WIRELESS COMMUNICATIONS AND MOBILE COMPUTING
Wirel. Commun. Mob. Comput. (2012)
Published online in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/wcm.2262

RESEARCH ARTICLE

Dimensioning the packet loss burstiness over wireless channels: a novel metric, its analysis and application

Fangqin Liu[1]*, Tom H. Luan[2], Xuemin (Sherman) Shen[2] and Chuang Lin[1]

[1] Department of Computer Science and Technology, Tsinghua University, Beijing, 100084, China
[2] Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, Ontario, N2L 3G1, Canada

ABSTRACT

The packet loss burstiness over wireless channels is commonly acknowledged as a key impacting factor on the performance of networking protocols. An accurate evaluation of the packet loss burstiness, which reveals the characteristics and performance of the wireless channels, is crucial to the design of wireless systems and the quality-of-service provisioning to end users. In this paper, a simple yet accurate analytical framework is developed to dimension the packet loss burstiness over generic wireless channels. In specific, we first propose a novel and effective metric to characterize the packet loss burstiness, which is shown to be more compact, effective, and accurate than the metrics proposed in existing literature for the same purpose. With this metric, we then develop an analytical framework and derive the closed-form solutions of the packet loss performance, including the packet loss rate and the loss-burst/loss-gap length distributions. Lastly, as an example to show how the derived results can be applied to the design of wireless systems, we apply the analytical results to devise an adaptive packetization scheme. The proposed packetization scheme adaptively adjusts the packet length of transmissions based on the prediction of the packet loss rate and loss-burst/loss-gap lengths of the wireless channel. Via extensive simulations, we show that with the proposed packetization scheme, the channel throughput can be enhanced by more than 10% than the traditional scheme. Copyright © 2012 John Wiley & Sons, Ltd.

KEYWORDS

wireless channels; packet loss burstiness; correlation analysis

***Correspondence**

Fangqin Liu, Department of Computer Science and Technology, Tsinghua University, Beijing, 100084, China.
E-mail: fqliu@csnet1.cs.tsinghua.edu.cn

1. INTRODUCTION

Packet losses over wireless channels typically have strong correlations and occur in a bursty fashion. As the packet loss performance directly affects the effectiveness and efficiency of communications, the packet loss burstiness has a significant impact on the performance of on-top protocols in wireless networks [1–3]. For example, as reported in [4], the automatic repeat-request (ARQ) protocol performs much worse over links with burst packet losses than it does over links of independent packet losses. Because of its importance, an extensive body of research has been developed in the past decades to evaluate the packet loss burstiness in wireless channels.

To dimension the packet loss burstiness, the foremost issue is to devise the appropriate metrics, which can accurately characterize the burstiness. In [5], the Allan deviation [6] of packet reception rate (PRR) is used to measure the burstiness of packet losses in 802.11b mesh networks. With this metric, the packet-loss-burst length can be evaluated to improve the design of retransmission algorithms in wireless networks. In [7], a metric β is proposed based on the concept of conditional probability delivery function (CPDF) for measuring the burstiness in 802.15.4 wireless sensor networks. Such a metric is then used to devise an opportune transmission algorithm to boost the PRR.

In this work, we propose a new metric to characterize PRR, which distinguishes from the existing metrics in the following two aspects. First, the existing metrics are typically derived empirically through experiments using network datasets. As such, their effectiveness and efficiencies need to be validated before they can be applied in real-world scenarios. Our proposed metric is a theoretical and generic metric, which is applicable to general settings. Therefore, it is suitable for theoretical analysis to provide comprehensive insights on the features of wireless channels. Second, our proposed metric can be easily used to design and optimize the communication protocols. In this paper, we show an example by developing an

analytical framework based on the proposed metric to accurately evaluate and predict the packet loss performance, such as the packet loss rate and the distributions of the loss-burst/loss-gap lengths (defined as a maximum-length sequence of consecutive corrupted/correct packets). With the accurate evaluation of packet loss performance, we show that the optimal communication protocols can be devised according to the timely changing characteristics of wireless channels. To summarize, our major contributions are threefold.

First, we develop a new approach, namely *packet transmission segmentation*, to investigate on the correlations of packet losses mathematically, and propose a new and efficient metric to measure the packet loss burstiness. We show that this metric is more effective and accurate to gauge the burstiness of wireless channels than previous metrics when the channel is with intensive transmission errors or low packet loss burstiness.

Second, we show how to utilize the newly proposed metric to derive the packet loss performance. The closed-form solutions for the packet loss rate and loss-burst/loss-gap length distributions are developed accordingly. Via extensive simulations, we show the accuracy of the developed analysis in different scenarios of wireless channels. To the best of our knowledge, this work represents the first research to derive the closed-form solutions for the packet loss performance in bursty wireless channels. Notably, the analytical results bridge the channel conditions to the resultant packet loss performance using a simple yet accurate expression, which paves the way for the evaluation, design, and optimization of communication protocols.

Third, we demonstrate the performance gain brought by applying the derived analytical results in protocol designs. Using the bit error performance as the input to the proposed analytical framework, we predict the packet loss rate and loss-burst/loss-gap lengths. We then apply the prediction method to devise an enhanced packetization scheme. Using simulations, we show that the proposed analytical framework can work much faster than existing approaches to characterize the packet loss performance and the proposed packetization scheme can boost the channel throughput by more than 10% with guaranteed transmission delay and failure probability than the traditional scheme.

The remainder of this paper is organized as follows: Section 2 investigates on the correlations of packet losses and develops the metric. Section 3 derives the packet loss performance, including the average packet loss rate and the loss-burst/loss-gap length distributions. Section 4 designs a prediction mechanism and applies it to a packetization scheme. Section 5 reviews the metrics proposed in previous work to measure the packet loss burstiness. Finally, Section 6 concludes this paper.

2. MEASURING PACKET LOSS BURSTINESS

In this section, We first analyze the correlations of packet losses. With this analysis, we then introduce a new metric called the *average correlation length of packet loss rate* to characterize the packet loss burstiness in wireless channels.

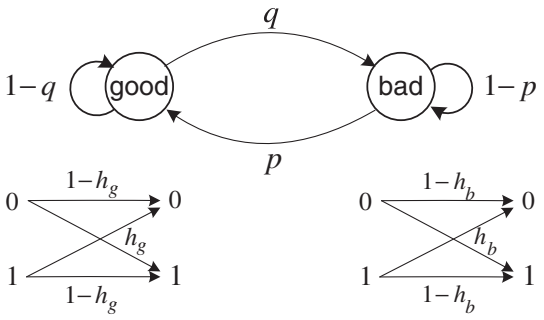


Figure 1. Transition diagram for the GE model.

2.1. Correlation analysis of packet losses

It has been shown that the Gilbert–Elliott (GE) model can capture the bit error statistics[†] with sufficient accuracy in Rayleigh fading channels [8,9], and the GE channel, referring to the wireless channel where the bit errors can be described by the GE model, is the most popular bursty channel used in existing literature. In this paper, we also focus on the GE channel. As shown in Figure 1, the GE channel has two states: $good$ and bad. The state changes from bad to $good$ with probability p and from $good$ to bad with probability q. h_g and h_b are the bit error probabilities in state $good$ and state bad, respectively.[‡]

Previous works evaluate the packet losses typically by enumerating the combinations of the bit states (bit is in $good$ or bad channel conditions) in a packet. As a packet typically consists of hundreds or thousands of bits, it is difficult to evaluate the packet losses efficiently and accurately using such a method. In what follows, we introduce a novel and much more efficient method, namely *packet transmission segmentation*, to analyze the packet loss performance.

(1) *B-run and G-run*: Considering a communication session where multiple packets are transmitted in the unit of bits over the wireless channel, we define a B-run as a maximum-length run (or sequence) of consecutive bits in the bad state and a G-run as a maximum-length run of consecutive bits in

[†]Although the GE model can capture the bit error statistics with sufficient accuracy in Rayleigh fading channels, its effectiveness and accuracy in evaluating the packet error performance still remain an open problem.

[‡]In [10], we have derived the packet loss rate in the simplified Gilbert channel (i.e., $h_g = 0$ and $h_b = 1$) and on the assumption of $p > q$. In this paper, our analysis is over the general GE channel and without the assumption. Moreover, we extend the work to study the correlation length of packet losses and the loss-burst/loss-gap lengths.

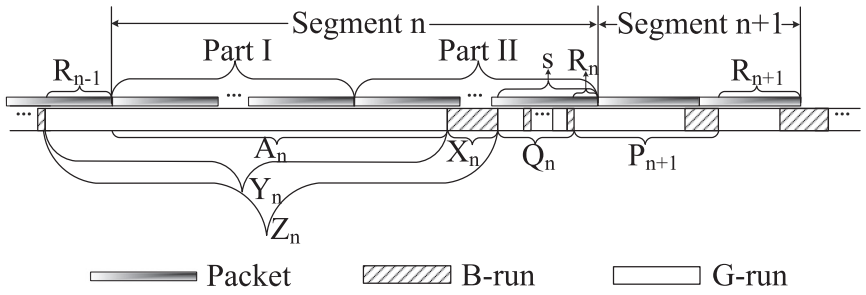


Figure 2. Transmission segments over a wireless channel.

