

Performance Analysis of TFRC over Wireless Link with Truncated Link-Level ARQ

Hong Shen, Lin Cai, *Member, IEEE*, and Xuemin (Sherman) Shen, *Senior Member, IEEE*

Abstract—In this paper, an analytical framework is proposed for evaluating the quality of service (QoS) of TCP-Friendly Rate Control protocol (TFRC) in hybrid wireless/wired networks. For the wireless network, a link-level truncated Automatic Repeat reQuest (ARQ) scheme is deployed to reduce the packet losses visible to the transport layer protocol. Two Discrete Time Markov Chains (DTMC) are introduced to analyze the network performance and the QoS of the TFRC protocol, in terms of wireless link utilization, flow throughput, packet loss rate, and probability of end-to-end delay exceeding a prescribed threshold. Extensive simulations are performed to verify the accuracy of the analytical results. It is concluded that the analytical results are useful for optimizing the system parameters, such as the interface buffer size and the maximum number of retransmissions of the truncated ARQ scheme. With the optimal parameters, wireless link utilization can be maximized, and the QoS requirements of multimedia applications can be statistically guaranteed.

Index Terms—Transport layer protocol, quality of service, multimedia, wireless networks, performance analysis.

I. INTRODUCTION

WIRELESS cellular networks and the IP-based Internet will converge to a ubiquitous information transport infrastructure, allowing users to access the hybrid networks for multimedia services anywhere, anytime. One of the key factors to the success of the Internet is the dominant transport layer protocol, Transmission Control Protocol (TCP). TCP is a reliable connection-oriented protocol with flow and congestion control mechanisms. Network congestion is signified by packet losses. To maintain network stability and integrity, a TCP sender halves its sending rate when a network congestion signal is captured.

Unfortunately, for many emerging multimedia applications, TCP is undesirable since halving the sending rate for a single packet loss will noticeably degrade the user-perceived quality of service (QoS). Thus, User Datagram Protocol (UDP), an unreliable and connectionless transport layer protocol, has been used by some emerging multimedia applications that have more stringent timely delivery requirements rather than object integrity. However, the UDP sender does not respond to network congestion signals. When unresponsive UDP flows occupy a large portion of the network traffic mix, not only the throughputs of responsive TCP flows will be severely

degraded, but also the network will be driven to congestion collapse [1]. Therefore, it is important to have appropriate traffic regulations which are suitable for multimedia applications, and compatible with TCP [2].

Recently, several TCP-friendly protocols [3]–[8] have been proposed for Internet-based multimedia applications, where *TCP-friendliness* is defined as the average throughput of non-TCP-transported flows over a large time scale does not exceed that of any conformant TCP-transported ones under the same circumstance [2]. A popular TCP-friendly protocol supporting multimedia streams is the equation-based TCP-Friendly Rate Control (TFRC) protocol [4], which uses a stochastic model to estimate the throughput of a TCP flow under the same circumstance and sets the maximum allowable sending rate accordingly. According to the stochastic model, TCP throughput is a function of round-trip time (RTT), retransmission timeout value (T_O), and loss event¹ rate p [9]:

$$\lambda = \frac{1}{RTT\sqrt{2p/3} + T_O(3\sqrt{3p/8})p(1 + 32p^2)} \quad (1)$$

According to the protocol specification [10], p is measured and calculated by the TFRC receiver, and RTT and T_O are estimated and calculated by the TFRC sender. The TFRC receiver periodically sends feedback messages to the TFRC sender at least once per RTT or whenever a new loss event is detected. Initially, the TFRC sender sets its rate to one packet² per second and doubles the rate every RTT until loss occurs. Thereafter, the sending rate is set according to (1).

Since TFRC is less aggressive in probing for available bandwidth, and more moderate in responding to transient network congestion, TFRC can achieve TCP-friendliness, and its sending rate is much smoother than that of TCP. In addition, the TFRC sender has no obligation to retransmit lost or corrupt packets. The unreliable service is desirable for many time-sensitive applications.

On the other hand, wireless channels are characterized by limited bandwidth, high bit error rate, time-varying and location dependent channel condition, *etc.* Similar to TCP, the TFRC sender treats any packet losses as congestion signals and sets its sending rate inversely proportional to the square root of loss event rate, which leads to under-utilization of the wireless link.

¹One or more packet losses during one RTT are counted as one loss event.

²Link layer fragmentation is not considered here, since time-sensitive multimedia traffic usually has small packet size. Modern transport layer protocols can negotiate for the maximum segment size on their connection establishment to avoid IP fragmentation. In the sequel, the term *packet* is used generically to represent the link *frame*, network *packet*, and transport *segment*.

Manuscript received June 3, 2004; revised January 18, 2005; accepted February 19, 2005. The associate editor coordinating the review of this paper and approving it for publication was Z. Zhang.

H. Shen and X. Shen are with the University of Waterloo, Dept. of Electrical and Computer Engineering, Waterloo, ON N2L 3G1, Canada (e-mail: h2shen@bbcr.uwaterloo.ca; xshen@bbcr.uwaterloo.ca).

L. Cai is with the Dept. of Electrical and Computer Engineering, University of Victoria, Victoria, BC V8W 3P6, Canada (e-mail: cai@uvic.ca).

Digital Object Identifier 10.1109/TWC.2006.04351

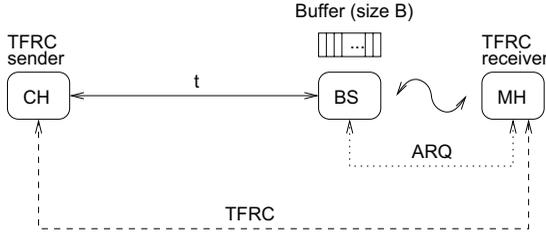


Fig. 1. TFRC connection in hybrid wireless/wired networks.

To improve the performance of congestion controlled protocols over lossy wireless links, three categories of schemes have been proposed: link layer enhancement schemes [11]–[16], transport layer strategies with the assistance from the interface node [17]–[20], and end-to-end solutions [21]. Among them, the link-level Automatic Repeat reQuest (ARQ) scheme can efficiently recover transmission errors and is comparatively simple to deploy. In addition, there is no competitive retransmissions between the TFRC and the link layer, because the TFRC sender is not obligated to retransmit lost packets. Therefore, the link-level ARQ scheme is a good candidate to improve the TFRC performance over wireless links and is anticipated to be deployed in wireless networks [26].

In the literature, different models have been proposed to analyze TCP performance over wireless links with the link-level ARQ scheme [11]–[14]. In [11] and [13], it is assumed that the link layer can retransmit corrupted packets with unlimited times. In [12] and [14], the buffer at the interface node is assumed large enough to avoid buffer overflow. In practice, for time-sensitive multimedia applications, a truncated ARQ scheme (with finite retransmission times) and limited interface buffer size are used to reduce queuing delay, since delay and delay jitter are very important QoS indexes for time-sensitive applications. To the best of our knowledge, there is no systematic analysis on delay performance of TFRC over wireless links.

