

A Stochastic Model of TCP and Fair Video Transmission

Stephan Bohacek

Department of Electrical and Computer Engineering

University of Delaware

Newark, DE 19716

bohacek@eecis.udel.edu

Abstract—A stochastic model of TCP is developed. Unlike many other models, this model accounts for variations in latency and loss probability. A major strength of this model is that it easily produces the probability distribution of the congestion window. Thus, the mean as well as the median and percentiles can be found. It is shown that the mean congestion window can be far larger than the median. Other new insights include the effect of the rate of change of the latency on the performance of TCP. Specifically, this model predicts the well-known TCP-friendly formula only if the round-trip time rapidly varies. However, if the round-trip time does not vary quickly, then the TCP-friendly formula may not hold. Both rapidly and slowly varying round-trip times have been observed in real networks. As an application of the model, the question as to when a video can be fairly transmitted is addressed. If it is possible to transmit the video, the model yields the distribution of the size of the receiving buffer required to avoid underflow. Since the distribution can be found, it is possible to select a buffer size so that a specified percentage of users will view the video without interruption.

I. INTRODUCTION

There has been extensive work focused on determining sending rates for non-TCP flows that are compatible with TCP flows. The basic approach is to send the non-TCP traffic at a rate similar to the rate that TCP would send data if the TCP flow was sending data in the same environment. By environment, we mean round-trip time and loss probability. Implementations determine the environment through estimation and determine sending rate through modeling. The focus of this paper is on modeling and has three objectives.

The first objective of this paper is to develop a new model of TCP. The difference between the work presented here and previous work is that models developed here provide a more detailed understanding of TCP's performance and incorporates often ignored aspects. In particular, the model includes a dynamic model of the round-trip time and loss probability. In most other modeling efforts, the round-trip time and loss probability are taken to be constant. It will be shown that the rate at which the round-trip time varies plays an important role in the performance of TCP.

A second objective of this paper is to demonstrate the utility of stochastic differential equation (SDE) modeling and probability distributions. While others have utilized SDEs, most modeling approaches have focused on average behavior (an exception is [1]). While the mean does provide insight, one should be careful when making implementation decisions

based on mean alone. For example, it has been suggested that under a fixed environment, a flow that sends data at the ensemble average rate of TCP flows is compatible to TCP flows. However, a more reasonable alternative is that the fair sending rate is the *median* TCP sending rate. As can be noticed by examining Figure 9, the distribution of the TCP congestion window may have a large tail. Hence, the mean may be far larger than the median (e.g., Figure 11). In this case, the mean TCP sending rate will be overly aggressive, whereas if the median sending rate is utilized, then it is assured that 50% of the flows would send faster and 50% of the flows would send slower.

The final objective of this paper is the application of this modeling approach to TCP-friendly unicast video transmission. Much of the work in this direction [2], [3], [4], [5] assumes that video requires constant bit-rate and hence, the objective has been to determine a constant rate that is compatible with TCP. However, modern video compression techniques (e.g., MPEG) use variable bit-rate (VBR) compression. Hence, these streams do not require a constant bit-rate. Indeed, squeezing the VBR video stream into a constant bit rate stream is technically challenging and typically leads to a decrease in quality. As an alternative, this paper directly focuses on the compatibility question; is it possible to send the video while remaining fair to TCP flows? In order to maintain compatibility with TCP flows, the video is transmitted at the median TCP sending rate given the round-trip time; a quantity that varies with round-trip time. Then, the question as to whether the video can be fairly transmitted is formulated as a queueing problem where the input to the queue is governed by the VBR video stream and the output is governed by the median TCP sending rate. If the queue fills without bound, then the video stream is not compatible with the network conditions.