Table I. Notations used in the paper.

p	Transition probability from the *bad* state to the *good* state in the GE model.
q	Transition probability from the *good* state to the *bad* state in the GE model.
h_g	Bit error probability in the *good* state of the GE model.
h_b	Bit error probability in the *bad* state of the GE model.
L	The packet length in terms of bits.
X_n (or X)	The *B*-run length in segment *n*.
Y_n (or Y)	The *G*-run length in segment *n*.
R_n (or R)	The *n*th residual length.
Z_n (or Z)	The sum of a *G*-run length and a adjacent *B*-run length in segment *n*.
A_n (or A)	The remainder of Y_n minus R_{n-1}.
CL	The correlation length of the packet loss rate.
PLR	The average packet loss rate of the communication session.
E_n (or E)	The length of the *n*th loss gap.
F_n (or F)	The length of the *n*th loss burst.
M_n (or M)	$A_n \bmod L$ (or $A \bmod L$).

the *good* state.[§] The communication session is hence composed of the iterative B-runs and G-runs, as shown in Figure 2. Let X and Y denote the lengths of a randomly selected B-run and G-run, respectively. As indicated in [11], in the GE model, both X and Y are independent and identically distributed (IID) random variables and geometrically distributed with the probability mass functions (PMFs) $\Pr(X = k) = p(1-p)^{k-1}$ and $\Pr(Y = k) = q(1-q)^{k-1}$, respectively. Table I summarizes the primary notations used in this paper.

(2) *Residual length*: We divide the whole transmission session into multiple segments for analysis on the basis of the residual length of packets. The residual length of a packet is defined as the length (in the unit of bits) from the bit immediately after the last B-run in a packet to the end of the packet. Let R_n, where $n = 1, 2, \ldots$, denote the nth residual length during the communication session, as shown in Figure 2. R_n varies within $[0, L-1]$, where L is the packet length.[¶]

(3) *Packet transmission segmentation*: We define segment n ($n = 1, 2, \ldots$) as the length from the bit right after R_{n-1} to the last bit of R_n (R_0 is the start of the communication session), as shown in Figure 2. Apparently, the segmentation does not break any single packet, and each segment can be divided into two parts: in *Part I*, all the bits in the packet are in the *good* state; whereas in *Part II*, the packet must contain at least 1 bit in the *bad* state. We first present the properties of segmentations to reveal the correlations of packet losses.

Lemma 1. *R_{n-1} and R_n are independent, and R_n has the following probability distribution.*

$$\Pr(R_n = k) = \frac{p(1-q)^{k+1} - q(1-p)^{k+1}}{\frac{p(1-q)(1-(1-q)^L)}{q} - \frac{q(1-p)(1-(1-p)^L)}{p}}, \quad k \in [0, L) \tag{1}$$

Proof. Denote the sum of a G-run length and an adjacent B-run length in segment n by a new random variable Z_n, that is, $Z_n = Y_n + X_n$. Z_n is memoryless as X_n and Y_n are both memoryless. Intuitively, R_{n-1} and R_n are independent according to the memoryless property of Z_n. See Appendix A for the detailed proof. □

[§]The adjacent bits ahead of and after a B-run (resp. G-run) are in the $good$ state (resp. the bad state).

[¶]R_n equals to 0 when the last bit of the packet is in the bad state and the following bit is in the $good$ state.

Let A_n denote the portion that Y_n minus R_{n-1}, as shown in Figure 2. From Lemma 1, we have Lemma 2 as follows.

Lemma 2. *X_{n-1}, Y_{n-1}, and A_{n-1} are independent of A_n, and A_n has the probability distribution as*

$$\Pr(A_n = k) = \begin{cases} \frac{1}{\frac{p(1-q)(1-(1-q)^L)}{q} - \frac{q(1-p)(1-(1-p)^L)}{p}} (1-q)^{k-1} \left(\frac{q^2(1-p)(1-((1-p)(1-q))^L)}{pq-p-q} + \frac{p(1-q)(1-(1-q)^{2L})}{2-q} \right), & k \in [1, \infty), \\ \frac{1}{\frac{p(1-q)(1-(1-q)^L)}{q} - \frac{q(1-p)(1-(1-p)^L)}{p}} \left(\frac{q^2(1-p)^{2-k}(1-((1-p)(1-q))^{L+k-1})}{pq-p-q} + \frac{p(1-q)^{2-k}(1-(1-q)^{2(L+k-1)})}{2-q} \right), & k \in [2-L, 0] \end{cases} \tag{2}$$

Proof. Intuitively, X_{n-1} is independent of R_{n-1} as X_n and Y_n are both memoryless. X_{n-1} and A_n are then independent because X_{n-1} is also independent of Y_n and $A_n = Y_n - R_{n-1}$. More strictly, we can derive $\Pr(A_n = a) = \Pr(A_n = a | X_{n-1} = x)$ similarly as in Appendix A to prove that X_{n-1} is independent of A_n (the derivation is omitted here because of the space limit). Likewise, we can also prove that Y_{n-1} and A_{n-1} are independent of A_n. The PMF of A_n can be derived from the PMFs of Y_n and R_{n-1}. □

For simplicity, in this paper, we do not consider the error-recovery (e.g, ARQ) or error-resilient (e.g, forward error correction (FEC) codes) mechanisms. In other words, once a bit error occurs in a packet, the packet will be corrupted. Then, from Lemma 2, we obtain the following theorem.

Theorem 1. *The average packet loss rate in segment $(n-1)$ is independent of that in segment n.*

Proof. Let N_n and $N_{e,n}$ denote the total number of packets and the number of error packets in segment n, respectively. The average packet loss rate in segment n thus equals to $\frac{N_{e,n}}{N_n}$. According to the definitions of R_n and segment n, the length of segment n minus $(A_n + X_n)$ is no longer than L, so N_n can be derived as

$$N_n = \sum_{x=1}^{\infty} \sum_{a=2-L}^{\infty} \Pr(X_n = x) \Pr(A_n = a) \left\lceil \frac{x+a}{L} \right\rceil \tag{3}$$

As shown in Figure 2, $N_{e,n}$ can be derived as

$$N_{e,n} = \sum_{x=1}^{\infty} \sum_{a=2-L}^{\infty} \Pr(X_n = x) \Pr(A_n = a) f(a, x) \tag{4}$$

where $f(a, x)$ is the number of error packets in segment n when $A_n = a$ and $X_n = x$.

As both A_{n-1} and X_{n-1} are independent of A_n and X_n, N_n is independent of N_{n-1}. Similarly, $N_{e,n}$ is also independent of $N_{e,n-1}$ because $f(a, x)$ is only related to A_n, X_n, and Y_n. Then, the length of a segment and the number of error packets in a segment are both independent. As a result, Theorem 1 holds. □

From Theorem 1, we know that the correlations of packet losses only last for a segment.

2.2. Average correlation length of packet loss rate

The correlation length of packet loss rate, denoted as CL, is defined as the minimum interval to guarantee that the average packet loss rates are independent in different intervals. From Theorem 1, we know that the average packet loss rates in different segments are independent. Thus, the correlation length of packet loss rate (in the unit of packets) equals to the number of packets in a segment. From here onwards, for simplicity we remove the subscript n of notations in Table I. The PMF of CL is then derived as $\Pr(CL = k) = \Pr\left(\left\lceil \frac{A+X}{L} \right\rceil = k \right)$. From the PMFs of X and A, we can derive that

$$\Pr(CL = k) = \left(\sum_{x=k \cdot L}^{(k+1) \cdot L-2} \sum_{a=2-L}^{k \cdot L-x} + \sum_{x=1}^{k \cdot L-1} \sum_{a=(k-1) \cdot L+1-x}^{k \cdot L-x} \right) \Pr(A = a) \Pr(X = x), \quad k \geq 1 \tag{5}$$

From Equation (5), we can obtain the average correlation length of packet loss rate, that is, $E[CL]$ (the closed-form solutions for $\Pr(CL = k)$ and $E[CL]$ are represented in [12]). $E[CL]$ is the metric that we propose to measure the packet loss burstiness. The larger $E[CL]$ is, the longer the correlations of packet losses stay, indicating that the wireless channel is more bursty. If $E[CL] = 1$, the packet losses are memoryless.

2.3. Comparisons to other metrics

In previous literature, a variety of metrics have been proposed to measure the packet loss burstiness, including the metric μ [13], the Allan deviation of PRR [5], and the

metric β [7]. This section compares our proposed metric with the existing metrics.

If the packet errors can be described by the GE model in wireless networks, the metric μ [13] is calculated as

$$\mu = 1 - p^{(p)} - q^{(p)} \tag{6}$$

where $p^{(p)}$ and $q^{(p)}$ are the probabilities of the state transiting from $good$ to bad and from bad to $good$ in the GE model, respectively. The superscript (p) denotes that it is the packet-level quantity. The larger μ is, the more bursty the channel is.

The formula of the Allan deviation of PRR [5], denoted as D, is

$$D = \sqrt{\frac{1}{2n}\sum_{2}^{n}(r_i - r_{i-1})^2} \tag{7}$$

where r_i is the PRR in interval i. In the cases of different time scales, we can obtain different Allan deviations. For the case that the interval lengths are approximately equal to the packet-loss-burst length, the PRRs of different intervals deviate a lot, and accordingly, the Allan deviation is large. For the case of smaller intervals, the PRRs of adjacent intervals change slowly, making the Allan deviation small. In the case of large intervals, the PRRs of different intervals tend to be the long-term average, and the Allan deviation is small.

The metric β [7] is calculated from CPDFs. The CPDF, denoted as $C(n)$, is the probability that the next packet will succeed given n consecutive packet successes (for $n > 0$) or failures (for $n < 0$). The formula of β is

$$\beta = \frac{KW(I) - KW(E)}{KW(I)} \tag{8}$$

where $KW(\cdot)$ is the Kantorovich–Wasserstein (KW) distance [14] from the CPDF of the ideal bursty link.[‖] E is the CPDF of the link in question, and I is the CPDF of an independent link with the same average PRR. An ideal bursty link has a $\beta = 1$, and a link with independent packet losses has a $\beta = 0$. Negative β values are permitted, and this happens when there is a negative correlation in packet losses; that is, the next packet is more likely to be corrupted when more packets are received, and the next packet is more likely to be delivered successfully when more packets are corrupted.