In this paper, we develop an analytical framework for evaluating the QoS of TFRC in hybrid wireless/wired networks. We consider the wireless network with the link-level truncated ARQ scheme and limited interface buffer size. We develop one Discrete Time Markov Chain (DTMC) to investigate the wireless bandwidth utilization and packet loss rate of TFRC flows over wireless links, and another DTMC to study the delay outage probability, the probability of packet delay exceeding a prescribed threshold. Furthermore, by using the Network Simulator (*ns-2*) [25], extensive simulations have been conducted to verify the analytical results. It is concluded that the analytical results are useful for optimizing the network parameters, such as the interface buffer size and the maximum number of retransmissions of the truncated ARQ scheme. With the optimal parameters, wireless link utilization can be maximized, and the QoS requirements of multimedia applications can be statistically guaranteed.

The remainder of the paper is organized as follows. Section II introduces the system model. In Section III, the performance of TFRC over wireless links is studied, and the

TABLE I
NOTATIONS

1. B : BS buffer size
2. C : wireless link capacity (packets per second)
3. F : fading margin
4. f_c : carrier frequency
5. f_d : Doppler frequency
6. l_b : average length of burst errors
7. M : maximum number of transmissions allowed for each packet in the truncated ARQ scheme
8. p : loss event rate
9. p_e : average packet error rate in transmission
10. p_l : packet loss rate
11. q : queue length
12. t : deterministic end-to-end delay
13. T : packet transmission time over the wireless link
14. T_O : timeout value
15. t_q : queuing delay
16. U : link utilization
17. v : velocity of the MH
18. λ : flow throughput (packet per second)

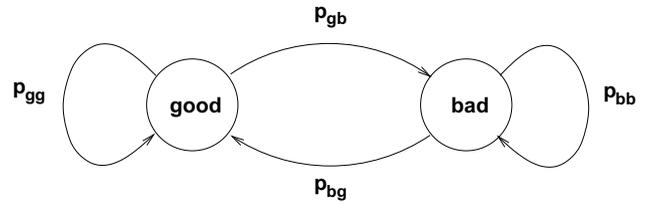


Fig. 2. Two-state wireless channel model.

wireless link utilization, packet loss rate, and delay outage rate are analytically obtained. Section IV presents the simulation results, followed by concluding remarks in Section V.

II. SYSTEM MODEL

As shown in Fig. 1, a TFRC connection is established between a correspondent host (CH) and a mobile host (MH). The MH is connected with a base station (BS) through a last-hop wireless link. We assume that all packets of the target TFRC flow have the same packet size and travel through the same route. The BS, the interface node, allocates a dedicated wireless link with a capacity of C packets per second and a buffer of size B for the TFRC flow. End-to-end packet delay contains a deterministic part and a random part: the former is denoted by t ; the latter is referred to as delay jitter. For easy reference, the notations used throughout the paper are listed in Table I.

A. Truncated ARQ Scheme

To improve the TFRC performance over lossy wireless links, a truncated ARQ scheme is deployed in the link layer. Each packet sent by the TFRC sender (at the CH) is first queued in the BS buffer. Then, the BS forwards the packet at the front of the queue to the MH, where the TFRC receiver resides. In the link layer, if a packet is received by the MH successfully, the MH sends back a local acknowledgment to the BS immediately. After that, the BS removes the packet from the buffer. If the MH fails to receive the packet, the BS retransmits the corrupt packet. M is the maximum number of transmissions allowed for each packet. Packets are removed from the buffer after being received by the MH successfully, or after M failed attempts. Assume that there is no packet

loss in the wired domain, and the packet losses visible to the TFRC receiver are those due to BS buffer overflow and those discarded after M failed attempts.

B. Wireless Channel Model

In a typical wireless channel with multipath propagation and non-line-of-sight (NLOS) frequency-nonselctive (flat) fading, the received signal envelope can be modeled as a Rayleigh fading channel [22].

Finite-state Markov models have been proposed in the literature to characterize Rayleigh fading channels in the packet level, which are widely acknowledged as a reasonably accurate and mathematically tractable approach [23]. In [24], the behavior of packet errors in fading channels is investigated. It is concluded that a two-state Markov model is a good approximation of wireless channels for a wide range of parameters. With the two-state model, time is discretized into slots, and the duration of each slot is the transmission time of a packet; the channel alternates between a *good* state (in which all packets are transmitted error free) and a *bad* state (in which all transmissions are failed). The residence time of each state follows a geometric distribution, as shown in Fig. 2. Furthermore, the state transition probabilities are insensitive to the coding and modulation schemes and only depend on the normalized Doppler frequency. This observation simplifies the performance analysis of the upper-layer protocols.

Let v denote the velocity of the MH, and f_c represent the carrier frequency. The Doppler frequency f_d equals $f_c v / v_c$, where v_c is the light speed. Let F be the fading margin. Given the physical layer coding and modulation schemes, when the received Signal to Noise Ratio (SNR) is below a threshold, $E[SNR]/F$, channel is in the *bad* (b) state; otherwise, it is in the *good* (g) state. $E[SNR]$ is the mean of the received SNR. As derived in [24], the average packet error rate is given by

$$p_e = 1 - e^{-1/F}. \quad (2)$$

The state transition probabilities are

$$p_{bb} = 1 - \frac{Q(\theta, \rho\theta) - Q(\rho\theta, \theta)}{e^{1/F} - 1}, \quad (3)$$

$$p_{gg} = \frac{1 - p_e(2 - p_{bb})}{1 - p_e}, \quad (4)$$

where $\theta = \sqrt{\frac{2/F}{1-\rho^2}}$ and $\rho = J_0(2\pi f_d T)$ [24]. ρ is the Gaussian correlation coefficient of two samples of the complex amplitude of a fading channel with Doppler frequency f_d , sampled T seconds away. T is the packet transmission time over the wireless link, *i.e.* $T = 1/C$. $J_0(\cdot)$ is the Bessel function of the first kind and zero order. $Q(\cdot, \cdot)$ is the Marcum Q function.

The correlation coefficient ρ depends on $f_d T$. If $f_d T$ is small, the channel fading process is highly auto-correlated, and vice versa. Since f_d is proportional to v , the slower the MH moves, the burstier the transmission errors are.

III. PERFORMANCE ANALYSIS

The performance indexes considered are the wireless link utilization, packet loss rate, and delay outage rate. Wireless link utilization is the number of packets transmitted successfully per T . Packet loss rate is the ratio of lost packets over the packets being sent by the TFRC sender. Delay outage rate is the ratio of received packets with delay jitter exceeding a prescribed threshold.

