Performance of TCP over wireless links with hybrid ARQ

Article

Sammy Chan
City University of Hong Kong

Li Ping
City University of Hong Kong

King-Tim Ko
City University of Hong Kong

ITC19/ Proceedings of Poster
Beijing University of Posts and Telecommunications Press
2267-2276

Performance of TCP over wireless links with hybrid ARQ*

Qian Huang [1], Sammy Chan [1], Li Ping [1], King-Tim Ko [1] and Moshe Zukerman [2]

[1] Department of Electronic Engineering, City University of Hong Kong. No. 83, Tat Chee Avenue, Kowloon, Hong Kong, China
{qhuang, schan}@ee.cityu.edu.hk; {eeliping, eektko}@cityu.edu.hk
[2] the ARC Special Research Centre for Ultra-Broadband Information Networks, an affiliate program of National ICT Australia, Department of Electrical and Electronic Engineering, The University of Melbourne, Melbourne, Vic. 3010, Australia
m.zukerman@ee.mu.oz.au

Abstract: This paper studies TCP performance over a wireless link where hybrid Automatic Repeat reQuest (ARQ) schemes are adopted for combating packet loss due to wireless channel errors. A Markov chain is developed to analytically evaluate the TCP throughput over an arbitrary hybrid ARQ scheme. Base on this model, we can learn how the underlying coding scheme, the link layer retransmission method and the TCP packet buffer size at the access point affect the TCP performance. We demonstrate by analysis and verify by simulation that the CT-TCM hybrid ARQ scheme provides more efficient spectrum utilization and higher TCP throughput than other existing hybrid ARQs.
Keywords: Wireless TCP, hybrid ARQ, TCM code.

1 Introduction

The increasing demand for high data rate services in wireless networks promotes Transmission Control Protocol (TCP) to be extensively used in the Third Generation (3G) systems. The congestion control mechanism of TCP dominates the wired networks under the assumption that packet losses are mainly due to network congestion. However, previous studies have demonstrated that TCP will suffer serious performance degradation, e.g., spurious end-to-end retransmission and timeouts, when it is directly used over wireless links [1, 4]. The reason is that the above assumption is not suitable for wireless networks, where packet losses are mainly incurred by wireless channel errors instead of congestion. Various solutions have been proposed to improve the TCP performance over wireless networks [2]-[6]. Among these solutions, error control schemes, e.g., local retransmission by ARQ, forward error recovery (FEC) or their hybrid (hybrid ARQ), are proposed to enhance the reliability of wireless links [4]-[6]. These error control schemes have demonstrated the following benefits to wireless TCP applications: 1) They can cooperate with

* This work was jointly supported by grants from the Research Grants Council of the Hong Kong SAR, China [Project No., CityU 1190/02E] and the Australian Research Council.

TCP without requiring any changes to the current congestion control mechanism. 2) If the cross-layer configuration is considered between the link layer and TCP, they can provide useful physical or link layer information to TCP for optimization, or they can cooperate with some existing wireless TCP solutions, e.g., snoop TCP [4], explicit-loss-notification based TCP [4], to maximize the throughput over wireless links.

With new error control schemes used in the link layer, one would naturally be interested in what are their impacts on the TCP level throughput, because these impacts have implications for performance perceived by end users. In [4]-[6], simulations or analyzes based on a simple channel model with two states (good and bad) are used to evaluate the TCP level throughput. However, when hybrid ARQ schemes are used in wireless links, the channel condition is characterized by the symbol error rate (SER) or frame error rate (FER) achieved by the hybrid ARQ under a given signal-to-noise ratio (SNR). It is therefore desirable to know the resulting TCP level throughput for a given SNR. In this paper, we develop an analytical model to study the performance of TCP over a wireless link, where hybrid ARQ schemes are deployed to combat wireless channel errors.

The remainder of this paper is organized as follows. A Markov chain is developed to evaluate the TCP throughput over an arbitrary hybrid ARQ in Section 2, which is followed by a survey of recently proposed hybrid ARQ schemes in Section 3. Analytical and simulation results of TCP performance over these hybrid ARQs are presented in Section 4. The conclusions are presented in Section 5.

2 An analytical model for TCP over hybrid ARQ

We consider the TCP performance in a common scenario in practice, as shown in Fig. 1, where a mobile user downloads information from a server. A hybrid ARQ scheme is used for local error control in the wireless link from the access point (AP) to the receiver. In Fig. 1, d and D are round trip propagation delays over the wireless link and the wired path from the sender to the AP, respectively; B is the TCP buffer size at the AP for queueing TCP packets from the sender; T and T_w are random variables representing service time and queueing delay of a TCP packet at the AP, respectively; t_a is the transmission delay of a link layer acknowledgement (ACK) or negative acknowledgement (NAK) packet over the wireless link; C is the wireless link capacity in bits/s; p is FER over the wireless link.

