On the autocorrelation structure of TCP traffic [☆]

Daniel R. Figueiredo [*], Benyuan Liu, Vishal Misra, Don Towsley

Department of Computer Science, University of Massachusetts, Amherst, MA 01003-9264, USA

Abstract

The statistical characteristics of network traffic—in particular the observation that it can exhibit long range dependence—have received considerable attention from the research community over the past few years. In addition, the recent claims that the TCP protocol can generate traffic with long range dependent behavior has also received much attention. Contrary to the latter claims, in this paper we show that the TCP protocol can generate traffic with correlation structures that spans only an analytically predictable finite range of time-scales. We identify and analyze separately the two mechanisms within TCP that are responsible for this scaling behavior: timeouts and congestion avoidance. We provide analytical models for both mechanisms that, under the proper loss probabilities, accurately predict the range in time-scales and the strength of the sustained correlation structure of the traffic sending rate of a single TCP source. We also analyze an existing comprehensive model of TCP that accounts for both mechanisms and show that TCP itself exhibits a predictable finite range of time-scales under which traffic presents sustained correlations. Our claims and results are derived from Markovian models that are supported by simulations. We note that traffic generated by TCP can be misinterpreted to have long range dependence, but that long range dependence is not possible due to inherent finite time-scales of the mechanisms of TCP.
© 2002 Elsevier Science B.V. All rights reserved.

Keywords: TCP mechanisms; Network traffic characterization; Long range dependence

1. Introduction

The existence of non-degenerate correlation structures over a range of time-scales in network traffic has been observed in a variety of network environments, such as Ethernet traffic [1], wide-area networks traffic [2], and World Wide Web traffic [3]. Since then there have been many research efforts to investigate various aspects of this correlation, for example, traffic models [4–7], impact on network performance [8–10], and explanations for the presence of such correlation structures [3,11,12].

Several efforts to explain this correlation structure in the network traffic have been provided. Crovella et al. [3] conjectured that the correlation behavior in network traffic is caused by the heavy-tailed distribution of WWW document sizes, the effect of caching, and human think time. The

[*] Corresponding author.
E-mail addresses: ratton@cs.umass.edu (D.R. Figueiredo), benyuan@cs.umass.edu (B. Liu), misra@cs.umass.edu (V. Misra), towsley@cs.umass.edu (D. Towsley).

authors of [3] claim that these causes, which are all located in the application or user level, can generate long range dependence in the data traffic being sent by underlying transport protocols. In [12], the authors studied the chaotic behavior of TCP congestion control and concluded that the assumptions in [3] are not necessary to explain the origin of self-similarity in TCP data traffic. The authors argue that the "chaos" created by the TCP congestion control mechanism alone can generate self-similarity.

Stimulated by [12], other studies have tried to corroborate or explain why traffic generated by TCP exhibits self-similarity. In [13] the authors claim that TCP itself can generate self-similar traffic. Their study is based on traffic measurements of a high bandwidth WAN. The analysis is based on the coefficient of variation to determine the burstiness of the traffic; however this does not necessarily indicate the presence of self-similarity. In [14], Guo et al. claim that the TCP congestion control mechanism (i.e., exponential back-off of the timeout, TO) generates heavy-tailed off periods in the traffic transmission pattern when the loss probability is high. The authors use a Markovian model to show that the heavy-tailed silence periods would lead to traffic with self-similarity. We note that in reality, the off periods in the TCP TO mechanism are not heavy-tailed, since implementations of the protocol impose a maximum value in the retransmission timer after a TO occurs. In a similar work, the authors of [15] attempt to show that TCP congestion control mechanism can lead to self-similar traffic. However, this result is not consistent with that obtained through simulation in [11,16], which indicates that the correlation in TCP traffic exists only over a finite set of time-scales.

A few studies have advocated the existence of an upper bound in the time-scale associated with the correlation structure of TCP traffic and beyond which correlation becomes negligible. In a pioneering paper [11], Manthorpe et al. use simulation to point out that network traffic is not strictly self-similar, but that it exhibits a correlation structure only over a finite range of time-scales. To properly describe this phenomena, which is clearly not consistent with the concept of self-similarity, the authors introduce the terms *pseudo self-similarity* and *local Hurst parameter*. However, the authors only conjecture as to the causes for this limited scaling behavior. In separate work, [17] used simulation results to claim that general simple retransmission schemes of network protocols can make traffic appear self-similar over a certain range of time-scales. However, the author does not precisely identify the range of time-scales of interest nor explores specific TO mechanisms. In [16] the authors perform a careful statistical analysis of the same data used in [12] and evaluate a Markovian model of TCP to show that the protocol does not generate self-similar traffic, but instead presents a correlation structure only over a finite range of time-scales. In our earlier work [18], we identified and modeled the mechanisms of TCP that are responsible for the traffic correlation structure over a finite range of time-scales, pointing out it was not self-similar. We relate the correlation behavior with the packet loss probabilities, and show that the correlation is present over a wide range of packet loss probabilities. In subsequent work [19], Guo et al. also point out the existence of finite time-scales in the TCP traffic pattern. They show that the exponential back-off in the TO mechanism under relatively high loss rate can generate pseudo self-similar traffic.

To date, then, there is confusion regarding the correlation structure of traffic generated by TCP. Furthermore, the TCP protocol is composed of several mechanisms and little is known about the contribution of each of them to the correlation of TCP traffic. The goal of this paper is to carefully analyze the TCP protocol and identify how the TO and congestion avoidance (CA) mechanisms give rise to correlation structures in the traffic sending rate. We provide separate Markovian models for each mechanisms that, under the proper loss probabilities, accurately predict the range of time-scales and the strength of the sustained correlation structure of the traffic sending rate of a single TCP source. A simple simulation scenario is also used to support our conjecture and validate the models. Our analysis applies to a wide range of loss probabilities, including very low and high loss probabilities. We show that, depending on the loss probability, each of TCP's internal mechanisms

has a dominant effect on the correlation structure of the traffic generated by that session. For example, at a low packet loss probability of 0.01 a sustained correlation appear in time-scales ranging from 2 to 64 round trip times (RTTs) due to CA, while at a high loss probability of 0.2 the range is from 2 to 512 RTTs due to the TO mechanism. We present an analytical technique for predicting the largest time-scale associated with the correlation structure of each Markovian model. We also analyze a comprehensive model of TCP that accounts for both mechanisms and show that the traffic generated by this model also exhibits sustained correlation only over a finite range of time-scales and that the TO and CA mechanisms are primarily responsible for this characteristic in the traffic. Finally, we simulate and analyze a realistic scenario with many non-identical TCP flows to validate our results.

There are also some studies that have investigated the behavior of aggregate wide-area TCP/IP traffic over a broad range of time-scales. Their results indicate that at very fine time-scales, the network traffic tends to exhibit multi-fractal scaling [20]. The cut-off time-scale up to which this phenomena is present was empirically observed to be on the order of a round-trip time. The focus of our work is on the correlation structure of the traffic sending rate of a *single* TCP session, which has its internal mechanisms (TO and CA) governed by multiples of round-trip times. Therefore, in our study, there is no need to investigate this correlation at smaller time-scales since a TCP session does not change its sending rate at time-scales finer than one RTT.

Many publications in the literature use the term self-similar and long range dependence loosely and in many cases this has led to much confusion. Moreover, the Hurst parameter has also been used as a parsimonious measure to describe the correlation structure of analyzed data, which is also misleading since it must assume that the underlying process is self-similar. There have been attempts to generalize the concept of the Hurst parameter to models that are not self-similar but exhibit a sustained correlation structure over a finite range of time-scales. Among these, the work of [6] introduced the idea of pseudo self-similarity.

However, these generalizations, and the resultant terminology, also suffer pitfalls and will not be used in this paper. Instead, we will turn to the frequency domain to classify such processes, and use the classical family of $1/f$ processes [21] as the basis for our definition.

We define sustained correlation structure in terms of the power spectral density (PSD) $\psi(f)$ of a process, which will be detailed in Section 3. A $1/f$ process is generally defined as a process whose empirical PSD is of the form $\psi(f) \approx k/f^{\nu}$ for some $k > 0$ and $0 < \nu < 2$ over several decades of frequency f. For a long range dependent (LRD) processes, this relationship is true for arbitrary small frequencies, i.e., there is no low frequency roll-off. However, there is a class of processes that behave like LRD processes over a certain range of frequencies and the relationship $\psi(f) \approx k/f^{\nu}$ with ν previously defined, does *not* hold for some frequency $f < F_1$, but it does hold for all frequencies $F_1 < f < F_2$. We say such processes (which are also $1/f$ processes) have a *sustained correlation structure* over a finite range of time-scales, where the range of time-scales is determined by the frequency interval $[F_1, F_2]$ and is given by $[1/F_2, 1/F_1]$. [1] A classical example of such a process is the Markov on–off process. The PSD of this process resembles that of a Weiner process, with a ν value of 2, bounded by the on–off time periods on the low frequency range (F_1). However, unlike the Weiner process its spectrum flattens out for frequencies below F_1 and it does not suffer from the infinite variance problem associated with LRD processes. The Markov on–off process is a well known short range dependent process, however one could easily be misled if only a short finite frequency range is observed, as the process can possess sustained correlation structure over the corresponding time-scales.

It is worth mentioning that neither the TO nor the CA mechanisms can alone give rise to traffic with LRD characteristics, due to the inherent finiteness in their time-scales. However, the traffic

---

[1] Note that a continuous time process under our definition can have $F_2 = \infty$, but the inertia inherent in any physical system makes F_2 a finite value for any observable process.

induced by these mechanisms exhibits sustained correlation structure with similar ranges of time-scales to that of measured Internet traffic of previous studies, that have identified traffic to be self-similar. For example, in [22] the authors claim that the measured TCP packet traces over a lossy link is consistent with self-similar behavior. However, the range time-scales for which their results exhibits sustained correlations (Fig. 1(c)), which is from 2 to about 512 RTT, overlaps almost entirely with the time-scales generated by the TO mechanism. In a more general sense, since most applications use TCP as their data transport protocol and the majority of the network traffic is carried by TCP, one can easily misinterpret the origin of these correlation structures and the existence of self-similarity itself. The efforts made to explain self-similarity from an upper level perspective should take into consideration the behavior of TCP, since characteristics of the application layer are not necessary in order to generate the observed correlation structures in the network traffic if TCP is used as the transport protocol. As a result, we emphasize that one needs to be careful when claiming that network traffic is self-similar or when potentially identifying the cause of such correlation structures.

The rest of the paper is organized as follows. Section 2 describes the TCP TO and CA mechanisms and, respectively, presents Markov models of their behaviors. Section 3 presents the mathematical framework used to analyze these models, and shows that both models give rise to sustained correlation structures in the traffic over certain ranges of time-scales. Section 4 provides a simple simulation scenario and results from that scenario to validate our models. Section 5 presents the comprehensive model of TCP and its analysis. In Section 6 we present a more realistic simulation scenario composed of many non-identical TCP

flows and present the results obtained. Finally, Section 7 summarizes the paper.