While an outgoing queue forces the VBR video stream to be fair, a receiving buffer is necessary to accept data at the fair TCP rate and output it to the monitor at the VBR video rate. If this buffer empties, then frames are skipped leading to a greatly degraded video. In order to avoid this underflow condition, the receiving buffer is filled before the video playback begins. In section IV, the model of TCP and a model of VBR video are combined to determine the distribution of the receiving buffer size required to avoid

underflow

It must be emphasized that the mean required buffer size is not appropriate. Typically, the objective is for a large percentage of the viewers to view the video without underflow. Thus, the receiver buffer is prefilled to a size so that Q percent of the viewers will not experience underflow. If the mean required buffer size is used instead of the Q^{th} percentile, it is not possible to guarantee even half of the viewers could receive the transmission without underflow. Figure 12 shows that case where the 99% percentile required buffer size is two orders of magnitude larger than the mean required buffer size.

The model presented is far from the first model of TCP. The most popular models are the so called TCP-friendly models [6], [7], [8] which yield a simple formula that relates average loss probability, average round-trip time and average throughput. A principle advantage of these models is their simplicity. However, these models provide only average behavior of TCP. The models presented here yield the entire distribution of TCP's window size. In [1], the distribution of TCP's window size is also investigated. However, in [1], it is assumed that the loss probability varies with the congestion window. While this is true for a single flow over a single bottleneck, our experience is that in general the round-trip time is a far stronger indicator of loss probability than congestion window. Furthermore, the model developed here directly addresses the stochastic nature of large networks. In [9], SDEs are also employed. However, the object in [9] is the average behavior and variation of round-trip time and loss probability are again neglected. Other modeling approaches to TCP include [10], [11], and [12]. The last is particularly relevant since it also takes a stochastic approach. However, the approach in [12] is substantially different from the one taken here. The model developed in this paper is applicable to both short-term flows as well as long-term flows. However, the three-way hand shake is not modeled. A model of short-lived connections along with the three-way hand shake can be found in [13].

The paper consist of three principle sections. In the first section, the model of TCP is developed. To this end, two submodels are presented; a model of round-trip time and a model for drops. While the model of round-trip time has been developed elsewhere, the model for loss probability is unique. These models are combined into the TCP model in Section II-C. Section III discusses the TCP model. Specifically, Section III-A validates the model with ns-2 simulations while Section III-B utilize the model to develop some new insights into TCP's performance. Section IV applies the TCP model to determine when a video can be fairly and successfully transmitted.

Due to space limitations, no proofs are included and as little mathematics is presented as possible. Details of the mathematical technicalities are left to a technical report. Furthermore, because of space limitations, the details of some techniques are not discussed and a discussion of the measurement methods utilized is skipped. A longer version of this paper will include these critical details.

II. A DIFFUSION MODEL OF TCP

No attempt is made to exactly model a particular variant of TCP, e.g. TCP-SACK or RENO. Instead, the approach is to model the basic principles of TCP congestion control and develop an idealized model of TCP congestion control. The principles of TCP are that the window size increases by one every round-trip time, divides by two when a drop event is detected, goes to one when a time-out occurs and exponentially increases after a time-out until it reaches its size before the time-out divided by two. Features of TCP that are ignored include delayed ACKs, the details of fast recovery, the subtleties of time-out and interaction with the application layer. From this idealized point of view, the only inputs to TCP are packet loss and the round-trip delay experienced by data packets and their ACKs. Thus, in order to model TCP, round-trip time and packet loss must be modeled.

A. A Diffusion Model of round-trip Time

There has been extensive work focused on understanding the distribution of latency, e.g. [14], [15], [16]. However, there has been less work on developing a dynamical model of latency. Work that has investigated the dynamics of the round-trip time includes [17] and [18]. In [18], diffusion models of round-trip delay are developed. These models are validated with extensive real network measurements. These models and their merits will be briefly reviewed.

