TCP Over Packet Radio Link With Adaptive Channel Coding

Dongjie Huang and James J. Shi

Ericsson Inc.
740 East Campbell Road MP-8, Richardson, TX 75081, USA
E-mail: dongjie.huang{james.shi}@ericsson.com

Abstract – A wireless packet data communications system that uses adaptive channel coding on the radio link is studied. A channel-coding adaptation mechanism based on recommendation from the receiver is proposed. A traffic model that exhibits self-similarity is considered for packet data applications. The radio channel is modeled as a Rayleigh fading channel. The performance of TCP over such an adaptive radio link is obtained through analysis and simulation. The result indicates significant improvement in throughput and delay performance for TCP over adaptive packet radio link. To gain insight to dimensioning of wireless packet networks, TCP throughput and delay are studied with respect to the number of users and file size. System utilization is analyzed and theoretical bounds are derived and compared with simulation results.

I. Introduction

Performance of Transmission Control Protocol (TCP) over wireless links can easily degrade due to handoff, high bit error rate and long roundtrip delay on the air interface. The TCP congestion control measures would cause unnecessary reduction of throughput in wireless networks [1]. Radio Link Protocols (RLPs) have been used in circuit-switched data to alleviate the effect of non-congestion-related loss over wireless links [2][3].

It has been shown [4] [5] that adaptive channel coding can significantly improve end-to-end TCP performance over fixed channel coding in a circuit-switched wireless data system. For bursty data, packet transmission is preferred because it utilizes radio resource more efficiently. General Packet Radio Service (GPRS) is an example of wireless packet data communications [6].

In this paper, we investigate end-to-end TCP performance over packet radio link with adaptive channel coding. We study the TCP user throughput and transmission delay with internet traffic and dynamic radio channel. To accomplish this, we consider a traffic model that exhibits self-similarity. We model radio channels as Rayleigh fading channels in terms of bit error rate. The main interest is to study performance improvement from adaptive channel coding on packet data services at various radio conditions. The study also analyzes the TCP performance with respect to system load and traffic profile.

The rest of the paper is organized as follows. Section II reviews basic concepts of adaptive channel coding and describes a protocol model, a fading channel model and a traffic model for wireless data. Section III proposes an adaptation scheme to update the channel code. Section IV presents theoretical analysis of system load, user throughput and delay. Section V presents simulation results. Finally, Section VI concludes the paper.

II. Basic Concepts

In this section, we briefly review adaptive channel coding and then describe a protocol model for TCP over packet radio. We also discuss a Markov model for Rayleigh fading channel and an internet traffic model.

II.1 Adaptive Channel Coding

There are two methods to adapt channel coding to channel condition. The first method starts with the least redundant code and transmits more redundant bits and/or retransmits as needed until the information can be decoded error free at the other end [4]. Punctured convolutional codes can be used. The second method changes the channel coding used in both transmission and retransmission based on the current channel condition [5]. The punctured convolutional codes can also be used in the second method. We consider the second method in our study.

When a frame is detected to contain errors after decoding, it is discarded and an acknowledgment is sent back to the sender requesting a retransmission of the frame. This is called selective automatic retransmission request (ARQ) and is considered in our study.

Probability of retransmission is the probability that the frame contains at least one uncorrectable error. The latter is called frame error rate (FER). FER can be calculated from probability of bit error, $P_b(E)$, of the decoded frame. Assuming a frame length of l, we have

$$FER = 1-(1-P_b(E))^l . \qquad (1)$$

It follows that the relative throughput (T_r) of a data communications system using a rate-R convolutional code and selective ARQ can be written as

$$T_r = R(1-FER) = R(1-P_b(E))^l . \qquad (2)$$

II.2 TCP Over Packet Radio

A simplified model of protocol stack for packet data communications is shown in Fig. 1. In GPRS, TCP packets are first added with IP, Subnetwork Dependent Convergence Protocol (SNDCP) and Logical Link Control (LLC) related overhead and then forwarded to Radio Link Control and Medium Access Control (RLC/MAC) layer. RLC/MAC layer adds its own overhead and then forwards it to the physical layer supported by radio frequency. For this model, we assume the Round Trip Time (RTT) between the base station (BS) and TCP server is negligible compared to the time spent on the radio link. The GPRS protocol generates relatively fixed overhead, so SNDCP and LLC layers are not included in the model. The RLC layer uses selective ARQ, error correction and detection to provide reliable data transmission. When sending data, RLC uses an NAK-based selective repeat to eliminate detectable errors. Additionally, the RLC is enhanced with adaptive channel coding. Section III provides detail on the selection of adaptive codes. The physical layer below RLC uses packet data traffic channel (PDTCH).