In the analysis, without loss of generality, we make the following assumptions:

- 1) The bottleneck of the cross-domain connection is the wireless link, and the packet losses due to congestion between the CH and the BS are negligible. This tends to be acceptable, since in general, wired links have larger bandwidth and lower bit error rates than wireless ones.
- 2) The delay jitter of a packet is mainly due to queuing delay in the BS buffer, since the wireless link is presumed to be the bottleneck link, and the throughput of the wireless link with ARQ is highly dynamic. Therefore, the delay outage rate equals the probability that a packet's queuing delay at the BS exceeds the maximum tolerable delay jitter.
- 3) A perfect and instantaneous wireless uplink channel is assumed so that, at the end of each time slot, the BS knows whether the transmission was successful or not (*e.g.* in a TDMA system, a logic link corresponds to a time slot of each frame, and the time duration between a link's two consecutive slots is long enough to get the link-level acknowledgment [26]). Packets that were not successfully received are immediately retransmitted in the next time slot to reduce the jitter.

A. Link Utilization and Packet Loss Rate

Since TFRC is a rate-based protocol, the TFRC sender spaces outgoing packets evenly by setting inter-packet interval to $1/\lambda$ second to produce a smooth flow with rate λ packets per second. The arrival process of TFRC packets at the BS buffer can be approximated by a stationary Bernoulli process, and the probability of a packet arrival in a given time slot is λT . Simulation results show that the Bernoulli model is acceptable.

When the buffer is non-empty, a packet is sent at the beginning of the slot. It is assumed that the channel condition remains at its current state during the transmission of a packet. If the transmission is successful, the packet leaves the buffer at the end of the slot.

At the beginning of each slot, let n denote the number of packets in the buffer, s represent the channel state, and m denote how many times the currently served packets being transmitted. A three-dimensional DTMC is defined as (n, s, m) , where $0 \leq n \leq B$, $s \in \{b, g\}$, and $1 \leq m \leq M$. The DTMC is called the *queue-associated* DTMC to differentiate it from the other DTMC developed in Section III-B. Obviously, states with both $n = 0$ and $m > 1$ are unreachable. The state transition diagram is shown in Fig. 3, and the state transition probabilities are listed in Table II.

Given λT and the state transition probability matrix \mathbf{P} , the steady state probability $\pi(n, s, m)$ can be derived according

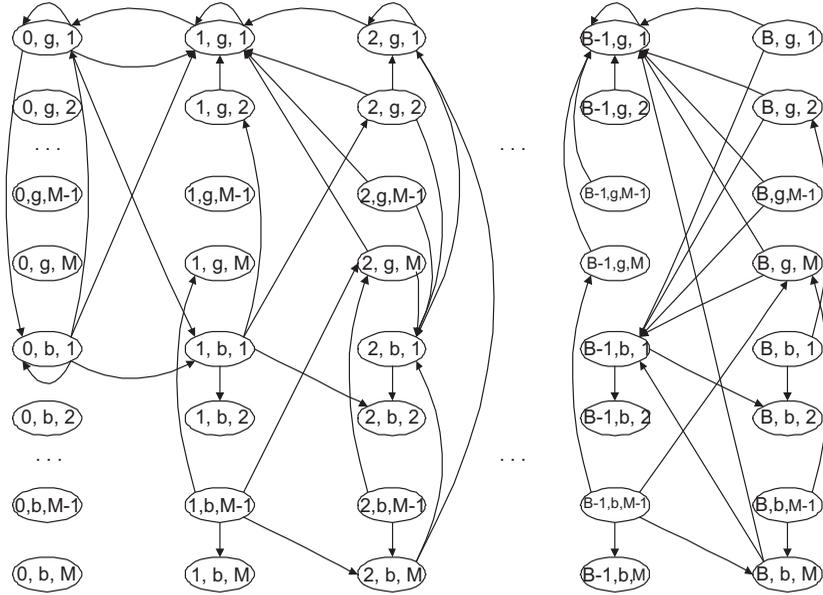
Fig. 3. State transition diagram of the *queue-associated* DTMC.

TABLE II

STATES TRANSITION PROBABILITIES OF THE QUEUE-ASSOCIATED DTMC

current state	next state	transition probability
$(0, g, 1)$	$(0, g, 1)$	$(1 - \lambda T)p_{gg}$
	$(0, b, 1)$	$(1 - \lambda T)p_{gb}$
	$(1, g, 1)$	$\lambda T p_{gg}$
	$(1, b, 1)$	$\lambda T p_{gb}$
(n, g, m) $1 \leq n \leq B-1$ $1 \leq m \leq M$	$(n-1, g, 1)$	$(1 - \lambda T)p_{gg}$
	$(n-1, b, 1)$	$(1 - \lambda T)p_{gb}$
	$(n, g, 1)$	$\lambda T p_{gg}$
	$(n, b, 1)$	$\lambda T p_{gb}$
(B, g, m) $1 \leq m \leq M$	$(B-1, g, 1)$	p_{gg}
	$(B-1, b, 1)$	p_{gb}
$(0, b, 1)$	$(0, g, 1)$	$(1 - \lambda T)p_{bg}$
	$(0, b, 1)$	$(1 - \lambda T)p_{bb}$
	$(1, g, 1)$	$\lambda T p_{bg}$
	$(1, b, 1)$	$\lambda T p_{bb}$
(n, b, m) $1 \leq n \leq B-1$ $1 \leq m \leq M-1$	$(n, g, m+1)$	$(1 - \lambda T)p_{bg}$
	$(n, b, m+1)$	$(1 - \lambda T)p_{bb}$
	$(n+1, g, m+1)$	$\lambda T p_{bg}$
	$(n+1, b, m+1)$	$\lambda T p_{bb}$
(n, b, M) $1 \leq n \leq B-1$	$(n-1, g, 1)$	$(1 - \lambda T)p_{bg}$
	$(n-1, b, 1)$	$(1 - \lambda T)p_{bb}$
	$(n, g, 1)$	$\lambda T p_{bg}$
	$(n, b, 1)$	$\lambda T p_{bb}$
(B, b, m) $1 \leq m \leq M-1$	$(B, g, m+1)$	p_{bg}
	$(B, b, m+1)$	p_{bb}
(B, b, M)	$(B-1, g, 1)$	p_{bg}
	$(B-1, b, 1)$	p_{bb}

to the following balance equations:

$$\begin{cases} \mathbf{\Pi}^T \mathbf{P} & = \mathbf{\Pi}^T \\ \sum_{n=0}^B \sum_{s \in \{b, g\}} \sum_{m=1}^M \pi(n, s, m) & = 1, \end{cases} \quad (5)$$

where $\mathbf{\Pi} = [\pi(n, s, m)]$.