Let R_t be the variable part of the round-trip time experienced by a packet send at time t . Thus, the actual round-trip time is $RTT_t = R_t + T$ where T accounts for fixed delays such as propagation delay, transmission delay, etc. R_t is dominated by queuing delay but may also include effects such as address lookup. However, the effect of delays between network layers and transport or application layers is not modeled. The first model presented in [18] is a simple three parameter mean-reverting diffusion model

$$dR_t = \frac{\sigma^2}{2} (\lambda - \phi R_t) dt + \sigma \sqrt{R_t} dB_t, \quad (1)$$

where ϕ , λ and σ are scalar parameters and B_t is Brownian motion. This system is known as CIR model of interest rates and has been widely studied and utilized in finance [19]. The stationary density of the system is given by the gamma density

$$p(r) = \frac{(\phi)^\lambda}{\Gamma(\lambda)} r^{\lambda-1} e^{-\phi r}, \quad (2)$$

where Γ is the Gamma function. In the case where there are λ M/M/1 queues between source and destination, and each queue has an arrival rate a and a departure rate b , with $a < b$, $a \approx b$ (i.e., heavy traffic) and $b - a = \phi$, then the probability density of the latency is given by (2) [20]. Thus, there is a bit of rational behind the model. A very useful feature of this model is that the transition probabilities are known,

$$p(R_t | R_0, \phi, \lambda, \sigma) = c \exp(-u - v) \left(\frac{v}{u}\right)^{q/2} I_q(2\sqrt{uv}), \quad (3)$$

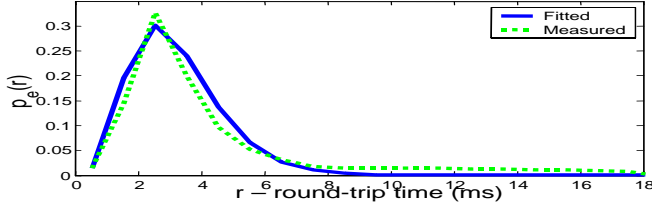


Fig. 1. Observed and Fitted Density of the Round-trip Time.

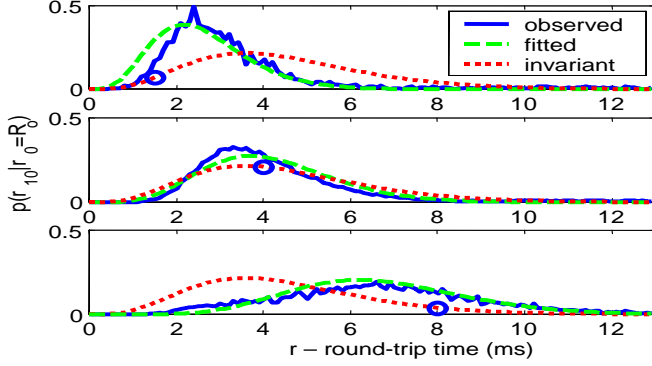


Fig. 2. The above plots show the probability density of the increments of the round-trip time, $R_{10} - R_0$, for various initial conditions R_0 . The solid curve shows the observed density, the dashed curve shows the density with the estimated σ . The dotted curve shows the invariant density for reference purposes. Along the invariant density curve are circles that indicate the initial condition R_0 .

where

$$\begin{aligned} c &:= \frac{\phi}{1 - \exp\left(-\frac{\sigma^2}{2}\phi t\right)}, & u &:= cy \exp\left(-\frac{\sigma^2}{2}\phi t\right) \\ v &:= cx, & q &:= \lambda - 1 \end{aligned}$$

and I_g is the modified Bessel function of the first kind of order g . A typical stationary distribution is shown in Figure 1. The transition probabilities are shown in Figure 2. These figures show both the fitted and observed round-trip times. This data was collected from a connection with source in San Jose and destination in Los Angeles. This data was collected at 2pm on March 26, 2001. At this time, this connection experienced its heaviest congestion of the day. Note that while the quality of the fit is good, Figure 1 show a problem with fitting the tail. Thus, while large round-trip times are rare, this model under predicts their rate of occurrence. At times of less congestion (most of the day), the tail is far smaller and this model performs quite well.