To simplify our analysis on TCP performance in the considered scenario, the wired path is assumed error-free, the wireless link is assumed to be an additive white Gaussian noise (AWGN) channel, where the packet errors are independently and identically distributed (i.i.d.). With i.i.d. packet errors, the TCP sender is not likely to experience two or more consecutive packet losses. Hence, it is reasonable to assume that the probability of TCP timeouts due to consecutive packet losses is small enough to be negligible. It is also reasonable to assume that most of the traffic from the sender to the receiver is made of aggressive large flows. Under these circumstances, we ignore the slow-start phase and only consider the congestion avoidance phase of TCP.

The TCP packets from the sender are queued in the buffer at the AP. Each TCP packet is fragmented into N smaller link layer frames with fixed length L_f and transmitted in the wireless link. Since the link layer ACKs and NAKs usually have fairly smaller size than TCP packets, t_a can be ignored. For the same reason, we also ignore any delay that

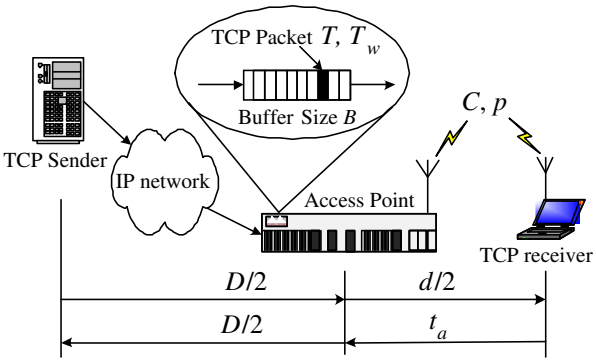

Fig. 1. Wireless TCP system.

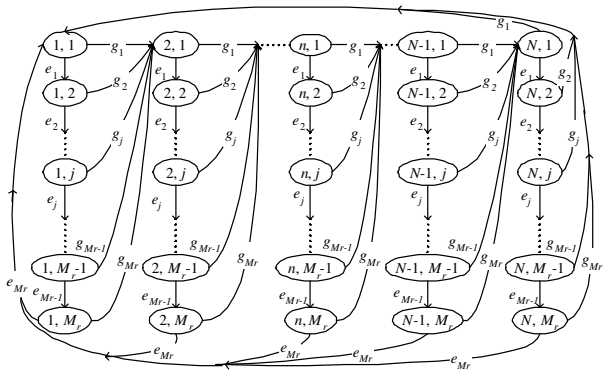

Fig. 2. The DTMC.

results from retransmissions of the link layer ACKs and NAKs. As usually, $D \gg d$, we assume $D + d = D$. The arrival of TCP packets at the AP is assumed to be a Poisson process with rate λ. It has been demonstrated to be reasonable when the network is large and shared by many traffic sources [7]. We consider that the link layer frames are orderly transmitted by stop-and-wait ARQ to the receiver. It leads to one-by-one transmission of the TCP packets from the AP to the receiver. Observing the number of TCP packets in the buffer at a set of special epochs when a TCP packet departs the buffer, we can model the TCP buffer as an M/G/1 queue [8]. From the M/G/1 queueing system, the average waiting time per TCP packet at the AP, denoted $\bar{T}_w$, is given by

$$\bar{T}_w = \lambda E[T^2] \big/ [2(1 - \lambda \bar{T})] \tag{1}$$

where $\bar{T}$ is the average service time per TCP packet, $E[T^2]$ is the second moment of T. Let $\bar{\tau}$ be the average round trip time (RTT) per TCP packet. It can be calculated by

$$\bar{\tau} = D + \bar{T} + \bar{T}_w = D + \bar{T} + \lambda E[T^2] \big/ [2(1 - \lambda \bar{T})]. \tag{2}$$

