Comprehensive Performance Analysis of a TCP Session Over a Wireless Fading Link with Queueing

By
Alhussein A. Abouzeid
Summit Roy
Murat Azizoglu
IEEE Transactions on Wireless Communications
Vol. 2 No.2 March 2003

Presented By
Eswin Anzueto

WPI

Outline

- Introduction
- A Model for TCP Congestion Avoidance
- Channel Model
- Analysis
- Model Validation and Discussion
- Conclusion

WPI

Introduction

- TCP is currently the most widely used transport protocol in packet networks, and is largely responsible for end to end congestion control.
- This paper provides a comprehensive performance (goodput) analysis of a single TCP connection for a mobile host/client communicating with a server in a backbone (wired).
- TCP goodput can be limited due to losses both at the wireless (last) hop, as well as the network backbone due to congestion.
- Literature on TCP modeling to date has typically concentrated on one of the two loss mechanisms:
  - Channel loss: Errors due to the wireless channel
  - Congestion loss: Packet drops at the queues or buffers

WPI

Introduction (cont.)

- This papers provides a framework where both losses can be treated as is appropriate for wireless Internet access where the channel loss due to (time-varying) multi-path fading on the wireless hop can be dominant (or comparable to) congestion losses.
- A useful abstraction for wireless loss scenarios is a bulk TCP transfer over a single link between a source/destination pair that is subject to channel-induced packet loss between TCP sender and receiver.
- On the other hand the typical abstraction used for congestion loss is equivalent per flow random loss model that randomly drops packets at the buffer
- Lakshman and Madhow work provides a complete analytical description of the TCP congestion window evolution over ideal Channels with finite buffer size, this work provides a point of departure for our work.

Introduction (cont.)
Unique contributions from this work

1. Channel Driven Model for Packet Loss:

   Previous wireless channel models for packet loss implicitly assume that the TCP sources are continuously transmitting and ignore the fundamental bursty nature of TCP.

   This is tie to the alternating two state channel model currently implemented in ns-2 where the state evolution is clocked by packet transmissions. (link does not advance in time when TCP sender is idle). This contradicts the actual physical behavior since channel state evolution is not predicated by packet transmissions.

   Previous TCP wireless channel models impose a distribution on the number of consecutive packets drops in a bad state regardless of the time duration between the transmission of the packets. This assumption is only valid for continuous packet transmissions. (TCP is a bursty protocol)

   This paper present a stochastic channel-driven model of TCP over wireless links that does not make any prior assumptions on the packet level loss statistics, this instead are derived from the underlying physical channel behavior.

Introduction (cont.) Unique contributions from this work

2. Modeling the Queue Behavior:

Previous wireless loss models neglect the effect of the queue behavior, all previous correlated loss models consider the case of very high wireless channel loss rates and hence neglect queue losses.

Our model accounts for both of the above two characteristics; it allow the RTT to be a function of the window size (rather than fixed) and it model both types of packet loss, queue loss (buffer overflows) and channel (wireless) losses.

Introduction (cont.) Unique contributions from this work

3. A unified Model for Various Systems Parameters

Our model is applicable to a wider range of network parameters than previous related work.

"Comparative performance Analysis of versions of TCP in a local network with a lossy link" neglects the effects of possible queueing losses.

"Comparative performance Analysis of versions of TCP in a local network with a lossy link part II" assumes a buffer with infinite capacity.

An finally our analysis results are validated against the public-available network simulator, contrary to other models that used a custom made simulator for validation.

Outline

- Introduction
- A Model for TCP Congestion Avoidance
- Channel Model
- Analysis
- Model Validation and Discussion
- Conclusion

A model for TCP Congestion Avoidance

This section introduces our abstract system model, the model TCP congestion avoidance, previous well known analytic characterization of TCPs behavior for ideal channels and our model for TCPs bursty nature

A. System Model

Consider a mobile host connected to a host in the wired backbone via last-hop wireless link. (wireless link constitute the bottleneck on the path and the rest of path connection can be modeled as a constant delay)

- γ = round trip time
- μ = wireless link raw capacity in packets/s
- $T = \gamma + 1/\mu$ = Time (seconds) between the start of the packet transmission and reception of ACK (excluding queue delay in buffer)
- μT = Bandwidth-delay product
- $\beta = B/(\mu T)$ normalized buffer size by the bandwidth - delay
- Wp = maximum number of packets possibly in transit
- $Wp = B + \mu T = \mu T(\beta + 1)$