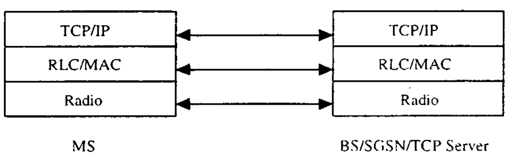

Fig. 1. Model for TCP/IP over GPRS.

II.3 Markov Channel Model

We consider a two state Markov model [7] for representing a wireless channel. In Fig. 2, state 0 represents the quiet state in which the BER (ε_0) is low and single bit errors are produced independent of other errors. State 1 represents the noisy or burst state, where the BER (ε_1) is high. The following parameters are derived from the model, the average burst length (b), the average BER ($\bar{\varepsilon}$) and the duty cycle (d).

$$b = \frac{1}{1 - p'} \tag{3}$$

$$\bar{\varepsilon} = \frac{(1 - p')\varepsilon_0 + p\varepsilon_1}{(1 - p') + p} \tag{4}$$

$$d = \frac{p}{(1 - p') + p} \tag{5}$$

As shown in [7], we take $\varepsilon_0 = d\bar{\varepsilon}$. If b, $\bar{\varepsilon}$ and d are defined, we can calculate all the other parameters of the model using the four equations. In this work, we discuss the case of burst channel which has an average BER ($\bar{\varepsilon}$) of 1.8% and a duty cycle (d) of 0.28. This translates to $\varepsilon_0 = 0.5\%$ and $\varepsilon_1 = 5\%$. The same analysis can be done for other kinds of channels. The minimum time for the channel to stay in the same state is determined by the data frame duration (for example, 20 ms).

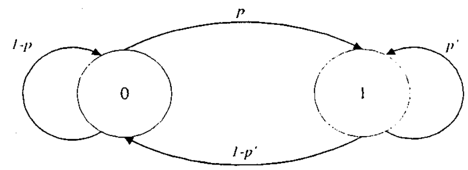

Fig. 2. Two-state Markov channel model.

II.4 Wireless Internet Traffic Model

Recent research demonstrated that Poisson process and its variations do not capture some of the important characteristic of internet date traffic. For example, the duration of the ON and OFF periods are found to have a heavy-tail distribution [8]. The heavy-tail distribution leads to the self-similar behavior of the aggregated traffic. As the wireless internet becomes more popular, it is important that the underlying traffic model capture the unique characteristic of packet date traffic.

For the analysis and simulation in this paper, a WWW traffic model that exhibits the ON/OFF pattern is considered. A similar model has been described in [9].

The ON period represents the file transmission time on the downlink. The OFF period represents the interval when there is no data transmission. The ON time is a function of the file size, available system bandwidth and the number of simultaneously transmitting users. The file size follows a Pareto distribution of the first kind

$$F(x) = 1 - \left(\frac{k}{x}\right)^{\alpha} \tag{6}$$

where the shape parameter $\alpha = 1.5$ and the minimum file size $k = 4000$ bytes. This leads to an average file size of 12000 bytes. These values are taken from traditional internet traffic and may need adjustment when new wireless internet technologies are available.

The OFF period follows an exponential distribution with a mean idle time of 90 seconds. The default parameters used in this paper are summarized in Table 1.

Table 1 Traffic Parameters

Number of PDTCHs	8
Max. number of PDTCHs a terminal can access	8
WWW file size	Pareto with $\alpha = 1.5$ and $k = 4000$
Idle time	Exponential with mean 90 seconds

III. Channel Adaptation

We consider a set of punctured convolutional codes derived from the rate 1/2 convolutional code. They are rate 1/2, 2/3, and 3/4, respectively. For comparison, we also include the uncoded case. They correspond to the four coding schemes (CS-1 through CS-4) defined in the GPRS standard. The length of the coded frame is 456 bits, also from the GPRS standard. Fig. 3 shows the normalized throughput of these codes with respect to the BER. It is seen that each code offers optimum performance for a certain range of BER. Considering the complexity, we use the combination of rate 3/4 and 1/2 codes. The corresponding adaptive algorithm is to use rate 3/4 code for BER in [0, 0.029] and rate 1/2 code for BER in [0.029, 0.1], respectively.