2. TCP timeout and congestion avoidance models

In this section we describe the behavior of the TCP TO and CA mechanisms and introduce two corresponding Markov chains that capture their functionality. Our purpose here is not to provide a detailed or complete model of TCP, but rather to highlight and investigate the TO and CA mechanisms in isolation. The analysis of these two simple models will demonstrate that each mechanism can lead to a sustained correlation structure over a finite range of time-scales in the traffic rate generated by a TCP session.

TCP is a window based network transport protocol that provides reliable end-to-end data communication [23]. It performs flow control and congestion control by regulating its sending window size through an additive increase/multiplicative decrease mechanism. TCP includes several mechanisms, among which TO and CA can have a significant effect on the traffic pattern.

2.1. Timeout model

TCP uses a TO mechanism to achieve reliable transmission of data and to avoid congestion collapse [23]. For every packet sent by the source, TCP starts a retransmission timer and waits for an acknowledgment from the receiver. The duration of the retransmission timer (R_{TO}) is based on an estimate of the running average and variance of the RTT. TCP maintains a running average and variance of the RTT that is updated based on the arrival time of acknowledgment packets (ACKs). The retransmission timer expires if the ACK for the corresponding data packet is not received and
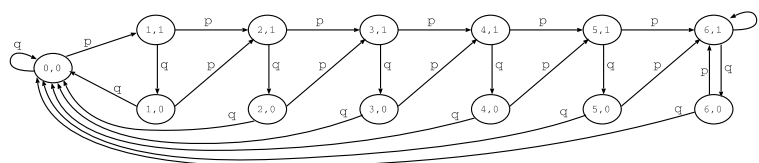

Fig. 1. TCP TO model.

if there are no three duplicate (TD) ACKs. Multiple ACKs for the same data packet can be generated by TCP, since a data packet that arrives out of sequence will trigger an ACK with the sequence number that was actually expected (duplicate ACK). When three consecutive duplicate ACK packets are received, the sender assumes that the corresponding data packet was lost and retransmits it without waiting for the retransmission timer to expire. If no TD ACKs are received, the retransmission timer will expire (TO) which will trigger the retransmission of the data packet and set the sending window size to one. This abruptly reduces the sending rate, hopefully relieving possible network congestion. Moreover, the retransmission timer value (R_{TO}) for the retransmitted packet is set to twice the value of the previously used timer value. TCP doubles the timer value for each subsequent retransmission to further reduce the sending rate and adapt to network congestion. This exponential back-off continues for each potential retransmission after the first TO occurs. After a TO occurs, all TCP implementations impose a maximum value in the retransmission timer. However, different methods are used to enforce this bound. In some TCP implementations the maximum retransmission timer value is predefined (e.g., 120 s in Linux [24]); while in others the maximum number of back-off stages (E_{max}) is fixed (e.g., 6 in netBSD and in ns simulator), which leads to a maximum retransmission timer of $2^{E_{max}} R_{TO}$. In the TO model described below, we assume the latter criteria; that the retransmission timer can be doubled up to a maximum of E_{max} times. After the TO value reaches its maximum value, it does not further increase and remains unchanged if loss continues. When the subsequent data packet is successfully transmitted, TCP recomputes the RTT estimate, resets the TO value accordingly and starts to increase its sending window.

For both the TO model and the following CA model, we assume that packet loss is described by a Bernoulli process with parameter p and that ACKs are never lost in the network. Even though the packet loss process in real networks is not given by a Bernoulli process, we will later argue that under our assumptions this process is sufficient to gener-

ate losses and can be easily tuned. We will relax this constraint in later sections of the paper when presenting both the comprehensive model and simulation results. The assumption that ACKs are never lost is mainly for simplicity and should not have significant impact on our results. Based on the TCP TO mechanism described above, we construct a discrete time Markov chain illustrated in Fig. 1 ($E_{max} = 6$) to model the TO mechanism. The single parameter of the model is the packet loss probability p, where $q = 1 - p$. The detailed explanation of the chain is given as follows.

- *State definition:* The finite state Markov chain is defined by the tuple (E_i, R_i), $i = 0, 1, \ldots$, where $\{E_i\}$ is a sequence of random variables indicating the value of the back-off exponent after the ith transition; and $\{R_i\}$ is a sequence of random variables indicating if the packet being sent after the ith transition is a retransmission ($R_i = 1$) or a new packet ($R_i = 0$). Note that $0 \leqslant E_i \leqslant E_{max}$, where E_{max} is the maximum number of times the retransmission timer doubles.

- *Transition probability matrix:* Fig. 1 illustrates all possible transitions in the model associated with packet transmissions for the case $E_{max} = 6$. Note that every time a packet is lost, the back-off exponent is increased by one to double the TO value. Moreover, it takes two consecutive successful packet transmissions in order for TCP to resume its normal mode of operation and reset its back-off exponent. Let $p_{e,r;e',r'} = P(E_{i+1} = e', R_{i+1} = r' | E_i = e, R_i = r)$ be the probability that the TO mechanism transits to state $(E_{i+1}, R_{i+1}) = (e', r')$ after the $(i+1)$th transition given that it was previously in state $(E_i, R_i) = (e, r)$. The transition probabilities are given by

$$\begin{aligned}
p_{e,1;e+1,1} &= p, & 0 < e < E_{max}, \\
p_{e,0;e+1,1} &= p, & 0 \leqslant e < E_{max}, \\
p_{e,1;e,0} &= 1 - p, & 0 < e \leqslant E_{max}, \\
p_{e,0;0,0} &= 1 - p, & 0 \leqslant e \leqslant E_{max}, \\
p_{E_{max},r;E_{max},1} &= p, & r = 0, 1.
\end{aligned}$$

- *Time:* Each state of the Markov chain has an associated holding time, which is dependent on the outgoing transition. Time is in units of RTT. We will assume that the RTT is constant and

has value R. Thus, if the packet is successfully transmitted, then the state holding time is equal to one R. If a packet is lost, then the elapsed time is equal to $2^E R_{\mathrm{TO}}$, where $R_{\mathrm{TO}} = \alpha R$ and represents the initial value of the retransmission TO timer. Note that α must be an integer greater than zero (in the analysis we assume $\alpha = 1$). Formally, let T_i, $i = 1, 2, \ldots$, be the state holding time for state (E_{i-1}, R_{i-1}), which is the state holding time after the ith transition. Thus, we can define T_i, $i = 1, 2, \ldots$ as

$$
T_i = \begin{cases}
2^e R_{\mathrm{TO}}, & E_{i-1} = e,\ R_{i-1} = 1;\ E_i = e + 1, \\
 & R_i = 1;\ 0 < e < E_{\max}, \\
2^e R_{\mathrm{TO}}, & E_{i-1} = e,\ R_{i-1} = 0;\ E_i = e + 1, \\
 & R_i = 1;\ 0 \leqslant e < E_{\max}, \\
R, & E_{i-1} = e,\ R_{i-1} = 1;\ E_i = e, \\
 & R_i = 0;\ 0 < e \leqslant E_{\max}, \\
R, & E_{i-1} = e,\ R_{i-1} = 0;\ E_i = 0, \\
 & R_i = 0;\ 0 \leqslant e \leqslant E_{\max}, \\
2^{E_{\max}} R_{\mathrm{TO}}, & E_{i-1} = e,\ R_{i-1} = r;\ E_i = e, \\
 & R_i = 1;\ e = E_{\max},\ r = 0, 1.
\end{cases}
$$

- *Traffic sending rate:* Associated with each state of the Markov chain is a packet sending rate, which depends on the outgoing transition. Note that in each state of the chain exactly one packet is sent. More formally, let S_i, $i = 1, 2, \ldots$ be the packet sending rate associated with state (E_{i-1}, R_{i-1}). Thus, S_i can be defined as

$$
S_i = \frac{1}{T_i}, \quad i = 1, 2, \ldots \tag{1}
$$

Current available techniques to analyze the correlation structure of an arbitrary discrete time Markovian model require the state holding times to be identical. In order to use such techniques, we modify the TO model by trivially expanding the state space so that all states have identical holding times of R. A state M in the original chain that has holding time kR, $k = 1, 2, \ldots$, is expanded to a sequence of states N_i, $i = 1, \ldots, k$, each having a holding time of R, where transitions into state M are now associated with state N_1, transitions out of state M are now associated with state N_k, and $p_{N_i, N_{i+1}} = 1, i = 1, \ldots, k - 1$. The traffic sending rate of state M is associated with state N_k and all

other states in the sequence have a sending rate of zero.

2.2. Congestion avoidance model

We now focus on the CA mechanism and ignore the presence of timeouts to better capture the behavior of the additive-increase/multiplicative-decrease mechanism for governing the window size. Two distinct mechanisms, namely slow start and congestion avoidance, control the window growth of TCP and consequently, its traffic sending rate. TCP exits the slow start phase and enters CA after the window size exceeds a certain threshold. This threshold is dynamically adjusted and is set to half of the largest window size when congestion is detected. We will ignore the slow start phase and focus solely on the CA mechanism since most of the traffic of a TCP source is transmitted while in this phase. Our simulation results show that for loss probabilities of 0.01, 0.1 and 0.2, the percentage of packets sent in the CA phase is 0.92, 0.95 and 0.99, respectively. This result was obtained using the simulation scenario that will be presented in Section 4, which basically consists of a single TCP-SACK flow traversing a single link that drops packets according to a Bernoulli process. This result together with the fact that the vast majority of the traffic in the Internet is carried by a few long TCP sessions [25], justifies our focus on the congestion avoidance phase. Moreover, the correlation strucucture of the traffic sending rate obtained with the model agrees well with simulation results, as we will demonstrate in Section 4. We conjecture that the impact of the slow start phase in the traffic correlation is very small and can be ignored.

In the CA phase, the window size increases by one packet when all packets in the current window are successfully acknowledged. Hence, the window size grows linearly in time during the CA phase. When a packet is lost, TCP reduces the size of the current congestion window to reduce the amount of traffic it can inject into the network. In most versions of currently deployed TCP, such as TCP-Reno and TCP-Sack, the window size is reduced by half when TD acknowledgments are received. If a TO occurs before that, the window size is
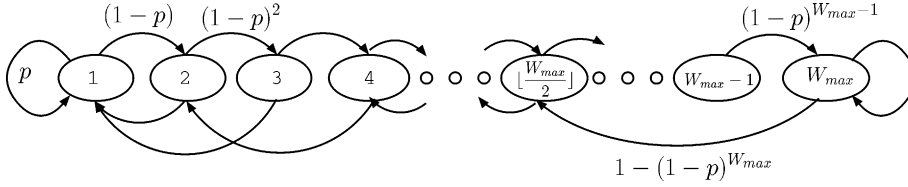
Fig. 2. TCP CA model.

reduced to one and, after exiting from the TO mode of operation, TCP starts the window growth cycle again.

In the CA model, we only consider window reduction events due to triple duplicate ACKs, ignoring the TO events and the slow start phase. Hence, the window grows linearly when no loss occurs and is reduced by half when a packet is lost. We also make the assumption that the RTT is longer than the time required to send all packets in a window [26], which is reasonable in a wide-area network. The discrete time Markov chain for this model is illustrated in Fig. 2 and has only two parameters: the maximum window size ($W_{\max}$) and the packet loss probability (p).