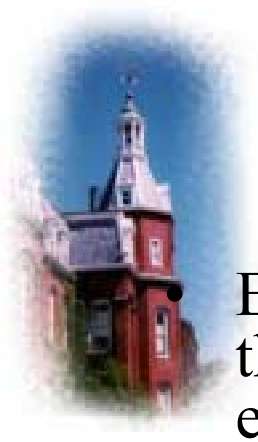
WPI

A model for TCP Congestion Avoidance

B. TCP Operation in Ideal Channels

The window-base congestion avoidance mechanism in TCP/IP acts as a self clocking regulator based on receiver Feedback. (Assumption: reader familiar with TCP Reno) It is well known that for ideal channels (i.e. no random packet loss) TCP exhibits a periodic evolution. Let t` = 0 denote the time of establishment of the TCP-Reno session under consideration. W(t`) and W_{th}(t`) represent the congestion window size and the slow start phase threshold at time t`and Δ(t`) is the current timeout value. The main obsevations for TCP Reno window evolution can be summarized as follows.

WPI



TCP Operation in Ideal Channels(cont.)

- Except for an initial slow start phase at the beginning of the session, the congestion window size shows a periodic evolution. Each cycle starts with window size W(t\`) = $W_p/2$ and continues in congestion avoidance phase until a buffer overflow at W_p, whereupon the window is halve to $W_p/2$
- Assuming fast recovery and retransmit option implemented, no delayed ACKs, and sufficient large window size, packet(s) lost due to a buffer overflow is (are) detected via Triple duplicate or TD. Hence no TOs take place, and the TCP Reno session remains in congestion avoidance during the life time of the session.
- In congestion avoidance, the window increase is either sub linear if $W_p/2 > \mu T$, or a combination of linear followed by a sub linear increase if $W_p/2 < \mu T$

TABLE I

DEFINITIONS OF VARIOUS PARAMETERS USED FOR COMPUTING THE REWARD AND TRANSITION PROBABILITIES

Parameter	$H = A$	$H = B$
$t(w_I, w_F, H)$	$T(w_F - w_I)$	$\frac{w_F^2 - w_I^2}{2\mu}$
$t_n(w_I, H)$	$T(\sqrt{w_I^2 + 2n} - w_I)$	$\frac{n}{\mu}$
$n(w_I, w_F, H)$	$\frac{w_F^2 - w_I^2}{2}$	$\frac{w_F^2 - w_I^2}{2}$

TABLE II

DURATION AND NUMBER OF PACKETS SUCCESSFULLY TRANSMITTED DURING THE TYPICAL CYCLE

Parameter	$\beta < 1$	$\beta > 1$
t_A	$t(w_p/2, \mu T, A)$	0
t_B	$t(\mu T, w_p, B)$	$t(w_p/2, w_p, B)$
N_A	$n(w_p/2, \mu T)$	0
N_B	$n(\mu T, w_p)$	$n(w_p/2, w_p)$

TCP Operation in Ideal Channels(cont.)

C. TCP Burst Model

For clarity of presentation we make the following definitions

Def. 1 – ***Window Round***: Round for a window of size W_1 is the time elapsed between the congestion window successive increments for $W_1 + 1$. If μ is the packet transmission time, then the duration of the round during the congestion avoidance phase A is equal to $\gamma + 1/\mu$ and congestion avoidance phase B is equal to $(W_1)/\mu$.

Def. 2 – ***Busy Period of a Round***: A busy period of a Round W_1 is the time duration from the beginning of the round, during which TCP sender is busy transmitting packets; equivalently the duration when all the packets during this round are transmitted.

TCP Operation in Ideal Channels(cont.)

Def. 3 ***Idle Period of a Round:*** In a window round W_1, if $W_1 < \mu T$ then an idle period of duration $T - W_1/\mu$ terminates the round.

Outline

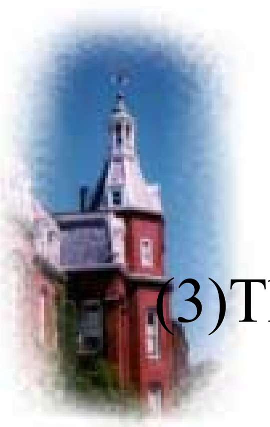
- Introduction
- A Model for TCP Congestion Avoidance
- Channel Model
- Analysis
- Model Validation and Discussion
- Conclusion

Channel Model

The wireless channel is modeled by continuous time, two state alternating process {S(t`)|t`≥0} Taking values k=0 (good) and 1 (bad) with properties that:

(1) When the channel leaves one state, it will enter the other state with probability 1 (alternating process)

(2) The durations of time in good state denoted $\{X_i, i = 1,2,\ldots\}$ are i.i.d. with known exponential cumulative distribution function $F_x(x)$ and mean $E[X_i] = 1/\lambda_0$



Channel Model(cont.)