We observe that the process of TCP packets transmission over the hybrid ARQ can be modeled by a discrete-time Markov chain (DTMC), from which we can obtain $\bar{T}$ and $E[T^2]$ per TCP packet at the AP. Let (n, j) be a state of the DTMC, representing the nth frame of a TCP packet transmitted the jth time by the hybrid ARQ. Let T_f be the frame transmission time over the wireless link. We have $T_f = L_f/C$. The sojourn time at each state is a fixed time slot given by T_f. Let M_r be the maximum number of transmissions per frame over the wireless link. The DTMC is shown in Fig. 2, where $1 \le n \le N$, $1 \le j \le M_r$, g_j and e_j are the probabilities that a frame succeeds and fails in the jth transmission, respectively. $g_j + e_j = 1$. Note that g_j and e_j are obtained from individual hybrid ARQ schemes. Let Ω be the set of states of the DTMC. The total number of states in Ω is NM_r. Let $s = (n-1)M_r + j$ be the index of state (n, j), $1 \le s \le NM_r$. After a TCP packet finishes its transmission due to either failure in M_r attempts (i.e., at state $s \in \{s = nM_r : 1 \le n \le N\}$) or successful transmission (i.e., at state $s \in \{s = (N-1)M_r + j : 1 \le j \le M_r\}$), the state transits to $(1, 1)(s = 1)$, representing a new transmission start of the next TCP packet in the buffer. Let $p_{ss'}$ be the transition probability from any state $s \in \Omega$ to any other state $s' \in \Omega$. The values of

$\{p_{ss'} : 1 \leq s, s' \leq NM_r\}$ according to Fig. 2 are given by

$$p_{ss'} = \begin{cases} e_{M_r}, \text{for } s = nM_r, s' = 1 \\ g_j, \quad \text{for } s = (N-1)M_r + j, s' = 1 \bigcup s = (n-1)M_r + j, s' = nM_r + 1 \\ e_s, \quad \text{for } s = (n-1)M_r + j, s' = s + 1 \\ 0, \quad \text{for } s = s'. \end{cases}$$

From $p_{ss'}$, we can form the transition probability matrix $\mathbf{P}_{NM_r \times NM_r}$=$[p_{ss'}]$. Let π_s be the probability of the system being in state s. Define $\Pi = [\pi_s]$ a row vector representing the state probabilities of the DTMC. We now have the set of steady state equations $\pi_s = \sum_{s' \in \Omega} \pi_{s'} p_{s's}$ for all $s \in \Omega$ and the normalizing equation $\sum_{s \in \Omega} \pi_s = 1$. By solving these equations, we obtain the set of stationary state probabilities Π. However, on calculating $\bar{T}$ and $E[T^2]$, we are also interested in the transient state probabilities of the model.

Define $\Pi^{(0)}$ as a row vector representing the initial probabilities for state $s \in \Omega$ with elements ${\pi_1}^{(0)} = 1$ and ${\pi_s}^{(0)} = 0$ for $s \in \Omega - \{1\}$, because the state transition always starts from state $(1,1)$ $(s = 1)$ for every TCP packet. Let $\mathbf{\Phi}$ be a NM_r-by-NM_r matrix, representing the state transition probabilities of our Markov chain, excluding the transitions to state $(1,1)$ $(s = 1)$. Let $\phi_{ss'}$, $s, s' \in \Omega$, be the element of $\mathbf{\Phi}$. Its values are given by $\phi_{ss'} = p_{ss'}$ for $s \neq s', s' \neq 1$ and $\phi_{ss'} = 0$ for other $s, s' \in \Omega$. Define $X = [x_s]$ and $Y = [y_s]$ as two column vectors representing the transition probabilities from state $s \in \Omega$ to state $(1,1)$ due to success and failure in a TCP packet transmission, respectively. From Fig. 2, we have $x_s = g_j$ for $\{s = (N-1)M_r + j : 1 \leq j \leq M_r\}$ and $x_s = 0$ for all other s; $y_s = e_{M_r}$ for $\{s = nM_r : 1 \leq n \leq N\}$ and $y_s = 0$ for all other s. Let l be the number of state time slots experienced in one TCP packet transmission (including success and failure). Based on the property of Markov chains [8], the transient state probabilities for state $s \in \Omega$ at discrete time l is given by $\Pi^{(l)} = \Pi^{(0)} \mathbf{\Phi}^{l-1}(X + Y)$. From Fig. 2, l is given from $\min(M_r, N)$ to $N \times M_r$. Thus, the average service time of a TCP packet over the wireless link and its second moment can be calculated by

$$\bar{T} = \sum_{l=\min(M_r,N)}^{NM_r} \Pi^{(0)} \mathbf{\Phi}^{(l-1)}(X + Y) l T_f, \tag{3}$$

$$E[T^2] = \sum_{l=\min(M_r,N)}^{NM_r} \Pi^{(0)} \mathbf{\Phi}^{(l-1)}(X + Y)(l T_f)^2. \tag{4}$$

Substituting Eq. (3) and Eq. (4) into Eq. (2), we can obtain the average RTT $\bar{\tau}$ of a TCP packet through the wireless link deploying a general hybrid ARQ scheme.

