Improving Link Layer Performance on Satellite Channels With Shadowing Via Delayed Two-Copy Selective Repeat ARQ

Jing Zhu, *Student Member, IEEE,* and Sumit Roy, *Senior Member, IEEE*

Abstract—This paper focuses on improving performance of land mobile satellite channels (LMSCs) at high band (*Ka*-band or EHF band), where shadowing is the primary impediment to reliable data transmission. Compared with multipath fading, shadowing exists on a longer time scale; hence, interleaving to combat shadowing introduces unacceptably large decoding delay. We use Lutz's model to investigate bit-error rate/packet-error rate (BER/PER) performance of interleaving with various forward error correction (FEC) coding as a function of different channel parameters to demonstrate its limited effectiveness for combatting burst errors whose mean duration significantly exceed a link layer (LL) packet. We propose a *delayed two-copy selective repeat ARQ* (DTC-SR-ARQ) scheme, whereby two copies of a packet are sent-the second with a delay relative to the first-in every transmission or retransmission. Closed-form expressions for *mean transmission time, success probability*, and *residual loss probability* are provided and simulations used to validate the analysis. Furthermore, the issue of optimum *delay* is addressed as well, and a simple yet effective strategy is suggested to support transmission control protocol (TCP) traffic over this data link layer. DTC-SR-ARQ is shown to achieve much shorter additional delay than interleaving and compared with normal SR-ARQ, reduces mean transmission time at expense of a small increase in residual packet loss probability. Furthermore, ns2 simulation results show that for TCP traffic, DTC-SR-ARQ acquires higher end-to-end throughput than normal SR-ARQ.

Index Terms—ARQ/forward error correction (FEC), Internet, satellite networks.

I. INTRODUCTION

CURRENT SATCOM systems at lower bands (L-band: 1.5 GHz, such as Globalstar and Iridium) only support very low data rate services (e.g., Globalstar supports voice at rates of $n \times 1.2$ kb/s, where $n = \{1, 2, 3, 4\}$). However, emerging mobile broadband services (MBS) [1] including tele-medicine, distance learning, BOD (bandwidth-on-demand) applications, and Internet-in-the-sky require data rates of the order of megabits per second. Therefore, there is a great interest in extending satellite communications to higher bands (*Ka*-band or EHF band) to avail of greater bandwidth to support increased data transmission rates. Astrolink, iSky, Spaceway, SkeBridge, and Teledesic [1] are examples of such *Ka*-band SATCOM systems.

It was shown in [9] that the primary impairment on performance of land mobile satellite channel (LMSC) at *Ka*-band (20 GHz) or EHF band (30 GHz and above) is shadowing due to blockage rather than multipath fading. Generally, a shadowing channel can be represented by a two-state Markov process as in [2]. In bad (shadowed) states, the average signal-to-noise ratio (SNR) is too low for reliable detection even with powerful forward error correction (FEC) codes, while in good (unshadowed) states, the large value of Rice factor guarantees reliable signal transmission even without much FEC protection. As is well known, interleaving is widely used with FEC to resist fading and improve the reliability of a wireless channel with burst errors. However, with increasing average error burst length as in shadow fading, needed interleaving depth adds unacceptably large delays for time sensitive applications such as streaming media.

The challenge is, thus, to architect a *flexible* link/data link layer for a satellite channel that can adapt to varying channel conditions (from largely additive noise dominated situation to those with fading and degrees of shadowing). Adaptive hybrid ARQ/FEC approaches [4]–[7] provide the general context of approaches to this problem. Generally, FEC with reasonable code rate is used for random error correction and ARQ is adopted for blockage (shadowing) mitigation. Link level adaptive methods (e.g., adaptive ARQ [10]–[18], adaptive coded modulation [19]–[21], etc.) are a natural fit for wireless communications since wireless channels are time-varying. With such techniques, transmit parameters such as modulation, coding rate, packet length, ARQ type are adjusted (typically based on feedback from a cooperative receiver that provides current channel state information) to achieve optimized performance over time-varying channels. However, utility of all these adaptive proposals depends crucially on the efficiency and accuracy of channel state information (CSI), which requires feedback at least as fast as the variation of channel state. We classify the state variation of a satellite shadowing channel into two levels: short-term variations [shorter than round-trip time (RTT)] that are typical in the unshadowed state and the long-term variations (much longer than RTT) typical of shadowing. Naturally, the former cannot be tracked over a satellite channel, making it impossible to adopt any adaptive coded modulation scheme to enhance performance. However, it is feasible to track the longer-term variation of a shadowed channel. In this paper, a key parameter of our scheme is tuned according to the estimated mean length of error bursts obtained

based on the number of contiguous negative ACKs (NACKs) as in [18].

To further improve the efficiency, [6] proposed a new acknowledgment mechanism, by which not only the status of one packet but the entire receive window (such as noncontiguous blocks of packets) is returned into an ACK packet. Reference [7] showed that TCP end-to-end throughput is significantly improved by *jointly optimizing* the maximum retransmission number of selective repeat ARQ (SR-ARQ) and FEC code rate. Reference [4] proposed a novel adaptive hybrid ARQ for low earth orbit (LEO) systems, in which the side information of turbo decoder is utilized to choose different modes of link layer protocol—Go-Back-N ARQ for large-scale burst errors and FEC with interleaving for short-scale (random) errors.

In this paper, we introduce a *delay* transmission technique for the shadowed channel where the average length of error bursts is shorter than satellite segment RTT but longer than a link layer (LL) packet duration. This technique was first proposed in [8] with adaptive number of retransmissions; however, the mode of increasing the copy number with retransmission used in [8] was inefficient and the issue of optimizing delay was left unexplored. In this paper, we reduce the capacity loss caused by multiple-copy transmission by fixing the copy number to two in each transmission/retransmission. To further improve the efficiency, we suggest that the link layer protocol dynamically switch among three modes: Reed–Solomon (RS) code with interleaving, delayed two-copy (DTC) SR-ARQ, and normal SR-ARQ according to a long-term channel state estimate so that good link layer performance can be achieved for all scenarios. To study the effect of *delay*, we propose an analytical model by which closed-form expressions are obtained for the mean transmission time, success probability of first transmission, and residual loss probability after finite number of retransmissions. Furthermore, optimizing the *delay* is investigated and a simple yet effective strategy is suggested. Simulation results show that compared with normal SR-ARQ, delayed two-copy selective repeat ARQ (DTC-SR-ARQ) dramatically reduce the mean transmission time at the expense of small increase in residual packet loss probability, implying that our proposal may be suited for realtime services. By using ns2, we also compare the end-to-end throughput of TCP traffic over our suggested protocol *vis-a-vis* a normal selective-repeat ARQ data link layer, showing great performance improvement for a useful range of system parameters.

The paper is organized as follows. In Section II, we introduce our simulation model for the satellite channel physical layer. We study output bit-error rate/packet-error rate (BER/PER) performance of interleaving method with EEC in Section III to show the ineffectiveness of packet interleaving on satellite channels with shadowing. In Section IV, a model for DTC-SR-ARQ is introduced and expressions for metrics of interest, such as success probability of each transmission, mean transmission time, and residual loss probability are obtained. In Section V, we show numerical results to validate our analysis. We also compare the achieved end-to-end throughput for TCP traffic with our proposal to normal SR-ARQ and conclude the paper with some final remarks in Section VI.
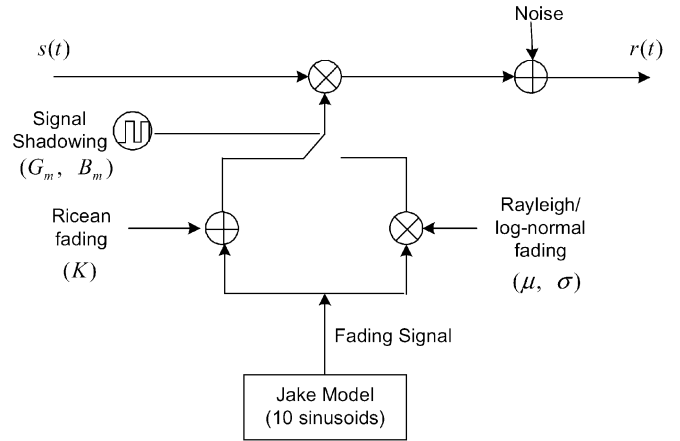

Fig. 1. Lutz's model for satellite channels.