(3)The durations of time in bad state denoted $\{Y_i, i = 1,2,\ldots\}$ are i.i.d. with known exponential cumulative distribution function $F_Y(y)$ and mean $E[Y_i] = 1/\lambda_1$

(4) In each of the states, the packet loss mechanism is that of a discrete memory less (packet) channel, with respective loss probabilities p_0 and p_1. For simplicity in this work we assume that

$p_0 = 0$ and $p_1 = 0$

Note that a difficulty with the above model is that it Cannot be directly translated to packet loss statistics. For example the average good duration $E[X_i]$ does not correspond to the average successful packet transmission (similarly for the average bad duration).

Channel Model(cont.)

- For future reference, a (long-term) fraction of the time the channel spends in the bad state f is given by

$$f = \frac{E[Yi]}{E[Xi] + E[Yi]}.$$

Outline

- Introduction
- A Model for TCP Congestion Avoidance
- Channel Model
- Analysis
- Model Validation and Discussion
- Conclusion

Analysis

- We start the analysis by identifying the key characteristics of TCP behavior as they relate to an underlying lossy channel motivated by ns-2 simulation traces. Subsequently, we describe a semi-Markov process that closely model this TCP behavior within a number of well motivated approximations. A Markov renewal-reward approach is then applied toward the deriving steady state window distribution and goodput.

Key Characteristics of TCP Over a Channel-Driven Loss Model

- Correlated loss channels are characterized by a finite bad state duration; consequently, packet loss can occur in burst as well as in isolation.
- Channel visits to the good state during which TCP transmits packets are sample good state.
- Channel visits during which TCP does not transmit packets are unsample good state.
- A bad state visit may result in a TO, TD or not packet loss.
- In addition to channel induced losses, buffer overflow may also take place if the window reaches Wp.

Key Characteristics of TCP Over a Channel-Driven Loss Model (Cont.)

- Most of the above characteristics are evident in the ns-2 simulation trace in Figure 1. Initially the channel is in a (sample) good state and TCP is successfully transmitting packets. A sufficiently long bad state duration follows causing a loss of enough packets to trigger a TO and, hence a retransmission of these packets. However, these packets are also lost since the channel remains in the bad state. Notice that TCP has to wait for the next TO before attempting to retransmit again. During this time, a channel visit to the good state may not be detected resulting in an un-sampled good state. In the simulation trace in Fig. 1 three good states are un-sample, the next good state has sufficient duration to be sample, and TCP enters again into congestion avoidance with an initial window size equal to one. The bad state visit labeled visit 1 is of a very short duration and results in a short bust of packet losses which is detected by a TD. The subsequent bad visit (visit 2) is long enough to cause a TO

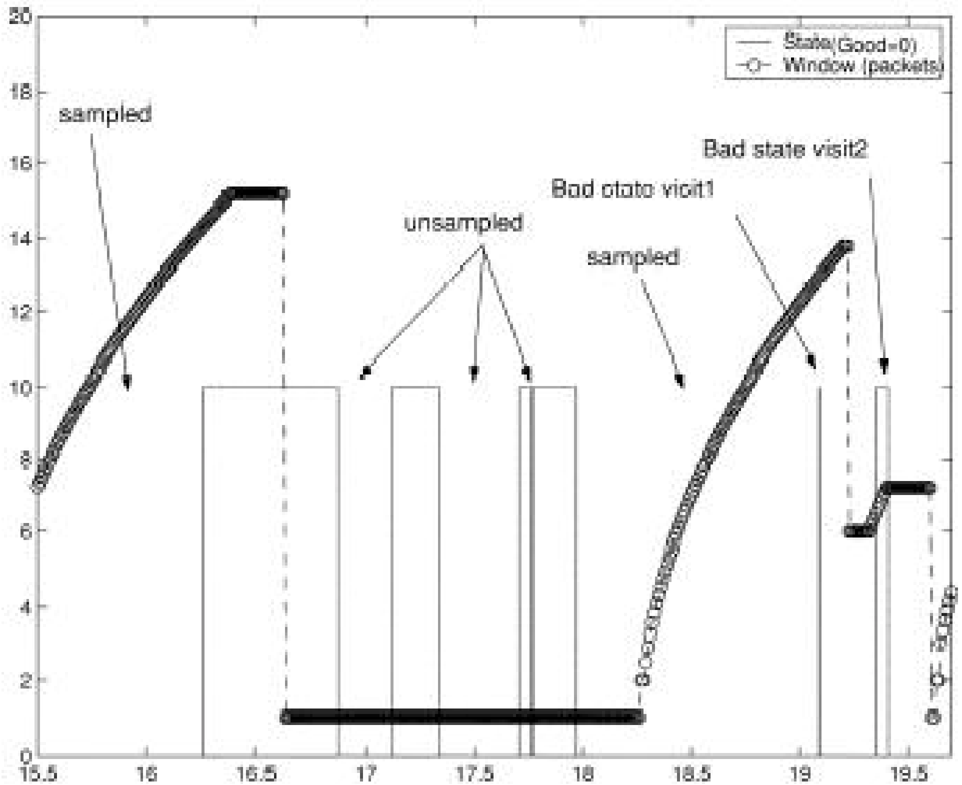


Fig. 1. A simulation trace for the window size and channel state against time using the channel-driven loss model implementation in *ns*-2. ($\mu = 100$, $\beta = 8.0$, $\tau = 0.01$ s, $B = 16$ packets, $w_p = 18$).