In order to obtain the average arrival rate λ of TCP packets at the AP, we need estimate the average size of congestion window (*cwnd*) of a TCP connection. There are following two cases related to the variations of the *cwnd* of a TCP connection over wireless links.

Firstly, if the detected packet losses related to wireless channel errors are very small, i.e., the wireless channel errors have been mostly or thoroughly combated by the link layer hybrid ARQ schemes, the size of *cwnd* can additively increase every RTT (since we have already assumed no congestion-related losses) till the maximum buffer size B at the AP is reached. Then, the subsequently arrived TCP packets will be dropped due to buffer overflow and the size of *cwnd* will be halved (e.g. for TCP Reno [9]) due to packet losses. Let w_p denote the maximum size of *cwnd* of a TCP connection. In such cases, we have

$w_p = (B + \mu\bar{\tau} + 1)$. Here, μ denotes the TCP packets transmission rate over the wireless link, and $\mu = 1/\bar{T}$. Let $\bar{K}$ represent the average number of TCP packets successfully transmitted when the size of *cwdn* increases from $w_p/2$ to w_p. We have $\bar{K} = 3w_p^2/8$. Then, the average TCP packet arrival rate at the AP can be derived by

$$\lambda = \bar{K} \Big/ \sum_{w_p/2}^{w_p} \bar{\tau} = (3w_p)/(4\bar{\tau}). \tag{5}$$

Secondly, if the packet losses related to wireless channel errors cannot be combated within M_r transmissions per frame over the wireless link, the size of *cwnd* is mainly decided by the detected packet losses [5]. Given the TCP packet loss rate q over the wireless link, the average number of packets successfully transmitted between two consecutively corrupted packets is $1/q$. For TCP Reno [9], once a corrupted TCP packet is detected, the fast retransmission mechanism is invoked, leading to halve the size of *cwnd*. Thus, the average number of packets transmitted when the *cwnd* decreasing from w_p to $w_p/2$ can be approximated to be $1/q$, then we have $w_p = \sqrt{8/(3q)}$. For a TCP connection over a general hybrid ARQ scheme, its TCP packet loss rate q in the case of maximum M_r transmissions per frame is calculated by $q = (\Pi Y)/[\Pi(X+Y)]$, based on the stationary state probability Π obtained from the above Markov chain. Substituting q into $w_p = \sqrt{8/(3q)}$, we have

$$w_p = \sqrt{[8\Pi(X+Y)]/(3\Pi Y)}. \tag{6}$$

Considering the effects of the above two factors (limited buffer size B and limited link layer retransmission M_r) on the *cwnd*, the average TCP packet arrival rate λ at the AP can be calculated from Eq. (2)–(6) as follows,

$$\lambda = \Big\{\frac{3}{4}\cdot\min\{\sqrt{[8\Pi(X+Y)]/(3\Pi Y)},\ [B+(D+\lambda E[T^2]/[2(1-\lambda\bar{T})])/\bar{T}+2]\}\Big\}\Big/\bar{\tau}\ . \tag{7}$$

How to solve for λ in Eq. (7) is a fixed point problem [10]. Rearranging Eq. (7) to a function of λ, i.e. $f(\lambda) = 0$, we see that $f(0) > 0$, $f(1/\bar{T}) < 0$ and $\frac{df}{d\lambda} < 0$ for $0 \le \lambda \le 1/\bar{T}$. It shows that there is a unique solution to Eq. (7). Thus, we can obtain the value of λ by fixed point iteration. Finally, the TCP level efficiency is obtained by $\lambda N L_f/C$.

To validate the analytical model, we simulate in OPNET [17] a TCP Reno [9] session with a saturated source in the system shown in Fig. 1. A simple hybrid ARQ which repeats the same codeword in all retransmissions is deployed in the wireless link. The received codewords failed in decoding are dropped. In such a case, given p (FER) and M_r, the TCP packet loss rate over the wireless link, denoted q, is obtained by $q = 1-(1-p^{M_r})^N$. Then, we can work out the analytical TCP throughput over such a hybrid ARQ from Eq. (7), where q of w_p is replaced by $1-(1-p^{M_r})^N$, $\bar{T}$ and $E[T^2]$ are calculated from Eq. (3) and Eq. (4) respectively. A comparison of the TCP throughput between the analytical results obtained from Eq. (7) and the simulation results is presented in Fig. 3, for the cases of $B = 5$, 20 TCP packets, $D = 100$ ms, $C = 2.048$ M bits/s and $M_r = 4$. Here, we assume that the values of FER p vary from 0.001 to 0.5. In these simulations, the radius of the 95% confidence intervals was kept below 1% of the values obtained. The results demonstrate strong agreements between the analytical and the simulation results. Additionally, for a given FER over the wireless link, the end-to-end TCP throughput can be improved by increasing the buffer size B at the AP, provided that the resultant queueing delay at the buffer will not violate the required QoS.