Figure 3 compares the inverse of our metric $\frac{1}{E[CL]}$ to previous metrics μ and β in different wireless channels. We set $h_g = 0$, $h_b = 1$, and the other parameters in Figure 3(a)–(c) respectively as follows: $p = 10^{-4}, q = 5 \times 10^{-4}, L \in [2000, 20000]$; $p \in [10^{-4}, 10^{-3}], q = 5 \times 10^{-4}, L = 8000$; and $p = 10^{-4}, q \in [10^{-4}, 10^{-3}], L = 8000$. We use the GE model to model the packet errors and calculate $p^{(p)}$ and $q^{(p)}$ to obtain μ[**]. In Figure 3(a), we see that the packet loss burstiness decreases when L increases, and the burstiness is vanished when L is long enough. This matches with our intuitive understanding. Figure 3(b) shows that the burstiness also decreases when p or q increases. This is because in the GE model, the memory will fade away with p or q increasing (i.e., $p + q$ approaches to 1) because the memory will disappear when $p = 1 - q$. In Figure 3(b), the values of μ are close to 0 almost all the time, whereas the variation of $\frac{1}{E[CL]}$ is more obvious, implying that our metric $E[CL]$ can reveal the burstiness more effectively than μ when the packet loss burstiness is low.

In Figure 3(a), β decreases monotonically with L increasing when L is small. However, when L is large, we have $\beta < 0$ and we cannot observe any patterns that β follows. This is because more transmission errors would be encountered when increasing the packet length L, and intensive transmissions errors would cause the negative correlations in packet losses. In addition, the metric β cannot reveal the packet loss burstiness when $\beta < 0$ as that when $\beta \geq 0$ (we can get that from the calculation of β). In Figure 3(b) and (c), we also show the negative β values in different p and q, and we cannot find the pattern of β values either. Therefore, our metric $E[CL]$ can characterize the packet loss burstiness more effectively than β when the transmission errors are severe.

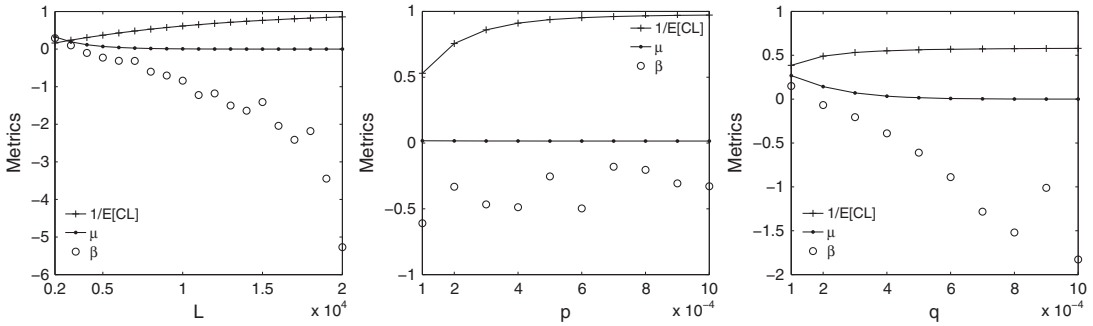
Figure 4 shows the Allan deviation of PRR (D) in different time scales (in the unit of packets) and with different average correlation lengths of packet loss rate ($E[CL]$) and different average loss-burst lengths ($E[F]$). The four combinations of $E[CL]$ and $E[F]$ values are obtained with the settings of p,q, and L, respectively: $p = 10^{-6}, q = 10^{-5}, L = 8000$; $p = 10^{-5}, q = 10^{-5}, L = 8000$; $p = 10^{-5}, q = 10^{-4}, L = 8000$; and $p = 10^{-6}, q = 10^{-4}, L = 8000$. We can see that the Allan deviation of PRR is maximized when the time scale is near the average loss-burst length, indicating that we can obtain the average loss-burst length approximately by analyzing the Allan deviations of PRR in different time scales. However, the average loss-burst length is not a good metric to measure the packet loss burstiness because it cannot tell whether the wireless channel is of burst losses or independent losses. In addition, we cannot obtain $E[CL]$ or other clues to characterize the packet loss burstiness from the Allan deviations. Therefore, the Allan deviation of PRR is not suitable to measure the packet loss burstiness.

3. CLOSED-FORM SOLUTIONS FOR PACKET LOSS PERFORMANCE

This section derives the packet loss performance, including the average packet loss rate and the loss-burst/loss-gap

[‖]In the ideal bursty link, $C(n) = 0$ if $n < 0$ and $C(n) = 1$ if $n > 0$.

[**]In Section 3.2.2, we demonstrate that the packet errors also follow the GE model if the bit errors can be captured by the GE model, and derive the closed-form solutions for $p^{(p)}$ and $q^{(p)}$.



(a) L. (b) p. (c) q.

Figure 3. Comparisons of $\frac{1}{E[CL]}$, μ, and β with different L, p, and q.

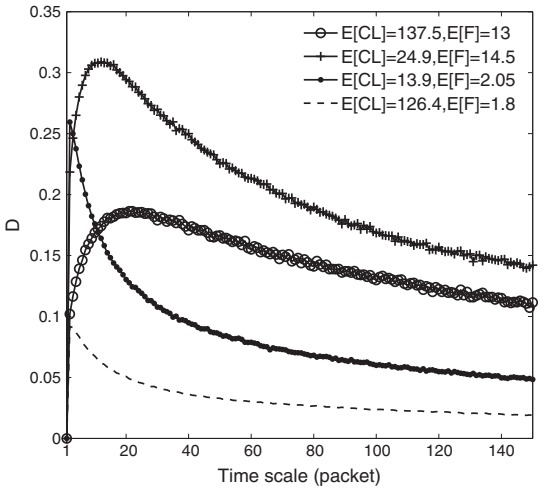


Figure 4. Allan Deviation of PRR (D) with different $E[CL]$ and $E[F]$.

length distributions based on our correlation analysis and the metric CL proposed earlier.

3.1. Average packet loss rate

Let PLR denote the average packet loss rate in the communication session. It equals to the average packet loss rate in a single segment because the length of a segment is just the correlation length of packet loss rate CL.

As shown in Figure 2, each segment can be divided into two parts. $f_1(a,x)$ and $f_2(a,x)$ are denoted as the numbers of error packets in *Part I* and *Part II*, respectively. Then, $f(a,x)$ in Equation (4) equals to $f_1(a,x) + f_2(a,x)$.

For a packet, n_b is denoted as the number of bits in the bad state, then the packet loss probability is $1-(1-h_g)^{L-n_b}(1-h_b)^{n_b}$. Therefore, $f_1(a,x) = \lfloor \frac{a}{L} \rfloor \cdot (1-(1-h_g)^L)$. In the last packet of *Part II*, denote the length of the bits after the first B-run as s, as shown in Figure 2. The number of bits in the bad state in s cannot be derived easily because the number of B-runs cannot be determined. In the calculation of $f_2(a,x)$, we use the average number of bits in the bad state, that is, $s \cdot \frac{q}{p+q}$, as an approximation.[††] Then by denoting that $a = q_a \cdot L + r_a$ and $x = q_x \cdot L + r_x$ ($0 \le q_x, q_a;\ 0 \le r_x, r_a < L$),[‡‡] we obtain

$$\begin{aligned} f_2(a,x) = & H\left(\lceil \frac{r_a + x}{L} \rceil - 2\right) \cdot \left(1-(1-h_g)^{r_a}(1-h_b)^{L-r_a}\right) \\ & + \max\left(0, \lceil \frac{r_a + x}{L} \rceil - 2\right) \cdot (1-(1-h_b)^L) \\ & + \left(1-(1-h_g)^{L-n_b}(1-h_b)^{n_b}\right) \end{aligned} \tag{9}$$

where $H(n) = \begin{cases} 0, & n < 0, \\ 1, & n \ge 0, \end{cases}$ and n_b can be derived as

$$n_b = \begin{cases} r_x + \frac{q}{p+q}(L - r_a - r_x), & r_x + r_a \le L, q_x = 0, \\ L - \frac{p}{p+q}(r_a + r_x), & r_x + r_a \le L, q_x \ge 1, \\ L - \frac{p}{p+q}(2L - r_a - r_x), & r_x + r_a > L, q_x \ge 0 \end{cases}$$

The three components in Equation (9) are the number of packets that contain the bits of the first G-run in a segment, the number of packets that are totally located in the B-run, and the number of packets that contain the bits of s, respectively.

Then, we derive the closed-form expression of PLR. The expression is omitted here as it is lengthy. See [12] for the details.

When the wireless channel is not in a very bad condition, the period of good channel state is usually longer than the period of bad channel state; that is, p is larger than q.

[††]This approximation will not affect the accuracy of PLR too much because it only exists in the last packet of a segment and s is usually not long. The simulation results in Section 3.3.1 also validate that.

[‡‡]q_a (resp. q_x) is the quotient of a (resp. x) divided by L and r_a (resp. r_x) is the remainder.