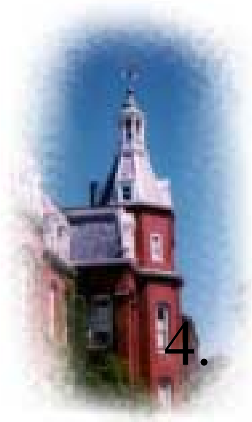
Modeling Approach

- The dynamics exhibit a large amount of memory. For example TCP algorithm for setting the timer-expiry period Δ(t`) retains all RTT samples since the start of the TCP session. The memory is also evident in the delay in TCPs TD algorithm which depends in a nontrivial way on the specific locations of the packet(s) lost within a window round.
- Our analysis is based on applying Markov renewal-reward theory based on the following approximations to the actual TCP/channel evolution

Modeling Aproach (cont.)

1. The delay between a packet loss event and its detection- of the order of one half of a RTT for a symmetric link is neglected.
2. We assume only one bad state visit per round which suffices when the good and/or bad state durations are sufficiently long. For multiple transitions between good and bad states during the same round, our model provides a lower bound since such multiple bad state visits are reflected in our model as visits to different rounds
3. We assume the timer expiry value at the end of a sample good state i.e., $\Delta(W_2)$ is completely determined by the window size W_2.

WPI



Modeling Aproach (cont.)

4. Instead of keeping track of the exact instant within a window round at which a sampled good state ends, we assume an origin of the bad state duration uniformly distributed over the duration of the round.
5. Following a TO, TCP-Reno reverts to slow start, where the window is set to one and the slow-start threshold is set to half the window size of which TO is detected. If further TOs occur, the threshold window will be successively reduce toward the limit 1. We assume that upon a single TO event, the window is set to one and that TCP-Reno immediately enters the congestion avoidance phase.

WPI

The Embedded MCs

- Let W_n denote the window size at the instant good state X_n is first sampled. Then the remaining time in the good state V_n is also exponential distributed with parameter λ_0, by the memory less property of the exponential distribution. Conditioned on the value of W_n, the window size evolution until the sampled good terminates is completely independent of the past. Let the window size at the end of the sample good state $W`_n$. Then similarly conditioned on the value of $W`_n$. the future evolution of the congestion window is independent of the past. The duration between successive sample good state may include multiple un-sample good states in addition to bad states.

The Embedded MCs

The later can be lumped into a single "effective bad duration" denoted U_n for analytical purposes. Since X_i and Y_i are independent, V_n and U_n are also independent Thus, the sequence $\{W_n\}_{n=1}$ of the process defined just after a good state is first sample form an embedded MC (of the semi-Markov process). Similarly the sequence $\{W`_n\}_{n=1}$ of the process defined just before a sample good state terminates form an embedded MC. The former is called the "good" MC and the latter is called the "bad" MC. Each of previously mentioned approximations to the actual TCP process is necessary for the above embedded process to be Markov

- Calculating Conditional Reward $E[n|W_n]$
- The Transition Probability $Pr[W`_n \mid W_n]$
- Outcome of a Visit to a Bad State
- Transmission Probability $Pr[W_{n+1} \mid W_n]$ and $Pr[W`_{n+1} \mid W`_n]$
- Steady-State Window Size Distribution $\prod_g(w)$ and $\prod_b(w)$
- The Average effective Bad duration
- Proportional of Sampled Good States Terminating with NL, TD, or TO
- Average Goodput $\rho = \dfrac{E[N]}{E[V] + E[U]} \dfrac{1}{\mu}$

Outline

- Introduction
- A Model for TCP Congestion Avoidance
- Channel Model
- Analysis
- Model Validation and Discussion
- Conclusion

Model Validation and Discussion

- In the previous section, a number of approximations were perform to the actual TCP process in order to arrive at what we called and ‘idealized’ process, which is semi-Markov – hence enabling the application of tools from the Markov renewal-reward theory. In this section, we evaluate the accuracy of the resulting analytical model by comparing its results against measurements from the actual TCP process, as implemented in the well know network simulator ns-2

Model Validation and Discussion (cont.)