II. SIMULATION SYSTEM OF THE SATELLITE PHYSICAL LAYER

A comprehensive model for the land mobile satellite communication channel (LMSC) proposed in [2] uses a total shadowing model that statistically combines Ricean fading and Rayleigh fading for received signal at a satellite, i.e.,

$$P(S) = X p\left(S \mid A, a_d^2\right) + (1 - X) \int_0^\infty p(S \mid S_0) p(S_0) \, dS_0$$
$$= X \frac{1}{\sigma_d^2} e^{-\frac{A^2 + 2s}{2\sigma_d^2}} I_0\left(\sqrt{2S} \frac{A}{\sigma_d^2}\right) + (1 - X) \int_0^\infty \frac{1}{S_0} e^{-\frac{S}{S_0}}$$
$$\times \frac{10}{\sqrt{2\pi} \ln 10} \frac{1}{S_0} e^{-\frac{(10 \log S_0 - \mu)^2}{2\sigma_d^2}} \, dS_0 \qquad (1)$$

where $0 \leq X \leq 1$. According to (1), the fading behavior of the channel consists of two dominant states. In the unshadowed or "good" channel state, the channel is characterized by the presence of a line-of-sight (LOS) component, implying higher average received power and Ricean fading. In the shadowed or bad state, the channel has no LOS component, implying low received power and Rayleigh fading. The parameter X is a long term measure that describes the fractional amount of time spent in good state. The characteristics of the channel state switching process are assumed to be described by a two-state Markov model.

Fig. 1 shows a schematic diagram for generating the faded/shadowed signal from the Lutz model. In our work, we do not use filtered Gaussian noise method to generate Rayleigh-fading signal envelope—instead, the sum of random sinusoids method suggested by Jake's model is used with ten sinusoids. For Ricean fading, a LOS component is added based on a a given Rice factor K. For Rayleigh fading, the fading envelope is scaled by log-normal statistics. The switching between Ricean and Raleigh fading is controlled by a two-state continuous time Markov process. The input signal $s(t)$ is multiplied by the fading signal and observed in additive Gaussian noise to produce the received signal $r(t)$.

We use MATLAB to simulate the satellite physical layer (see Fig. 2) with the Lutz channel model. An outer (n, k) RS code over GF (n) is symbol-interleaved to depth I (coding rate $r = k/n$) by using block interleaving. The output of the interleaver is convolutionally encoded with a binary code with code rate R,
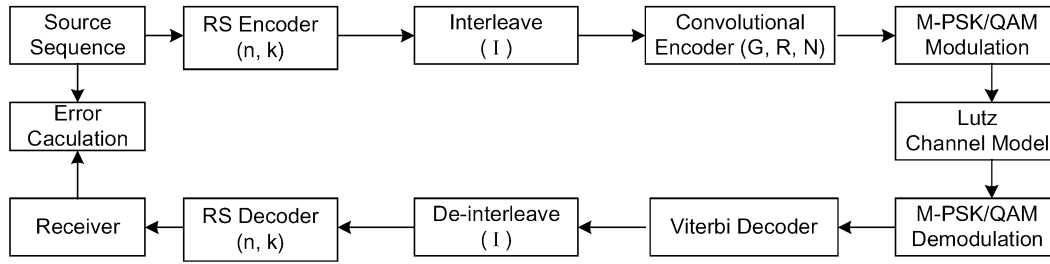
Fig. 2.   Block diagram of simulating the satellite physical layer.

constraint length N, and generator matrix G. Finally, M-PSK or M-QAM is utilized to modulate the signal. After demodulation, the received signal is hard limited followed by Viterbi decoding of the convolutional code, following the consultative committee on space data systems (CCSDS [24]) coding standard for deep-space communication.[1]

As is well known, concatenation of inner RS and outer convolutional codes (CCs) are a canonical FEC combination to achieve both coding gains in additive white Gaussian noise (AWGN) channel (CC), as well as some burst error correction capability (RS) when fading occurs. However, regular moderate-to-high rate [$r = (1/2)$ or higher] convolutional codes are largely ineffective in a shadowed state as they provide insufficient coding gain due to the very low SNR at the CC decoder input [6]. In this paper, the encoding/decoding function of outer convolutional codes is prohibited, leading to a FEC method using only RS codes with interleaving. Given a code rate (say $r = 0.2$) and codeword length (say $n = 63$), the corresponding information length $k = \lfloor n \times r \rfloor$, where $\lfloor x \rfloor$ is the largest integer contained in x).[2]

The system parameters for the model are tabulated as follows.

$f_d T$ Product of the maximum Doppler frequency f_d and symbol duration T that indicates a normalized measure of channel state time variations.

SNR   Signal noise power ratio at receiver input.

Bw Maximum symbol rate over the satellite channel (symbols per second).

T_{sim} Simulation run time (seconds).

m_b Mean duration of bad state (seconds).

m_g Mean duration of good state (seconds).

C_t Mean duration of shadowing period, given by $m_b + m_g$ (seconds).

X Good state time-share parameter, given by $(m_g)/(m_b + m_g)$.

III. INEFFECTIVENESS OF INTERLEAVING FOR SHADOWED CHANNELS

Here, we only consider slow fading signal, which is characterized by $f_d T \ll 1$. We fix $f_d T = 0.01$ as in the [9] study with 4800 symbols/s[3] with binary phase-shift keying (BPSK) modulation for a net bit-rate of 4800 b/s and SNR of 6 dB. In bad

[1]Note that CCSDS system was intended for an AWGN channel, and not a shadowed satellite channel characterized by [2].

[2]MATLAB functions encode/decode can be used to perform RS encoding/decoding with arbitrary code rate.

[3]4.8 k symbols/s is the minimum symbol rate for many satellite modems, such as P400 series (http://www.paradisedata.com/) and SDM series (http://www.comtechefdata.com/)
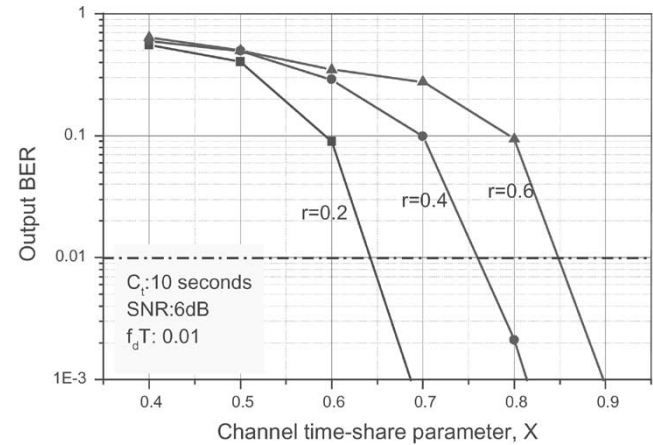

Fig. 3.   Effect of coding rate ($T_{\text{sim}} = 300$ s.)

states of the channel, the log-normal Rayleigh fading is characterized by $\mu = -20.8$ dB and $\sigma = -0.09$ dB, and in good states of the channel, the Rice factor K is 20 dB. C_t is fixed to 10 s and RS (63, k) code is used as FEC, where each symbol has 6 bits.