- *State definition:* The state of the Markov chain is given by W_i, $i = 0, 1, \ldots$, which indicates the window size of the TCP sender after i transitions. Note that $1 \leqslant W_i \leqslant W_{\max}$ and that the initial state W_0 depends on the value of the CA threshold, and can range from 1 to $\lfloor W_{\max}/2 \rfloor$.
- *Transition probability matrix:* During the CA phase, the window size increases by one when all packets in the current window have been successfully acknowledged, provided that the current sender window lies below its maximum value. Thus, the transitions occur after the window is fully transmitted. If the window size is w, the probability that the entire window is successfully transmitted is just $(1-p)^w$. Once the window size reaches its maximum value, it remains at this value until a packet is lost. Whenever a packet is lost, the TCP sender reduces the window size to half, causing a transition from state w to state $\lfloor w/2 \rfloor$. The probability that at least one packet is lost among the w packets in the current window is just $1 - (1-p)^w$. Let $p_{w;w'} = P(W_{i+1} = w' | W_i = w)$, where $0 \leqslant w, w' \leqslant$

$W_{\max}$, which is the probability associated with this state transition. The transition probabilities for the chain are then given by

$$
\begin{aligned}
p_{w;w+1} &= (1-p)^w, & 1 \leqslant w < W_{\max}, \\
p_{w;\lfloor w/2 \rfloor} &= 1 - (1-p)^w, & 1 < w \leqslant W_{\max}, \\
p_{W_{\max};W_{\max}} &= (1-p)^{W_{\max}}, \\
p_{1;1} &= p.
\end{aligned}
$$

- *Time:* We associate a state holding time with each state of the Markov chain. As with the previous model, time will be measured in units of RTT and we assume that RTT is constant with value R. Since we also assume that R is larger than the time required to transmit a window of packets, every transition in the Markov chain requires the same amount of time. Let T_i, $i = 1, 2, \ldots$ be the state holding time for state W_{i-1}. Thus, $T_i = R$ for all $i > 0$.
- *Traffic sending rate:* Associated with each state of the Markov chain is a packet sending rate. The sending rate for a state is defined to be the number of packets transmitted divided by the state holding time. Let S_i, $i = 1, 2, \ldots$ be the packet sending rate associated with state W_{i-1}. Thus, it is defined as

$$
S_i = \frac{W_{i-1}}{R}, \quad i = 1, 2, \ldots \tag{2}
$$

3. Model analysis

In this section we first describe the mathematical framework used to analyze the correlation structure of both the TO and CA models. We then present various analytical results supporting our claim of sustained correlation in TCP traffic over a finite range of time-scales.

We are interested in the correlation structure of the traffic sent by a TCP source. Using the TO and

CA models and S_i defined in Eqs. (1) and (2), respectively, we can obtain the correlation structure of the packet sending rates. However, in order to compare different models and simulations, we adjust the time series associated with packet sending rate so that it has zero mean and unit variance. Thus, let μ be the mean and σ^2 be the variance of the sequence S_i, $i = 1, 2, \ldots$ The normalized sequence is $S_i' = (S_i - \mu)/\sigma$, $i = 1, 2, \ldots$ Both μ and σ^2 can be easily obtained from the steady state probability distribution of the Markov chain.

Assume that the initial state distribution of the Markov models is the steady state probability distribution. The autocorrelation function of the traffic rate is given by

$$\rho(\tau) = E[S_1' S_{\tau+1}'], \quad \tau = 0, 1, 2, \ldots$$

Note that $E[S_1']$ is the expected steady state traffic sending rate of the model, which is zero due to normalization. The power spectral density (PSD) of a discrete time (stationary) stochastic process is defined as the discrete Fourier transform of its autocorrelation function and is given by

$$\psi(f) = \sum_{n=-\infty}^{\infty} \mathrm{e}^{-2\pi f n i} \rho(n), \quad \text{for } -\infty < f < \infty$$

where $i = \sqrt{-1}$. If $\rho(n)$ is real, this transforms to

$$\psi(f) = \rho(0) + 2 \sum_{n=1}^{\infty} \cos(2\pi f n) \rho(n) \qquad (3)$$

To construct and analyze both models, we use the TANGRAM-II modeling tool [27], which allows us to obtain numerically, among other measures, the autocorrelation function $\rho(\tau)$ defined above (see [28] for a description of the technique). The PSD $\psi(f)$ is then numerically computed from the autocorrelation function using Eq. (3). Note that Eq. (3) applies since $\rho(\cdot)$ is real for any real-valued process, which is the case for the traffic sending rate.

3.1. Analytical techniques

A technique frequently used in the literature for analyzing the correlation behavior of a process over different time-scales is based on wavelet transforms. Wavelet-based analysis [29,30] is compu-

tationally very efficient. It is also robust under certain types of non-stationary components sometimes present in the data that generally cause problems for other estimation techniques. This analysis estimates the variance of the wavelet coefficients of the analyzed time series at particular time-scales. This estimate is then plotted in a double $\log_2$ scale and a linear asymptotic regime over all time-scales above a certain threshold is taken as evidence of self-similarity in the data. The slope α of this asymptotic linear region yields an estimate of the Hurst parameter through the relation $H = (\alpha + 1)/2$.

The estimate of the variance of the wavelet coefficients measures the "energy" in the signal at the given time-scale. This measure corresponds to an estimate of the PSD of the process at a frequency determined by the given time-scale [29]. In particular, the time-scale 2^j corresponds to the frequency $2^{-j} f_0$, where f_0 is determined by the sampling rate of the time series. Therefore, the wavelet estimator is essentially the PSD of the process associated with the analyzed time series. However, the estimate of the PSD via the variance of the wavelet coefficients suffers from a bias that is dependent on the time-scale being analyzed. In the case the process is truly LRD (i.e., $\psi(f) \approx 1/f^v$, with $0 < v < 2$), this bias can be removed giving rise to an unbiased estimator, as pointed out in [29]. However, if the process is not LRD, then the power spectral density estimate is biased.

To maintain consistency of presentation, we plot the analytically obtained PSD using the same axes as the wavelet analysis. That is, the x-axis is in a $\log_2$ scale of increasing time-scales (decreasing frequencies) and the y-axis is in a $\log_2$ of increasing energy (increasing PSD). This also retains the intuitive nature of plots, in which quantities increase from left to right. Thus, the reader should visually interpret these plots the same way as the energy-scale plots commonly used in wavelet literature and elsewhere in this paper. An illustration of this mapping is shown in Fig. 3. We note that the computationally efficient wavelet analysis method developed in [30] only yields estimates at time-scales 2^j, $j = 0, 1, \ldots, n$, where n depends on the length of the time series, while the PSD can be computed at any time-scale (frequency). In the
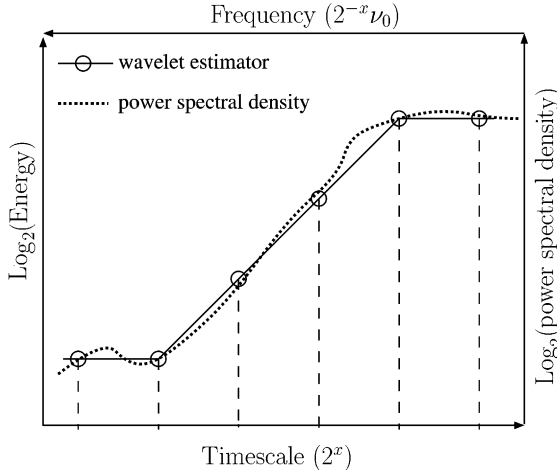
Fig. 3. Relating the wavelet estimator plot to the PSD.

wavelet analysis the energies estimated at each time-scale are connected by straight lines, as illustrated in Fig. 3. This plot also has an interpretation as a laterally inverted periodogram, with the axes scaled to match the axes in the wavelet plots.

3.1.1. Wavelet analysis of simulated traces

All simulations in this paper generate packet traces sent by each source. The traces are then aggregated, using fixed length bins, into a packet rate time series. Let P_i, $i = 1, 2, \ldots, n$ be the packet time series generated from a packet trace using bin sizes of length M time units. Thus, P_i indicates the number of packets sent during the time interval $[M(i-1), Mi)$. In all time series generated, M is set to the average RTT of the corresponding TCP session. Without loss of generality, the time series is normalized to have zero mean and unit variance before being analyzed. Let $\bar{\mu}$ be the sample mean and $\bar{\sigma}^2$ be the sample variance. Thus, the normalized time series is $P_i' = (P_i - \bar{\mu})/\bar{\sigma}$, $i = 1, 2, \ldots, n$. This normalization allows a direct comparison between all results obtained, since the same normalization is performed for the models.

In order to understand the correlation structure of the traffic sending rate, we use the wavelet-based estimator described above to analyze the packet rate time series. The publicly available wavelet analysis algorithms developed by Veitch and Abry [30] were used and we chose the Daubechies-1 as

the mother wavelet. We did not find any need for higher order mother wavelets, since the results obtained were very similar. As mentioned above, the output of the wavelet analysis is a graph where the x-axis represents time-scales. It is important to note that these time-scales are a multiple of the aggregation time unit used for generating the time series (i.e., the sampling frequency). Thus, j in the x-axis represents a time of $2^j M$, where M in our time series is the average RTT. The estimates of the variance of the wavelet coefficients together with a a 95% confidence interval of such estimates are computed by the algorithm.

3.1.2. Time-scales and power spectral density

In this section, we derive an analytical procedure to obtain the lower boundary of the frequency range over which the process exhibits sustained correlation structure. Recall that sustained correlation structure is defined over a range of time-scales $[1/F_2, 1/F_1]$. Here we are interested in deriving $1/F_1$, which corresponds to the largest time-scale associated with the correlation structure.

The PSD function of a discrete-time Markov process can also be expressed in term of the eigenvalues of the transition probability matrix. This result was derived in [31] and is given by

$$\psi(\omega) = \sum_l \frac{\beta_l(1 - \lambda_l^2)}{1 - 2\lambda_l \cos(\omega T) + \lambda_l^2}$$

where ω is the frequency in radians, λ_l is the lth eigenvalue of the transition probability matrix of the Markov chain, and β_l is the average power contributed by λ_l as defined in [31]. The parameter T represents the time unit of the process, which in our case equals R, one RTT.

Note that each eigenvalue λ_l contributes a component to $\psi(\omega)$ over all frequencies. Let λ_0 be the largest positive real eigenvalue smaller than 1. It is known that the contribution of λ_0 dominates $\psi(\omega)$ for low frequencies [31]. The contribution of λ_0 is given by

$$\psi_0(\omega) = \frac{\beta_0(1 - \lambda_0^2)}{(1 - \lambda_0)^2 + 4\lambda_0 \sin^2(\omega T/2)}$$

where the identity $\cos(\omega T) = 1 - 2\sin^2(\omega T/2)$ was applied.

