] \end{aligned} \tag{34}$$

Some simple calculations and noting Lemma 3.2 and 3.3 follows:

$$\begin{aligned} E[\tilde{s}^2] = \frac{N^2 e^{-\lambda T}}{R^2 \alpha^H}\Big[&\big(2\lambda T \alpha^{-N(2+\eta)} + 2\lambda^2 T^2 \alpha^{-N(4+\eta)} \\ &+ 4\eta\lambda T \alpha^{-N(2+\eta)} + 2\eta^2 \alpha^{-H}\big) e^{\lambda T \alpha^{-2N}} \\ &- \big(\lambda T \alpha^{-N} + \lambda^2 T^2 \alpha^{-2N} + 2\eta\lambda T \alpha^{-N} + \eta^2\big) e^{\lambda T \alpha^{-N}} \\ &- \big(2\eta^2 \alpha^{-H} - \eta^2\big)\Big]. \end{aligned} \tag{35}$$

Since no packet is generated for zero symbol accumulation during an interval, we exclude the packets with zero symbols and length H according to (32) to obtain the moments for service time as follows:

$$\begin{aligned} & E[\tilde{s}^n] = E[\tilde{s}^n | k = 0] p(k=0) + E[\tilde{s}^n | k \neq 0] p(k \neq 0) \\ & \Rightarrow E[s^n] = E[\tilde{s}^n | k \neq 0] = \frac{E[\tilde{s}^n]}{1 - e^{-\lambda T}}. \end{aligned} \tag{36}$$

Substituting (33) and (34) in (36) completes the proof.

REFERENCES

[1] M. Y. Cheung, W. Grover, and W. Krzymien, "Combined framing and error correction coding for the ds3 signal format," *Communications, IEEE Transactions on*, vol. 43, no. 234, pp. 1365 –1374, feb/mar/apr 1995.

[2] B. Hong and A. Nosratinia, "Overhead-constrained rate-allocation for scalable video transmission over networks," in *Data Compression Conference, 2002. Proceedings. DCC 2002*, 2002, p. 455.

[3] T. Nage, F. Yu, and M. St-Hilaire, "Adaptive control of packet overhead in xor network coding," in *2010 IEEE International Conference on Communications (ICC)*, may 2010, pp. 1 –5.

[4] J. H. Hong and K. Sohraby, "On modeling, analysis, and optimization of packet aggregation systems," *Communications, IEEE Transactions on*, vol. 58, no. 2, pp. 660 –668, february 2010.

[5] A. Razi, F. Afghah, and A. Abedi, "Power optimized DSTBC assisted DMF relaying in wireless sensor networks with redundant super nodes," *IEEE Transactions on Wireless Communications*, vol. 12, no. 2, pp. 636–645, 2013.

[6] S.-S. Tan, D. Zheng, J. Zhang, and J. Zeidler, "Distributed opportunistic scheduling for ad-hoc communications under delay constraints," in *INFOCOM, 2010 Proceedings IEEE*, march 2010, pp. 1 –9.

[7] M. Krishnan, E. Haghani, and A. Zakhor, "Packet length adaptation in wlans with hidden nodes and time-varying channels," in *Global Telecommunications Conference (GLOBECOM 2011), 2011 IEEE*, dec. 2011, pp. 1 –6.

[8] M. Yasuji, "An approximation for blocking probabilities and delays of optical buffer with general packet-length distributions," *Lightwave Technology, Journal of*, vol. 30, no. 1, pp. 54 –60, jan.1, 2012.

[9] E. Gelenbe and G. Pujolle, *Introduction to Queuing Networks*. John Willey, Second Edition, 1998.

[10] D. E. K. Ronald L. Graham and O. Patashnik, *Concrete Mathematics*. AddisonWesley, 1998.

[11] L. Kleinrock, *Computer Applications, Volume 2, Queueing Systems*. Wiley-Interscience, Second Edition, 1976, vol. 2.

[12] R. L. Tweedie, "Sufficient conditions for ergodicity and recurrence of markov chain on a general state space space," *Stoch. Proc. and their Appl*, vol. 3, pp. 385–403, 1975.

[13] D. V. Lindley, "The theory of queues with a single server," *Mathematical Proceedings of the Cambridge Philosophical Society*, vol. 42, no. 2, pp. 227–289, 1952.

[14] J. F. C. Kingman, "On the algebra of queues," *J. Appl. Probability*, vol. 3, pp. 258–326.

[15] B. Rennie and A. Dobson, "On stirling numbers of the second kind," *Journal of Combinatorial Theory*, vol. 7, no. 2, pp. 116–121, Sep. 1969.