A preferred link-layer method to resist burst errors over a fading channel is interleaving. However, choosing an optimal value of the interleaving depth I in a time-varying fading channel is not straightforward. If the interleavers are designed for worst case scenarios, it may lead to unnecessarily long channel delays in good channel states. The maximum value of I corresponds to interleaving all codewords into one block, denoted as *sufficient interleaving* implying that errors at the deinterleaver output can be considered to be independent. For convenience, the interleaving depth in the following sections is normalized by the maximum value, i.e., $I = 1$ and $I = 0$ correspond to sufficient interleaving and no interleaving, respectively.

A. BER Performance of FEC With Sufficient Interleaving

First, we assume sufficient interleaving and focus on the issues of choosing coding scheme and appropriate coding rates. Fig. 3 illustrates the performance of FEC with different coding rates with increasing good state time-share parameter X. Obviously, the larger X leads to the smaller BER.[4] Moreover, for $X < 0.65$, even the lowest coding rate ($r = 0.2$) cannot achieve BER lower than 0.01, which is required for acceptable packetized voice communication. One solution to this problem is increasing the codeword length (say from 6 to 8), however, this

[4]If an RS block codeword has an error after decoding, all k information bits of the code are considered as error bits.
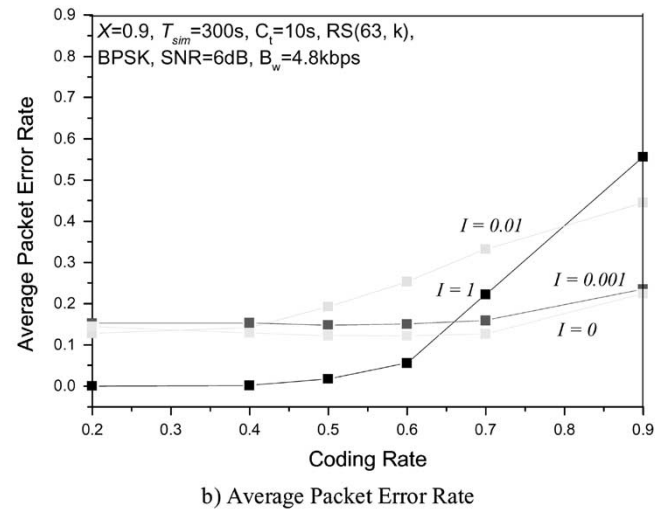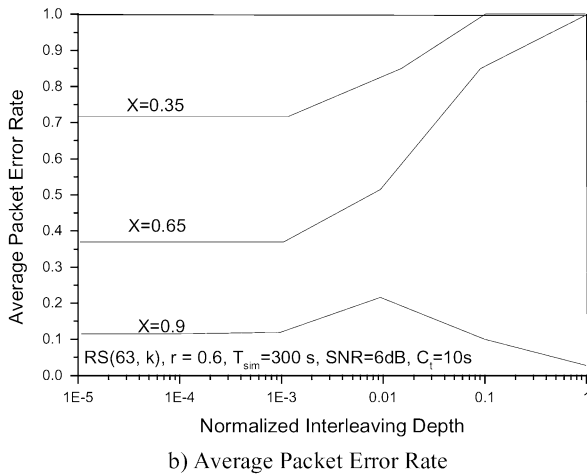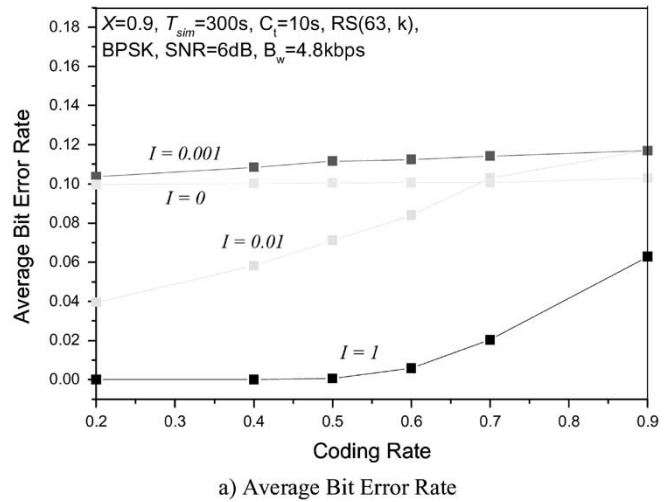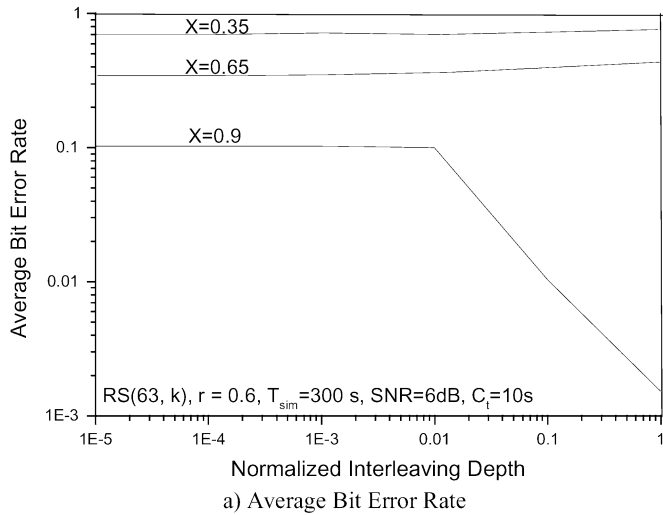
a) Average Bit Error Rate

b) Average Packet Error Rate

Fig. 4. Effect of interleaving depth with different X on output BER and PER.

a) Average Bit Error Rate

b) Average Packet Error Rate

Fig. 5. Effect of interleaving depth with different code rate on output BER and PER.

greatly increases the complexity of RS coding/decoding and is not a desirable option.
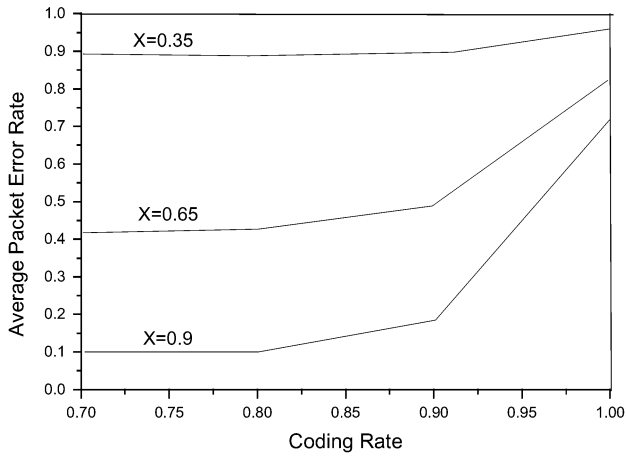
B. Insufficient Interleaving

The above results were based on the assumption of sufficient interleaving which is idealistic since the instantaneous error bursts may exceed any design value I with some finite probability. Therefore, we next investigate the effect of insufficient interleaving on output BER performance. Fig. 4 compares the effect of interleaving depth on both BER and PER, where 75 bytes length LL packets are assumed. We see that for the worse channels ($X = 0.65, 0.35$), BER does not change much with increasing interleaving depth, but PER increases—the main reason being insufficient interleaving. For the same BER, longer interleaving results in more PER as it disperses the same bit errors over greater number of packets. For $X = 0.9$, the average BER improves significantly after the normalized interleaving depth exceeds 0.01, implying that interleaving is long enough to take effect. Accordingly, the average PER increases first with interleaving depth and then improves after interleaving depth is longer than 0.01. For our simulation time of 300 s, the delay caused by interleaving with normalized depth of 0.01 is approximately 3 s, which is unacceptable as

it is corresponds to several RTTs [approximately six for the geosynchronous earth orbit (GEO) channel].
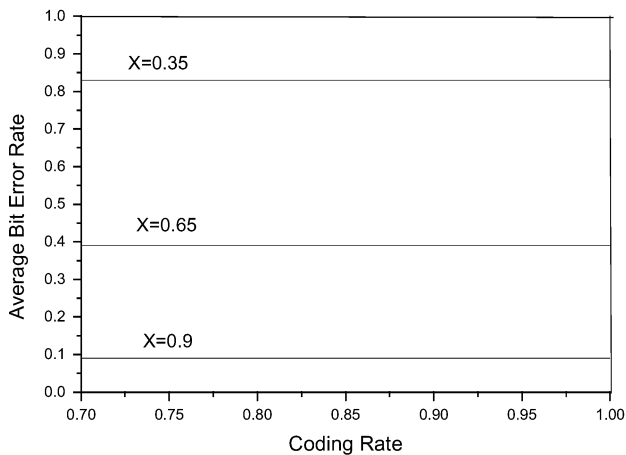
Fig. 5 shows the impact of interleaving on the error performance of FEC. We see that coding rate must be low enough to achieve BER/PER improvements with interleaving. Fig. 5(b) indicates that with code rate $r < 0.65$, full interleaving $(I = 1)$ scheme performs best; otherwise, no interleaving $(I = 0)$ is the best.