Assume $wT \ll 1$, thus we have $\sin(\omega T/2) \approx \omega T/2$. We will observe that this assumption holds in the analysis of our models. Thus, we have

$$\psi_0(\omega) = \frac{\beta_0(1 - \lambda_0^2)/\lambda_0 T^2}{\left(\dfrac{1 - \lambda_0}{\sqrt{\lambda_0}T}\right)^2 + \omega^2}.$$

In [4] the authors relate the boundary of the frequency range to the poles of the PSD. This observation is also discussed by Franks in [32,33]. Using this approach, the lower boundary of the frequency range is the pole of $\psi_0(\omega)$. Therefore, since $\omega = 2\pi f$, we have that the lowest frequency is given by

$$\frac{1}{F_1} = \frac{1}{2\pi} \frac{1 - \lambda_0}{\sqrt{\lambda_0}T}. \qquad (4)$$

Another interesting metric that can also be obtained analytically from our models is the strength of the correlation of the process. The strength of the correlation is given by $\psi(0)$, which can be determined from Eq. (3). At frequency zero, Eq. (3) reduces to the sum of the correlation function over all lags, which can be easily computed numerically. By comparing the values of $\psi(0)$ for different normalized processes, we can reason about the correlation strength of each respective model.

3.2. Analysis of the TO model

We begin by presenting the results of the TO model. Fig. 4 illustrates the wavelet plots of the correlation structure of the traffic sending rate for different loss probabilities. The *x*-axis represents the time-scales in units of RTT and in $\log_2$ scale (i.e., $R2^x$). The *y*-axis represents the energy, or $\log_2$ of the PSD as discussed in Section 3.1.2. Note that the first data point in all curves occurs at twice the RTT, since the basic time unit of the model is one RTT. From the figure, we observe that the envelope of the curves at integer powers all have a linear increasing part which gradually become a flat horizontal line. The linear increasing part indicates the presence of sustained correlation
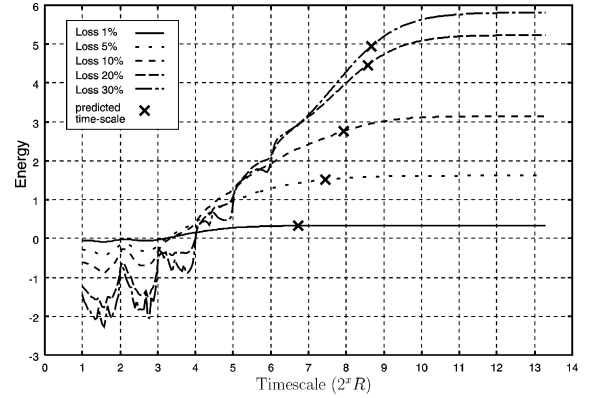

Fig. 4. Analysis of the TO model.

structure over the respective time-scales. A horizontal line in this plot indicates the absence of any correlation in those time-scales. The fine correlation structure at lower time-scales and within the integer powers is due to the interaction of the expanded states of the TO model. However, our focus is on larger time-scales and in particular on the time-scales at which the energy levels off.

From Fig. 4 we observe that different packet loss probabilities lead to different correlation behavior that span across different time-scales. In particular, as the loss probability increases, the range of time-scales for which sustained correlations are present increases. For example, when the loss probability is 0.1 the time-scales range from $2R$ to $128R$, while for a higher loss probability of 0.3, the range extends from $2R$ to $512R$. Table 1 depicts the time-scales $1/F_1$ (in $\log_2$), given by Eq. (4), at which the process should start to lose its correlation structure. These values are also marked with an $\times$ cross in each corresponding curve in

Table 1
Times-scales and correlation strength of TO mechanism

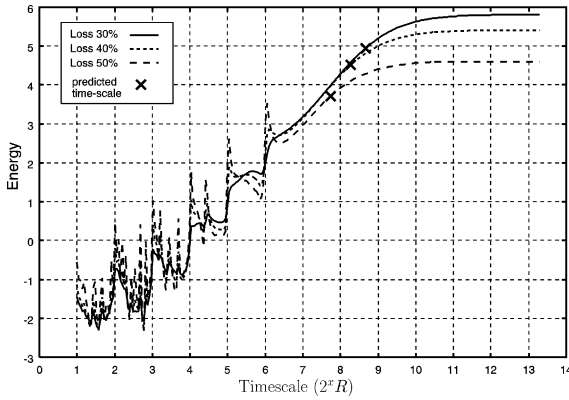| Packet loss probability	$\log_2$ (max time-scale)	$\log_2(\psi(0))$— "energy"
0.01	6.69	0.34
0.05	7.44	1.62
0.10	7.92	3.15
0.20	8.52	5.24
0.30	8.62	5.82
0.40	8.26	5.42
0.50	7.73	4.60

Fig. 5. Analysis of the TO model under high loss probability.

Fig. 4. We observe from Figs. 4 and 5 that the time-scales predicted by Eq. (4) agrees well with the TO mechanism. Table 1 also depicts $\psi(0)$, which measures the strength of the correlation of the process and corresponds to the "energy" at larger time-scales in the graphs of Fig. 4.

An intuitive explanation for the increase in the range of time-scales with loss probability, is that for higher loss the system is more likely to reach higher values of the back-off exponent, which introduces longer delays between packet transmission times, significantly reducing the traffic rate. This phenomenon introduces more correlation in the generated traffic, as shown by the increase in $\psi(0)$ in the table. Moreover, this also directly affects the range of time-scales for which sustained correlation appears.

For low loss probabilities the model exhibits almost no correlation structure, which is observed by the presence of an almost horizontal line over all time-scales. We observe that even though the range of time-scales predicted is high, the strength of correlation of the TO mechanism is very low, having values of 0.34 for loss probability of 0.01, as compared with 5.82 for a loss probability of 0.3.

One could ask what would happen if the TO mechanism was subjected to severe packet loss conditions and conjecture that the correlation structure would be more pronounced. Fig. 5 illustrates this situation and the results show that this hypothesis is not true. At higher packet loss probabilities, the range of time-scales over which the model generates sustained correlations de-

creases slowly, as well as the strength of correlation. This can also be observed in Table 1. Intuitively, when the loss probability is high, the dynamics of our model drifts to the states with high inter-packet delay. However, due to the finite state space of the model, the largest inter-packet delay is $R_{TO}2^6$ which limits the time-scales for which sustained correlation is present. The peaks in the linear rise of the curves are due to periodic behavior of the autocorrelation function. The autocorrelation function exhibit periodic oscillations with a period of $64R$ that is being caused by the structure of the Markov chain. These peaks would not appear in the wavelet analysis of measured traces since the wavelet analysis computes a smoothed value for the estimate of the PSD around each time-scale.

3.3. Analysis of the CA model

We now focus on the analysis of the CA model. Fig. 6 illustrates the energy–time-scale plot of the traffic sending rate for different loss probabilities. Here the maximum window size, W_{max}, was set to 30 packets, which is a typical value used in real TCP connections.

From the results, we observe that the curves have a linear increasing part which stops rising at a certain time-scale, and becomes horizontal. Again, we observe that the CA model exhibits very different behavior under different loss probabilities. In particular, as the loss probability decreases,
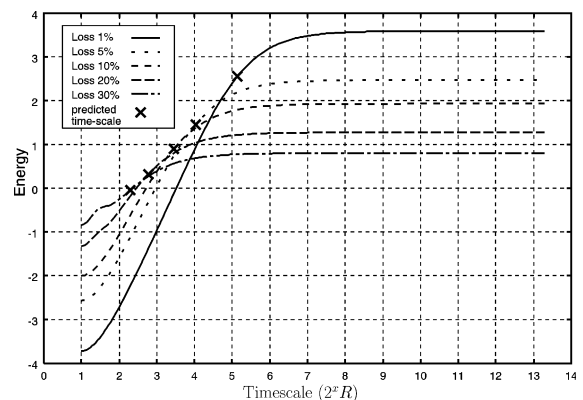

Fig. 6. Analysis of the CA model.

Table 2
Times-scales and correlation strength of CA mechanism

Packet loss probability	$\log_2$ (max time-scale)			$\log_2(\psi(0))$—"energy"		
	$W_{max} = 30$	$W_{max} = 60$	$W_{max} = 120$	$W_{max} = 30$	$W_{max} = 60$	$W_{max} = 120$
0.001	5.11	6.34	6.75	3.80	4.89	5.25
0.01	5.10	5.14	5.14	3.59	4.11	4.11
0.05	4.00	–	–	2.48	3.62	3.62
0.10	3.45	–	–	1.93	–	–
0.20	2.77	–	–	1.27	–	–
0.30	2.25	–	–	0.80	–	–

both the range of time-scales for which sustained correlations are present and the strength of the correlation increases. For example, when the loss probability is 0.3 the time-scales range from $2R$ to $8R$, and strength of correlation is 0.80. These values increase to a range of $2R$ to $64R$ and strength of 3.59 when the loss probability is 0.01. Table 2 illustrates the time-scales $1/F_1$ associated with the CA mechanism that were predicted by Eq. (4) and the strength of correlation ($\psi(0)$). The values predicted by the equation are also marked with an $\times$ in the corresponding curves in Fig. 6. We observe the trend in the range of time-scales and in the strength of correlation as loss probability decreases.

An interesting question is what happens to the range of time-scales when the model is exposed to very low loss probabilities. Fig. 7 illustrates the results for the CA model with loss probabilities of 0.01 and 0.001 and values of W_{max} of 30, 60 and

Fig. 7. Analysis of the CA model under very low loss probability.

120. We start by inspecting the results when the loss probability is 0.01. In this case, the correlation behavior of a TCP session does not depend on W_{max} once it exceeds 30. When the loss probability is 0.001, we observe that the correlation behavior of a TCP session is sensitive to the value of W_{max}. For $W_{max} = 30$, the time-scales and the correlation behavior are similar to the case that the loss probability is 0.01. However, for $W_{max} = 60, 120$ the correlation behavior spans a greater range of time-scales than their counterparts when the loss probability is 0.01. Moreover, the curve with $W_{max} = 120$ has a slightly larger range of time-scales in which the TCP session exhibit sustained correlations. This discussion is also supported by the time-scales predicted by Eq. (4) and by the strength of the correlation, both depicted in Table 2.

Intuitively, this behavior arises from the fact that for very low loss probabilities the CA mechanism increases the window size as much as possible. However, due to the limitation on the maximum window size imposed by W_{max}, this growth is inherently limited. Thus, a larger value of W_{max} allows the mechanism to reach larger window sizes and, consequently, reach higher time-scales. However, this drift to higher window sizes is only relevant under very low loss probabilities, since in this case the probability that the window increases past some larger value is not negligible ($(1-p)^w$).

We end this section by summarizing the behavior of both TCP mechanisms under different loss probabilities. The range of time-scales over which the model for the TO mechanism exhibits sustained correlations increases as the loss probability *increases*. In contrast, the model for the CA
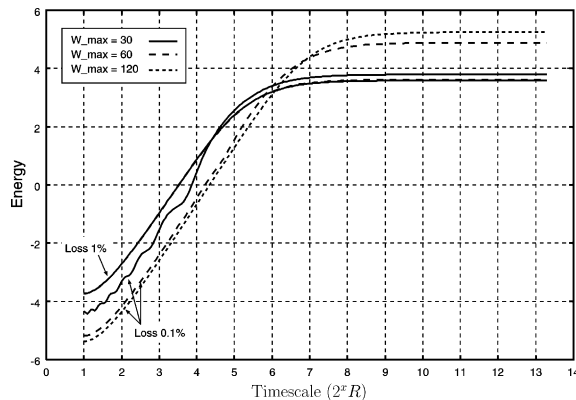
mechanism predicts that the range of time-scales increases as the loss probability *decreases*. Another observation is that in the case of $W_{max} = 30$, the range of time-scales of the TO model is much larger than that of the CA model. Thus, these two mechanisms combined play important roles at different time-scales of the TCP protocol.