A. Our implementation of the Channel-Driven Error Model in ns-2

The original two state error model implemented in ns-2 uses packet reception as the clock for advancing the (two state) MC representing channel state. Hence when packets are not received the channel state is frozen.

In the ns-2 implementation the receive procedure:

- Decides whether the packet is to be dropped, depending on the current state
- Advance the two state MC in time by an amount equal to a packet transmission interval.

Model Validation and Discussion (cont.)

- Our implementation modifies the ns-2 procedure as follows:
  - Advance the channel state from the time the last packet was processed until the time the first byte of the current packet is received.
  - Advance the channel state in time until the packet is completely received.
  - The packet is successfully received only if the channel remains in the good state during the entire duration of the packet reception, otherwise is dropped.

WPI

Model Validation and Discussion (cont.)
Model Validation

- The figure shows that the model provides accurate goodput predictions for a wide range of E[X]. We node that for high loss rate (small E[X]), the analysis results slightly underestimates the goodput. This is consistent with the expected impact of the approximation 2 and 5 made in the Modeling Approach.

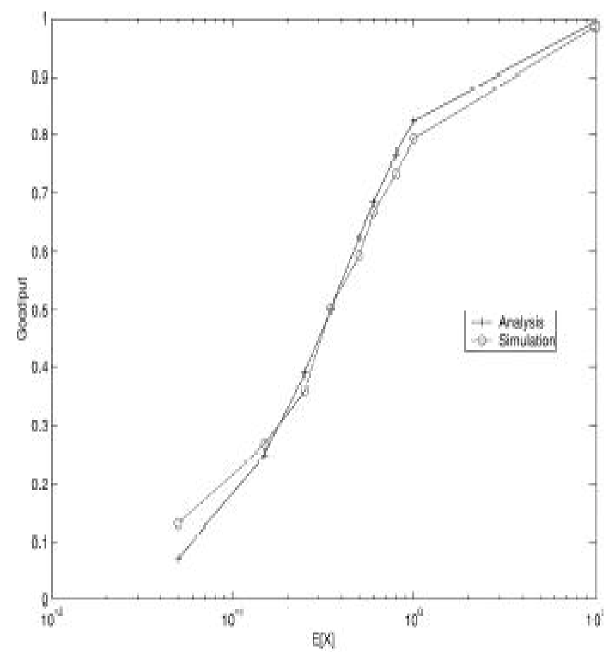


Fig. 4. Throughput comparison for the random packet loss case ($\mu = 100$ packets/s, $\beta = 1.2$, $\tau = 0.1s$, $B = 13$ packets, $w_p = 25$ packets, $K = 8 * 10^3 b$).

WPI

Model Validation and Discussion (cont.)
Model Validation

- The figure contains a representative comparison of goodput for the case of correlated packet loss showing plots of the goodput against the average holding time in the bad state E[Y]

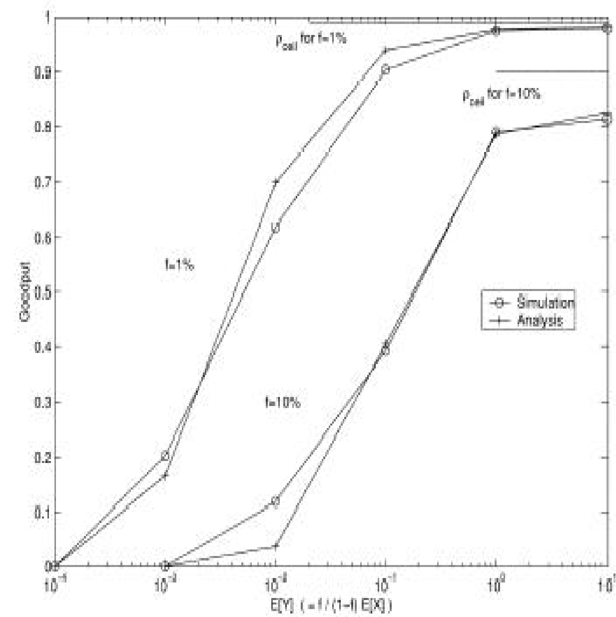
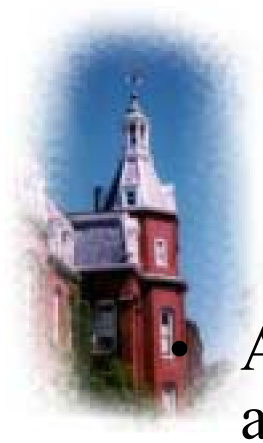


Fig. 5. Throughput comparison for the correlated packet loss case ($\mu = 100$ packets/s, $\beta = 4.0$, $\tau = 0.1$s, $B = 44$ packets, $w_p = 55$ packets, $K = 8 * 10^3$b).