In summary, interleaving to correct burst errors in the bad (shadowed) channel state is only effective with *large depths* implying long delays of several seconds. Moreover, the code rate must be low enough to obtain the required error correction capability; otherwise, interleaving results increased PER.

Fig. 6 illustrates the effect of reducing coding rate on BER and PER (link layer PER) *without* interleaving. We see that the BER is nearly invariant but PER improves significantly with reasonably low coding rate. This is because errors occur in long bursts in the bad state that overwhelm the FEC capability. Since the errors in bad states constitute the majority of the total errors, any BER improvement by reducing code rate is slight. However, the lost packets in bad states includes more error bits than those in good states. In other words, assuming the same number of
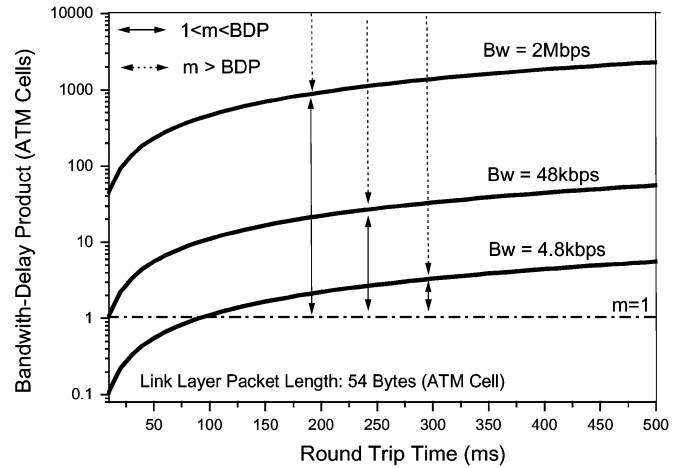
(a) Average Link Layer Packet Loss Rate

(b) Average Bit Error Rate

Fig. 6. Output BER/PER of a correlated channel.

Fig. 7. Classification of *bit*, *packet*, and *round-trip* level fading.

higher bands, packet level fading becomes more and more important. In the next section, we introduce a novel DTC-SR-ARQ scheme to combat packet loss in a packet level fading channel.

IV. DELAYED TWO-COPY SR-ARQ (DTC-SR-ARQ)

The protocol employs basic SR-ARQ, except that two identical copies of a packet with a delay D is sent at each attempt (an attempt consists of both transmissions of new packets or retransmissions of previously unsuccessful packets). Only when *both* copies in an attempt are lost, a NACK is produced and retransmission occurs. Compared with a normal one-copy scheme, the equivalent code rate is thus 0.5.

The parameters used in the subsequent analysis are as follows.

X Good state time-share parameter.
Peg   Packet error rate in good states.
Peb   Packet error rate in bad states.
m Mean length of bad states.
RTT   Round-trip time of satellite channel.
Bw Bandwidth of satellite channel (bits per second).

Both m and RTT are in seconds, however, for convenience we measure them with the number of transmitted LL packets.

The usual alternating two-state Markov model is assumed to represent the channel state evolution in time; the duration of a "bad" or shadowed state (i.e., the error burst length) is exponentially distributed with mean m. In the shadowed state, it is reasonable to assume $\text{Peb} \cong 1$, implying that no successful transmission is possible. Further, $\text{Peg} \ll \text{Peb}$ and for analytical purposes, may be assumed equal to zero, i.e., all packet transmissions during "good" or unshadowed state are successfully received. This leads to the important simplification that the sequence of packet success and failures is a Markov process.[5]

First, we consider transmission of any two successive packets with a general delay (d) in between. For transmission of two copies (of the same packet), we identify two important scenarios 1) the success/failure of the second copy in an attempt is dependant on that of the first copy in the same attempt and 2) the success/failure of the first copy in current attempt is dependant

error bits in bad and good states, the number of corresponding lost packets in good states is more than that in bad states. In conclusion, link layer performance of a satellite shadowing channel can be greatly improved by using FEC only (without interleaving) to correct random errors in unshadowed (good) states.

In a channel with large bandwidth-delay product (BDP), a fading process can be classified into three levels: *bit, packet*, and *round trip*, defined as $m < 1$, $1 < m < \text{BDP}$, and $m > \text{BDP}$, respectively, where m is the mean length of error bursts given by $C_t(l - X)$. Note that all quantities here are in units of LL packets. Most shadowing channels exhibit packet to round-trip level fading, and interleaving does not work well in such channels as we have demonstrated. However, in packet level fading where the average length of error bursts is shorter than RTT, additional options other than just retransmission are available. Fig. 7 illustrates the BDP [measured in terms of asynchronous transfer mode (ATM) cells] as a function of RTT, which clearly shows the demarcations of regions for bit, packet, and round-trip level fading for different RTT and bandwidth. The range of RTT covers 10 ms (LEO) to 500 ms (GEO), and three values of bandwidth, 4.8, 48 kb/s, and 2 Mb/s, are considered. Obviously, the region of packet level fading (i.e., $1 < m < \text{BDP}$) expands with increasing BDP, implying that as SATCOM systems migrate to

---

[5]In general, the sequence of packet successes and failures is a *hidden* Markov process and not strictly a Markov process.

on success/failure of the second copy in the *previous* attempt. Hence, there are two relevant values of d for consideration, D and RTT.

Since the PER is negligible in good states, packet losses take place in only bad states. If the previous copy is lost, the probability of correctly receiving the current one can be written as

$$P\{S \,|\, F\} = P\{B_E, G\}P\{S_G\} = P\{B_E\}P\{G \,|\, B_E\}P\{S_G\} \tag{2}$$

where the notations represent the following events.

S Success on the current copy.
F Failure on the previous copy.
B_E The bad state during transmission of the previous copy is completed before transmission of the current copy.
G The current copy is transmitted in a good state.
S_G Success of the current copy in a good state.

Since the duration of bad state is exponentially distributed with the mean of m and the interduration between two copies is d, we have

$$P\{B_E\} = 1 - e^{-\frac{d}{m}} \tag{3}$$

and clearly

$$P\{S_G\} = I - \text{Peg}. \tag{4}$$

An exact expression for $P\{G \,|\, B_E\}$ is difficult and, therefore, an intuitive approximation is given next. We know that $P\{G \,|\, B_E\}$ is a function of d. Let us consider two extreme cases.

Case 1) For $d \rightarrow 0$, i.e., two copies are sent successively with no delay. Since the end of the bad state for the previous transmission is followed by a good state, the current copy will be sent in the good state with probability 1, leading to $P\{G \,|\, B_E\} \rightarrow 1$. (Here, we ignore the event that the good state ends before one LL packet duration because the good state mean duration is significantly longer than a LL packet duration).