4. Simulation scenario

In this section we use simulation to support our claim that the internal mechanisms of TCP can generate data traffic with sustained correlation structures over a finite range of time-scales under different loss probabilities.

All simulations were performed using the ns-2 simulator [34]. A simple network topology, consisting of a single source, a packet queue and a receiver, was simulated to investigate the traffic correlation structure generated by a single TCP session over a lossy link.

Fig. 8 illustrates the model being simulated. The sender has an infinite amount of data to transmit; thus it always wants to send as much data as possible. The queue stores packets from the source and forwards them to the receiver. We assume that the queue has an infinite size buffer to remove any packet loss correlation that might occur due to buffer overflow. A loss agent is attached to the queue and drops packets randomly at the time of their arrival according to a Bernoulli process with parameter p. The receiver acts as a sink and simply collects the data packets sent. The TCP transport protocol is used to transfer the data between sender and receiver. We assume that the ACKs sent by the receiver are never lost.

In this simulation scenario, we investigated the behavior of a single TCP flow in the absence of any background traffic. All packets traversing the queue belong to the TCP session being analyzed.
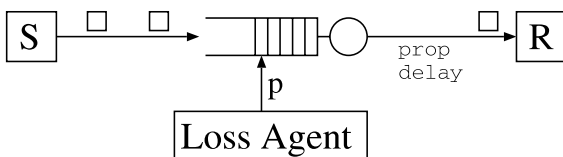

Fig. 8. Simulation scenario.

Thus, packet losses were generated solely by the Bernoulli loss process. This allows a direct comparison with the TO and CA models that capture the behavior of a single TCP flow.

During a simulation run, we capture two packet traces: one at the link between the sender and the queue (before the loss agent); the other between the queue and the receiver (after the loss agent). In all results presented, we analyze the first trace. For purposes of our study, we verified that both traces yield similar behavior, thus our conclusions also hold for the latter trace.

The parameters varied during the simulations were the packet loss probability, the link propagation delay (which is an important contributor to the RTT), and W_{max}, the maximum window size of the TCP protocol. Our results showed that the propagation delay and W_{max} had no significant impact on our conclusions. For the results presented here the link propagation delay was set to 50 ms, the link bandwidth to 1000 packets/s and W_{max} to 30. Notice that the time to transmit the maximum window (30 ms) is much smaller than the RTT (around 100 ms), which agrees with previous assumption made in both TO and CA models. The results shown are for the SACK version of TCP, which is becoming increasingly the dominant TCP used in the Internet [35]. However, we also simulated and analyzed TCP-Tahoe, which exhibited results very similar to the ones shown by SACK.

The simulations were usually executed for 100–500 h of simulated time, corresponding roughly to the transmission of 2–40 million packets, depending on the loss probability. The reason for such long simulation runs is to obtain tight confidence intervals when performing the wavelet analysis.

4.1. Observations from simulation

The packet rate time series was generated using the simulation packet trace with a bin size of the average RTT ($R = 100$ ms). Fig. 9 shows the results of the wavelet analysis of the normalized time series for different loss probabilities. The solid line indicates the simulation results with a 95% confidence interval denoted by a vertical line over integer values of the time-scale.
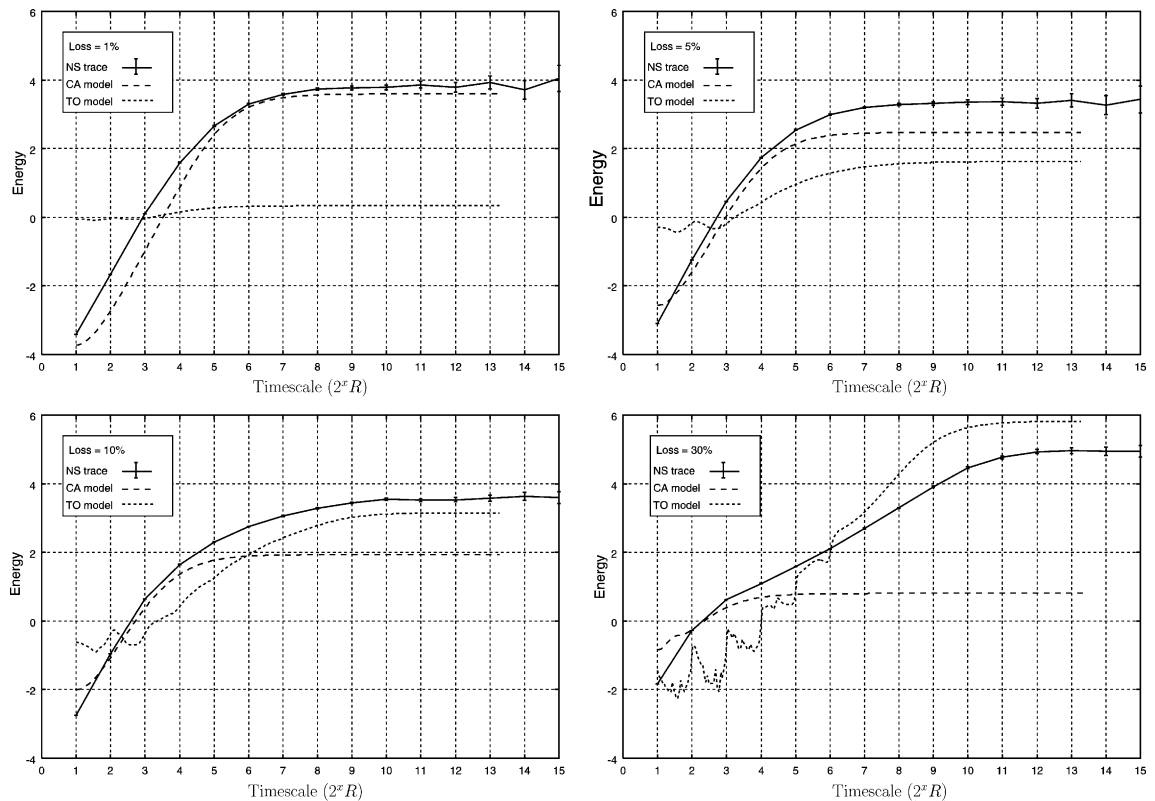
Fig. 9. Wavelet analysis of the simulation traces.

Our first observation from Fig. 9 is that under a Bernoulli loss process, a single TCP flow exhibits sustained correlation over a finite range of time-scales. We note that this correlation structure is present across all loss probabilities. A similar observation was also made in [11] where the authors use a different simulator, a more complex network model (i.e., more protocol layers) and no artificial loss process. Note that as the loss probability increases, the time-scales over which sustained correlation structure is present increases. For a loss probability of 0.01, the time-scales range from $2R$–$64R$, while for a loss probability of 0.3 it ranges from $2R$ to $1024R$, which corresponds to a range from 0.2 to 102.4 s (almost 2 min).

A second observation is that for low loss probability (0.01) the simulation result matches closely with the CA model under the same loss rate, as illustrated in the upper left graph of Fig. 9. Moreover, under high loss probability (0.3) the

simulation matches the TO model, as illustrated in the lower right graph of Fig. 9. This clearly indicates that under low and high loss probabilities, the CA and TO mechanisms, respectively, dominate the traffic pattern generated by TCP. Under low loss probability the CA mechanism dominates and produces correlations over a shorter range of time-scales. In this loss regime the TO mechanism has very little impact. However, for high loss probability the effect of the CA mechanism diminishes and the traffic correlation is dominated by the TO mechanism, producing a larger range of time-scales.

The results in Fig. 9 also validate the CA and TO models since under the proper loss regime the results from the model agrees well with simulation results. We point out that the correlation structure in the data traffic is caused by a mixture of the effect of the TO and CA mechanism. Thus, the simulation results obtained for medium loss

probability (0.1) cannot be directly compared to either the TO or CA models. However, the upper right and lower left graphs show that indeed the correlation behavior is a mix of both mechanisms. These plots show the simulation results together with the results for both mechanisms in isolation for the same loss probability, and we observe that the simulation curve is almost a combination of the other two curves.

5. Comprehensive TCP model

So far we have seen that both models for TO and CA mechanisms in isolation can produce sustained correlation structures in the data traffic over some analytically predictable range of time-scales. In this section we analyze a comprehensive Markovian model of TCP that includes both TO and CA mechanisms as well as a bursty packet loss process.

Many Markovian models have been developed to model the various mechanisms of TCP [19,36–39]. In the following analysis we will use a modified version of the model proposed in [36], which captures the TO and CA behavior of TCP-Reno using a discrete-time Markov process.

The TCP behavior is modeled in terms of "rounds", where a round represents the back-to-back transmission of the current congestion window. The assumption, as used for both of the TO and CA models, is that the round-trip time is larger than the time required to transmit all packets in the congestion window. The packet loss process is independent among different rounds. However, if a packet is lost in a round, then all remaining packets in that window are also lost. This loss process is motivated by packets overflowing a bottleneck queue with a drop-tail policy in the network. Moreover, since the round-trip time is larger than the time to send all packets in the window, it is reasonable to assume that the loss process is independent between rounds. The state of the model is described by the tuple of $(W_i, C_i, L_i, E_i, R_i)$, $i = 0, 1, \ldots$, where i is the number of the current round. W_i represents the window size for round i; C_i helps model the delayed ACK behavior of the TCP receiver, $C_i = 0$ indicates the

first of the two rounds and $C_i = 1$ indicates the second; L_i is the number of packets lost in the $(i-1)$th round; E_i denotes whether the connection is in a TO state and the value of the back-off exponent in round i; R_i indicates if a packet being sent in the TO phase is either a retransmission ($R_i = 1$) or a new packet ($R_i = 0$). The transition matrix of the chain is obtained from the behavior of TCP-Reno by examining all possible transitions from a given state. A state holding time that is dependent on the next transition is associated with each state of the model. Time values are in units of RTT and are denoted by an integer multiple of the RTT. Let R be the value of the RTT. As in Section 2, the model assumes that the retransmission timer R_{TO} is an integer multiple of R. This ratio is also a parameter of the model.

The original model proposed in [36] was validated using simulation and measurement results which indicate its adequacy in capturing the essence of TCP traffic [36,40,41]. In particular, it was shown that the model can predict the TCP throughput and packet sending rate under different scenarios quite well. The model described above, which will soon be analyzed, differs slightly from the one proposed in [36]. The original model allowed the sender to exit the TO phase and return to the CA phase immediately after a successful packet transmission. However, as discussed in the TO model, the TCP protocol requires the successful transmission of the subsequent data packet for the TCP sender to exit the exponential back-off phase and resume normal mode of operation. We modified the model and introduced this proper behavior. Note that the modified model requires two consecutive successful transmissions (one for the lost packet and the other for the new packet) before allowing the TCP sender to return to congestion avoidance phase. This modification was accomplished by introducing the state variable R_i into the state space. Apart from this extension, the model used here is exactly the same as the one proposed in [36]. A complete description of the modified Markov chain is now given.