By Little's law, the average round-trip time of the TFRC

flow can be derived as follows:

$$RTT = 2t + \frac{1}{\lambda} \sum_{n=1}^B \sum_{s \in \{b, g\}} \sum_{m=1}^M n \pi(n, s, m). \quad (6)$$

Packet losses are due to BS buffer overflow and M failed transmissions. Therefore, packet loss rate, p_l , can be calculated by

$$p_l = \sum_{s \in \{b, g\}} \sum_{m=1}^M \pi(B, s, m) + \sum_{n=0}^{B-1} \pi(n, b, M). \quad (7)$$

The link utilization, U , can be obtained as follows:

$$U = \lambda(1 - p_l)/C. \quad (8)$$

Burst error length measures the number of consecutive packet losses due to transmission errors. Since i packets will be discarded consecutively due to transmission errors if the channel remains at the bad state for iM slots, the average length of burst errors in transmission is given by

$$l_b = \sum_{i=1}^{+\infty} i(p_{bb}^M)^{i-1}(1 - p_{bb}^M) = 1/(1 - p_{bb}^M). \quad (9)$$

One or more packet losses during one RTT are counted as one loss event by the TFRC receiver. Assume that there is only one burst of losses during one RTT and the length of burst losses is approximated by l_b . The loss event rate can be approximated by

$$p \approx p_l/l_b. \quad (10)$$

We can obtain p as a non-decreasing function of λ , according to the above analysis; on the other hand, the TFRC sender adjusts λ as a non-increasing function of p , according to (1), as shown in Fig. 4. We can obtain the coordinates of

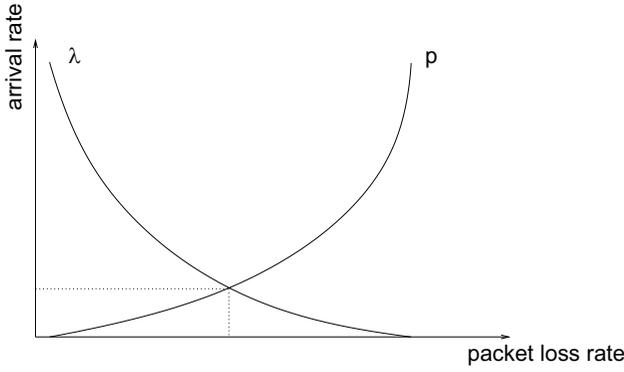


Fig. 4. TFRC sending rate vs. packet loss rate.

the intersection point of the two curves, which correspond to the p and λ in the steady state.

The procedure to obtain p_l and λ in the steady state is as follows. We set λ to $C/2$ initially, and calculate RTT , p_l , and p according to (6), (7), and (10), respectively. Then, we substitute the calculated p and RTT into (1) to obtain an updated λ . By repeating this procedure, λ converges to a constant, which is the steady state sending rate. Thereafter, p_l can be obtained by substituting λ into (5) and (7), and the link utilization can be obtained by substituting p_l into (8).

TFRC performance over wireless link without the ARQ scheme can be obtained by setting M to 1.

B. Delay Performance Analysis

To study the delay performance, we develop another DTMC in this section. When a packet arrives at the BS buffer, the queue length (including the target packet) is denoted by q . We use a triplet of indexes (n_d, s_d, m_d) to represent the current state of a packet, where $1 \leq n_d \leq q$, $s_d \in \{b, g\}$, and $1 \leq m_d \leq M$. $n_d - 1$ is the number of packets currently being queued in the BS buffer before the target packet. s_d is the current wireless channel state. m_d denotes how many times the currently served packet has been transmitted. The state evolves every slot, and the sequence of the states is a *packet-associated* DTMC. For instance, if $n_d = q$ and the packet currently being served is transmitted successfully or discarded by the BS, in the next slot, $n_d = q - 1$; otherwise, $n_d = q$ in the next slot.

Figure 5 shows the state transition trajectories of the *packet-associated* DTMC. At the end of each trajectory, there is an absorbing state, either (*success*) or (*failure*), corresponding to the target packet being successfully received by the MH or being discarded after M failed attempts, respectively.

The one-step state transition probabilities of the *packet-associated* DTMC are listed in Table III, and the k -step transition matrix of the Markov chain \mathbf{P}_k equals \mathbf{P}_1^k . From \mathbf{P}_k , the probability that the packet is transmitted successfully at the k -th slot after its arrival, denoted by $p_{k(q,s,m),(success)}$, can be obtained, where (q, s, m) is the initial state associated with the target packet. The minimal queuing delay at the BS is $q - 1$, if the wireless channel is in the *good* state for q consecutive slots. The maximum queuing delay is $qM - 1$, if the wireless channel is in the *bad* state for $(qM - 1)$

TABLE III

STATES TRANSITION PROBABILITIES OF THE PACKET-ASSOCIATED DTMC

current state	next state	transition probability
(q, g, m)	$(q - 1, g, 1)$	p_{gg}
$1 < q \leq B, 1 \leq m \leq M$	$(q - 1, b, 1)$	p_{qb}
(q, b, m)	$(q, g, m + 1)$	p_{bg}
$1 < q \leq B, 1 \leq m < M$	$(q, b, m + 1)$	p_{bb}
(q, b, M)	$(q - 1, g, 1)$	p_{bg}
$1 < q \leq B$	$(q - 1, b, 1)$	p_{bb}
$(1, g, m)$	success	1
$1 \leq m \leq M$		
$(1, b, m)$	$(1, g, m + 1)$	p_{bg}
$1 \leq m < M$	$(1, b, m + 1)$	p_{bb}
$(1, b, M)$	failure	1

consecutive slots and changes to the *good* state at the qM -th slot. Let D be the maximum delay jitter that an application can tolerate. The conditional probability that the queuing delay of a successfully received packet is larger than D/T slots is given by:

$$\Pr\{t_q > D/T | (q, s, m)\} = \frac{\sum_{k=D/T+2}^{q \cdot M} p_{k(q,s,m),(success)}}{\sum_{k=q}^{q \cdot M} p_{k(q,s,m),(success)}}, \quad (11)$$

where t_q represents the queuing delay at the BS buffer.

Obviously, $(1, g, 2), \dots, (1, g, M)$ and $(1, b, 2), \dots, (1, b, M)$ are invalid initial states, and the probability of being in one of these states initially is zero. The probability of the initial state associated with the packet can be calculated as the steady state probability of the *queue-associated* DTMC normalized by the sum of all possible initial states, *i.e.* $\Pr\{q, s, m\} = \pi(q - 1, s, m) / \Pi_v$, where $\Pi_v = \sum_{s \in \{b, g\}} \pi(0, s, 1) + \sum_{q=2}^B \sum_{s \in \{b, g\}} \sum_{m=1}^M \pi(q - 1, s, m)$, and $\pi(q - 1, s, m)$ has been derived in Section III-A.