During periods of high congestion, a better fitting model can be had at the expense of adding more parameters. A six parameter model of round-trip time is

$$\begin{aligned} dR_t &= \sqrt{\sigma^2 R_t^\rho} dB_t \\ &+ \frac{\sigma^2}{2} R_t^{\rho-1} \left(\frac{(\delta R_t^\delta + \gamma \beta^\gamma R_t^\gamma)}{(R_t^\delta + \beta^\gamma R_t^\gamma)} + \rho - \phi 2 \ln(R_t) \right) dt. \end{aligned} \quad (4)$$

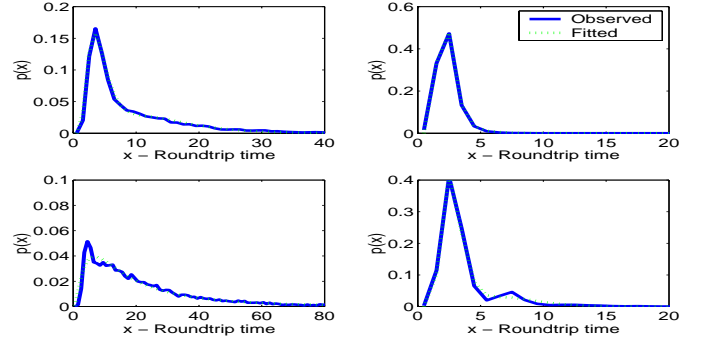


Fig. 3. Observed and fitted stationary distributions of the round-trip time. The top left hand plot shows the round-trip time for an intra-California connection at 3:30pm on March 26, 2001. The right hand plot show the round-trip time for the same connection at 1am on March 27 2001. The lower two plots show the round-trip times at the same times as the upper plots but for a transcontinental connection.

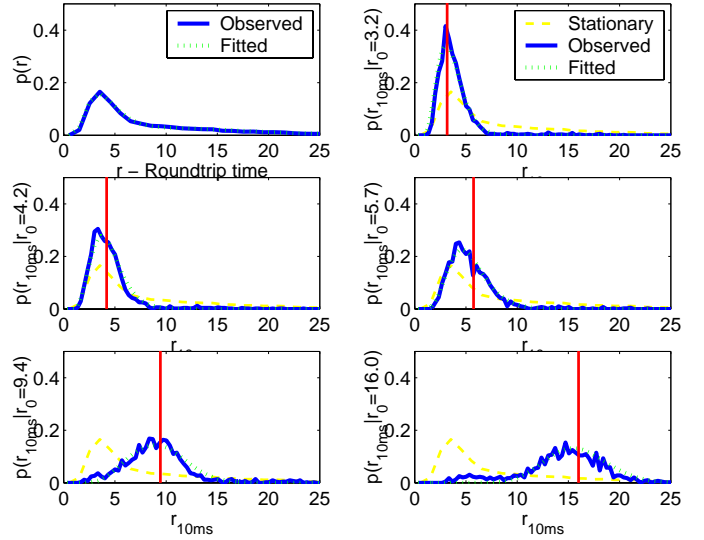


Fig. 4. Observed and fitted transition probabilities for a intra-California round-trip time on March 26 at 3:30pm. The *upper left* shows the stationary density. The remaining plots show the observed transition probability (solid line), the fitted transitions probability (dotted line) and, for reference, the stationary density. The vertical solid line indicates the initial condition.

In this case the invariant density is given by

$$h(r) = C (r^\delta + \beta^\gamma r^\gamma) \exp\left(-\phi (\ln(r))^2\right). \quad (5)$$

Figure 3 shows some observed and fitted stationary distributions while Figure 4 shows the observed and fitted transition probabilities.

While it may be argued whether the above models are accurate over all connections, the above and other measurements strongly suggests that these models perform well over a number of connections. Clearly, there is a need to understand the validity of these models over large portions of the network. While a large scale study is underway, its results will be left for the future.

Remark 1: As discussed in [18], the parameters of these

model vary. Our data indicates that the parameters vary slowly and can be taken to be constant over time intervals less than 1 or 2 minutes. Furthermore, the parameters may remain valid for a very long time, sometimes for a few hours. This behavior is to be expected since the models are dynamic and hence allow for variation of the round-trip time.