Then, we have $\frac{(1-p)^L}{(1-q)^L} \approx 0$ and can obtain a simplified expression for PLR accordingly as

$$\begin{aligned} PLR = 1 - \left(1 - (1 - \frac{q}{t_2}) \cdot \frac{1 - PLR_0}{(1-q)^L}\right) \\ \cdot p \cdot t_1 - \left(1 - (1 - \frac{p}{t_2}) \cdot \frac{q(1-h_b)}{t_3}\right) \\ \cdot (1-h_g)^L \cdot (1 - PLR_0) \end{aligned} \quad (10)$$

where PLR_0 is the packet loss rate when $h_g = 0$ and $h_b = 1$, and

$$\begin{aligned} PLR_0 &= 1 - \frac{(p-q)(1-q)^L}{(p-q) - q(1-q)^L(1-p)} \\ t_1 &= \frac{(1-h_b)^{\frac{qL}{p+q}}(1-h_g)^{\frac{pL}{p+q}}}{p - 1 + (\frac{1-h_g}{1-h_b})^{\frac{p}{p+q}}} \\ t_2 &= 1 + (q-1) \cdot (\frac{1-h_g}{1-h_b})^{\frac{q}{p+q}} \\ t_3 &= (1-q)h_g - (1-p)h_b + q - p \end{aligned} \quad (11)$$

3.2. Loss-burst/loss-gap length distributions

We first study the loss-burst/loss-gap length distributions in the simplified Gilbert channel, which is a special case of the GE channel when $h_g = 0$ and $h_b = 1$.

3.2.1. In the simplified Gilbert channel.

From Lemma 2, we have Theorem 2 as follows.

Theorem 2. *In the simplified Gilbert channel, both the loss-gap length and the loss-burst length are IID, and the loss-gap length is independent of the loss-burst length.*

Proof. In the simplified Gilbert channel, the bits are error-free in the *good* state and are always error in the *bad* state. Therefore, as shown in Figure 2, for a segment, in *Part I*, $\left\lfloor \frac{A_n}{L} \right\rfloor$ packets are transmitted correctly, whereas in *Part II*, $\left\lceil \frac{M_n + X_n}{L} \right\rceil$ packets are corrupted, where M_n equals to $A_n \bmod L$, that is, the remainder of A_n divided by L. According to the definition of a segment, apparently, a loss gap never crosses multiple segments because a segment always ends with an error packet, and a loss burst may cross different segments when all the packets in a segment are corrupted. For example, as shown in Figure 2, the packets in segment $(n+1)$ are all corrupted. With these observations, we derive the proof as follows.

Let E_i and F_i denote the lengths of the ith loss gap and the ith loss burst, respectively. Assuming that E_i is in segment n, the PMF of E_i is thus

$$\Pr(E_i = k) = \Pr\left(\left\lfloor \frac{A_n}{L} \right\rfloor = k\right) \quad k \geq 1 \quad (12)$$

As different loss gaps are within separate segments and A_n is IID, the loss-gap length is IID.

Assume that F_i crosses m ($m \geq 1$) segments (starting from segment n to segment $n+m-1$); denote the number of error packets in segment n as $N_{e,n}$ ($N_{e,n} \geq 1$), and denote by

$$\begin{aligned} C = \{&N_{e,n} + N_{e,n+1} + \ldots + N_{e,n+m-1} = k : \\ &N_{e,n+j} \geq 1 (0 \leq j < m)\}, \end{aligned}$$

$$\begin{aligned} W = \Bigg\{ &\left(\bigcup_{j=0}^{m-1} \left\lceil \frac{M_{n+j} + X_{n+j}}{L} \right\rceil \right. \\ &\left. = N_{e,n+j} \right) \bigcup \left(\bigcup_{j=1}^{m-1} A_{n+j} < L \right) \bigcup (A_{n+m} \geq L) \Bigg\} \end{aligned}$$

The probability that $F_i = k$ equals to the probability that there are k packets in the m segments. The PMF of F_i can then be derived by

$$\Pr(F_i = k) = \sum_{m=1}^{k} \sum_{C} \Pr(W) \quad k \geq 1 \quad (13)$$

where $\left(\bigcup_{j=1}^{m-1} A_{n+j} < L \right)$ in W indicates that the error packets span to m consecutive segments, and $(A_{n+m} \geq L)$ in W indicates that the loss burst terminates after m segments.

Apparently, different loss bursts will appear in separate segments, and A_n, X_n, and M_n are all IID. Therefore, the loss-burst length is also IID.

Let $W' = W - \left\{ \left\lceil \frac{M_n + X_n}{L} \right\rceil = N_{e,n} \right\} - \{A_{n+m} \geq L\}$. From (12), we have

$$\begin{aligned} \Pr(W | E_i = k_0) = \Pr\left(\left\lceil \frac{M_n + X_n}{L} \right\rceil = N_{e,n} | E_i = k_0 \right) \\ \Pr(W') \Pr(A_{n+m} \geq L) \end{aligned} \quad (14)$$

It can be easily proved that $\Pr(M_n = m | E_i = k_0) = \Pr(M_n = m)$. As a result, $\Pr\left(\left\lceil \frac{M_n + X_n}{L} \right\rceil = N_{e,n} | E_i = k_0 \right) = \Pr\left(\left\lceil \frac{M_n + X_n}{L} \right\rceil = N_{e,n} \right)$, which implies that E_i is independent of F_i.

As F_i spans from segment n to segment $(n+m-1)$, E_{i+1} would appear in segment $(n+m)$. Consequently, $\Pr(E_{i+1} = k) = \Pr\left(\left\lfloor \frac{A_{n+m}}{L} \right\rfloor = k \right)$. Similarly, we can prove that F_i is independent of E_{i+1}.

Therefore, Theorem 2 holds: both the loss-gap length and the loss-burst length are IID, and the loss-gap length is independent of the loss-burst length. □

For simplicity of notation, we remove the subscript i in the remaining analysis of this section. From Equations (12)

and (13), we have the PMFs of the loss-gap length and the loss-burst length approximately as

$$\begin{aligned}\Pr(E=k) &= (1-(1-q)^L)((1-q)^L)^{k-1}, && k \geq 1,\\ \Pr(F=k) &= \frac{(1-q)^{L-1}(1-(1-q)^L)(1-(1-p)^L)}{qL}\left(1-\frac{(1-q)^{L-1}(1-(1-q)^L)(1-(1-p)^L)}{qL}\right)^{k-1}, && k \geq 1\end{aligned} \tag{15}$$

where

The detailed derivation is given in Appendix B. Therefore, in the simplified Gilbert channel, the loss-gap length and the loss-burst length follow geometric distributions with parameters $1-(1-q)^L$ and $\frac{(1-q)^{L-1}(1-(1-q)^L)(1-(1-p)^L)}{qL}$, respectively.

With the loss-burst/loss-gap length distributions, we can develop a model to characterize the packet losses in the simplified Gilbert channel. Let γ_n denote the status of the nth packet transmitted, where $\gamma_n = 0$ if the nth packet is corrupted and $\gamma_n = 1$ if the packet is correctly transmitted. From Theorem 2, we know that in the simplified Gilbert channel, both the loss-gap length and the loss-burst length are IID, and the loss-gap length is independent of the loss-burst length. Therefore, we have that

$$\begin{aligned}\Pr(\gamma_{n+1}=1|\gamma_n=1) &= \frac{\Pr(E \geq i+1)}{\Pr(E \geq i)},\\ \Pr(\gamma_{n+1}=0|\gamma_n=1) &= \frac{\Pr(E=i)}{\Pr(E \geq i)}\end{aligned} \tag{16}$$

where i refers to that there are i consecutive correct packets before the $(n+1)$th packet, and

$$\begin{aligned}\Pr(\gamma_{n+1}=1|\gamma_n=0) &= \frac{\Pr(F=j)}{\Pr(F \geq j)},\\ \Pr(\gamma_{n+1}=0|\gamma_n=0) &= \frac{\Pr(F \geq j+1)}{\Pr(F \geq j)}\end{aligned} \tag{17}$$

where j refers to that there are j consecutive corrupted packets before the $(n+1)$th packet.

By substituting Equation (15) into Equations (16) and (17), we have

$$\begin{aligned}&\Pr(\gamma_{n+1}=1|\gamma_n=1) = (1-q)^L,\\ &\Pr(\gamma_{n+1}=0|\gamma_n=1) = 1-(1-q)^L,\\ &\Pr(\gamma_{n+1}=1|\gamma_n=0)\\ &\quad = \frac{(1-q)^{L-1}(1-(1-q)^L)(1-(1-p)^L)}{qL},\\ &\Pr(\gamma_{n+1}=0|\gamma_n=0)\\ &\quad = 1-\frac{(1-q)^{L-1}(1-(1-q)^L)(1-(1-p)^L)}{qL}\end{aligned} \tag{18}$$

Therefore, we develop a simplified Gilbert model for packet losses in the simplified Gilbert channel with the transition probability matrix between the $good^{(p)}$ state and the $bad^{(p)}$ state as

$$\begin{pmatrix} 1-q^{(p)} & q^{(p)} \\ p^{(p)} & 1-p^{(p)} \end{pmatrix} \tag{19}$$

where

$$\begin{aligned}p^{(p)} &= \Pr(\gamma_{n+1}=1|\gamma_n=0),\\ q^{(p)} &= \Pr(\gamma_{n+1}=0|\gamma_n=1)\end{aligned} \tag{20}$$

The packet loss probability in the $good^{(p)}$ state, denoted by $h_g^{(p)}$, is 0, and the packet loss probability in the $bad^{(p)}$ state, denoted by $h_b^{(p)}$, is 1.

3.2.2. In the GE channel.

The loss-burst/loss-gap length distributions in the general GE channel cannot be derived so easily as in the simplified Gilbert channel. In this section, we first develop a model to characterize the packet losses in the GE channel based on the model in the simplified Gilbert channel, and then derive the loss-burst/loss-gap length distributions.