The probability of queuing delay exceeding D/T is given by

$$\begin{aligned} & \Pr\{t_q > D/T\} \\ &= \frac{1}{\Pi_v} \left[\sum_{s \in \{b, g\}} \Pr\{t_q > D/T | (0, s, 1)\} \cdot \pi(0, s, 1) \right. \\ & \left. + \sum_{q=2}^B \sum_{s \in \{b, g\}} \sum_{m=1}^M \Pr\{t_q > D/T | (q - 1, s, m)\} \cdot \pi(q - 1, s, m) \right]. \end{aligned} \quad (12)$$

Although with a larger M , the packet loss rate due to transmission errors can be reduced and the wireless link utilization can be increased, more delay and delay jitter are introduced. Similarly, with a larger BS buffer size, B , the packet loss rate due to buffer overflow can be reduced; however, it also results in more delay and delay jitter. Therefore, M and B should be chosen appropriated by making the tradeoff between system utilization and QoS requirements. For time-sensitive applications, the delay outage probability given by (12) should be bounded. Thus, the system parameters, M and B , can be optimized according to the wireless channel profile, to maximize the link utilization and TFRC flow throughput under the QoS constraint. This will be discussed in detail in the following section.

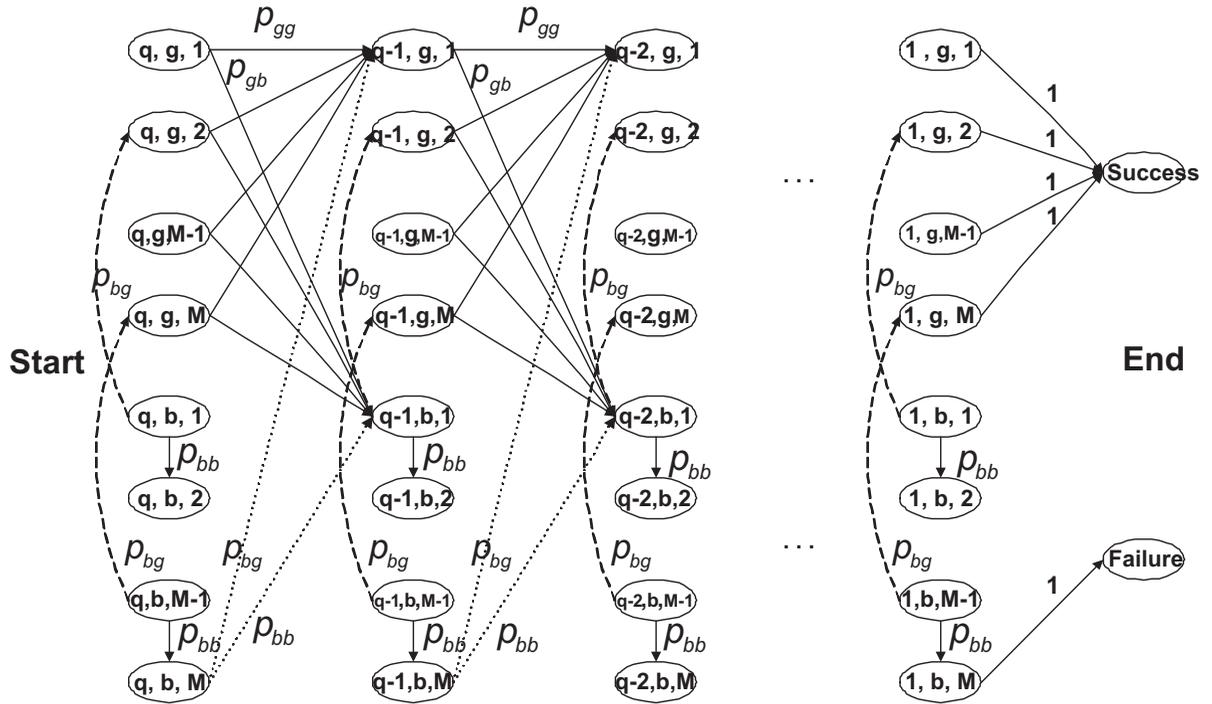


Fig. 5. State transaction diagram of the *packet-associated* DTMC.

IV. SIMULATION RESULTS

To verify the analysis in Sections III, extensive simulations are performed by using the Network Simulator (*ns-2*) [25]. The simulation topology is shown in Fig. 1.

We assume that the underline physical layer power control and error control schemes can be adaptive to transmission distance and channel conditions, such that the fading margin F is maintained at $10dB$. In other words, the packet transmission error rate p_e remains a constant ($p_e = 1 - e^{-1/F} = 0.0952$). Although the packet error rate can be guaranteed by the physical layer protocol, with different Doppler frequency and channel coherence time, the channel state transition probabilities are different, which will affect the performance of the upper layer protocols. Since the Doppler frequency is determined by the MH velocity, the analytical and simulation results are compared with respect to the MH velocity. The numerical results can be used as a guidance to select M and B appropriately to efficiently utilize the wireless links and statistically guarantee the packet loss rate and delay outage rate.

In the simulation, the following parameters are used unless otherwise explicitly stated. The downlink and uplink bandwidth allocated to the target TFRC flow is 384 Kbps and 96 Kbps, respectively. The TFRC flow has packet size 240 bytes, and the downlink capacity is equal to 200 packets per second. The bandwidth of the wired link between the CH and the BS is 100 Mbps. The carrier frequency of the wireless transmission is 900 MHz. The deterministic end-to-end delay t is 20 ms. The transmission time of one packet over the wireless

link (the duration of a time slot) is $T = 1/C = 5$ ms. The maximum tolerable delay jitter of the multimedia application is 50 ms. Since the transmission time of 10 packets equals 50 ms, the buffer size B is set to 10 packets.

Since the TFRC flow is allocated a dedicated wireless link, a simple drop tail queue management scheme is deployed at the BS. The TFRC sender is saturated so that it always has data to send. Each simulation lasts for 500 seconds. We repeat the simulation with different random seeds, and the simulation results are the average ones with a high degree of confidence.