Case 2) For $d \rightarrow \infty$, the correlation between the channel states for the two copies vanishes, leading to $P\{G \,|\, B_E\} \rightarrow P\{G\}$, where $P\{G\} = X$.

We define the above two extreme cases as two mutually exclusive events: the channel state for the current copy is *completely correlated* (CC) or *completely uncorrelated* (CU) to that for the previous one. The probability of CC is given by the correlation function of $e^{-(d/m)} = \{ \begin{smallmatrix} 1, & d=0 \\ 0, & d=\infty \end{smallmatrix}$, and the probability of CU is $1 - e^{-(d/m)}$. The desired result for $P\{G \,|\, B_E\}$ is given by the statistical average of the above, i.e.,

$$P\{G \,|\, B_E\} \approx 1. e^{-\frac{d}{m}} + X \cdot \left(1 - e^{-\frac{d}{m}}\right). \tag{5}$$

In conclusion, the probability of correctly receiving the current copy given that the previous is lost is expressed as

$$P_{\text{Se}}(d) = \left(1 - e^{-\frac{d}{m}}\right) \left(e^{-\frac{d}{m}} + X \left(1 - e^{-\frac{d}{m}}\right)\right)$$
$$(1 - \text{Peg}), \quad d = \{D, \text{RTT}\}. \tag{6}$$

We next consider the success probability for transmission and retransmission, denoted by P_{St} and P_{Sr}, respectively. Here are the following events involved.

S_1 Success on the first copy.
F_1 Failure on the first copy.
S_2 Success on the second copy.
F_2 Failure on the second copy.
F_0 Failure on all previous attempts.

Obviously

$$P\{S_2 \,|\, F_1\} = P_{\text{Se}}(D) \tag{7}$$

and since the channel state for the current copy only depends on that for the previous copy, it yields

$$P\{S_1 \,|\, F_o\} = P_{\text{Se}}(\text{RTT}). \tag{8}$$

The first copy of the transmission has probability of $X(1 - \text{Peg})$ of being correctly received. Therefore

$$P_{\text{st}} = P\{S_1\} + P\{F_1\}P\{S_2 \,|\, F_1\}$$
$$= X(1 - \text{Peg}) + (1 - X(1 - \text{Peg}))P_{\text{Se}}(D). \tag{9}$$

Then

$$P_{\text{Sr}} = P\{S_1 \,|\, F_o\} + P\{F_1 | F_o\}P\{S_2 \,|\, F_1, F_o\}$$
$$= P\{S_1 \,|\, F_o\} + P\{F_1 \,|\, F_o\}P\{S_2 \,|\, F_1\}$$
$$= P_{\text{Se}}(\text{RTT}) + (1 - P_{\text{Se}}(\text{RTT}))P_{\text{Se}}(D). \tag{10}$$

If the maximum number of attempts allowed is N, the packet loss probability $P_l^{(N)}$ after N attempts is

$$P_l^{(N)} = (1 - P_{\text{sr}})^{N-1}(1 - P_{\text{st}}). \tag{11}$$

Denote by T the time from the first transmission to receipt of the ACK. We have the following results of probability distribution function of T.

1) Correctly receiving the second copy

$$\mathbf{P}[T = i(\text{RTT} + D)]$$
$$= \begin{cases} (1 - X(1 - \text{Peg}))P_{\text{Se}}(D) & i = 1 \\ (1 - P_{\text{St}})(1 - P_{\text{Sr}})^{i-2}(1 - P_{\text{Se}}(\text{RTT}))P_{\text{Se}}(D) & i \geq 2 \end{cases}. \tag{12}$$

2) Currently receiving the first copy

$$\mathbf{P}[T = i\text{RTT} + (i-1)D]$$
$$= \begin{cases} X(1 - \text{Peg}) & i = 1 \\ (1 - P_{\text{St}})(1 - P_{\text{Sr}})^{i-2}P_{\text{Se}}(\text{RTT}) & i \geq 2 \end{cases}. \tag{13}$$

Thus, the mean value of T is given below

$$E[T] = \frac{1}{\left(1 - P_l^{(N)}\right)} \sum_{i=1}^{N} \{(i\text{RTT} + (i-1)D)$$
$$\times \mathbf{P}[T = i\text{RTT} + (i-1)D]$$
$$+ i(\text{RTT} + D)\mathbf{P}[T = i(\text{RTT} + D)]\}. \tag{14}$$

As $N \rightarrow \infty$, $P_l^{(N)} \rightarrow 0$, (14) converges to

$$\lim_{N \to \infty} E[T] = \frac{(1 - P_{\text{St}})}{P_{\text{Sr}}}[\text{RTT} + D + \text{RTT} \, P_{\text{Sr}}$$
$$+ D(1 - P_{\text{Se}}(\text{RTT}))P_{\text{Se}}(D)]$$
$$+ X(1 - \text{Peg})\text{RTT}$$
$$+ (1 - X(1 - \text{Peg}))P_{\text{Se}}(D)(\text{RTT} + D). \tag{15}$$

The analytical expressions of (11) and (15) are most useful in providing performance estimates. Of course, the mean error
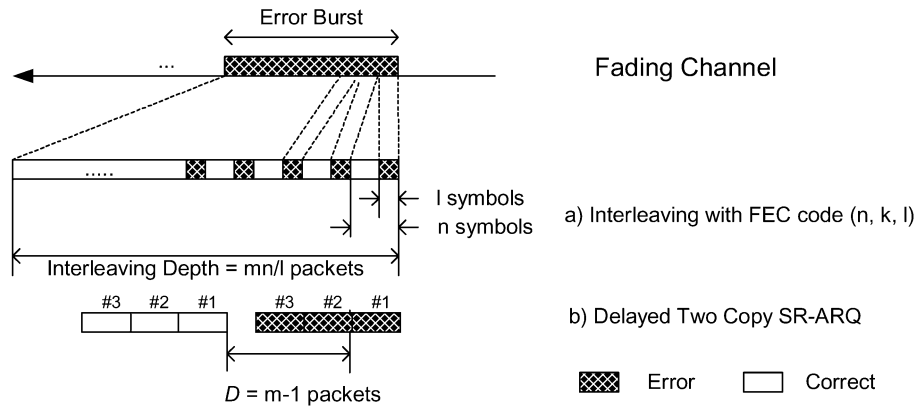
Fig. 8.   Isolated burst error model performance comparison.

burst length must be obtained *a priori* using a suitable long-term channel estimator. In the following, we assume that m is known.

V. NUMERICAL RESULTS AND DISCUSSION

In all simulations and analysis reported here, a LL packet is chosen as the unit of time and wireless channel bandwidth (Bw) is fixed at 1 Mb/s.

A. Performance Comparison (DTC-SR-ARQ to RS+Interleaving) With Isolated Burst Error Model (Fig. 8)

We first consider an isolated error burst with (fixed) length m (packets) followed by an long error-free period. This is called *isolated burst error model*, where the metric of interest is *additional delay* introduced by the method used.

For DTC-SR-ARQ, the minimum delay inserted between two copies of a packet for a successful transmission in a fixed model is $m-1$. The additional delay of the first copy is zero while that of the second one is m. Since either can be transmitted within an error burst with equal probability, the average additional delay introduced by DTC-SR-ARQ is $(m/2)$.

For interleaving with RS (n, k, l), the minimum interleaving length for error-free transmission is $(n/l)m$ (packets), where n is the length of codeword, k is the number of information symbols in a codeword, and l is the maximum number of revisable error symbols in a codeword. Deinterleaving starts at receiver side only after receiving all $(n/l)m$ packets if block interleaving is used. Therefore, additional delay is $(n/l)m$ for the first packet, and zero for the last for an average additional delay introduced equal to $(n/2l)m$. Fig. 9 demonstrates the results with m increasing and it is evident that DTC-SR-ARQ has much shorter additional delay than interleaving, especially for longer error burst lengths.