Let $p_{w,c,l,t,r;w',c',l',t',r'} = P(W_{i+1} = w', C_{i+1} = c', L_{i+1} = l', E_{i+1} = t', R_{i+1} = r' | W_i = w, C_i = c, L_i = l, E_i = t, R_i = r)$ be the probability associated with this state transition. Moreover, associated with

each state of the Markov chain is a packet sending rate, which is dependent on the outgoing transition. Let the sequence S_i, $i = 1, 2, \ldots$ be the packet sending rate associated with state $(W_{i-1}, C_{i-1}, L_{i-1}, E_{i-1}, R_{i-1})$. Let $S_i = s_{w,c,l,t,r;w',c',l',t',r'}$ for two successive states. Thus, the transition probabilities and the packet sending rate are given by

- No packets are lost

$$p_{w,0,0,0,0;w,1,0,0,0} = (1-p)^w, \qquad 1 \leqslant w < W_{\max},$$
$$p_{w,1,0,0,0;w+1,0,0,0,0} = (1-p)^w, \quad 1 \leqslant w < W_{\max},$$
$$p_{w,0,0,0,0;w,0,0,0,0} = (1-p)^w, \qquad w = W_{\max},$$
$$s_{w,0,0,0,0;w,1,0,0,0} = w/R, \qquad 1 \leqslant w < W_{\max},$$
$$s_{w,1,0,0,0;w+1,0,0,0,0} = w/R, \qquad 1 \leqslant w < W_{\max},$$
$$s_{w,0,0,0,0;w,0,0,0,0} = w/R, \qquad w = W_{\max}.$$

- One or more packets are lost in a round

$$p_{w,c,0,0,0;w-l,0,l,0,0} = p(1-p)^{w-l}, \quad 2 \leqslant w \leqslant W_{\max},$$
$$\qquad\qquad\qquad c = 0,1, \ 1 \leqslant l < w,$$
$$p_{w,c,0,0,0;1,0,0,1,1} = p, \qquad\qquad 1 \leqslant w \leqslant W_{\max},$$
$$\qquad\qquad\qquad\qquad c = 0,1,$$
$$s_{w,c,0,0,0;w-l,0,l,0,0} = w/R, \qquad 2 \leqslant w \leqslant W_{\max},$$
$$\qquad\qquad\qquad\qquad c = 0,1, \ 1 \leqslant l < w,$$
$$s_{w,c,0,0,0;1,0,0,1,1} = w/R_{\text{TO}}, \qquad 1 \leqslant w \leqslant W_{\max},$$
$$\qquad\qquad\qquad\qquad c = 0,1.$$

- One or more packets are lost in a short round

$$p_{w,0,l,0,0;1,0,0,1,1} = 1, \quad 1 \leqslant w < 3,$$
$$p_{w,0,l,0,0;1,0,0,1,1} = \sum_{i=0}^{2} p(1-p)^i, \quad 3 \leqslant w < W_{\max},$$
$$p_{w,0,l,0,0;\lfloor (w+l)/2 \rfloor,0,0,0,0} = \sum_{i=3}^{w-1} p(1-p)^i + (1-p)^w,$$
$$\quad 3 \leqslant w < W_{\max},$$
$$s_{w,0,l,0,0;1,0,0,1,1} = w/(R_{\text{TO}} - R), \quad 1 \leqslant w < 3,$$
$$s_{w,0,l,0,0;1,0,0,1,1} = w/(R_{\text{TO}} - R), \quad 3 \leqslant w < W_{\max},$$
$$s_{w,0,l,0,0;\lfloor (w+l)/2 \rfloor,0,0,0,0} = w/R, \quad 3 \leqslant w < W_{\max}.$$

- Exponential Back-off

$$p_{1,0,0,i,r;1,0,0,\min(i+1,7),1} = p, \qquad 1 \leqslant i \leqslant 7,$$
$$\qquad\qquad\qquad\qquad\qquad r = 0,1,$$
$$p_{1,0,0,i,1;1,0,0,i,0} = 1-p, \qquad\qquad 1 \leqslant i \leqslant 7,$$
$$p_{1,0,0,i,0;2,0,0,0,0} = 1-p, \qquad\qquad 1 \leqslant i \leqslant 7,$$
$$s_{1,0,0,i,r;1,0,0,\min(i+1,7),1} = 1/(2^{(i-1)}R_{\text{TO}}), \quad 1 \leqslant i \leqslant 7,$$
$$\qquad\qquad\qquad\qquad\qquad r = 0,1,$$
$$s_{1,0,0,i,1;1,0,0,i,0} = 1/R, \qquad\qquad 1 \leqslant i \leqslant 7,$$
$$s_{1,0,0,i,0;2,0,0,0,0} = 1/R, \qquad\qquad 1 \leqslant i \leqslant 7.$$

All other transitions not defined above have probability zero. A detailed description and explanation of the original model can be found in [36]. In the analysis that follows, we will present results for the case $R_{\text{TO}} = R$ and $R_{\text{TO}} = 2R$. The value of $2R$ was chosen to be comparable with the simulation results that will follow in the next section.

In order to apply a technique to analyze the correlation structure of this model we need all states of the chain to have identical holding times. As discussed in Section 2, we expand the state space of the model so that each state holding time is equal to R. The expansion used here is exactly the same as the one presented in Section 2 for the TO model.

The number of states in the Markov chain depends directly on the R_{TO} value. For the case $R_{\text{TO}} = R$ and $R_{\text{TO}} = 2R$, the state space of the comprehensive model has 567 and 663 states, respectively. For such state spaces, the numerical computation of the autocorrelation function for a maximum lag of a 1000 ($\rho(\tau); \ \tau = 1, \ldots, 1000$) and of the PSD for a minimum frequency of 10^{-5} having 1024 points, is quite fast. The computation of both of these functions takes under 2 min for the larger state space on a Pentium III—730 MHz CPU.

5.1. Model analysis

This detailed model of TCP was analyzed using the same techniques described in previous sections to characterize the correlation behavior of traffic generated by TCP. Since this model captures the behavior of both TO and CA mechanisms, we expect the correlation to be the combination of the correlations for each of the individual mechanisms. Moreover, each mechanism is expected to dominate the correlation structure of the TCP model for loss probabilities where they have greater impact. Fig. 10 illustrates the results of the comprehensive TCP model for different loss probabilities for the case $R_{\text{TO}} = R$ together with the results for the TO and CA models. We indeed observe that for low loss probabilities (0.01) the correlation structure is completely dominated by the CA mechanism, while for higher probabilities
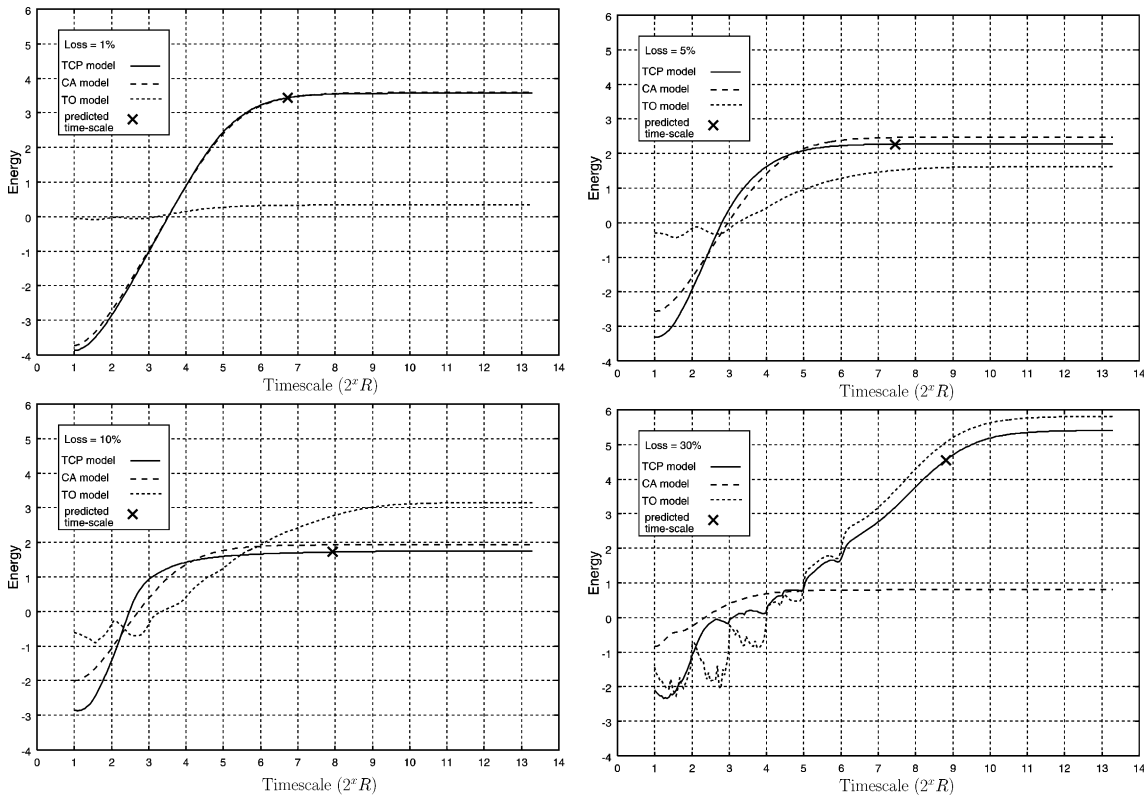
Fig. 10. Analysis of the comprehensive Markovian model of TCP with comparisons to the TO and CA model ($R_{\text{TO}} = R$).

(0.3) the structure is dominated by the TO mechanism. Under these loss probabilities, the results from the TCP model agree very well with both CA and TO results, respectively.

We also note that as loss probability increases the correlation structures change in a non-monotonic fashion, indicating the interplay of the CA and TO mechanisms. For intermediate and high loss probabilities (above 0.1), we notice a phase shift, characterized by a knee, in the correlation structure at time-scales between $4R$ and $16R$. This knee indicates the time-scale at which the CA mechanism diminishes its impact on the correlation structure, while the TO has its effect magnified. We also observe that the range of time-scales for which the model exhibits sustained correlation structure varies according to the loss probability as well as the R_{TO} to R ratio. In the case $R_{\text{TO}} = R$, for low loss the time-scales ranges are from $2R$ to $128R$, while for high loss it ranges from $2R$ to

$1024R$. This is consistent with the time-scales associated with the CA and TO models under low and high loss, respectively, as illustrated in Fig. 10.