Remark 2: One aspect that is missing from these round-trip time models is the long-term correlation that have been observed by [17]. There are three reasons to not include such long-term correlations. 1. Our measurements have not shown long-term correlations. One explanation of [17] findings is that [17] was chiefly concerned with delay observed at the application layer. This delay is principally the sum of the network delay and delay induced by the operating system. Since operating systems experience periodic stalls (periodic signals have infinitely long-term correlations), it is not clear if [17] was actually measuring the operating system. 2. Short-term correlations are modeled. Making use of model parameter variation, this correlation can be extended in such a way that long-term correlation can be approximated. 3. The model is an approximation and is greatly simplified by using neglecting long-term correlations. The simplicity increases the utility of the model while not substantially degrading its accuracy.

B. A Probabilistic Model of Packet Loss

There has been extensive work on modeling packet loss. In [21], a small network is considered and a deterministic model for packet drops is developed. In [8] and [13], drops are assumed to be highly correlated over short time scales and independent over longer time scales. In [22], drops are assumed to be bursty. Furthermore, [22] makes no distinction between drop events and packet drops. Similarly, in this paper a drop event is when the congestion window divided by two, which can be made up of many lost packets, is referred to as a packet loss. In [12], drops events are modeled as a renewal process with various distributions; deterministic, Poisson, i.i.d and Markovian. A specific example of the model in [12] is developed in [9], where drop events are modeled by a Poisson process and, hence, the time between drop events are exponentially distributed and independent of the sending rate. In [23] and [1], this approach is generalized and drops events are modeled as a Poisson process where the intensity depends on TCP's congestion window size. In [24], a dynamic model of loss is developed. The model presented here is also dynamic. The difference between the two is that our model recognizes a strong dependence on the round-trip time.

Our loss model is along the lines of [23] and [1] in that packet loss is modeled as a Poisson process. However, we allow the loss rate to be a function of round-trip time and the sending rate. As a result of the loss rate depending on round-trip time, the model is able to remain accurate for a large set of operating conditions; light congestion with small round-trip times and heavy congestion with longer round-trip times.

Define the sending rate at time t of be V_t . The loss events are modeled as a Cox process¹ N_t , with rate $n(R_t, V_t)$. We

decompose this rate into

$$n(R_t, V_t) = V_t \cdot g(R_t).$$

Thus, the drop rate is a function of the sending rate and $g(R_t)$, the conditional loss probability. It is commonly assumed that the loss rate can be decomposed in this way. However, it appears that the assumption that the loss rate depends on the round-trip time is novel. Indeed, many have assumed the loss probability to be constant (exceptions include [24], [23] and [1]). The dependence on round-trip time is reasonable since, as queues fill, the round-trip time increases and the probability of a packet loss increases.

While many models of g are appropriate, a spline representation has proven useful. There has been extensive work smoothing observed data with splines [25]. To this end, define

$$g(R_t) = \alpha_0 T_0(R_t) + \alpha_1 T_1(R_t) + \dots + \alpha_n T_n(R_t),$$

where T_i are a set of function. For example, these could be Taylor series function, 1, x , x^2, \dots , Chebshev polynomials, or, as we have chosen them, splines.

Since the round-trip times of dropped packets is not observed, the conditional loss probability is not directly observable. Define d_k to be one if the k^{th} packet is lost and zero otherwise. Then, through empirical averaging, the directly computable quantities are

$p(R_{t_k} | R_{t_{k-1}})$ – the transition probability of the round-trip time.

$P(d_k | R_{t_{k-1}})$ – the probability of packet being dropped given the last round-trip time.

The transition probability $p(R_{t_k} | R_{t_{k-1}})$ depends on the time interval $t_k - t_{k-1}$. For ease of presentation, we will assume that the packets are transmitted at a constant sending rate. The case of time varying sending rate is a straight forward extension.

A simple manipulation yields

$$\begin{aligned} p(d_k, R_{t_k} | R_{t_{k-1}}) &= P(d_k | R_{t_k}, R_{t_{k-1}}) p(R_{t_k} | R_{t_{k-1}}) \quad (6) \\ &= P(d_k | R_{t_k}) p(R_{t_k} | R_{t_{k-1}}) \\ &= g(R_{t_k}) p(R_{t_k} | R_{t_{k-1}}), \end{aligned}$$

where it is assumed that given the present round-trip time, packet loss is independent of the past round-trip time. Integrating (6) yields