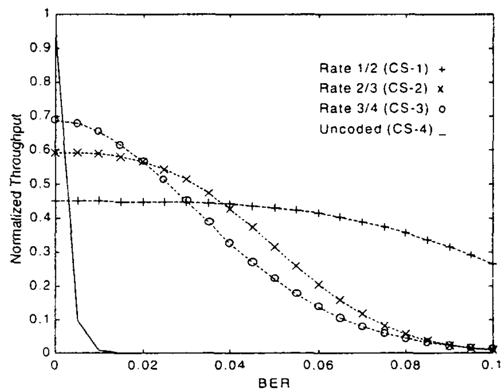

Fig. 3. Throughput with respect to the BER.

Most wireless systems monitor channel condition through channel quality feedback. In general, channel feedback includes received signal strength, BER and FER. As an example, the IS-136 standard defines a channel quality measurement scheme in which the mobile station calculates bit errors in every frame and reports back to the base station the moving average of BER every second.

Obviously, the existing channel feedback mechanism introduces a large amount of transmission overhead when the feedback has to be sent frequently to adapt the coding quickly to the changing radio condition. To reduce the overhead, we propose a new receiver-based channel adaptation scheme (patent pending) as follows. In stead of sending channel quality information back to the transmitter and letting the transmitter calculate the appropriate channel coding, the receiver calculates and sends the recommended coding to the transmitter. The receiver also filters the channel condition to eliminate unnecessary changes of coding due to fading that lasts only a small duration such as a frame length. The transmitter simply uses the recommended coding.

IV. Throughput and Delay Analysis

In this section, we analyze throughput and delay for packet radio. We consider two packet arrival scenarios, corresponding to the highest and lowest possible system load for a given user application traffic profile.

In general, after downloading a file a user will stay idle for a period of time before initiating another download. Following assumptions on mobile station (MS)'s access to a radio channel are made in the analysis. An MS can access any number of (up to eight for GPRS) time slots of a downlink channel and MS are all in the same class. That is, if only one MS is receiving at a time, the MS will occupy all eight time slots. If more than one MS are receiving at a time, they will share the time slots in a round-robin manner.

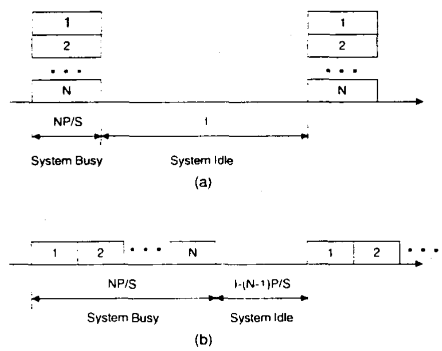

Fig. 4. Packet arrival scenarios.

Let us consider a packet arrival scenario as shown in Fig. 4(a). The packets from N users arrive at the base station's transmitter simultaneously with same average packet size of P (bits). The system then transmits these packets to MS's at the same rate and all packets are delivered at the same time. After delivering the packets, the system stays idle until the next batch of packet arrival. If the system has a throughput of S and the packet arrival has an interval of I, it will be busy for an interval of $B = NP/S$ and idle for $T_1 = I$. The system load U_1, user throughput S_1 and file transmission delay D_1 in this scenario can be expressed as

$$U_1 = \frac{B}{T_1 + B} = \frac{NP}{SI + NP} \tag{7}$$

$$S_1 = \frac{S}{N} \tag{8}$$

$$D_1 = \frac{NP}{S} \tag{9}$$

In the other scenario, packet arrivals from the N users are sequential and do not overlap with each other during packet transmission. It means that at most one packet is served by the system at any time as shown in Fig. 4(b).

The system busy time (B) in the cycle remains as NP/S. Between ending of first user's packet and beginning of its next arrival, there are (N-I) packets transmitted separately. The system idle time in this cycle is $T_2 = I - (N-1)P/S$. The system load, user throughput and transmission delay can be expressed as

$$U_2 = \frac{B}{T_2 + B} = \frac{NP}{SI + P} \tag{10}$$

$$S_2 = S \tag{11}$$

$$D_2 = \frac{P}{S} \tag{12}$$

The two scenarios actually represent the worst and best cases. In the first scenario (worst), all users have their packet arrivals at the same time thus have lowest throughput and longest delay. In the second scenario (best), all users have arrivals at different times thus have highest throughput and shortest delay. From the system point of view, the first scenario carries the minimum load and the second maximum.