We now extend the aforementioned packet-level simplified Gilbert model to the GE channel. In the GE channel, the transition probability matrix between the $good^{(p)}$ state and the $bad^{(p)}$ state remains the same as that in the simplified Gilbert channel, whereas the packet loss probability in the $good^{(p)}$ state ($h_g^{(p)}$) is nonzero and equals to $(1-(1-h_g)^L)$. The packet losses in the $bad^{(p)}$ state is no longer independent, and the packet loss probability depends on the number of error bits in a packet. In that case, to develop a precise model for packet losses according to these different loss probabilities, we need to add a multitude of new states, and the correlations between these states are strong because of the big number of bits in a packet and the correlation between the states of two adjacent bits in the GE channel. Here, in the $bad^{(p)}$ state, we use the average packet loss rate as the packet loss probability ($h_b^{(p)}$) and construct a first-order Markov model (i.e., a GE model) in the GE channel approximately.

According to the GE model, the average packet loss rate equals to $\frac{h_g^{(p)}p^{(p)}+h_b^{(p)}q^{(p)}}{p^{(p)}+q^{(p)}}$. Recall that in Section 3.1, we have derived the average packet loss rate (PLR) in the communication session, then $PLR = \frac{h_g^{(p)}p^{(p)}+h_b^{(p)}q^{(p)}}{p^{(p)}+q^{(p)}}$. Given the knowledge of PLR, $h_g^{(p)}$, $p^{(p)}$, and $q^{(p)}$, we can derive $h_b^{(p)}$ by $\frac{PLR\cdot(p^{(p)}+q^{(p)})-h_g^{(p)}p^{(p)}}{q^{(p)}}$.

The simplified Gilbert model is a special case of the GE model. As indicated in [15], the simplified Gilbert model is of renewal nature; that is, both the loss-burst length and the loss-gap length are independently distributed. The loss-burst length and the loss-gap length are then derived to be geometric distributions with parameters $q^{(p)}$ and $p^{(p)}$, respectively. Yet, the GE model is non-renewable, and the loss-burst/loss-gap length distributions are difficult

to derive. In [15], Pimentel and Blake gave a recursive solution for them using the enumeration theory. In the GE channel, we then can use such recursive solution to calculate the loss-burst/loss-gap length distributions.

In what follows, we discuss on the values of $h_g^{(p)}$ and $h_b^{(p)}$ on the basis of the typical values of p, q, h_g, and h_b. As indicated in [16], with vehicle speeds to be 50 km/h, modulation frequency $f = 900$ MHz, and Rayleigh fading envelope $\rho = -10$ dB, we have $\frac{p}{q} = 9.3$. With pedestrian speeds to be 2 km/h, the same modulation frequency, and $\rho = -20$ dB, we have $\frac{p}{q} = 99.4$. h_g is usually set to $1 \cdot 10^{-5}$, and h_b is set to 0.1. In Figure 5, $h_b^{(p)}$ is plotted with $h_g = 1 \cdot 10^{-5}$, $h_b = 0.1$, and different q and L as (1) $q = 10^{-3}$ and $L = 5000$, and (2) $q = 10^{-5}$ and $L = 10\,000$, respectively. We see that if p is small, $h_b^{(p)}$ equals to 1. Meanwhile, $h_g \ll 1$ makes $h_g^{(p)}$ approach to 0, which implies that the GE model for packet losses degrades to the simplified Gilbert model when p is small. p being small means that the channel is in bad condition as $\frac{1}{p}$ is the average period length of bad channel state. In addition to that, the GE model will be memoryless if p approaches to 1 as we have $p^{(p)} \approx 1 - q^{(p)}$ in Equation (19). This matches the analysis in Section 2.3: using the average correlation length of the packet loss rate, we demonstrate that the correlation between packet losses decreases when increasing p.

In summary, the GE model for the packet losses will regress to the simplified Gilbert model when the GE channel regresses to the simplified Gilbert channel or the channel is in bad condition, and the GE model will be memoryless when the channel is in good condition.

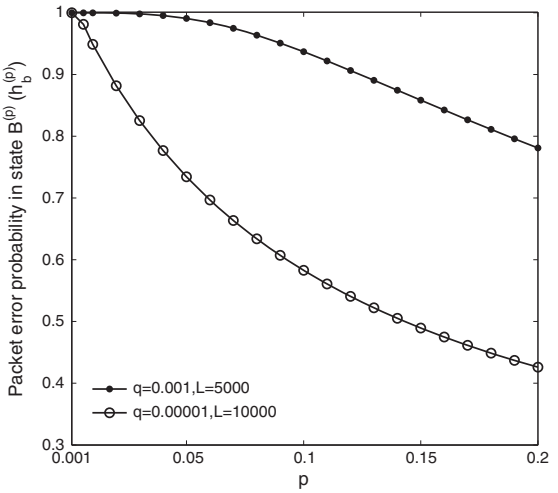


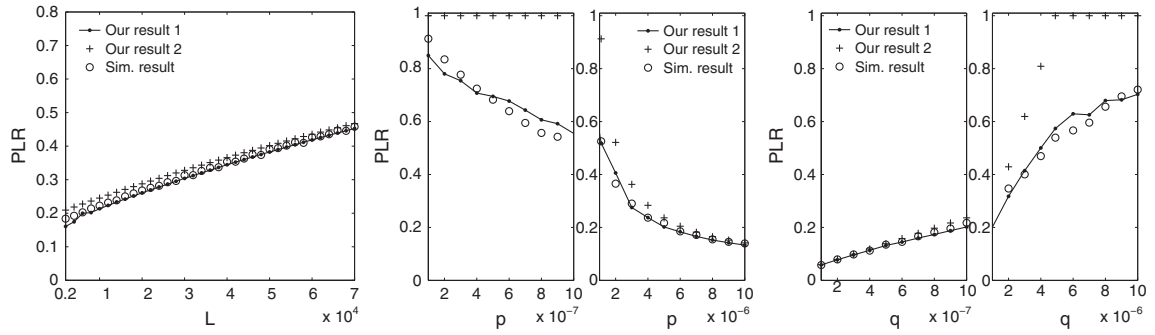
Figure 5. The packet loss probability in the *bad*$^{(p)}$ state ($h_b^{(p)}$) with the increase of p.

3.3. Simulation results

This section demonstrates the accuracy of the derived closed-form models for the average packet loss rate and the loss-burst/loss-gap length distributions in different settings.

3.3.1. Average packet loss rate.

Figure 6 shows the effects of L, p, and q on the packet loss rate when $h_g = 5 \times 10^{-6}$ and $h_b = 0.9$. The parameters p, q, and L, in Figure 6(a)–(c) are, respectively, $p = 5 \times 10^{-6}, q = 10^{-6}, L \in [2000, 70\,000]$; $p \in [10^{-7}, 10^{-5}], q = 10^{-6}, L = 8000$; and $p = 5 \times 10^{-6}$, $q \in [10^{-7}, 10^{-5}]$, $L = 8000$. We observe that the average packet loss rate increases with L increasing, indicating that increasing the packet length would incur more transmission errors; and the average packet loss rate increases monotonically with p decreasing or q increasing as the channel condition becomes worse in this case. *Our result 1* and *Our result 2* refer to the derived analysis without simplification and with the simplification in Equation (10), respectively. It can be seen that our results without simplification are accurate for different L, p, and q. In Figure 6(b) and 6(c), the results with the simplification in Equation (10) are inaccurate when p is close to or smaller than q, and they match the simulations well when $p > q$. This indicates that Equation (10) is accurate enough



(a) L. (b) p. (c) q.

Figure 6. Packet loss rate (PLR) with the increases of L, p, and q.

to predict the packet loss rate of the wireless channel in good channel condition.

3.3.2. Loss-burst/loss-gap lengths.

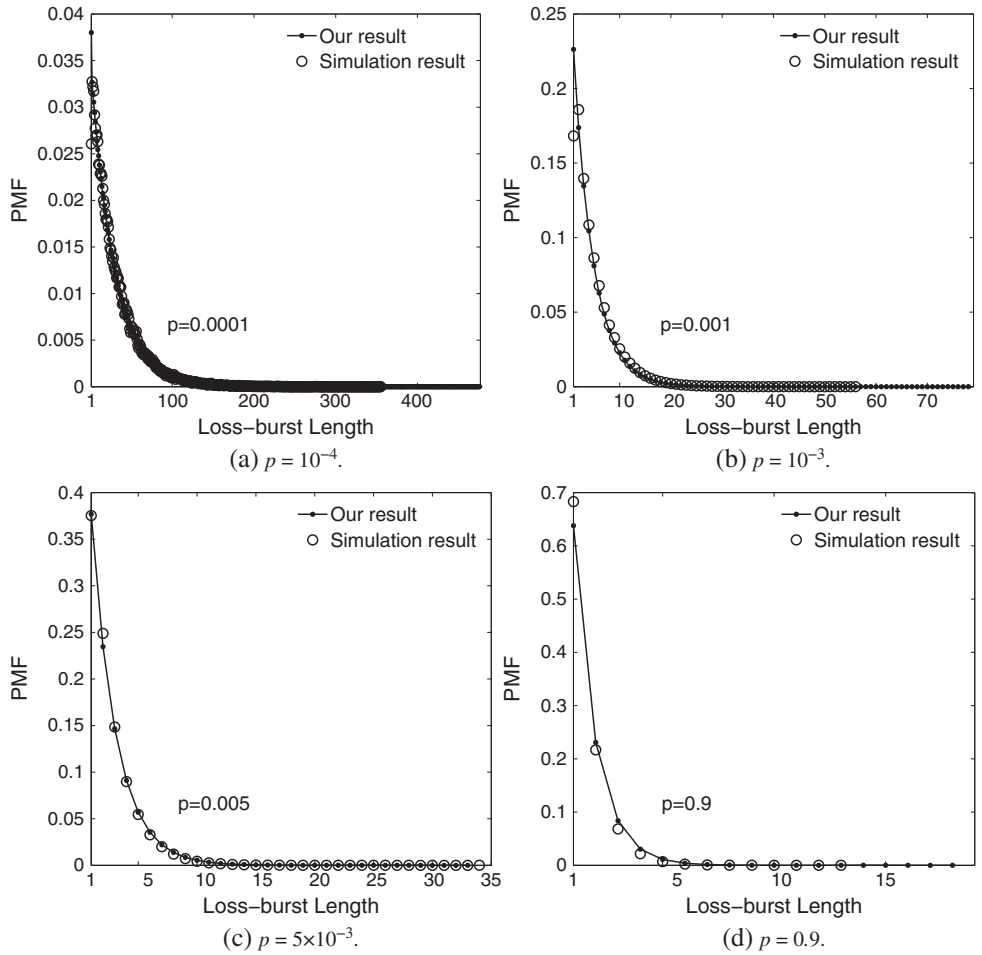
In this experiment, we evaluate the loss-burst/loss-gap length distributions and the average loss-burst/loss-gap lengths by changing p in the range of $[10^{-4}, 0.9]$ and setting other parameters as $q = 7 \times 10^{-4}, L = 1000, h_g = 10^{-5}, h_b = 0.5$.