$$\begin{aligned} P(d_k | R_{t_{k-1}}) &= \int p(d_k, R_{t_k} | R_{t_{k-1}}) dR_{t_k} \quad (7) \\ &= \int g(R_{t_k}) p(R_{t_k} | R_{t_{k-1}}) dR_{t_k} \\ &= \int \sum_{i=0}^n \alpha_i T_i(R_{t_k}) p(R_{t_k} | R_{t_{k-1}}) dR_{t_k} \\ &= \sum_{i=0}^n \alpha_i \left(\int T_i(R_{t_k}) p(R_{t_k} | R_{t_{k-1}}) dR_{t_k} \right), \end{aligned}$$

¹A Poisson process is a Cox process when the rate is constant.

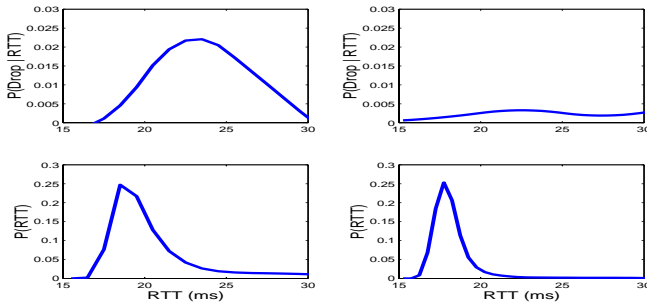


Fig. 5. Los Angeles to San Jose loss probability during the day (upper left) and night (upper right) and density of round-trip times for these connections (lower).

which is a linear system of equations. Therefore, given $p(d_k|R_{t_{k-1}})$ and $(\int T_i(R_{t_k})p(R_{t_k}|R_{t_{k-1}})dR_{t_k})$, it is straight forward to solve (7) for the coefficients a_i .

There is a minor difficulty with defining the loss probability as above. If the packet is dropped, then it is not initially clear how to define the round-trip time. The simplest solution to this is that models in the previous section determine $p(R_t|R_0)$ and these models, along with g , accurately model the loss probability. In this sense, loss probability is not dependent on round-trip time, but rather on the state of a model of round-trip time. Other rationalizations are possible. In the case that the packet is dropped, we could define the round-trip time as the latency that the packet would have experienced if had it not been drop. For example, if RED is used, then it is not difficult to conceive of the situation where a packet was not drop and the round-trip time it would have experience. With ECN, the packet is not actually dropped, so round-trip time can be measured.

Figure 5 shows observed round-trip times and the computed function g . Comparing the left to the right frames, it is seen that loss probability increases with congestion which increases with average round-trip time. However, these functions are peculiar in that the loss probability appears to decrease with round-trip time when the round-trip time is large. While this feature is interesting, its relevance is not clear as this decrease only occurs for large round-trip times and these round-trip times are rare.

More work is necessary to find a ubiquitous expression for the loss probability. For example, the dependence on the sending rate and the history of round-trip times has yet to be fully explored. Some may argue that the above methodology suffers from the fact that the actual round-trip time that would have been experienced by a dropped packet is poorly modeled by $p(R_t|R_0)$ since this round-trip time is never observed. However, by sending closely spaced packets (1 ms and less), we have been able to determine that this is a reasonable model. This result supports the assumption that over milliseconds time scales, the round-trip time varies at the same rate regardless if a packet is lost or not. Further discussion of loss models is not possible here. Indeed, this is a deep issue which requires study in its own right. The main point of this section is to present a

reasonable model for packet loss. The following work is not directly dependent on this model and can easily be adapted to other models.

C. A Diffusion Model of TCP

In this section, the above models are utilized to develop a model of TCP's congestion window. We start with a simply model where, when packets are not dropped, the congestion window increases by one in one round-trip and divides by two when a loss event is detected. Define the congestion window at time t to be X_t and recall that R_t is variable part of the round-trip time at time t and T is the fixed component of the round-trip time. In this case, a stochastic differential equation model of the congestion window is

$$\begin{aligned} dX_t &= \frac{1}{T+R_t}dt - \frac{1}{2}X_t dN_t \\ dR_t &= \mu(R_t)dt + \sigma(R_t)dB_t, \end{aligned} \quad (8)$$

where μ and σ are functions such as the ones given in Section II-A and N is a Cox process that counts the number of drops. This model is similar to and inspired by the one presented in [9]. However, here the fact that round-trip time is not constant is embraced. As will be seen, this variation plays an important role. Another difference is that the model of loss events is different; [9] assumes that loss the loss rate is independent of the sending rate as well as the round-trip time.