As the number of users and file size increase, the system load increases approximately linearly and the user throughput degrades from S to S/N. The transmission delay increases linearly with file size.

V. Simulation Results

In this section, we present simulation results of throughput and delay with adaptive channel coding for wireless packet network, and compare the system load from simulation with bounds from previous analysis.

The simulation is based on the model shown in Fig. 1. The TCP algorithms by Jacobson [10] and Karn [11] are implemented. The TCP behavior in dynamic radio environment is simulated, and throughput and delay are evaluated for the TCP link in an established state. The forward error correction is enhanced with adaptive coding. The system throughput of eight time slots is 112 kbps for 3/4 coding and 60.8 kbps for 1/2 coding. The channel fading rate is represented by the average burst length in the two-state Markov model. The average burst length used in the simulation is 500 ms, equivalent to 25 frames. Other parameters for the channel model are specified in Section II.3.

Fig. 5 compares the TCP user throughput with respect to the number of users for adaptive coding and fixed 3/4 and 1/2 coding. The default minimum file size of 4000 bytes is used. The result indicates that adaptive coding outperforms both fixed coding schemes. Fixed 3/4 coding has a higher throughput than fixed 1/2 coding and is close to adaptive coding for the current parameter setting. The positions of the two fixed coding schemes will change if the channel is made noisier by increasing the duty cycle of the noisy state. The throughput degrades approximately linearly with increasing number of users. Fig. 6 compares the file transmission delay with respect to the number of users for adaptive coding and fixed 3/4 and 1/2 coding. Among all three coding schemes, adaptive coding has the least transmission delay. It is seen that for all three coding schemes, the delay is small for small number of users (from one to sixteen). With small number of users, the possibility of multiple files arriving at the same time is small, therefore the system resource is dedicated to a single file, and the delay is only dependent on the file size.

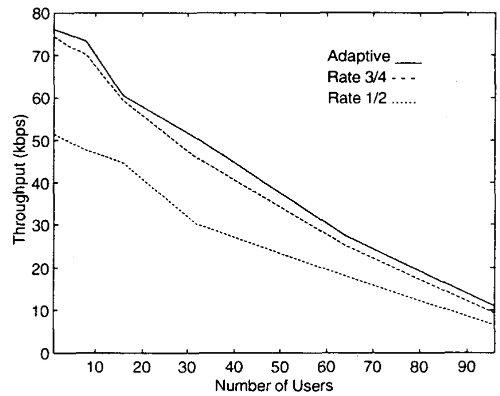

Fig. 5. User throughput with respect to number of users.

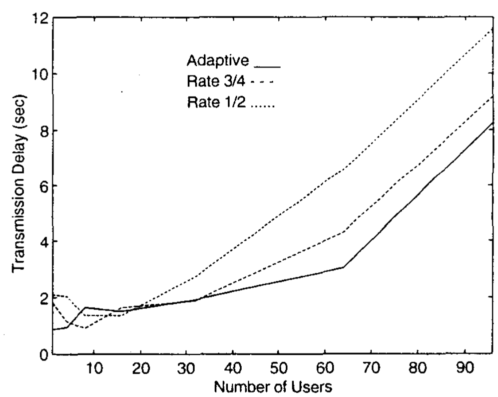

Fig. 6. Delay with respect to number of users.

Now let us consider a fixed number of users and variable file size. Only adaptive coding is considered. Figs. 7 and 8 show the user throughput and delay with respect to the file size for 16 users. For small file sizes, the user throughput is relatively constant, close to system throughput. The throughput is reduced for very small file size because of the relatively large overhead. As the file size increases, the throughput decreases because the load increases and multiple users share the radio channel. If the file size continues to grow, the throughput approaches the minimum given by (8).

It is seen in Fig. 8 that the delay increases approximately linearly with the file size. The slope is smaller at the beginning and increases with the file size. The delay is bounded by (9) and (12).