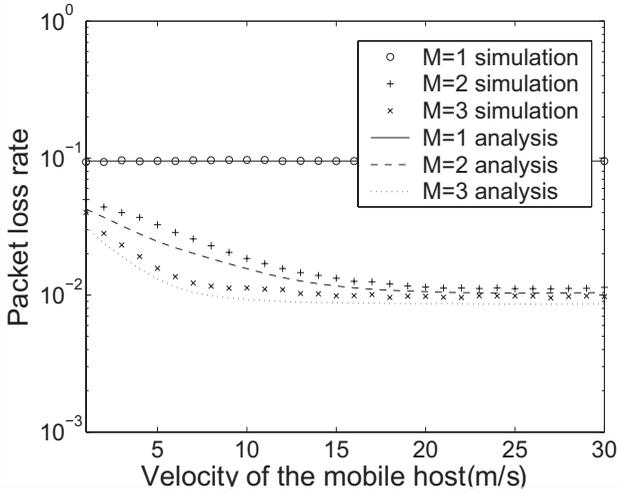
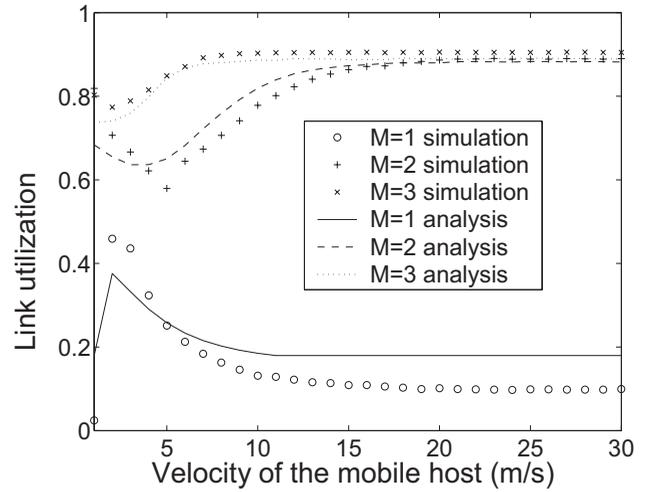
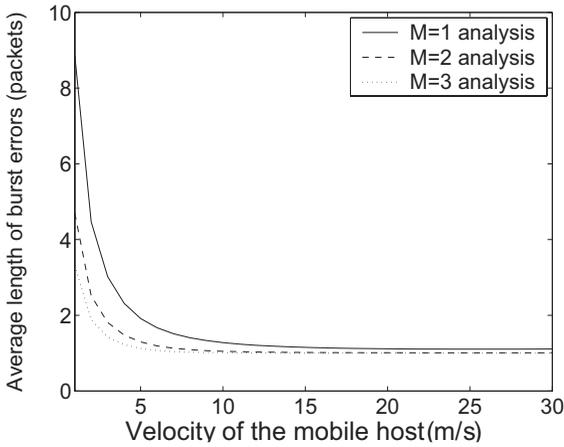
A. Packet Loss Rate

Figure 6 compares the analytical and simulation results of packet loss rates for $M = 1, 2, 3$, respectively.

For $M = 1$, since TFRC traffic is quite smooth and the sending rate is much less than the wireless link capacity due to the high transmission error rate, the BS queue does not build up and there is no buffer overflow. Thus, the packet loss rate is always equal to the transmission error rate p_e .

When the truncated ARQ is deployed, packet loss rate decreases as v increases. This is because, with higher velocity, the Doppler frequency is higher, and the channel coherence time is smaller. Given a *constant* transmission error rate, with higher velocity, packet loss rate is lower, since the length of burst errors is shorter and the truncated ARQ scheme can recover those burst errors with length no larger than M .

The average length of burst errors (l_b), determined by the wireless channel state transition probability and M , is shown in Fig. 7. It can be seen that, with higher velocity, l_b is


 Fig. 6. Packet loss rate, $t = 20ms$.

 Fig. 8. Link utilization, $t = 20ms$.

 Fig. 7. Average length of burst errors, $t = 20ms$.

smaller since the fading process becomes more uncorrelated. The packets will be discarded due to transmission errors only if l_b is larger than M . Therefore, when the velocity is low, packet loss rate decreases as M increases; when the velocity is quite high and l_b converges to one, there is no significant difference with $M = 2$ or $M = 3$. As shown in Fig. 6, when the velocity is larger than 20 m/s, the packet loss rates with $M = 2$ and $M = 3$ are almost the same.

B. Link Utilization

Figure 8 shows the analytical and simulation results of wireless link utilization, which can also be viewed as normalized flow throughput.

When $M = 1$, the throughput decreases as v increases. In other words, without the ARQ scheme, TFRC performs better in a burst error environment than in a random error environment. This is because, according to (1), TFRC sending rate is inversely proportional to the square root of loss event rate p , which is approximately p_l/l_b . Thus, at low velocity, larger l_b leads to smaller loss event rate and higher link

utilization. However, if l_b becomes so large that there is no acknowledgment received at the sender for a while, the sender timer will expire. According to the TFRC protocol specification, when timeout event occurs, the sending rate is cut by half, which results in an even lower flow throughput. As shown in Fig. 8, the link utilization is increased with v when $v < 2$ m/s due to less timeout events; after that, the utilization is decreased with the increased velocity due to smaller l_b .

When $M = 2$ and 3, the figure shows that the bandwidth utilization is high for low velocity, decreases for medium velocity, and then increases for high velocity. By comparing Fig. 6 and Fig. 7, for low velocity, although packet loss rate is high, due to large l_b , the loss event rate is small. Thus, the TFRC throughput and link utilization with low velocity is higher than that with the medium one. When the velocity becomes even higher, l_b decreases slower than packet loss rate does; therefore, the TFRC flow throughput and the link utilization become higher.

In Fig. 8, when $M = 2$ and $v > 20$ m/s, or $M = 3$ and $v > 10$ m/s, the link utilization can approach the upper bound, $1 - p_e$, *i.e.* the wireless link can be fully utilized by the TFRC flow.

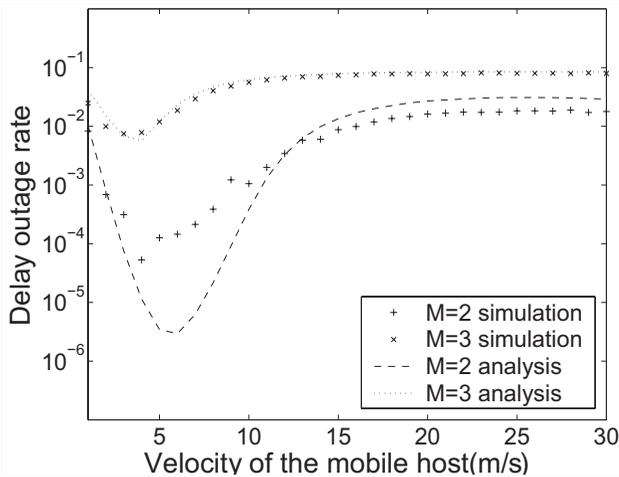
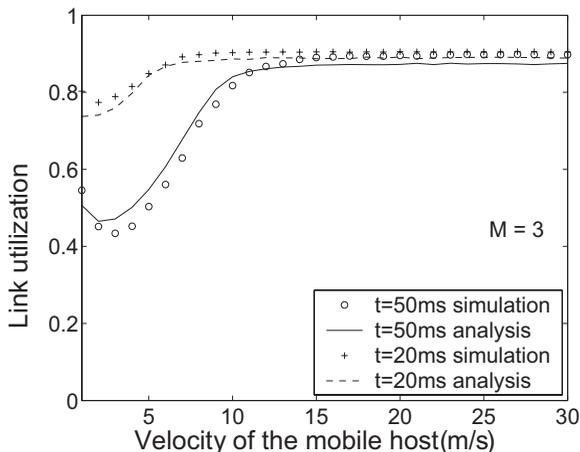
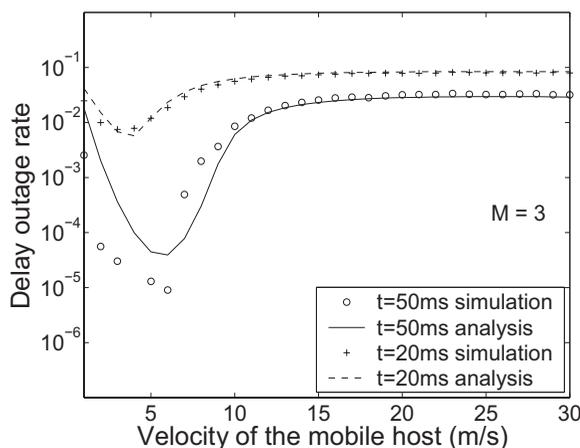
C. Delay Outage Rate