Model Validation and Discussion (cont.)
Model Validation

- After selecting 100 parameter sets according to the steps above, five simulations runs (with randomly selected simulation seeds) are performed for each parameter sets, during which the steady-state throughput is measured and the average of these five runs constitutes the result of each experiment which is plot in figure 6.
- One of the major reasons for pursuing such advance modeling as in our work is the significant savings (about two orders of magnitude) vis-à-vis ns-2 simulations that typically took hours.
- The validation figures show that the analysis matches those of the simulations with reasonable accuracy, since the points lie close to the ideal 45˚ line.

WPI

Model Validation and Discussion (cont.)
Model Validation

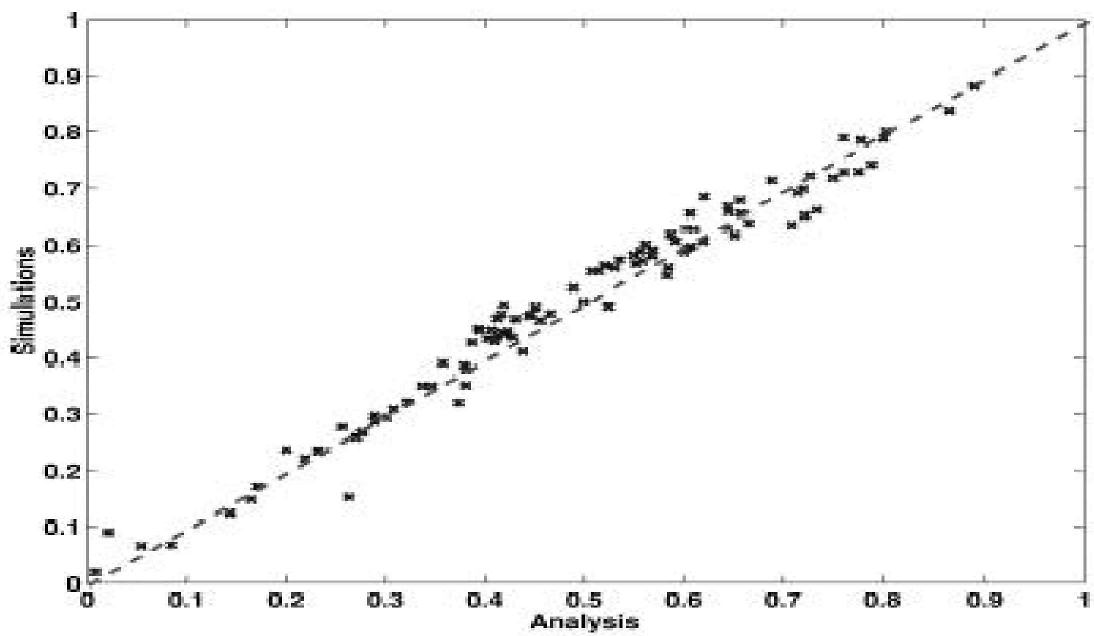


Fig. 6. Throughput from simulations versus analysis for the 100 randomly selected experiments.

WPI

Model Validation and Discussion (cont.) Observations

- Fig. 7: The behavior of the effective bad duration as a function of the actual channel bad duration is analyzed.
- Ideally, one would hope that the time wasted (due to idle sender or in transmission of packets that are lost) by TCP operating of a wireless link would not exceed that spent by the channel itself in the bad state, However, as can be seen from the figure 7 the effective bad duration can be as high as 30 times the average duration (for E[X] = 0.08 and even higher (74 times the average bad duration) for smaller E[X] ***as stated, this factor (skipping of good states as a result of timeouts coupled with the binary exponential backoff algorithm) is a key contributor to the deterioration of TCP performance over a wireless link***