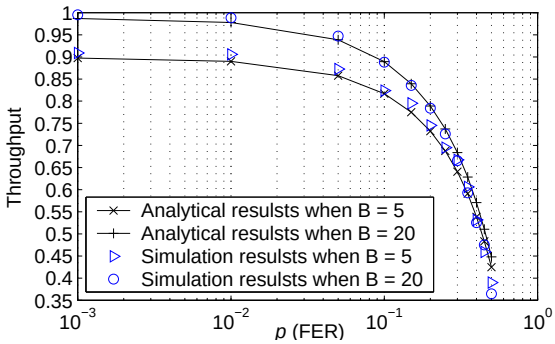

Fig. 3. TCP throughput by analysis and simulation.

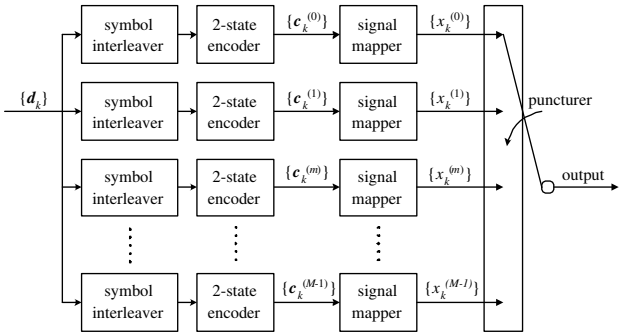

Fig. 4. The global structure of CT-TCM encoder.

3 Hybrid ARQ schemes

Using the model from the last section, we can compare the TCP throughputs over different hybrid ARQ schemes. Before we evaluate these performances, we first present several recently proposed hybrid ARQ schemes that adopt different coding algorithms to improve the spectrum utilization for possible applications in broadband wireless networks.

In hybrid ARQ schemes, type-II hybrid ARQ [11] can achieve efficient bandwidth utilization by sending parity-check bits in retransmission instead of repeating the same packet in all retransmissions as type-I hybrid ARQs do. However, it is difficult to combine a type-II hybrid ARQ with a bandwidth efficient trellis-coded modulation (TCM) scheme, because the information bits and redundancy bits of a TCM code cannot be separated. This nature of TCM codes confines their applications in hybrid ARQ schemes employing repetition method like type-I ARQs, e.g., the TCM-based schemes proposed in [12] and [13]. The scheme of [13] suggests to repeat the same TCM codeword in each retransmission but combine the multiple received versions of the original data packet for decoding at the receiver for better throughput performance. The scheme of [12] involves two different TCM codes that are alternatively used in retransmission. For the first transmission, the message is encoded by one TCM code. If the received copy contains uncorrectable errors, the message is encoded by the other TCM code for retransmission. After two transmissions, the scheme of [12] switches to the repetition mode like a type-I ARQ. Such a repetition method is not an optimized solution regarding transmission efficiency.

We proposed a new TCM-based hybrid ARQ for TCP applications over wireless in [14]. Our proposed scheme uses the concatenated two-state TCM (CT-TCM) algorithm [15] as the underlying coding algorithm to improve the bandwidth utilization over the wireless link. It overcomes the aforementioned limitation in the traditional TCM-based hybrid ARQ schemes by using different punctured parts of a CT-TCM code in retransmission, instead of repeating the same TCM code in all retransmissions as the schemes of [12] and [13] do. How to obtain the different punctured parts used in retransmissions is one key issue of our hybrid ARQ scheme in [14]. To answer this question, we first give an outline of the CT-TCM algorithm before we introduce the mechanism of our proposed hybrid ARQ. For more details of the CT-TCM algorithm, readers are referred to [15].