B. Delay Optimization of DTC-SR-ARQ

The above results are obtained for a fixed length of error burst. We next study the performance of DTC-SR-ARQ in more realistic case by modeling the channel as a two-state Markovian process, where the burst length is a random variable.

Fig. 10 shows the improvement in terms of success probability of the first transmission by increasing delay D. Our analytical results [given by (9)] match simulation results quite well
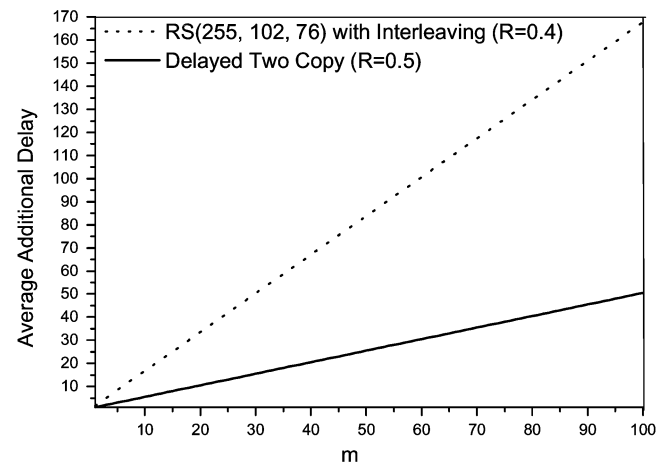

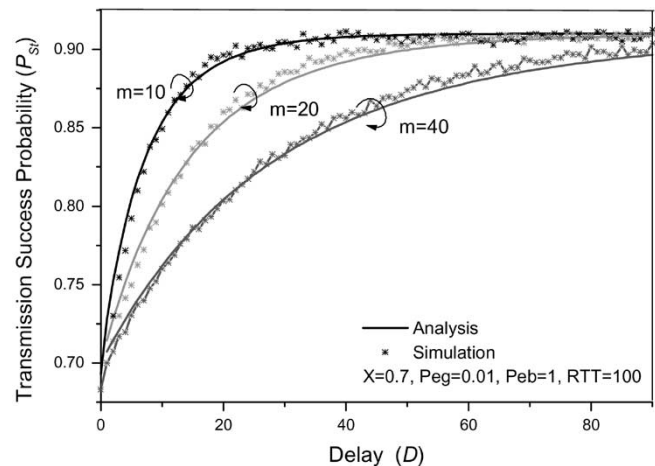Fig. 9.   Delay comparison of interleaving and DTC-SR-ARQ.

Fig. 10.   Success probability of transmission.

and we see that longer error bursts imply the need for longer delay D for same success probability as expected.

Fixing the maximum number of attempts at three, we study the residual loss probability after retransmission in Fig. 11 [equation (11) is used to obtain the analytical result]. The results show that the residual loss probability is dramatically reduced by increasing the delay. Fig. 12 investigates the mean transmission time with the analytical results from (14) showing
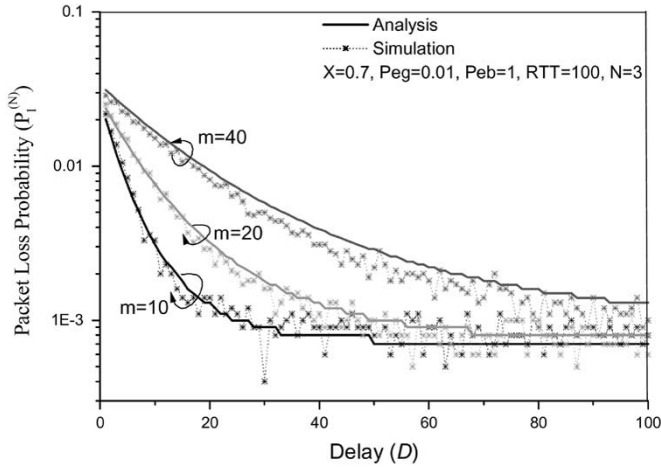
Fig. 11. Packet loss probability after N attempts ($N = 3$).

Fig. 13. Setting D with Km or $f_{\text{opt}}(m)$ ($\text{RTT} = 100, X = 0.7, P_{\text{eg}} = 0.01, N = \infty$).
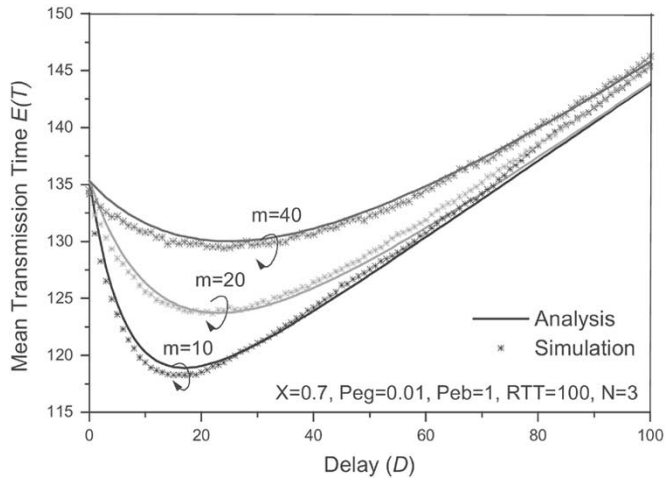

Fig. 12. Mean transmission time ($N = 3$).

a good match to simulation. There exists an optimal value of D yielding the minimum mean transmission time, which can be calculated by

$$\frac{dE[T]}{dD} = 0. \tag{16}$$

It is, however, tedious to explicitly solve (16), a numerical method is used to obtain D_{opt} as a function of m (denoted as $f_{\text{opt}}(m)$) for achieving the minimum mean transmission time with $N = \infty$, as shown in Fig. 13. For comparison, a simple linear function $D = Km$ is shown as well, where K is a constant.

Fig. 14 compares the link layer performance of DTC-SR-ARQ with $D = Km$ to $D = f_{\text{opt}}(m)$ in terms of residual packet loss probability and mean transmission time. Analytical results indicate that longer average burst error length leads to higher residual packet loss probability and longer mean transmission time. We compare the results with standard SR-ARQ equitably, thus, the maximum transmission number is set at MCN (say 4) for SR-ARQ and MCN/2 (say 2) for DTC-SR-ARQ. It is clearly seen that by using DTC-SR-ARQ, the mean transmission time is reduced at the expense of a small increase in residual packet loss probability (e.g., with
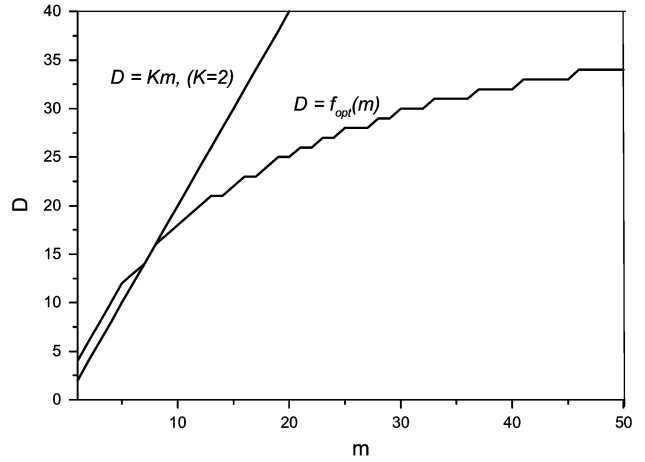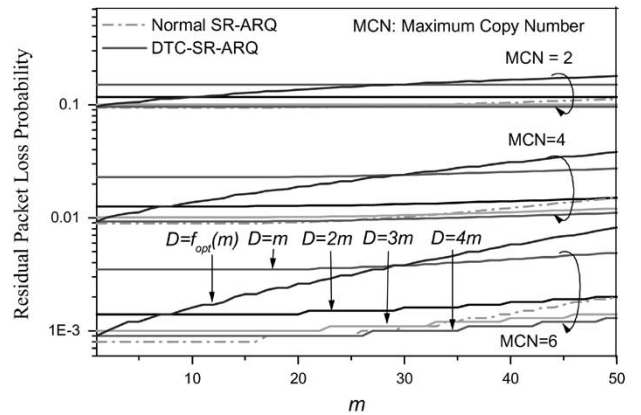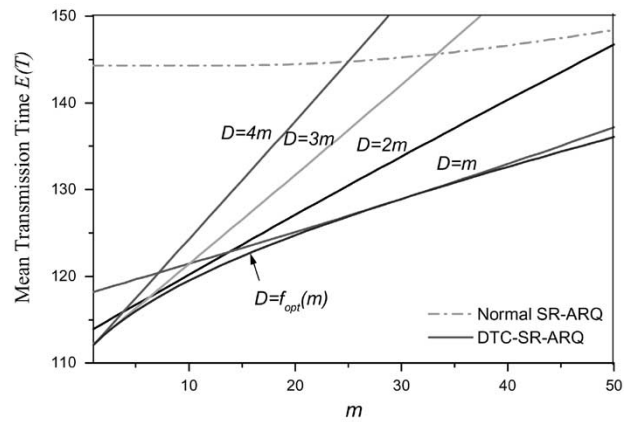