When $M = 2$ and 3, the analytical and simulation results of delay outage rates (the ratio of packets with delay jitter t_q larger than 50 ms), are shown in Fig. 9. For the TFRC flow, since higher sending rate (throughput) leads to higher buffer occupancy and longer queuing delay, the delay outage rate curve follows the utilization curve, as shown in both Fig. 8 and Fig. 9.

With $M = 1$, the queuing delay is negligible, since without the ARQ scheme, TFRC throughput is much lower than the link capacity, as shown in Fig. 8. Therefore, the delay outage rate is zero when $M = 1$.

D. Effect of Deterministic End-to-End Delay

To evaluate the effect of deterministic end-to-end delay, simulation results with $t = 20$ ms and $t = 50$ ms are

Fig. 9. Delay outage rate, $t = 20ms$.Fig. 10. Link utilization, $t = 20ms$ vs. $t = 50ms$.Fig. 11. Delay outage rate, $t = 20ms$ vs. $t = 50ms$.

compared and shown in Figs. 10 and 11, where the dashed lines correspond to $t = 20$ ms, and the solid lines correspond to $t = 50$ ms.

In Fig. 10, for $M = 3$, with low velocity, most of the losses are those packets being discarded by the BS when three attempts failed. Thus, TFRC flows with different propagation delays have approximately the same packet loss rate, which is determined by the wireless link statistics only. With the same loss rate, TFRC sending rate is inversely proportional to RTT , and the TFRC flow has lower throughput and link utilization with larger t .

With high velocity, most packets can be successfully transmitted in three attempts and packet losses are mainly due to buffer overflow. Therefore, the TFRC flow can fully utilize the wireless link and approach the maximum throughput, even with a large t . In other words, given appropriate buffer size, if there is no packet loss due to transmission errors, the sending rate of TFRC flows can always approach the link capacity no matter how large the t is.

As shown in Fig. 11, the delay outage curve follows the link utilization curve, and the analytical results match well with the simulation ones.

In summary, based on the wireless channel profile (transmission error rate, state transition probability, *etc.*) and the QoS requirements (the maximum tolerable delay jitter, delay outage rate, packet loss rate, *etc.*), the BS can determine suitable B and M to maximize the link utilization and TFRC flow throughput under the QoS constraint.

For example, a multimedia application requires that the delay outage rate (delay jitter exceeding 50 ms) should be less than 1% and the packet loss rate should be less than 2%. When $t = 20$ ms and v is around 15 m/s, the BS can set $M = 2$ and $B = 10$ packets since the delay outage rate will exceed the target one if $M = 3$. In this setting, the link utilization is around 0.84, and the efficiency does not change much if setting $M = 3$. On the other hand, if the velocity is around 5 m/s, the BS can set $M = 3$ and $B = 10$ packets, so the delay outage rate and packet loss rate can be guaranteed, and the link utilization is 20% higher than that with $M = 2$.

V. CONCLUSIONS

We have developed a framework for evaluating the QoS of TFRC in hybrid wireless/wired networks. In the wireless link, a truncated ARQ scheme has been deployed to reduce the packet losses visible to the higher layer protocols. Given the truncated ARQ scheme and wireless channel profile, we have derived the link utilization, TFRC flow throughput, packet loss rate, and delay outage probability. Extensive simulations have been performed to validate the analysis. The analytical results can be used to optimize the system parameters such as the BS buffer size and the maximum number of retransmission times for the truncated ARQ scheme, in order to maximize the TFRC flow throughput and wireless link utilization, and to satisfy the QoS requirements of time-sensitive multimedia applications.

Although a two-state Markov model has been used for wireless channel, our analytical framework can be extended to a finite-state Markov model. However, increasing the number

of channel states will significantly increase the computational complexity. How to simplify the computations with more channel states is under investigation.

ACKNOWLEDGMENT

This work has been supported by a research grant from Bell University Laboratories (BUL) under the sponsorship of Bell Canada and a Postgraduate Scholarship from the Natural Science and Engineering Research Council of Canada (NSERC).