The CT-TCM algorithm comprises multiple component encoders concatenated in parallel by symbol interleavers. Fig. 4 shows the structure of a CT-TCM encoder which com-

Table 1. Puncture patterns in case of $M = 4$

Γ_0	Γ_1	Γ_2	Γ_3
1 0 0 0	0 0 1 0	0 1 0 0	0 0 0 1
0 1 0 0	0 0 0 1	0 0 1 0	1 0 0 0
0 0 1 0	1 0 0 0	0 0 0 1	0 1 0 0
0 0 0 1	0 1 0 0	1 0 0 0	0 0 1 0

prises M component encoders. Each component encoder consists of a binary two-state trellis encoder followed by a multi-ary signal mapper. Let $\{\boldsymbol{d}_k,\ k \geq 0\}$ be an input sequence of information symbols to the encoder. For component encoder m $(0 \leq m < M)$, a sequence of modulated symbols $\{x_k^{(m)}, k \geq 0\}$ is produced.

For spectral efficiency, the CT-TCM coding scheme performs puncturing on the modulated symbols of all component codes on the basis that one and only one modulated symbol carrying the same $\boldsymbol{d}_k$ is transmitted, and the punctured symbols are uniformly distributed in each component code. In terms of these two constraints, a number of puncture patterns that satisfy these two constraints can be found and each of them can be represented as an $M \times M$ indication matrix. Table 1 exemplifies the puncture patterns, denoted as $\Gamma_j, 0 \leq j \leq 3$, which satisfy the two constraints in the case of 8PSK-based CT-TCM with $M = 4$. For a selected puncture pattern Γ_j, the puncturer at the sender works as follows. At time instant k $(k = 0, 1, 2, ...)$, $x_k^{(m)}$ from component encoder m is selected for transmission if the $(k \bmod M, m)$th entry of Γ_j is 1, otherwise it is punctured.

Based on the feature of the CT-TCM codes, we propose a simple but efficient ARQ scheme. In the proposed scheme, different punctured parts of a CT-TCM code are used for possibly subsequent retransmissions instead of being dropped. Let an 8PSK-based CT-TCM code with $M = 4$ be used in our scheme, which involves a set of puncture patterns that satisfy the above two constraints. Let s denote an link layer frame to be transmitted. Let γ_1 denote the puncture pattern used for the first transmission of s.

Step 1: Initially, puncture pattern γ_1 is used to generate sequence s_1 for the first transmission of s. The punctured parts are stored in the sender's buffer for retransmissions of s. If the received sequence s_1' (s_1 plus channel noise) is successfully decoded, an ACK is returned to the sender. Otherwise, s_1' is stored at the receiver's buffer and an NAK is returned, then the ARQ scheme moves to Step 2.

Step i $(2 \leq i \leq M)$: Another puncture pattern γ_i $(\gamma_i \neq \gamma_{i-1})$, is used to select another sequence, denoted s_i, out of the punctured parts stored in the sender's buffer for the ith transmission of s. If there is no error detected after the received sequence s_i' is decoded, an ACK is returned to the sender. Otherwise, s_i' is combined with the previously received sequences s_1', s_2', ..., s_{i-1}' of s for decoding. If no error is detected in decoding the combination, an ACK is returned. Otherwise, s_i' is stored in the buffer and an NAK is returned. Then, the ARQ scheme moves to Step $i + 1$ if $i < M$, or returns to Step 1 and the link layer receiver buffer is cleared if $i = M$.

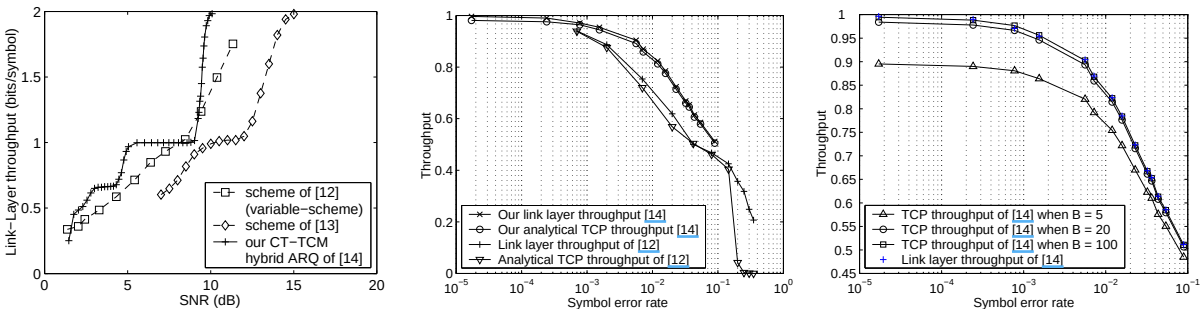

Fig. 5. Link layer throughput comparison.

(a) Under different SER. (b) Under different buffer size B.

Fig. 6. TCP performance comparison.