a) Residual Packet Loss Probability after N Attempts

b) Mean Transmission Time ($N=\infty$)

Fig. 14. The link layer performance comparison of DTC-SR-ARQ with $D = Km$ to $D = f_{\text{opt}}(m)$ ($\text{RTT} = 100, X = 0.7, \text{Peg} = 0.01$)

$D = 2$ m, the reduction in mean transmission time is about 20%, 10%, and $1 \sim 2\%$ for $m = 1, 25$, and 50, respectively). Obviously, shorter the average length of error bursts, the more the reduction in mean transmission time.

Comparing the results for the two rules $D = Km$ and $D = f_{\text{opt}}(m)$, we see that $D = f_{\text{opt}}(m)$ achieves the best performance in terms of mean transmission delay, but its residual packet loss probability increases dramatically with m; "$D =$
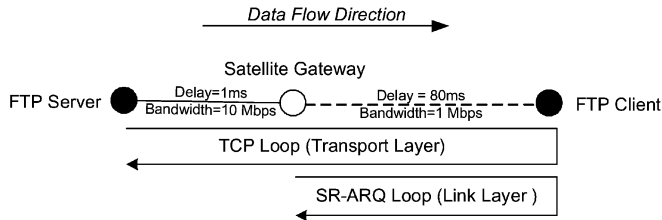
Fig. 15. Network simulation topology.

Km" has more stable performance in terms of residual packet loss probability, which is preferred especially for data traffic, although its mean transmission time is not optimized. The lack of a simple, explicit expression for $f_{\text{opt}}(m)$ is a potential implementation issue; thus, we suggest using $D = Km$ as a good pragmatic solution. To prevent values of the *delay* time exceeding desired bounds, the following equation may be used as a practical rule of thumb:

$$D = \min\left(Km, \frac{\text{RTT}}{2}\right). \qquad (17)$$

C. TCP Performance Comparison With DTC-SR-ARQ to SR-ARQ

In this section, we study the performance of TCP traffic (Reno) over DTC-SR-ARQ with $D = \min(Km, (\text{RTT}/2))$ ($K = \{1, 2, 3, 4\}$). Although TCP traffic is not delay sensitive, the long persistence time due to retransmission results in buffer overflow degrading the whole performance significantly, especially for the long RTT (equivalently large BDP) network. DTC-SR-ARQ greatly increases success probability of each transmission/retransmission with consequent improvement in TCP end-to-end throughput.

Fig. 15 shows the simulation network topology with ns2 simulator. A single TCP *Reno* connection with FTP application is considered. The bandwidth and delay of terrestrial part is set to 10 Mb/s and 1 ms, respectively. The buffer size of satellite gateway is 20 kB. The TCP packet length and the LL packet length is fixed at 500 bytes and 50 bytes, respectively. A 1 Mb/s satellite link is assumed that drops LL packets according to a two state Markovian channel model with associated parameters $-X = 0.7, P_{\text{eg}} = 0.01$, and $P_{\text{eb}} = 1$. The satellite RTT is 160 ms, leading to the BDP of 400 LL packets. We focus on the effect of average length of error burst (i.e., m) on TCP end-to-end throughput, ranging from 5 to 120 (measured in terms of LL packet transmission time). For a fixed maximum copy number (say 8), the maximum retransmission times for normal SR-ARQ is seven, and three for our proposed two-copy delayed SR-ARQ. Each point in Fig. 16 was obtained by averaging the results from six independent runs, each with 1000 s simulation run time.

Fig. 16 shows that the TCP end-to-end throughput is improved by using our scheme, especially when the average error burst length is short. When the error burst length increases, the performance improvement of our proposal is reduced-the main reason being that the mean burst length approaches the RTT where DTC-SR-ARQ is much less effective. In addition, $K = 2$ obtains the best performance for this scenario.
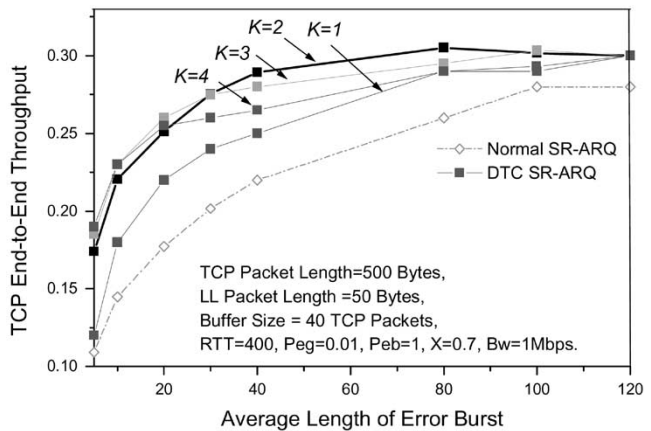

Fig. 16. TCP end-to-end throughput comparison (Buffer Size = 20 000 Bytes).

VI. CONCLUSION

In this paper, we used Lutz's model [2] to investigate BER/PER performance of shadowed channel with interleaving in conjunction with various FEC codes as a function of different channel parameters, and showed the failure of interleaving on such shadowing channels. According to the length of error burst, we classified the fading/shadowing into three levels: *bit, packet*, and *round trip*. Interleaving works very well on a bit level fading channel, while retransmission is the only effective recourse for round-trip level fading. The focus in this paper is packet level fading, with the average length of error burst more than one packet but far less than one RTT. A new method—DTC-SR-ARQ, rather than interleaving was suggested to combat burst errors in such packet level fading channels. Analytical results showed that success probability of each transmission is significantly improved, and mean transmission time is reduced as well. By comparing DTC-SR-ARQ with interleaving method in a isolated burst error model, we showed that DTC-SR-ARQ greatly reduces additional delay. Moreover, simulation was performed to evaluate TCP end-to-end throughput over DTC-SR-ARQ as compared with standard SR-ARQ, indicating that our proposal achieves performance improvement.