WPI

Model Validation and Discussion (cont.)
Observations

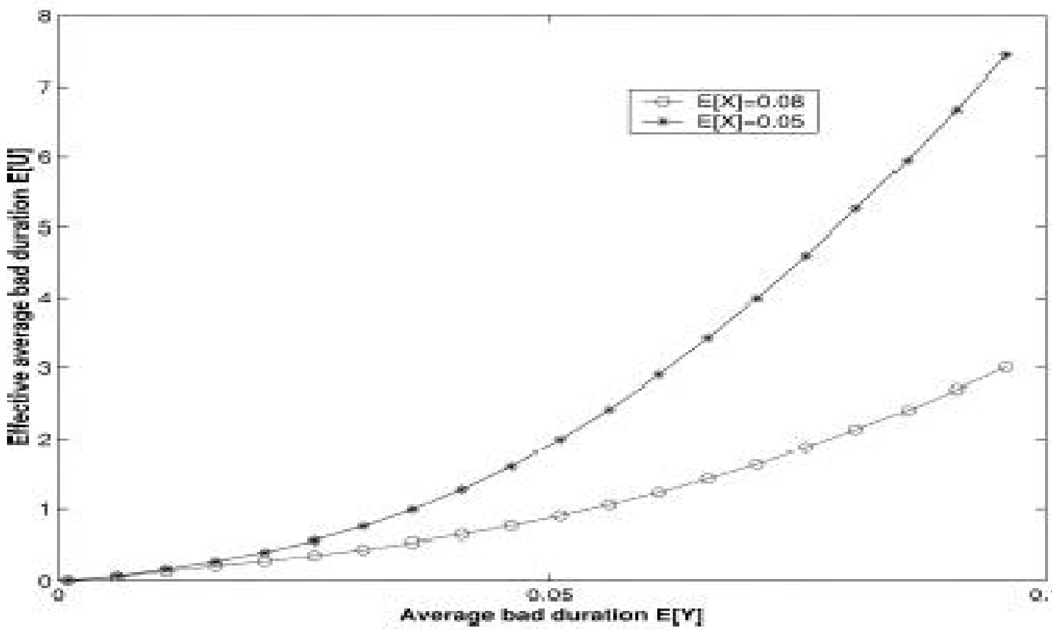


Fig. 7. The average *effective* bad duration can be many times the average value of the actual physical channel bad duration.

Model Validation and Discussion (cont.) Observations

- Fig. 9 presents an analysis of the effects of TCP packet size. It is show that, unlike previous results a larger packet size always results in a higher throughput, assuming the same channel fading rate. This can be explain as follows:
- The average window size at the end of a sample good state W2 in packets is inversely proportional to the packet size, as a direct consequence of using larger packets. However the window size in bytes (or bits) is directly proportional to the packet size, thus the direct reward must also be directly proportional to the packet size
- The effective bad duration is inversely proportional to the packet size.

Model Validation and Discussion (cont.) Model Validation

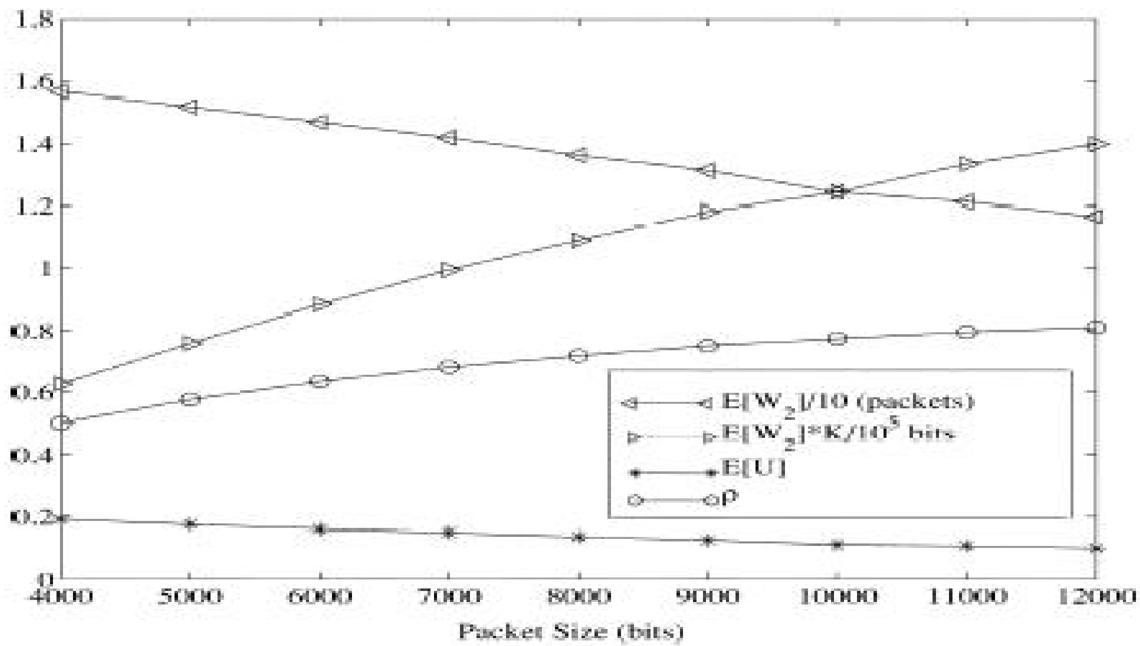


Fig. 9. The behavior of various parameters as a function of the packet size for $E[Y] = 0.01$ s and $f = 1\%$).