Figure 7 shows the PMFs of the loss-burst length with different p. From Figure 7(a)–(d), we can see that our results are slightly inaccurate when $p = 10^{-4}$ and with acceptable accuracy when $p = 10^{-3}$. The results accurately match the simulations when $p = 5\times10^{-3}$. However, the accuracy slightly decreases when $p = 0.9$. The gap between the simulations and analysis mainly attributes to the approximation we have made to calculate $\Pr(E = 1)$ in the analysis of the loss-burst length distribution (see Appendix B for details). Figure 8 shows the PMFs of the loss-gap length. The accuracy of the loss-gap length is not sensitive to p, and our results are accurate for all p. Figure 9 shows the effects of p on the average loss-burst/loss-gap lengths. The results for both the average loss-burst and the average loss-gap length are accurate for all p.

In summary, our results matches the simulations very well, and there exist a slight error when the probability of loss-burst length equals to 1.

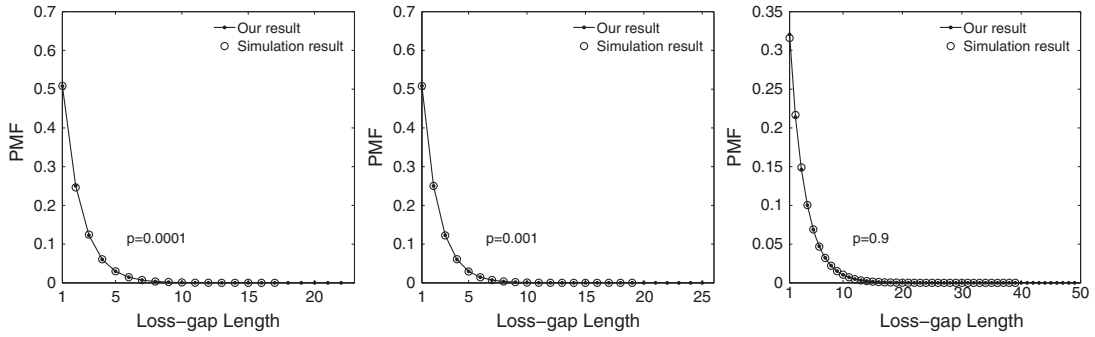
4. PROTOCOL IMPROVEMENTS

To know the packet loss rate and the loss-burst/loss-gap lengths is key for the evaluation or optimization of communication protocols and algorithms. For example, the packet loss rate is an important metric to design FEC codes, ARQ, and adaptive packetization schemes [17–19]; the loss-burst/loss-gap lengths can be applied to the design of the retransmission time and the route protocols [7]. The existing methods to predict the packet loss rate and the loss-burst/loss-gap lengths, however, are often complicate yet not accurate enough.



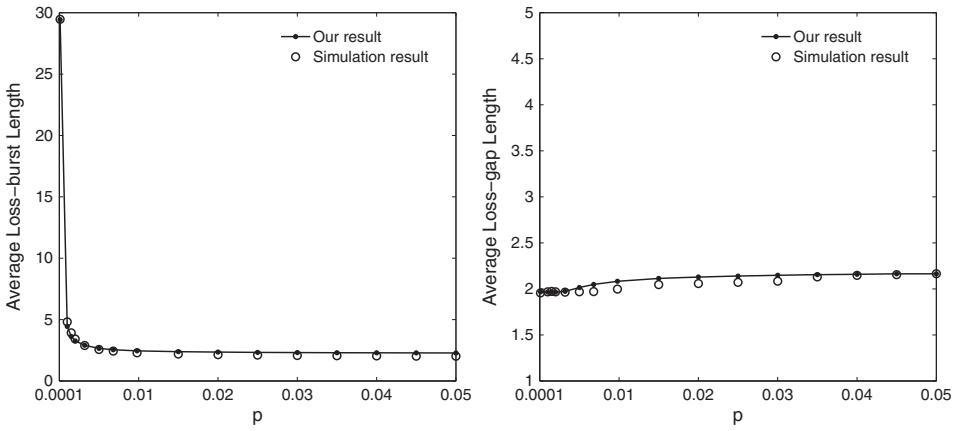
(a) $p = 10^{-4}$. (b) $p = 10^{-3}$. (c) $p = 5\times10^{-3}$. (d) $p = 0.9$.

Figure 7. The probability mass function (PMF) of the loss-burst length with different p.



(a) $p = 10^{-4}$. (b) $p = 10^{-3}$. (c) $p = 0.9$.

Figure 8. The probability mass function (PMF) of the loss-gap length with different p.



(a) The average loss-burst length. (b) The average loss-gap length.

Figure 9. The average loss-burst/loss-gap lengths with the increase of p.

In the previous sections, we have derived the closed-form expressions of the average packet loss rate and the average loss-burst/loss-gap lengths. On the basis of the results, in what follows, we propose an analytical method to predict the packet loss rate and the loss-burst/loss-gap lengths in wireless networks, and design a transmission-delay-constrained and failure-probability-constrained packetization scheme accordingly as an example to showcase the implementation of the analytical models.

4.1. Prediction of channel performance

As indicated in [16], with a specific digital modulation scheme and knowing the parameters of the channel fading model, such as the maximum Doppler frequency and the fading envelope, we can derive the parameters of the bit-level GE model, that is, p, q, h_g, and h_b (*Phase I*). In this paper, we bridge the bit-level GE model with the packet loss rate in Section 3.1 and the loss-burst/loss-gap lengths in Section 3.2 (*Phase II*). By combining the two phases, we can derive the packet loss performance given the parameters of the channel fading model. The computation complexity of this prediction method is constant as the computation complexities of the two phases are both constant. In addition, our update interval is adapted to the dynamics of the fading model parameters, and it is longer than the update interval in previous work, which typically updates the packet loss performance after the delivery of each packet [17,18]. Therefore, our prediction method is faster and can be used for the evaluation and optimization of wireless communications more conveniently than previous work.

4.2. Application for packetization design

Let t_d and p_f denote the expected transmission delay and failure probability, respectively. Our goal is to maximize the payload of transmission with bounded delay and failure probability, mathematically,

$$L_{\text{opt}}: \quad \underset{L}{\text{maximize}} \quad L \qquad \text{s.t.,} \quad t_d \le T_d, \; p_f \le P_f \tag{21}$$

where T_d and P_f denote the upper bounds of transmission delay and failure probability, respectively. Given the parameters of the channel fading model, we first derive the bit-level parameters (i.e., p, q, h_g, and h_b), then derive the packet loss rate PLR. Combining with the maximum retransmission number m and the channel capacity bandwidth B, the transmission delay and failure probability of packets are functions of the packet length L as

$$t_d = \frac{L}{B}\sum_{i=0}^{m}(i+1)(1-PLR(L_j))PLR(L_j)^i, \quad (22)$$

$$p_f = 1-\sum_{i=0}^{m}(1-PLR(L_j))PLR(L_j)^i. \quad (23)$$

By substituting Equations (22) and (23), L_{opt} can be optimally solved.

The packet length L is updated correspondingly at each interval T on the basis of L_{opt}. Here, the interval T is adapted to the dynamics of the fading model parameters.

Most previous work [17,18] to design packetization schemes assumes that the bit errors are IID for simplicity, and then to derive the packet loss rate, denoted as PLR_{IID}, by $PLR_{\text{IID}} = 1-(1-p_b)^L$, where p_b is the bit error probability. The packetization scheme [17,18], which uses PLR_{IID} to predict the packet loss rate and updates the packet length after the delivery of each packet, is called the IID-based scheme. In what follows, we compare our solution of L_{opt} with the IID-based scheme using simulations.

In the simulation, a node sends packets to a base station. We set $T_d = 2$ ms, $P_f = 10^{-3}$, the modulation frequency $f = 900$ MHz, and the Rayleigh fading envelope $\rho = -20$ dB. The moving speed of the node varies between 2 and 50 km/h. The protocol header (MAC+PHY header) size is 416 bits. We define the effective throughput of wireless channels as the transmission rate of payloads. From Figure 10, we can see that the effective throughput of the proposed scheme is more than 10% higher compared with the IID-based scheme. Figure 11 shows the PMFs of the transmission delay and the failure probability. As we can see, the results of both the schemes can meet the transmission-delay bound (2 ms) and the transmission-failure-probability bound (0.001). Most transmission failure probabilities from the schemes are much smaller than the bound, indicating that the main restriction on packet length is the delay bound in this simulation. Moreover, using our scheme the average transmission delay is more close to 2 ms compared with that achieved by using the IID-based scheme, and the distributions of transmission failure probability of the two schemes are very similar. In summary, by applying our prediction mechanism, the proposed packetization scheme can meet the requirements of the delay or packet-loss-sensitive services, and achieve a better balance among transmission delay, failure probability, and effective throughput to improve the service performance.