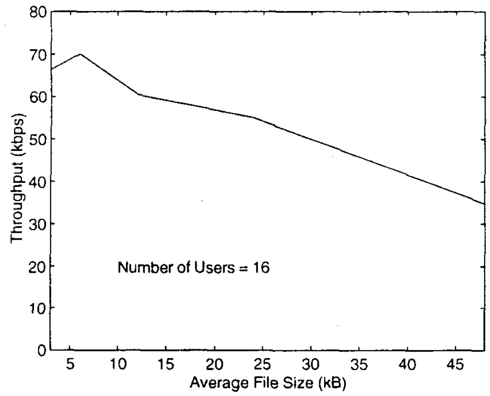

Fig. 7. User throughput with respect to file size.

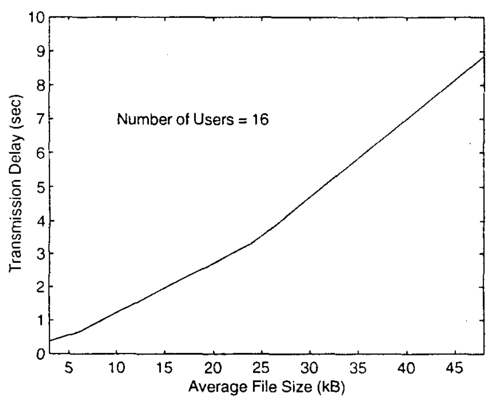

Fig. 8. Delay with respect to file size.

Table 2 Performance tradeoff between the number of users and file size (with load kept at the same level)

N_{user}	4	8	16	32	64	128
P_{min} (kB)	16	8	4	2	1	0.5
U_{upper}	0.17	0.17	0.18	0.18	0.18	0.18
U_{lower}	0.15	0.15	0.15	0.15	0.15	0.15
$U_{simulation}$	0.17	0.17	0.17	0.16	0.17	0.16
T (kbps)	71	71	57	63	64	57
D (sec)	4.8	2.0	1.3	0.64	0.34	0.18

Table 2 shows the balance between the number of users and application file size. The load is kept at the same level. It shows that doubling the number of users while reducing the file size in half does not change the system load and user throughput much but reduces the delay significantly. Overall, the throughput degrades slightly for smaller file sizes, possibly due to increased overhead. The result also indicates that the simulated system load fits in the upper and lower bounds.

VI. Conclusions

We have studied the performance of TCP over packet radio. It is shown that adaptive channel coding can significantly improve TCP performance. This approach does not require any modification at the TCP layer. Using a Markov fading channel model and an internet traffic model, user throughput and delay with different load conditions and user traffic profile have been simulated. The results give insight to dimensioning of wireless packet networks. Packet data transmission mechanisms that incorporate quality of services are currently under study.

VII. References

[1] V. K. Varma et al., "Architecture for Interworking Data with PCS," IEEE Commun. Mag., vol. 34, no. 9, Sept. 1996, pp. 124-130.

[2] TIA IS-130-A, TDMA Wireless Systems – Radio Interface – Radio Link Protocol 1, July 4, 1997.

[3] H. Balakrishana et al., "A Comparison of Mechanisms for Improving TCP Performance over Wireless Links," IEEE/ACM Trans. on Networking, Dec. 1997, pp. 756-769.

[4] S. Falahati and A. Svensson, "Hybrid Type-II ARQ Schemes for Rayleigh Fading Channels," In Proc. International Conference on Telecommunications, Greece, June 1998, pp. 39-44.

[5] D. Huang and J. Shi, "TCP with Adaptive Radio Link," 9th Virginia Tech Symposium on Wireless Personal Communications, Blacksburg, VA, June 1999.

[6] S. Hoff, M. Meyer and A. Schieder, "A Performance Evaluation of Internet Access via the General Packet Radio Service of GSM," IEEE VTC'98, Ottawa, May 1998.

[7] S. Yajnik, J. Sienicki and P. Agrawal, "Adaptive Coding for Packetized Data in Wireless Networks," in Proc. PIMRC'95, pp. 338-342.

[8] V. Paxson and S. Floyd, "Wide Area Traffic: The Failure of Poisson Modeling," IEEE/ACM Trans. on Networking, vol.3, no.3, June 1995.

[9] J. Ho, Y. Zhu and S. Madhavapeddy, "Throughput and Buffer Analysis for GSM General Packet Radio Services (GPRS)," WCNC99, 1999.

[10] V. Jacobson, "Congestion Avoidance and Control," in Proc. ACM SIGCOMM 88, Aug. 1988.

[11] P. Karn and C. Partridge, "Improving Round-Trip Estimates in Reliable Transport Protocols," in Proc. ACM SIGCOMM 87, Aug. 1987.