Model (8) can be extended to incorporated the effect of time-outs,

$$\begin{aligned} dX_t &= 1_{\{X_t \geq S_t\}} \frac{1}{T+R_t}dt - \frac{1}{2}X_t dN_t \\ &\quad - (X_t - 1) dM_t + 1_{\{X_t < S_t\}} \frac{\ln 2}{T+R_t} X_t dt \\ dS_t &= \frac{1}{2}X_t dM_t - S_t dN_t \\ dR_t &= \mu(R_t)dt + \sigma(R_t)dB_t, \end{aligned} \quad (9)$$

where M is a Poisson process that counts the number of time-outs and $1_{\{X_t < S_t\}}$ is the indicator function. Here, M and N are distinct processes and increment at the same time with probability zero. In this system S takes the role of slow start threshold (ssthresh). If X is less than S , then X increases at a rate such that it doubles every round-trip time. Note that when a time-out occurs, S is set to $\frac{1}{2}X$ and X is set to one. The congestion window then exponentially increases until it reaches S , at which time it increases at a rate $\frac{1}{T+R_t}$. After a time-out, exponential growth has completed and $S \neq 0$, the next drop sets $S = 0$, so $X > S$.

With these models, it is straight forward to determine the probability density functions of the congestion window. Let $p(x, r, t)$ be the probability density function for (8), i.e., let $p(x, r, t) = \frac{\partial^2}{\partial x \partial r} P(X_t < x, R_t < r)$. We will assume that

this and all necessary densities exists. Then $p(x, r, t)$ obeys

$$\begin{aligned} & \frac{\partial}{\partial t} p(x, r, t) \\ &= -\frac{\partial}{\partial r} (\mu(r) p(x, r, t)) + \frac{1}{2} \frac{\partial^2}{\partial r^2} (\sigma^2(r) p(x, r, t)) \\ & - n(x, r) p(x, r, t) \\ & - \frac{1}{T+r} \frac{\partial}{\partial x} p(x, r, t) + 2n(2x, r) p(2x, r, t), \end{aligned} \quad (10)$$

where $n(x, r)$ is the drop event rate discussed in subsection II-B. This partial differential equation representation of the density is known as Kolmogorov's forward equation and is a straight forward application of standard methods in stochastic differential equation [26]. It is not hard to show that the system is ergodic. Hence, with $p(x, r, t)$ found, the stationary density can be found by letting $t \rightarrow \infty$, i.e., $p(x, r) = \lim_{t \rightarrow \infty} p(x, r, t)$.

In the case of model (9), the density obeys

$$\begin{aligned} & \frac{\partial}{\partial t} p(x, r, s, t) \\ &= -\frac{\partial}{\partial r} (\mu(r) p(x, r, s, t)) + \frac{1}{2} \frac{\partial^2}{\partial r^2} (\sigma^2(r) p(x, r, s, t)) \\ & - (n(x, r) + m(x, r)) p(x, r, s, t) \\ & - (1_{\{s \leq x\}} + 1_{\{s > x\}} \ln 2) \frac{1}{T+r} \frac{\partial}{\partial x} p(x, r, s, t) \\ & + 1_{\{s=0\}} 2n(r, 2x) \int p(2x, r, u, t) du \\ & + 1_{\{x=1\}} 2m(r, 2s) \int p(2s, r, u, t) du. \end{aligned} \quad (11)$$

III. VALIDATION AND INSIGHTS

A. Validation

We briefly examine some ns-2 simulations. Thus, the topology utilized here is shown in Figure 6. The cross-traffic consisted of persistent TCP flows. Not all possible TCP flows were considered. Instead, only the following set of source destination pairs were considered

$$\begin{aligned} & (1, 5), (1, 6), (1, 7), (1, 8), (2, 6), \\ & (2, 7), (2, 8), (3, 7), (3, 8), (4, 8). \end{aligned}$$