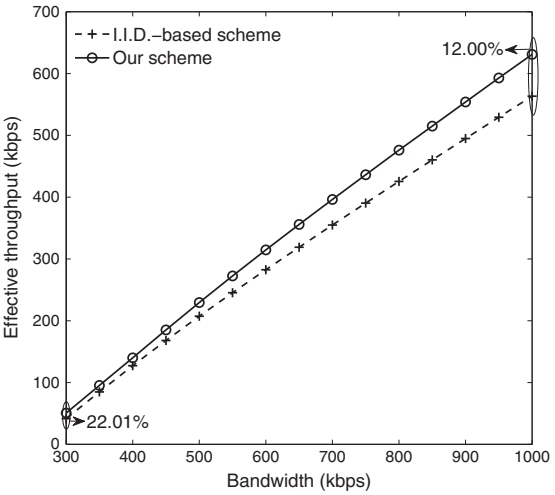


Figure 10. The throughput comparison of two packetization schemes.

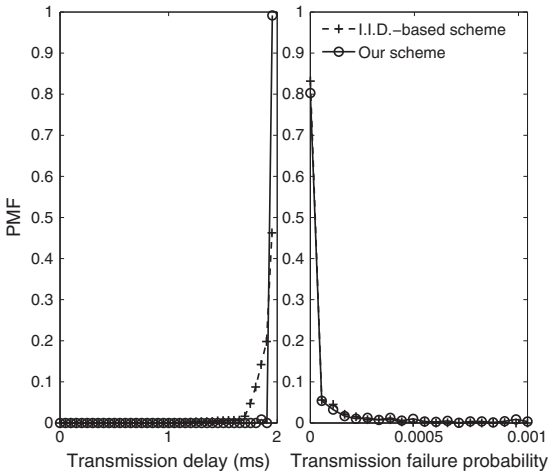


Figure 11. The PMFs of transmission delay and failure probability.

5. RELATED WORK

The phenomenon of packet loss burstiness in wireless channels has been observed in many early works, and how to measure the packet loss burstiness and address the burstiness in protocol designs are intensively pursued.

The typical mathematical metrics for describing the inter-correlations of a sequence, such as the autocorrelation and CPDF, can be used to measure the packet loss burstiness. However, the autocorrelation is inconvenient to help design protocols because it does not distinguish the loss gaps from the loss bursts. The CPDF is more informative and powerful than the autocorrelation, but cannot be directly applied for protocol designs because of its complexity. Consequently, some metrics to guide protocol

designs are developed on the basis of CPDF, such as β and μ described next.

Aguayo *et al.* [5] observe the packet loss burstiness in a 802.11b mesh network and use the Allan deviation [6] of PRR over different time scales to measure the burstiness of packet losses. The Allan deviation is large when the time interval is near the packet-loss-burst length, and it is small for other time intervals. Then, the Allan deviation is an indirectly metric through which we can obtain the packet-loss-burst length approximately. The packet-loss-burst length is very helpful for network designs, while it is not so powerful to represent the packet loss burstiness, because it is hard to distinguish the wireless channel with burst losses from the wireless channel with independent packet losses through their loss-burst lengths.

Srinivasan *et al.* [7] studied the bursty losses in various 802.15.4 wireless sensor networks and proposed a metric β that uses the concept of CPDF for measuring the burstiness. By analyzing the different values of β in different time intervals, we can find the smallest time interval with which the burstiness of packet losses is minimized. Then, we can pause for this smallest time interval to send the next packet after encountering a transmission failure. Such algorithm can break the packet loss correlations, and it is called opportune transmission, which can reduce the average transmission cost by 15%. Actually, the average packet-loss-burst length can also be used to design this algorithm instead of the smallest time interval obtained from the metric β.

All the aforementioned metrics are empirical; that is, they are more suitable to be derived from real network measurements. There are also some works for modeling packet losses using theoretical models, and such models could be extended to measure the packet loss burstiness. In the GE model [11,20], which is a popular model for wireless channels, the metric μ is proposed to reveal the burstiness [13]. The larger the μ is, the more bursty the wireless channel is. The metric μ is also an effective metric to reveal the channel capacity, and it can be used to improve the channel capacity.

Our proposed metric, *average correlation length of packet loss rate*, is also derived from the theoretical model. It, however, is more effective to represent the packet loss burstiness than previous metrics, and it provides a new means for the evaluation or optimization of communication protocols and algorithms, that is, to derive the closed-form solutions on the packet loss rate and the loss-burst/loss-gap length distributions.

6. CONCLUSION

In this paper, we have developed a new metric called *average correlation length of packet loss rate* to dimension the packet loss burstiness in a generic GE wireless channel. We have derived the closed-form analytical models for the packet loss performance, including the average packet loss rate and the loss-burst/loss-gap length distributions. These derived results can help to evaluate or optimize wireless communications conveniently. We have taken a packetization scheme as an example to show the improvements. In the future work, we will study the burstiness of the packet losses caused by on-top protocols, such as MAC protocols and FEC schemes.

APPENDIX A: PROOF OF *LEMMA 1*

In segment n, denote the length of the first pair of G-run and B-run as P_n, and the length of the left pairs of G-runs and B-runs as Q_n. Assume that there are C $(C \geq 0)$ packets, as shown in Figure 2. We then obtain Equation (A1) on the basis of the definitions of R_n and segment n.

$$\begin{aligned} R_{n-1} + C \cdot L &= R_n + Q_n + P_n, \\ P_n &> (C-1) \cdot L + R_{n-1}, \\ P_{n+1} &> R_n \end{aligned} \tag{A1}$$

Denote the sum of a G-run length and an adjacent B-run length in segment n as Z_n. Then, from Equation (A1) and the distribution of Z_n, we can derive the distributions of P_n and Q_n as in Equation (A2).

$$\begin{aligned} \Pr(P_n = z) &= \begin{cases} \Pr(Z_n = z), & (C-1) \cdot L + j + 1 \leq z \leq C \cdot L + j - r - 2 \cdot D, \\ 0, & \text{else}, \end{cases} \\ \Pr(Q_n = z) &= \begin{cases} \Pr(D \cdot Z_n = z), & z = C \cdot L + j - r - P_n, \\ 0, & \text{else} \end{cases} \end{aligned} \tag{A2}$$

where $R_n = r$, $R_{n-1} = j$, and D is the number of pairs of G-runs and B-runs in Q_n.

Thus, $\Pr(R_n = r | R_{n-1} = j)$ is derived by

$$\sum_{C=1}^{\infty} \sum_{D=0}^{(L-r-1)/2} \sum_{z=(C-1)\cdot L+j+1}^{C\cdot L+j-r-2\cdot D} \Pr(Z_n = z) \cdot \Pr(D \cdot Z_n = C \cdot L + j - r - z) \Pr(Z_{n+1} > r) \tag{A3}$$

On the basis of Equation (A3), we use mathematical induction to prove that the PMF of R_n is as Equation (1).

(I) We show that the PMF of R_1 is as Equation (1). $\Pr(R_n = r)$ can be calculated as

$$\sum_{j=0}^{L-1} \Pr(R_n = r | R_{n-1} = j) \Pr(R_{n-1} = j) \quad \text{(A4)}$$

By denoting that $R_0 = j$ (j can be any value between 0 and $L-1$), we have $\Pr(R_0 = j) = 1$. Combining with Equation (A3), we can then derive the PMF of R_1 as Equation (1).

(II) Assuming that the PMF of R_{n-1} can be derived by Equation (1) and substituting Equation (A3) into Equation (A4), we can derive the PMF of R_n as Equation (1).

Combining I and II, we conclude that the PMF of $R_n (n = 0, 1, 2 \ldots)$ is given by Equation (1).

Moreover, we obtain $\Pr(R_n = r) = \Pr(R_n = r | R_{n-1} = j)$ from Equation (A3); thus, R_{n-1} and R_n are independent.

APPENDIX B: DERIVATION OF THE LOSS-BURST/LOSS-GAP LENGTH DISTRIBUTIONS

From Equation (12), the PMF of E can be obtained as

$$\Pr(E = k) = ((1-q)^L)^{k-1}(1-(1-q)^L) \quad k \geq 1 \quad \text{(B1)}$$

Denoting by $W_j(k) = \left\{ \left\lceil \frac{M_{n+j} + X_{n+j}}{L} \right\rceil = k, A_{n+j} < L \right\}$, we have

$$\Pr(W_j(k)) = \begin{cases} \left(\sum_{a=2-L}^{0} \sum_{x=1-a}^{L-a} + \sum_{a=1}^{L-1} \sum_{x=1}^{L-a} \right) \Pr(A_{n+j} = a) \Pr(X_{n+j} = x) & k = 1, \\ \sum_{a=2-L}^{L-1} \sum_{x=(k-1)\cdot L+1-a}^{k\cdot L-a} \Pr(A_{n+j} = a) \Pr(X_{n+j} = x) & k > 1 \end{cases} \quad \text{(B2)}$$

The value of $\Pr(W_j(1))$ is a bizarre point. For ease of computation, it is necessary to ignore the difference between the expressions of $\Pr(W_j(k))$ when $k = 1$ and $k > 1$ in Equation (B2) and make an approximation that $\Pr(W_j(k))$ is calculated by

$$\sum_{a=2-L}^{L-1} \sum_{x=(k-1)\cdot L+1-a}^{k\cdot L-a} \Pr(A_{n+j} = a) \Pr(X_{n+j} = x) (k \geq 1) \quad \text{(B3)}$$

The accuracy of this approximation is investigated in Section 3.3.2. After normalization, $\frac{\Pr(W_j(k))}{\Pr(A_{n+j} < L)}$ is derived as $((1-p)^L)^{k-1}(1-(1-p)^L)$. Similarly, we can derive $\Pr\left(\left\lceil \frac{M_{n+j} + X_{n+j}}{L} \right\rceil = k \right) = ((1-p)^L)^{k-1}(1-(1-p)^L)$. Therefore, $\Pr(W_j(k)) = \Pr\left(\left\lceil \frac{M_{n+j} + X_{n+j}}{L} \right\rceil = k \right) \Pr(A_{n+j} < L)$.