Table 3 depicts both the time-scale values for $1/F_1$ obtained from Eq. (4) and the strength of the correlation given by $\psi(0)$ (both on $\log_2$), for the TCP protocol characterized by the comprehensive model. The corresponding values for $1/F_1$ for the

Table 3
Times-scales and correlation strength of TCP mechanism

Packet loss probability	$\log_2$ (max time-scale)		$\log_2(\psi(0))$—"energy"	
	$R_{\text{TO}} = R$	$R_{\text{TO}} = 2R$	$R_{\text{TO}} = R$	$R_{\text{TO}} = 2R$
0.01	6.69	7.69	3.57	3.61
0.05	7.44	8.44	2.28	2.34
0.10	7.92	8.92	1.74	1.90
0.20	8.57	9.54	3.47	3.99
0.30	8.78	9.67	5.41	5.64

comprehensive model are also marked with an $\times$ in the curves in Fig. 10. We note that increasing R_{TO} has a direct influence on the time-scales, in particular, it increases the range of times-scales by almost a constant factor of two. However, the correlation strength does not seem to follow any regular increase, even though a small increase is present. An interesting observation can be made on the correlation strength of the mechanism across the range of loss probabilities. The value is high for low loss probability, decreases and then increases again to a larger value, as the loss probability increases. We note that the trend of the $\psi(0)$ was either monotonically increasing or decreasing for the TO and CA model, respectively, as loss probability increases. The current trend in the strength of the correlation supports our claim that the correlation of TCP is the combination of the correlation of its internal mechanisms, with each mechanism dominating at particular loss probabilities.

5.2. Influence of packet loss process

It should be noted that the comprehensive TCP model assumes a loss process that is different than the one assumed in either of the individual models for CA and TO. In the former, the loss process is bursty within a round while in the latter it is independent and characterized by a Bernoulli process. However, our results in Fig. 10 suggest that the dependency of the traffic correlation structure on these two loss processes is negligible. To understand this result, recall our assumption that the time to transmit a window worth of packets is smaller than a round-trip time, which is more adequate for a wide-area network where the RTT tends to be larger. In separate work, the authors of [37] have shown that under the above assumption, the average throughput of a single TCP session when submitted to a correlated loss process and under a process where losses are iid is remarkably close to the actual measured throughput. This result shows that the loss processes investigated in [37] does not affect the average behavior of TCP.

A conjecture to explain this phenomena is that, under the above assumptions, packet loss correlations quickly vanish over short time-scales (on the order of one RTT) in real networks. We conducted calculations of conditional packet loss probabilities on our simulation traces and the results show that packet loss correlations disappear after an average window size of packets, which is smaller than one RTT. Thus, we believe that realistic packet loss processes for the scenario described here will tend to be uncorrelated for lags exceeding one RTT. Since this work focus on time-scales beyond a RTT, this explains the close match in the results of the two models which assume different packet loss processes. This result also justifies the use of a Bernoulli loss process for the TO and CA model, since losses tend to be independent beyond a RTT.

6. Realistic simulation scenario

In Section 4 we simulated a very simple network scenario consisting of a single TCP flow and a Bernoulli loss process, to show the behavior of the traffic correlation structure. We now consider a more realistic simulation scenario, where multiple TCP flows compete for bandwidth in a bottleneck link with a finite buffer. Losses are generated by overflow, each time a packet arrives to a full buffer queue.

This scenario is illustrated in Fig. 11. We will consider a fixed number n of infinite TCP sources and a variable number of short TCP sources that arrive according to a Poisson process with rate γ flows/s. The length of each short TCP flow is geometrically distributed with mean 50 packets, while the infinite sources always have packets to send. Each TCP source adds a uniformly distributed random delay between 0 and D before transmitting each packet. The value of D is on the order of the transmission time of a few packets and much smaller than an RTT. This random component models delay variations that would be introduced in a wide-area network. It also prevents the TCP flows from synchronizing in the simulation. The first link traversed by all flows is the bottleneck of the network, while the second adds different propagation delays. We monitor only a single infinite TCP flow which traverses the link with the highest propagation delay. In order for
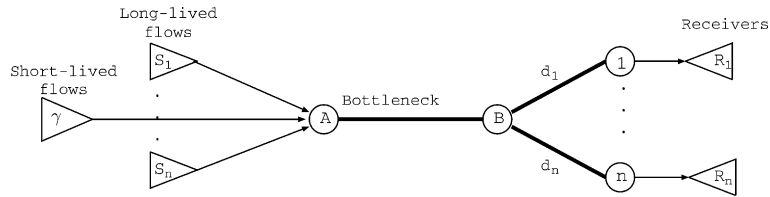
Fig. 11. A more realistic simulation scenario.

this TCP flow to perceive different loss probabilities, we vary the number of infinite TCP flows and the rate γ of short flows entering the bottleneck link, keeping constant all other parameters.

The bottleneck link speed was set to 2500 packets/s with zero propagation delay and assigned a buffer capacity of 160 packets. The second link had very high bandwidth and infinite buffer capacity. The propagation delay in the second link varied from 130 to 5 ms with increments of 2.5 ms for each link. To achieve an average loss probability of 0.01, 0.05, 0.1 and 0.2 the number of infinite TCP flows used was 27, 50, 80 and 110, while the arrival rate of short TCP flows was 5, 12.5, 20 and 35, respectively. As a result from the arrival rate, the average number of short TCP flows traversing the link at any point in time was 10.8, 40.6, 84.2 and 328.1, for the respective loss probabilities.

The above parameters ensure that the TCP flow being monitored complies with our earlier assumption that the average RTT is longer than the time to transmit a full window of packets. We obtained the TCP estimates for the RTT and R_{TO} in the ns simulator and calculated the average RTT and R_{TO} seen by our monitored flow for each simulation run. The ratio between R_{TO} and RTT, when rounded to the nearest integer, was 2 for all loss probabilities. To compare the simulation results with the Markovian model, we set this ratio to two in the model ($R_{\text{TO}} = 2R$).

Fig. 12 illustrates the wavelet analysis of the traffic sending rate of the infinite TCP source being monitored for different loss probabilities. We observe the same trend as in Sections 4 and 5: the presence of sustained correlation for a finite range of time-scales which varies according to the loss probability. Moreover, the dominance of the CA and TO mechanisms in the traffic correlation

structure for low and high loss probabilities, respectively. Due to the large scale of the simulation scenario, it is computationally expensive to obtain tight confidence intervals for the larger time-scales (all simulations ran for the longest possible time allowed by ns simulator). However, we strongly believe that the PSD is a flat curve for time-scales larger than 2^{11} in all plots.

Fig. 12 also plots the results obtained for the comprehensive TCP model to enable a direct comparison of the two correlation structures. We note that for low (0.01) and high (0.2) loss probabilities both the realistic simulation and the comprehensive model exhibit very similar traffic correlation structures. Under these loss probabilities the range of time-scales of both results is practically identical: $2R$ to $128R$ and $2R$ to $1024R$, respectively. The analytical curves are marked with an $\times$ at the predicted time-scale given by Eq. (4), which agrees reasonably well with the simulation curves. We note that for intermediate loss probabilities (between 0.1 and 0.2), the match between the simulation results and the comprehensive model is not as good. There are several possible explanations for this mismatch, including: (i) some behavior in the implementation of the TCP object in the ns simulator not captured by the comprehensive model; (ii) the assumption in the model that the RTT is constant, which does not hold true for the simulations; (iii) for intermediate loss probabilities, the packet loss process of the simulation is not similar to the loss process of the comprehensive model. Another source of discrepancy is the smoothing and the bias in the estimate used by the wavelet analysis, as discussed in Section 3.

Even though there are some discrepancies between the simulation and model results for
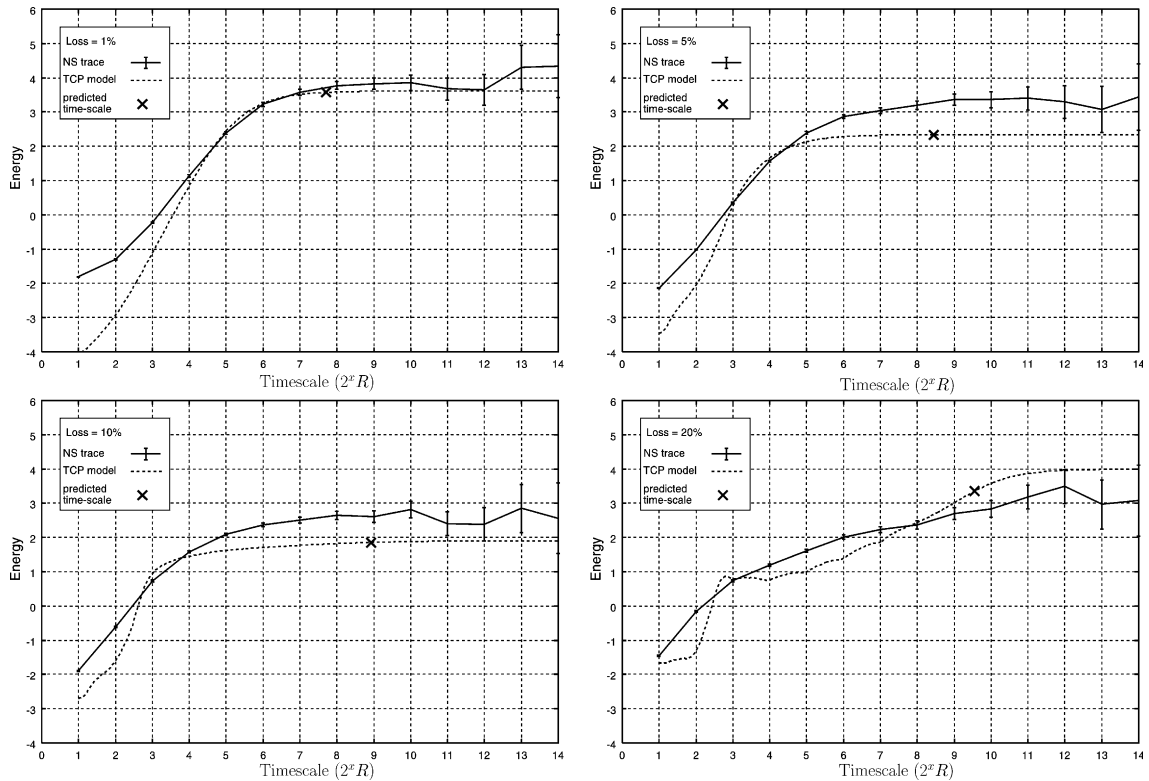
Fig. 12. Wavelet analysis of simulation traces from the realistic scenario and comparison with comprehensive TCP model ($R_{\mathrm{TO}} = 2R$).

particular loss probabilities, our observation that sustained correlation of TCP traffic spans a finite range of time-scales, which we can predict with reasonable accuracy, still holds. Moreover, the simulation results also show that the interplay between the CA and TO mechanism is the cause for this correlation. Note that the correlation structure observed is not an artifact of the Markovian model or the loss process, but is actually present and inherent to TCP; in particular, it arises from the CA and TO mechanisms.

7. Conclusion

In this paper we demonstrate that the TCP protocol can generate traffic with sustained correlation behavior over an analytically predictable finite range of time-scales. In particular, we point out that the TO and CA mechanisms are respon-

sible for generating this correlation structure. We show that under low loss probabilities (0.01) the traffic correlation structure is dominated by the CA mechanism while the TO mechanism has minimal impact. In contrast, under high loss probability (0.3) the TO mechanism has a dominant impact on the traffic correlation, while the effect of the CA mechanism is minimal. This sustained correlation structure in the traffic ranges from the time-scale of one RTT to a few orders of magnitude above the RTT (1024 RTT), and this range is dependent on both the loss probability and the value of R_{TO}. We provide separate Markovian models for each internal mechanisms which, under the proper loss probabilities, accurately predicts the range of time-scales and the strength of the sustained correlation structure of the packet sending rate of a TCP source. Our claim is supported by the analysis of a comprehensive TCP model that is available in the literature and has

been validated by others. The simulation results obtained also agree well with the results predicted by the models, which corroborates our conjecture and validates our models.