REFERENCES

- [1] V. Jacobson and M. Karels, "Congestion avoidance and control," in *Proc. ACM SIGCOMM*, pp. 314–329, Aug. 1988.
- [2] S. Floyd and K. Fall, "Promoting the use of end-to-end congestion control in the Internet," *IEEE/ACM Trans. Networking*, vol. 7, no. 4, pp. 458–472, Aug. 1999.
- [3] Y. R. Yang, and S. S. Lam, "General AIMD congestion control," University of Texas-Austin, Technical Report TR-2000-09, May 9, 2000.
- [4] S. Floyd, M. Handley, J. Padhye, and J. Widmer, "Equation-based congestion control for unicast applications," in *Proc. ACM SIGCOMM*, pp. 43–56, Aug. 2000.
- [5] D. Bansal and H. Balakrishnan, "TCP-friendly congestion control for real-time streaming applications," MIT Technical Report MIT-LCS-TR-806, May 2000.
- [6] I. Rhee, V. Ozdemir, and Y. Yi, "TEAR: TCP emulation at receivers-flow control for multimedia streaming," NCSU Technical Report, 2000.
- [7] R. Rajaie, M. Handley, and D. Esretin, "RAP: An end-to-end rate-based congestion control mechanism for realtime streams in the Internet," in *Proc. IEEE INFOCOM*, vol. 3, pp. 1337–1345, Mar. 1999.
- [8] D. Sisalem and A. Wolisz, "LDA+ TCP-friendly adaptation: A measurement and comparison study," in *Proc. ICME*, vol. 3, pp. 1619–1622, Aug. 2000.
- [9] J. Padhye, V. Firoiu, D. Towsley, and J. Kurose, "Modeling TCP throughput: A simple model and its empirical validation," in *Proc. ACM SIGCOMM*, pp. 303–314, Sep. 1998.
- [10] M. Handley, S. Floyd, J. Padhye, and J. Widmer, "TCP Friendly Rate Control (TFRC): Protocol specification," Network Working Group RFC 3448, Jan. 2003.
- [11] H. M. Chaskar, T. V. Lakshman, and U. Madhow, "TCP over wireless with link level error control: Analysis and design methodology," *IEEE/ACM Trans. Networking*, vol. 7, pp. 605–615, Oct. 1999.
- [12] C.-F. Chiasserini and M. Meo, "A reconfigurable protocol setting to improve TCP over wireless," *IEEE Trans. Veh. Technol.*, vol. 51, no. 6, pp. 1608–1620, Nov. 2002.
- [13] H. Shiu, Y. Chang, T. Hou, and C. Wu, "Performance analysis of TCP over wireless link with dedicated buffers and link level error control," in *Proc. ICC*, vol. 10, pp. 3211–3216, May 2001.
- [14] H. L. Vu and S. Hanly, "A study of TCP performance and buffer occupancy over a fading wireless link," in *Proc. IEEE Globecom*, vol. 6, pp. 3478–3482, Nov. 2001.
- [15] J. W. K. Wong, and V. C. M. Leung, "Improving end-to-end performance of TCP using link-layer retransmissions over mobile internetworks," in *Proc. ICC*, vol. 1, pp. 324–328, June 1999.
- [16] M. Miyoshi, M. Sugano, and M. Murata, "Performance improvement of TCP on wireless cellular networks by adaptive FEC combined with explicit loss notification," in *Proc. IEEE VTC Spring*, pp. 982–986, May 2002.
- [17] A. Bakre, and B. R. Badrinath, "ITCP: Indirect TCP for mobile hosts," in *Proc. 15th International Conference on Distributed Computing Systems*, pp. 136–143, May/June 1995.
- [18] H. Balakrishnan, S. Seshan, and R. H. Katz, "Improving reliable transport and handoff performance in cellular wireless networks," *ACM Wireless Networks*, vol. 1, no. 4, pp. 469–481, Dec. 1995.
- [19] J. A. Cobb, and P. Agrawal, "Congestion or corruption? A strategy for efficient wireless TCP sessions," *Computers and Communications*, pp. 262–268, June 1995.
- [20] W. Ding and A. Jamalipour, "A new explicit loss notification with acknowledgment for wireless TCP," in *Proc. PIMRC*, vol. 1, pp. B-65–B-69, Sep./Oct. 2001.
- [21] M. Mathis, J. Mahdavi, S. Floyd, and A. Romanow, "TCP selective acknowledgement options," Network Working Group RFC 1818, Apr. 1996.

- [22] J. W. Mark and W. Zhuang, *Wireless Communications and Networking*. Upper Saddle River, NJ: Prentice Hall, 2003.
- [23] H. S. Wang and N. Moayeri, "Finite-state Markov channel-A useful model for radio communication channels," *IEEE Trans. Veh. Technol.*, vol. 44, no. 1, pp. 163–171, Feb. 1995.
- [24] M. Zorzi, R. R. Raced, and L. B. Milstein, "ARQ error control for fading mobile radio channels," *IEEE Trans. Veh. Technol.*, vol. 46, no. 2, pp. 445–455, May 1997.
- [25] S. Floyd and S. McCanne, "Network Simulator," LBNL public domain software. [Online] Available via ftp from ftp.cc.lbl.gov; NS-2 available: <http://www.isi.edu/nsnam/ns/>
- [26] E. Esteves, "The high data rate evolution of the cdma2000 cellular system," *Multimedia, Mobility and Teletraffic for Wireless Communications*, vol. 5, pp. 61–72, Dec. 2000.



Hong Shen received the B.S. and M. S. degrees from the Department Electronic Engineering, Nanjing University of Aeronautics and Astronautics, China in 1995 and 1998, respectively, and the M.A.Sc. degree in Electrical and Computer Engineering from the University of Waterloo, Canada in 2004. From 1998 to 2001, she was with Shenzhen University, where she was a researcher and a lecturer. Her main research interests include broadband wireless and mobile communications and CDMA technologies.



Lin Cai (S'00-M'06) received the MASc and PhD degrees (with Outstanding Achievement in Graduate Studies Award) in electrical and computer engineering from the University of Waterloo, Waterloo, Canada, in 2002 and 2005, respectively. Since July 2005, she has been an Assistant Professor in the Department of Electrical and Computer Engineering at the University of Victoria, Victoria, Canada. Her research interests span several areas in wireless communications and networking, with a focus on network protocol and architecture design supporting emerging multimedia traffic over wireless, mobile, and ad hoc and sensor networks. She serves as the Associate Editor for EURASIP Journal on Wireless Communications and Networking (JWCN), and International Journal of Sensor Networks (IJSNet).



Xuemin (Sherman) Shen (M'97-S'00) received the B.Sc. (1982) degree from Dalian Maritime University (China) and the M.Sc. (1987) and Ph.D. degrees (1990) from Rutgers University, New Jersey (USA), all in electrical engineering. From September 1990 to September 1993, he was first with the Howard University, Washington, D.C., and then the University of Alberta, Edmonton (Canada). Since October 1993, he has been with the Department of Electrical and Computer Engineering, University of Waterloo, Canada, where he is a Professor. Dr. Shen's research focuses on mobility and resource management in interconnected wireless/wireline networks, UWB wireless communications systems, wireless security, and ad hoc and sensor networks. He is a coauthor of two books, and editor of 10 journal Special issues, and has published more than 150 papers in wireless communications and networks, control and filtering.

Dr. Shen was the Technical Co-Chair for IEEE Globecom '03 Symposium on Next Generation Networks and Internet, and ISPAN '04. He serves as the Associate editor for IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS; IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY; *ACM Wireless Networks*; *Computer Networks*; *Dynamics of Continuous, Discrete and Impulsive*, Series B: Applications and Algorithms; *Wireless Communications and Mobile Computing* (Wiley); and *International Journal Computer and Applications*. He also serves as Guest Editor for IEEE JSAC; IEEE WIRELESS COMMUNICATIONS; and IEEE COMMUNICATIONS MAGAZINE. Dr. Shen received the Premier's Research Excellence Award (PREA) from the Province of Ontario, Canada for demonstrated excellence of scientific and academic contributions in 2003, and the Distinguished Performance award from the faculty of engineering, University of Waterloo, for outstanding contribution in teaching, scholarship and service in 2003. Dr. Shen is a Senior Member of the IEEE, and a Registered Professional Engineer of Ontario, Canada.