One problem of DTC-SR-ARQ is that the saturation throughput is reduced by half compared with normal SR-ARQ because two copies are sent in each transmission/retranmission. In other words, reliability is improved at the cost of capacity loss. Therefore, it is important to determine whether such capacity loss is worthy or not before deploying DTC-SR-ARQ into a real system. Obviously, when channel is in good condition (e.g., large X or smaller m), the added overhead caused by DTC-SR-ARQ will greatly impact the performance. An *adaptive* mechanism can be used to solve this problem by allowing link layer protocol to switch among different modes according to the long-term estimate of channel condition (e.g., X and m). Thus, a flexible link layer protocol should operate in three modes: *FEC with interleaving, DTC-SR-ARQ*, and *normal SR-ARQ*, corresponding to three levels of fading, *bit, packet*, and *round trip* so that the optimal link performance can be achieved for all scenarios. Furthermore, since DTC-SR-ARQ

has much better delay performance than normal SR-ARQ, it can be used to support real-time services such as video phone. The larger bandwidth (megabits per second) provided by higher band satellite systems also gives a good opportunity of using implement DTC-SR-ARQ (equivalent to $1/2$ rate coding) to improve the transmission reliability, because FEC even with very low code rate (say $r < 1/2$) is not very useful in a shadowing channel due to the problem of insufficient interleaving.

Future work will focus on effective channel estimation methods for the probability density function (PDF) or important parameters (X, m) of the shadowing process online, so that it may be implemented within the flexible multimode link layer strategy mentioned above.

References

[1] J. Farserotu and R. Prasad, *IP/ATM Mobile Satellite Networks*. Norwood, MA: Artech House, 2002.

[2] E. Lutz, D. Cygen, M. Dippold, F. Dolainsky, and W. Papke, "The land mobile satellite communication channel-recording, statistics, and channel model," *IEEE Trans. Veh. Technol.*, vol. 40, pp. 375–386, May 1991.

[3] J. B. Schodorf and M. A. Gouker, "Performance evaluation of a hybrid ARQ protocol implementation for EHF SATCOM on the move systems," in *Proc. MILCOM'01*, vol. 1, 2001, pp. 311–315.

[4] W. Li, V. K. Dubey, and C. L. Law, "An adaptive hybrid-ARQ scheme combating burst-errors caused by power control lag in Ka-band LEO satellite systems," in *Proc. MILCOM'01*, vol. 1, 2001, pp. 694–699.

[5] F. Babich, E. Valentinuzzi, and F. Vatta, "Performance of hybrid ARQ schemes for the LEO satellite channel," in *Proc. GLOBECOM'01*, vol. 4, San Antonio, TX, 2001, pp. 2709–2713.

[6] L. Casone, G. Ciccarese, M. De Blasi, L. Patrono, and G. Tomasicchio, "An efficient ARQ protocol for a mobile geo-stationary satellite channel," in *Proc. GLOBECOM'01*, vol. 4, San Antonio, TX, 2001, pp. 2692–2697.

[7] J. Zhu and S. Roy, "Performance of land mobile satellite communication (LMSC) channel with hybrid FEC/ARQ," in *Proc. GLOBECOM'02*, 2002, pp. 2851–2854.

[8] J. Zhu, Z. Niu, and Y. Wu, "A delayed multiple copy retransmission scheme for data communication in wireless networks," in *Proc. 9th IEEE Int. Conf. Networks*, 2001, pp. 310–315.

[9] J. B. Schodorf, "EHF satellite communications on the move: Baseband considerations," MIT Lincoln Lab., Tech. Rep. 1055, Feb. 2000.

[10] S. Lin, D. J. Costello, and M. J. Miller, "Automatic-repeat-request error-control schemes," *IEEE Commun. Mag.*, vol. 22, pp. 5–17, Dec. 1984.

[11] Y.-D. Yao, "An effective go-back-N ARQ scheme for variable-error-rate channels," *IEEE Trans. Commun.*, vol. 43, pp. 20–23, Jan. 1995.

[12] H. Bruneel and M. Moeneclaey, "On throughput performance of some continuous ARQ strategies with repeated transmissions," *IEEE Trans. Commun.*, vol. 34, pp. 244–249, Mar. 1986.

[13] E. J. Weldon, "An improved selective-repeat ARQ strategy," *IEEE Trans. Commun.*, vol. 30, pp. 480–486, Mar. 1982.

[14] A. C. Martins and J. C. Alves, "ARQ protocols with adaptive block size perform better over a wide range of bit error rates," *IEEE Trans. Commun.*, vol. 38, pp. 737–739, June 1990.

[15] M. Rice and S. B. Wicker, "A sequential scheme for adaptive error control over slowly varying channels," *IEEE Trans. Commun.*, vol. 42, pp. 1533–1543, Feb./Mar./Apr. 1994.

[16] S. Kallel and C. Leung, "Analysis of memory and incremental redundancy ARQ schemes over a nonstationary channel," *IEEE Trans. Commun.*, vol. 40, pp. 1474–1480, Sept. 1992.

[17] B. Vucetic, "An adaptive coding scheme for time-varying channels," *IEEE Trans. Commun.*, vol. 39, pp. 653–663, Mar. 1991.

[18] A. Annamalai and V. K. Bhargava, "Analysis and optimization of adaptive multicopy transmission ARQ protocols for time-varying channels," *IEEE Trans. Commun.*, vol. 46, pp. 1356–1368, Oct. 1998.

[19] D. L. Goeckel, "Robust adaptive coded modulation for time-varying channels with delayed feedback," in *Proc. 35th Annu. Allerton Conf. Communication, Control, and Computing*, Oct. 1997, pp. 370–379.

[20] ——, "Adaptive coding for fading channels using outdated channel estimates," *IEEE Trans. Commun.*, vol. 47, pp. 844–855, June 1999.

[21] A. J. Goldsmith and S. Chua, "Adaptive coded modulation for fading channels," *IEEE Trans. Commun.*, vol. 46, pp. 595–583, May 1998.

[22] G. Fairhurst and L. Wood, Advice to link designers on link automatic repeat request (ARQ), Aug. 2001. draft-ietf-pilc-link-arq-issues-03.txt.

[23] W. S. Jeon and D. G. Jeong, "Improved selective repeat ARQ scheme for mobile multimedia communications," *IEEE Commun. Lett.*, vol. 4, pp. 46–48, Feb. 2000.

[24] S. G. Wilson, *Digital Modulation and Coding*. Englewood Cliffs, NJ: Prentice-Hall, 1999.
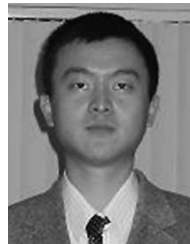
Jing Zhu (S'01) received the B.S. and M.Sc. degrees in electronic engineering from Tsinghua University, Beijing, China, in 2000 and 2001, respectively. Since August 2001, he has been working toward the Ph.D. degree in electrical engineering at the University of Washington, Seattle.

His main research interests are performance issues in mobile network (MONET) including land mobile satellite (LMS) system, next-generation cellular system, wireless LAN, ad hoc network, etc., and their applications, e.g., Internet-on-Move, Home Networking, and Intelligent Transport System, focusing on protocol design issues at different layers, i.e., transport layer (e.g., TCP), link layer (e.g., ARQ), and MAC layer (e.g., IEEE802.11). His current work is on MAC design for multihop ad hoc networks.

Sumit Roy (S'84–M'88–SM'00) received the B.Tech. degree from the Indian Institute of Technology, Kanpur, in 1983, and the M.S. and Ph.D. degrees from the University of California, Santa Barbara, all in electrical engineering, in 1985 and 1988, respectively, as well as the M.A. degree in statistics and applied probability in 1988.

His previous academic appointments were at the Moore School of Electrical Engineering, University of Pennsylvania, and at the University of Texas, San Antonio. He is presently a Professor of electrical engineering at the University of Washington, where his research interests center around analysis/design of physical and data link layer of communication systems, with a topical emphasis on next-generation mobile/wireless networks. He is currently on academic leave from Intel Wireless Technology Laboratory, Hillsboro, OR, working on high-speed UWB radios and next-generation wireless LANs. He is an Editor for the *Wiley Journal on Wireless Communications and Mobile Computing*.

Dr. Roy is an Editor for the IEEE Transactions on Wireless Communications. His activities for the IEEE Communications Society includes membership of several technical committees and technical program committees for conferences.