Model Validation and Discussion (cont.) Limitation and Extensions

- Computation Complexity: While the analysis in this paper did not lead to a closed form solution of the steady state throughput or average window size, it does provide simple expressions for the parameter needed to accurately compute them.
- Extensions: The analysis of this paper was directed to TCP Reno. Generalization to other TCP versions is nontrivial, and would succeed only if a suitable rewal-reward characterizations can be found.
  Another challenging extension of the model is to the case of multiple flows.

WPI

Model Validation and Discussion (cont.)
Limitation and Extension

- An appropriate Markovian description that accuratly captures the interaction between backbone links with multiple flows would be very useful; however the model complexity is expected to rise at least linearly with the number of flows, thus limiting the potential utility of this approach without further simplification.

Model Validation and Discussion (cont.)
Limitation and Extension

- Limitations: The analysis in this paper assumes that the window size can be expressed in terms of packets and hence are integer valued. The simulations have shown that this assumption does not have significant impact on the accuracy, mainly because we assume an infinite reservoir of reservoir of fixed-size packets equal to the maximum segment size allowed (500-1500 B)

WPI

Model Validation and Discussion (cont.)

Limitation and Extension

- Another limitation is that we assume TCP maximum advertised window size W_m is always greater than the congestion window. In reality the congestion window size is limited by the advertised window size.

Outline

- Introduction
- A Model for TCP Congestion Avoidance
- Channel Model
- Analysis
- Model Validation and Discussion
- Conclusion

Conclusion

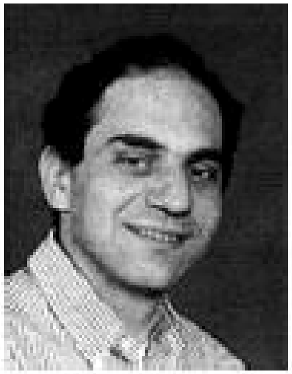
- A model for TCP operation over a wireless link was presented in which the channel evolution is independent of the TCP operation. A detailed Markow renewal-reward analysis for TCPs operation is performed that accounts for both channel lost and congestion (buffer) loss.
- The analysis captures key TCP aspects such as triple duplicate and timeout loss detection and the binary exponential back off algorithm. It was show that timeouts may cause TCP to skip some good states, resulting in a longer effective bad duration.
- Further, base on the model presented the result show that it is always recommended to uses a higher packet size for a given fading rate, unlike earlier work.

Questions

Alhussein A. Abouzeid (S'00–M'01) received the B.S. degree from Cairo University, Cairo, Egypt, in 1993 and the M.S. and Ph.D. degrees from University of Washington, Seattle, WA, in 1999 and 2001, respectively, all in electrical engineering.

He is currently an Assistant Professor in the Department of Electrical, Computer, and Systems Engineering (ECSE), Rensselaer Polytechnic Institute, Troy, NY. The industry interactions during his Ph.D. study included working with Microsoft Research, Redmond, VA (Ph.D. topic), Honeywell, Bellevue, WA (techniques for wireless access to airplanes), and at Hughes Research Labs, Malibu, CA (efficient networking protocol simulations for broadband satellite systems). He was a System Engineer with Alcatel Business Systems, Middle East Regional Office, Cairo, from 1994 to 1997 and a software engineer at the Information Technology Institute, Cairo, during 1993.

WPI



Murat Azizoglu (S'86–M'91) received the B.S. degree from Middle East Technical University, Ankara, Turkey, in 1985, the M.S. degree from Ohio State University, Columbus, in 1987, and the Ph.D. degree from Massachusetts Institute of Technology, Cambridge, in 1991, all in electrical engineering.

He is the Chief Network Architect of the Transport Business Unit at Sycamore Networks, Chelmsford, MA, where he leads an architecture group developing next-generation optical networking product architectures. Prior to joining Sycamore Networks in 1999, he was a tenured Associate Professor of Electrical Engineering at the University of Washington, Seattle, during 1994–1999. From 1991 to 1994, he taught and conducted research at the George Washington University, Washington, DC, as an Assistant Professor of Electrical Engineering and Computer Science. He was a Member of Technical Staff at Bellcore (now Telcordia), Morristown, NJ, in 1989. He has published extensively in a wide range of areas including optical networks, wireless networks, communication theory, and information theory.

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, in 1985 and 1988, respectively, all in electrical engineering, 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, Philadelphia, and at the University of Texas, San Antonio. He is presently Professor of Electrical Engineering, University of Washington, Seattle, where his research interests include analysis/design of communication systems/networks, with a topical emphasis on next-generation mobile/wireless networks. He is currently on academic leave at Intel Wireless Technology Lab working on high-speed UWB radios.

Prof. Roy is a Member of the IEEE Communications Society and is on several technical committees and TPC committees for conferences. He presently serves as an Editor for the IEEE Transactions on Wireless Communications.