4 Comparisons of TCP performance over hybrid ARQs

Realizing that the link layer throughput provides the capacity limit of TCP over the wireless link, we plot in Fig. 5 the simulation results of the link layer throughput of our hybrid ARQ scheme as a function of the average SNR in an AWNG channel. The puncture pattern used for the ith transmission of a link layer data frame is given by $\gamma_i = \Gamma_{(i-1) \bmod M}$, where $i \geq 1$, $M = 4$ and $\Gamma_{(i-1) \bmod M}$ is given by Table 1. The link layer throughput is measured by the number of information-bits-per-symbol (bits/symbol) transmitted over the channel. For a 8PSK-based TCM scheme, the maximum information-bits-per-symbol is 2. For comparison, we also plot in Fig. 5 the link layer throughput provided by the TCM-based ARQ schemes in [12] and [13][3] over an AWGN channel. We can observe from Fig. 5 that significantly improved link layer throughput is achieved by our scheme over the other two schemes, except for the small SNR range (8.2 dB $\sim$ 9.2 dB) where the performance of our scheme is worse than that of [12]. However, such stair-like phenomenon of our throughput can be eliminated and the link layer throughput can be further improved by the modified method presented in [14].

Now, we compare the performances of TCP over the wireless link using the above-mentioned hybrid ARQ schemes based on the following parameters. One FTP application over TCP Reno [9] is running in the system shown in Fig. 1, with the wireless link capacity $C = 2.048$ M bits/s. Each TCP packet contains 8192 bits. RTT over the wired link D is 100 ms. B is set to 20 TCP packets unless it is specified otherwise. $M_r = 4$. Fig. 6(a) presents the analytical results of the TCP throughput obtained by our hybrid ARQ scheme as a function of SER over the wireless link with SNR ranging from 1 dB to 10 dB. For comparison, we include in Fig. 6(a) the TCP throughput obtained by the scheme of [12], derived from the SER results provided in [12]. Since the results in Fig. 5 have shown that the link layer performance of [12] significantly outperforms that of [13], the TCP performance of [13] is not considered here. From Fig. 6(a), we observe that higher TCP throughput is achieved by our scheme than by that of [12]. For reference, we also include the link layer throughput of both schemes in Fig. 6(a). It demonstrates that the difference between the TCP throughput of the scheme of [12] and its link layer counterpart is higher

[3] The results from [12], [13] are cited by converting their E_b/N_0 and E_s/N_0 coordinates into corresponding SNR coordinates.

than that of our scheme. The large difference in the scheme of [12] is mainly attributed to its higher link layer losses than our scheme's. As a result, a TCP over the scheme of [12] requires larger buffer size than $B = 20$ at the AP to store the data frames for local retransmissions due to more corrupted packets over the wireless link.

To clearly demonstrate the effect of buffer size B at the AP, we plot in Fig. 6(b) the analytical results of the TCP throughput over our hybrid ARQ scheme under different values of B, where the SER coordinates correspond to the range of SNR from 1 dB to 10 dB. The link layer throughput of our scheme is also included in this figure as a reference. Similar to the conclusion obtained from Fig. 3, the end-to-end TCP throughput can be improved by increasing the buffer size B at the AP. It is consistent with the insights obtained from Eq. (7). However, we also observe that as the buffer size B increases from 20 packets to 100 packets, no significant improvement on TCP throughput is achieved. The reason for this effect is that larger buffer size incurs longer queueing delay to a newly arrived TCP packet at the AP and increases the end-to-end RTT for TCP packets, see Eq. (2). As a result, the TCP throughput cannot be further improved due to longer RTT under larger buffer size. Instead, it reaches the maximum link capacity that provides fundamental limit to TCP throughput. The above discussion is based on the assumption that a TCP source can accurately estimate its RTT and uses the results to update its timeout threshold. Otherwise, large buffer size leads to undesirable longer packet delay and even timeouts [16], which seriously degrade the end-to-end TCP throughput. These results provide arguments against the use of large buffers as the small benefit obtained in TCP throughput may not justify the longer RTT.

Finally, we verify our analytical results by simulations. To this end, we have developed a simulation model comprising a TCP Reno connection over our proposed hybrid ARQ scheme in OPNET [17]. The comparisons between the results from analysis and simulation are presented in Table 2, for the cases of $B = 5$, 20 TCP packets respectively. In the analytical model, the values of g_j and e_j fed into the Markov chain are obtained by the simulation of the CT-TCM decoding algorithm [15]. The comparisons again demonstrate strong agreements between the analytical and the simulation results.

Table 2. TCP throughput obtained by analysis and simulation.