In addition, $\left(\bigcup_{j=0}^{m-1} \left\lceil \frac{M_{n+j} + X_{n+j}}{L} \right\rceil = N_{e,n+j} \right)$ is the sum of m independent variables following the geometric distribution with parameter $1-(1-p)^L$. This is a negative binomial distribution, and the probability function is $f(m,k) = \frac{(k-1)!}{(k-m)!(m-1)!}(1-(1-p)^L)^m((1-p)^L)^{k-m}$ $(1 \leq m \leq k)$. From Equation (13), F is then derived by

$$\begin{aligned} &\Pr(F = k) \\ &= \sum_{m=1}^{k} f(m,k)(\Pr(A < L))^{m-1}(1 - \Pr(A < L)) \\ &= (\Pr(A < L) \cdot (1-(1-p)^L) + (1-p)^L)^{k-1} \\ &\quad \times (1 - (\Pr(A < L) \cdot (1-(1-p)^L) + (1-p)^L))\ k \geq 1 \end{aligned} \quad \text{(B4)}$$

where $\Pr(A < L) = \sum_{a=2-L}^{L-1} \Pr(A = a)$.

ACKNOWLEDGEMENTS

This work is supported in part by the National Basic Research Program of China (Nos. 2010CB328105 and 2009CB320505) and the National Natural Science Foundation of China (Nos. 60932003, 61070182, 60973144, 61173008, and 61070021).

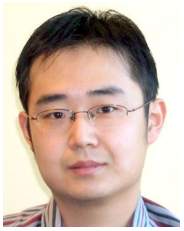
REFERENCES

1. Gandikota V, Tamma B, Murthy C. Adaptive FEC-based packet loss resilience scheme for supporting voice communication over ad hoc wireless networks. *IEEE Transactions on Mobile Computing* 2008; **7**(10): 1184–1199.
2. Kamer W, Nemethova O, Svoboda P, Rupp M. Link error analysis and modeling for video streaming cross-layer design in mobile communication networks. *ETRI Journal* 2007; **29**(5): 569–595.
3. Alizai M, Landsiedel O, Link J, Götz S, Wehrle K. Bursty traffic over bursty links, In *Proceedings of ACM*

Conference on Embedded Networked Sensor Systems (SenSys), Berkeley, CA, USA, ACM, 2009; 71–84.
4. Aggarwal R, Schniter P, Koksal C. Rate adaptation via link-layer feedback for goodput maximization over a time-varying channel. *IEEE Transactions on Wireless Communications* 2009; **8**(8): 4276–4285.
5. Aguayo D, Bicket J, Biswas S, Judd G, Morris R. Link-level measurements from an 802.11 b mesh network. *ACM SIGCOMM Computer Communication Review* 2004; **34**(4): 121–132.
6. Allan D. Time and frequency (time-domain) characterization, estimation, and prediction of precision clocks and oscillators. *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control* 1987; **34**(6): 647–654.
7. Srinivasan K, Kazandjieva M, Agarwal S, Levis P. The β-factor: measuring wireless link burstiness, In *Proceedings of ACM Conference on Embedded Networked Sensor Systems (SenSys)*, Raleigh, NC, USA, ACM, 2008; 29–42.
8. Wang H, Chang P. On verifying the first-order markovian assumption for a rayleigh fading channel miodel. *IEEE Transactions on Vehicular Technology* 1996; **45**(2): 353–357.
9. Zorzi M, Rao R, Milstein L. On the accuracy of a first-order markov model for data transmission on fading channels, In *Proceedings of IEEE International Conference on Universal Personal Communications (ICUPC)*, Tokyo, Japan, IEEE, 1995; 211–215.
10. Liu F, Fan Y, Shen X, Lin C, Zeng R. An analytical model to study the packet loss burstiness over wireless channels, In *Globecom 2010, 2010 IEEE Global Telecommunications Conference*, IEEE, 2010; 1–5.
11. Elliott E. Estimates of error rates for codes on burst-noise channels. *Bell Systems Technical Journal* 1963; **42**(9): 1977–1997.
12. [Online]. Available from: http://qos.cs.tsinghua.edu.cn/~fqliu/calculation.rar [Accessed on 2012].
13. Mushkin M, Bar-David I. Capacity and coding for the Gilbert–Elliot channels. *IEEE Transactions on Information Theory* 1989; **35**(6): 1277–1290.
14. Rubner Y, Tomasi C, Guibas L. A metric for distributions with applications to image databases, In *Proceedings of IEEE International Conference on Computer Vision (ICCV)*, Bombay, India, IEEE, 1998; 59–66.
15. Pimentel C, Blake I. Enumeration of Markov chains and burst error statistics for finite state channel models. *IEEE Transactions on Vehicular Technology* 1999; **48**(2): 415–428.
16. Lettieri P, Fragouli C, Srivastava M. Low power error control for wireless links, In *Annual International Conference on Mobile Computing and Networking (MobiCom)*, Budapest, Hungary, ACM, 1997; 139–150.
17. Setton E, Yoo T, Zhu X, Goldsmith A, Girod B. Cross-layer design of ad hoc networks for real-time video streaming. *IEEE Wireless Communications* 2005; **12**(4): 59–65.
18. van der Schaar M, Turaga D. Cross-layer packetization and retransmission strategies for delay-sensitive wireless multimedia transmission. *IEEE Transactions on Multimedia* 2007; **9**(1): 185–197.
19. Cai L, Shen X, Mark J. *Multimedia services in wireless internet: modeling and analysis*. John Wiley & Sons, 2009; 9–13.
20. Gilbert E. Capacity of a burst-noise channel. *Bell Systems Technical Journal* 1960; **39**(9): 1253–1265.

AUTHORS' BIOGRAPHIES

Fangqin Liu is in her final year of completing the dissertation research for her doctoral degree in Computer Science and Technology Department from Tsinghua University (China), where she received her BSc degree in 2006. Her research area is performance evaluation and transmission design of multimedia services in wireless networks.



Tom H. Luan received the BE degree in Xi'an Jiaotong University, China, in 2004 and the MPhil degree in electronic engineering from the Hong Kong University of Science and Technology, Kowloon, Hong Kong in 2007. He is now pursuing the PhD degree at the University of Waterloo, ON, Canada. His current research interests focus on wired and wireless multimedia streaming, QoS routing in multihop wireless networks, peer-to-peer streaming and vehicular network design.

Xuemin (Sherman) Shen received the BSc(1982) degree from Dalian Maritime University, China, and the MSc (1987) and PhD degrees (1990) from Rutgers University, NJ, USA, all in electrical engineering. He is a Professor and University Research Chair, Department of Electrical and Computer Engineering, University of Waterloo, Canada. Dr. Shen's research focuses on mobility and resource management in interconnected wireless/wired networks, UWB wireless communications networks, wireless network security, wireless body area networks, and vehicular

ad hoc and sensor networks. He is a co-author of three books and has published more than 400 papers and book chapters in wireless communications and networks, control, and filtering. Dr. Shen served as the Tutorial Chair for IEEE ICC'08, the Technical Program Committee Chair for IEEE Globecom'07, the General Co-Chair for Chinacom'07, and QShine'06, the Founding Chair for IEEE Communications Society Technical Committee on P2P Communications and Networking. He also serves as the Editor-in-Chief for Peer-to-Peer Networking and Application; Associate Editor for IEEE Transactions on Vehicular Technology; KICS/IEEE Journal of Communications and Networks, Computer Networks; ACM/Wireless Networks; and Wireless Communications and Mobile Computing (Wiley), etc. He has also served as Guest Editor for IEEE JSAC, IEEE Wireless Communications, IEEE Communications Magazine, and ACM Mobile Networks and Applications, etc. Dr. Shen received the Excellent Graduate Supervision Award in 2006, and the Outstanding Performance Award in 2004 and 2008 from the University of Waterloo, the Premier's Research Excellence Award (PREA) in 2003 from the Province of Ontario, Canada, and the Distinguished Performance Award in 2002 and 2007 from the Faculty of Engineering, University of Waterloo. Dr. Shen is a registered Professional Engineer of Ontario, Canada, an IEEE Fellow, and a Distinguished Lecturer of IEEE Communications Society.

Chuang Lin is a professor of the Department of Computer Science and Technology, Tsinghua University, Beijing, China. He is a Honorary Visiting Professor, University of Bradford, UK. He received the Ph.D. degree in Computer Science from the Tsinghua University in 1994. His current research interests include computer networks, performance evaluation, network security analysis, and Petri net theory and its applications. He has published more than 300 papers in research journals and IEEE conference proceedings in these areas and has published four books. Professor Lin is a senior member of the IEEE and the Chinese Delegate in TC6 of IFIP. He serves as the Technical Program Vice Chair, the 10th IEEE Workshop on Future Trends of Distributed Computing Systems (FTDCS 2004); the General Chair, ACM SIGCOMM Asia workshop 2005 and the 2010 IEEE International Workshop on Quality of Service (IWQoS 2010); the Associate Editor, IEEE Transactions on Vehicular Technology; the Area Editor, Journal of Computer Networks; and the Area Editor, Journal of Parallel and Distributed Computing.