There are many studies in the literature that try to explain the origin of "long range dependence" in network traffic, including attributing this phenomena to the behavior of the TCP protocol. Contrary to these latter studies, we show that TCP alone cannot generate traffic with sustained correlation structure that extends to arbitrary large time-scales. The existence of an upper bound on the time-scale is related to the inherent finiteness in the time-scales of TCP's internal mechanisms (TO and CA). However, we show that under proper circumstances, TCP can generate traffic with sustained correlation structure over a possibly significant range of time-scales, which can be analytically predicted.

We note that the effects of TCP should be considered when attributing the origin of statistical traffic characteristics to some particular phenomena, since the vast majority of network traffic is carried by TCP connections. We believe that statistical properties of network traffic is caused by different factors in different protocol layers, among which TCP definitely plays an important role.

References

[1] W.E. Leland, M.S. Taqqu, W. Willinger, D.V. Wilson, On the self-similar nature of ethernet traffic, IEEE/ACM Transaction on Networking 2 (1994) 1–15.

[2] V. Paxson, S. Floyd, Wide-area traffic: the failure of poisson modeling, IEEE/ACM Transactions on Networking 3 (1995) 226–244.

[3] M.E. Crovella, A. Bestavros, Self-similarity in world wide web traffic: evidence, possible causes, IEEE/ACM Transactions on Networking 6 (1997) 835–846.

[4] V. Misra, W. Gong, A hierarchical model for teletraffic, in: Proceedings of the 37th Annual IEEE CDC, 1998.

[5] M. Krunz, A. Makowski, Modeling video traffic using M/G/Infinity input processes: a compromise between markovian, Ird models, IEEE Journal on Selected Areas in Communications 16 (1998) 733–748.

[6] S. Robert, J.Y.L. Boudec, On a Markov modulated chain with pseudo-long range dependences, Performance Evaluation 27 & 28 (1996) 159–173.

[7] A.T. Andersen, B.F. Nielsen, An application of superpositions of two-state Markovian sources to the modelling of self-similar behaviour, in: Proceedings of IEEE INFOCOM, 1997, pp. 196–204.

[8] K. Park, G. Kim, M. Crovella, On the effect of traffic self-similarity on network performance, in: Proceedings of the SPIE International Conference on Performance Control of Network Systems, November 1997.

[9] Z. Liu, P. Nain, D. Towsley, Z.-L. Zhang, Asymptotic behavior of a multiplexer fed by a long-range dependent process, Journal of Applied Probability 36 (1999) 105–118.

[10] M. Grossglauser, J. Bolot, On the relevance of long-range dependence in network traffic, IEEE/ACM Transactions on Networking (1998).

[11] S. Manthorpe, I. Norros, J.Y.L. Boudec, The second-order characteristics of TCP, in: Proceedings of Performance '96, Lausanne, October 1996. Presented in the self-similarity hot-topic session.

[12] A. Veres, M. Boda, The chaotic nature of TCP congestion control, in: Proceedings of IEEE INFOCOM, April 2000.

[13] W. Feng, P. Tinnakornsrisuphap, The failure of TCP in high-performance computational grids, in: Proceedings of SC 2000: High-Performance Networking and Computing Conference, November 2000.

[14] L. Guo, M. Crovella, I. Matta, TCP congestion control and heavy tails, Tech. Rep. BUCS-TR-2000-017, Department of Computer Science, Boston University, 2000.

[15] B. Siklab, K.S. Vastola, The effect of TCP on the self-similarity of network traffic, in: Proceedings Conference on Information Science and Systems, The John Hopkins University, March 2001.

[16] A. Feldmann, D.R. Figueiredo, B. Liu, V. Misra, D. Towsley, W. Willinger, Does TCP produce self-similar traffic? Tech. Rep. 01-32, Department of Computer Science, University of Massachusetts, August 2001.

[17] J.M. Peha, Retransmission mechanisms and self-similar traffic models, in: Proceedings of IEEE/ACM/SCS Communications Network and Distributed Systems Modeling and Simulation, January 1997, pp. 47–52.

[18] D.R. Figueiredo, B. Liu, V. Misra, D. Towsley, On the autocorrelation structure of TCP traffic, Tech. Rep. 00-55, Department of Computer Science, University of Massachusetts, November 2000.

[19] L. Guo, M. Crovella, I. Malta, How does TCP generate pseudo-self-similarity?, in: Proceedings of MASCOTS '01, Cincinnati, Ohio, August 2001.

[20] R.H. Riedi, J.L. Véhel, TCP traffic is multifractal: a numerical study, Tech. Rep. RR-3129, INRIA Research Report, March 1997.

[21] G. Wornell, A. Oppenheim, Estimation of fractal signals from noisy measurements using wavelets, IEEE Transactions on Signal Processing 40 (1992) 611–623.

[22] A. Veres, Z. Kenesi, S. Molnár, G. Vattay, On the propagation of long-range dependence in the internet, in: Proceedings ACM SIGCOM 2000, Stockholm, Sweden, August 2000.

[23] D.E. Comer, Internetworking with TCP/IP, vol. 1, third ed., Prentice-Hall, Englewood Cliffs, NJ, 1995.

[24] S.T. Satchell, H. Clifford, Linux IP stack commentary, Coriolis (2000).

[25] K. Claffy, G. Miller, K. Thompson, The nature of the beast: recent traffic measurements from an internet backbone, in: Proceedings of Inet '98, Geneva, Switzerland, The Internet Society, July 1998, pp. 21–24.

[26] J. Padhye, V. Firoiu, D. Towsley, J. Kurose, Modeling TCP throughput: a simple model and its empirical validation, in: Proceedings ACM SIGCOMM'98, Vancouver, CA, September 1998, A longer version is available as UMass CMPSCI Tech. Rep. 98-08.

[27] E. de Souza e Silva, R.M.M. Leão, The TANGRAM-II environment, in: Proceedings of the 11th International Conference on Modelling Tools and Techniques for Computer and Communication System Performance Evaluation (TOOLs'2000), Schaumburg/Illinois, USA, Springer, March 2000, pp. 366–369, Lect. Notes in Comp. Sci. 1786.

[28] R. Leão, E. de Souza e Silva, S. de Lucena, A set of tools for traffic modeling analysis and experimentation, in: Proceedings of the 11th International Conference on Modelling Tools and Techniques for Computer and Communication System Performance Evaluation (TOOLs'2000), Schaumburg, USA, Springer, March 2000, pp. 40–55, Lect. Notes in Comp. Sci. 1786.

[29] P. Abry, D. Veitch, Wavelet analysis of long-range-dependent traffic, IEEE Transactions on Information Theory 44 (1) (1998) 2–15.

[30] D. Veitch, P. Abry, A wavelet based joint estimator of the parameters of long-range dependence, IEEE Transactions on Information Theory 45 (3) (1999) 878–897.

[31] S. qi Li, C.-L. Hwang, Queue response to input correlation functions: discrete spectral analysis, IEEE/ACM Transactions on Networking 1 (5) (1993) 522–533.

[32] L.E. Franks, Generating $1/f$ noise, unpublished report.

[33] L.E. Franks, Signal Theory, Dowden & Culver Inc., 1981.

[34] UCB, LBNL, VINT, Network simulator—NS (version 2). Lawrence Berkeley National Laboratory, University of California, Berkeley URL http://www-nrg.ee.lbl.gov/vat.

[35] M. Mathis, J. Mahdavi, S. Floyd, A. Romanow, TCP selective acknowledgement options. RFC 2018, April 1996.

[36] J. Padhye, V. Firoiu, D. Towsley, A stochastic model of TCP Reno congestion avoidance and control, Tech. Rep. 99-02, Department of CS, University of Massachusetts, February 1999.

[37] E. Altman, K. Avrachenkov, C. Barakat, A stochastic model of TCP/IP with stationary random losses, in: SIG-COMM, 2000, pp. 231–242.

[38] M. Marsan, E.S. e Silva, R.L. Cigno, M. Meo, A markovian model for TCP over ATM, Telecommunication Systems Journal 12 (4) (1999) 341–368.

[39] C. Casetti, M. Meo, A new approach to model the stationary behavior of TCP connections, in: Proceedings of IEEE INFOCOM, 2000, pp. 367–375.

[40] M. Handley, S. Floyd, TCP-friendly simulations. http://www.aciri.org/mjh/results.ps.gz.

[41] J. Bolliger, T. Gross, U. Hengartner, Bandwidth modelling for network-aware applications, in: Proceedings of IEEE INFOCOM, 1999.

Daniel Ratton Figueiredo received the BS degree and the M.Sc. degree in computer science from the Federal University of Rio de Janeiro, Brazil in 1996 and 1999, respectively. He is currently pursuing the Ph.D. degree in the Department of Computer Science at the University of Massachusetts, Amherst. His research interests include modeling and analysis of computer networks, and in particular, mechanisms used in the Internet.
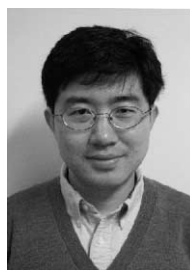
Benyuan Liu received the BS degree in physics from the University of Science and Technology of China in 1994. He received the MS degree in physics from Yale University in 1997. He is currently pursuing the Ph.D. degree in the Department of Computer Science at the University of Massachusetts, Amherst. His research interests include performance modeling and analysis of large scale heterogeneous computer networks.

Vishal Misra received a B.Tech from the Indian Institute of Technology, Bombay in 1992 and a Ph.D. from the University of Massachusetts, Amherst in 2000. He joined the Computer Science department of Columbia University as an Assistant Professor in 2001. He also holds a joint appointment in the Electrical Engineering department at the same University. His research interests include the modeling, analysis and development of algorithms for Communication Networks.

Don Towsley holds a B.A. in Physics (1971) and a Ph.D. in Computer Science (1975) from University of Texas. From 1976 to 1985 he was a member of the faculty of the Department of Electrical and Computer Engineering at the University of Massachusetts, Amherst. He is currently a Distinguished Professor at the University of Massachusetts in the Department of Computer Science. He has held visiting positions at IBM T.J. Watson Research Center, Yorktown Heights, NY (1982–1983); Laboratoire MASI, Paris, France (1989–1990); INRIA, Sophia-Antipolis, France (1996); and AT&T Labs—Research, Florham Park, NJ (1997). His research interests include networks, multimedia systems, and performance evaluation. He currently serves on the Editorial board of Performance Evaluation and Journal of the ACM has previously served on several editorial boards including those of the IEEE Transactions on Commu-

nications and IEEE/ACM Transactions on Networking. He was a Program Co-chair of the joint ACM SIGMETRICS and PERFORMANCE '92 conference. He is a member of ACM and ORSA, and Chair of IFIP Working Group 7.3. He has received the 1998 IEEE Communications Society William Bennett Paper Award and three best conference paper awards from ACM SIGMETRICS. Last, he has been elected Fellow of both the ACM and IEEE.