Throughput v.s. SNR (dB)	1	2	3	4	5	6	7	8	9	10
Analytical results (B = 5)	0.5×10^{-5}	0.12	0.314	0.32	0.46	0.475	0.475	0.475	0.485	0.905
Simulation results (B = 5)	1.12×10^{-4}	0.09	0.318	0.324	0.462	0.48	0.48	0.48	0.49	0.92
Analytical results (B = 20)	0.5×10^{-5}	0.124	0.324	0.330	0.479	0.495	0.495	0.485	0.507	0.982
Simulation results (B = 20)	1.15×10^{-4}	0.109	0.323	0.329	0.478	0.494	0.494	0.494	0.505	0.979

5 Conclusion

We have developed a Markov chain model to study TCP performance over the wireless link where hybrid ARQ schemes are used to combat wireless channel errors. Based the markov model, we have considered a recently proposed new TCM-based hybrid ARQ scheme that overcomes the limitation of the traditional TCM-based hybrid ARQ scheme

by using different punctured parts of a CT-TCM code in retransmission. We demonstrated by both analysis and simulation that significantly improved end-to-end TCP throughput can be achieved by adopting the CT-TCM hybrid ARQ in a wireless link. In addition, we have observed that an optimal buffer size exists above which no significant improvement on TCP throughput can be achieved.

References

1. Xylomenos, G. & Polyzos, G. C. & Mähönen, P. & Saaranen M. .: TCP performance issues over wireless links. IEEE Commun. Mag., vol. 39, pp. 52–58, April 2001.
2. Bakre, A. & Badrinath, B. R. .: I-TCP: indirect TCP for mobile hosts. Proceedings of 15th International Conf. on Dist. Comp. Sys., pp. 136–143, May 1995.
3. Ramakrishnan, K. & Floyd, S. .: The Addition of Explicit Congestion Notification (ECN) to IP. RFC 3168, Sept. 2001.
4. Balakrishnan, H. & Padmanabhan, V. N. & Seshan, S. & Katz, R. H. .: A comparison of mechanisms for improving TCP performance over wireless links. IEEE/ACM Trans. Networking., vol. 5, no. 6, pp. 756–769, Dec. 1997.
5. Chaskar, H. M. & lakshman, T. V. & Madhow, U. .:TCP over wireless with link level error control: analysis and design methodology. IEEE/ACM Trans. Networking., vol. 7, no. 5, pp. 605–615, Oct. 1999.
6. Chiasserini, C. F. & Meo, M. .: A reconfigurable protocol setting to improve TCP over wireless. IEEE Trans. Veh. Tech., vol. 51, no. 6, pp. 1608–1620, Nov. 2002.
7. Morris, R. & Lin, D. .: Variance of aggregated web traffic. Proceedings, IEEE INFOCOM 2000, Tel Aviv, Israel, 2000. pp. 360–366, Mar. 2000.
8. Kleinrock, L. .: Queueing systems. vol. 1. New York: Wiley, 1976.
9. Stevens, W. R. .: TCP/IP Illustrated. vol. 1. Reading, MA: Addison- Wesley, 1994.
10. Vu, H. L. & Hanly, S. .: A Study of TCP Performance and Buffer Occupancy Over a FadingWireless Link. Proceedings, IEEE GLOBECOM'01, vol. 6, pp. 3478–3482, Nov. 2001.
11. Lin, S. & Yu, P. S. .: A hybrid ARQ scheme with parity retransmission for error control of satellite channels. IEEE Trans. Commun., vol. COM-30, pp. 1701–1719, July 1982.
12. Yu, J. X. & Li, Y. & Murata, H. & Yoshida, S. .: Hybrid-ARQ scheme using different TCM for retransmission. IEEE Trans. Commun., vol. 48, pp. 1609–1613, Oct. 2000.
13. Schmitt, M. P. .: Hybrid ARQ scheme employing TCM and packet combining. IEE Electronics Letters, vol. 34, no. 18, pp. 1725–1726, Sept. 1998.
14. Huang, Q. & Chan, S. & Li, P. & Ko, K. T. & Zukerman, M. .: A Hybrid ARQ Scheme Based on CT-TCM Codes. To be published in IEEE Commun. Lett., July 2005.
15. Li, P. & Bai, B. & Wang, X. M. .: Low complexity concatenated two-state TCM schemes with near capacity performance. IEEE Trans. Inform. Theory, vol.49, pp. 3225–3234, Dec. 2003.
16. Appenzeller, G. & Keslassy, I. & McKeown, N. .: Sizing Router Buffers. Proceedings, SIGCOMM'04, pp. 281–292, 2004.
17. OPNET, www.opnet.com.