The flow under test had source S and destination D. Each link had a propagation delay of 10ms and bandwidth of 10MB. The queues employed RED. Figure 7 shows the results of this simulation as well as the fitted densities and the theoretical density of the congestion window and observed density. Here, the simpler model of round-trip time model (1) was used and yields a marginal fit. Nonetheless, the theoretical and observed densities of the congestion window appear to be similar (lower left in Figure 7). Figure 7 also shows the joint density of round-trip time and congestion window. Other simulations show similar results.

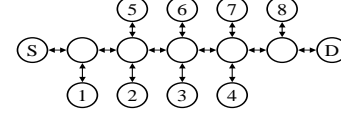


Fig. 6. The topology for the ns simulations. The flow of interest was a flow from S to D.

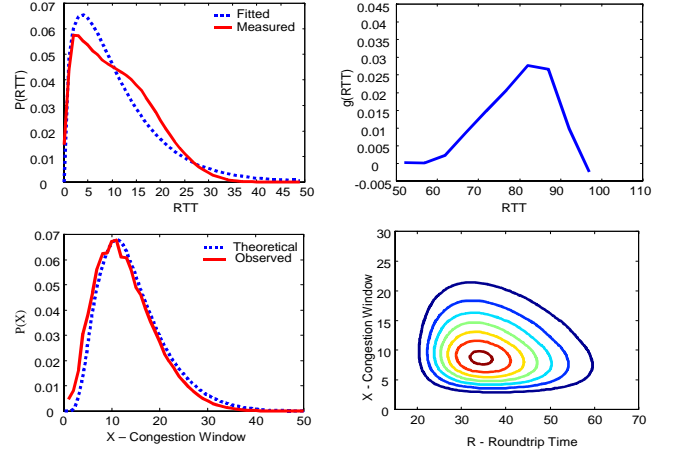


Fig. 7. ns-2 Validation. *Upper left*, Observed and Fitted Probability Density Function of the Round-trip Time; *right*, Fitted Conditional Packet Loss; *below left*, Observed and Theoretical Stationary Density of the Congestion Window; *right*, Contour Plot of the Joint Stationary Density of the Round-trip Time and Congestion Window, $p(x, r)$.

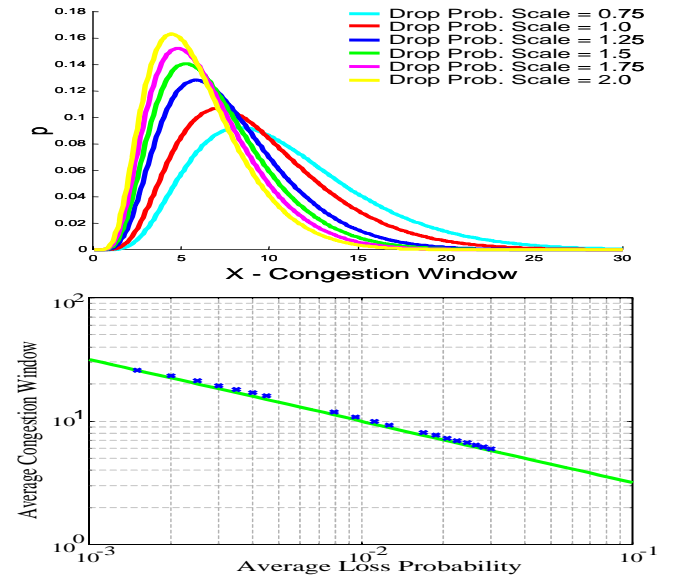


Fig. 8. *Upper*, The Stationary Distribution of the Size of the Congestion Window as the Loss Probability is Scaled; *lower*, The relationship between the Average Loss Probability and the Average Size of the Congestion Window.

B. The Effect of the Model Parameters

With this model it is trivial to vary the parameters and observe changes in the congestion window distribution. We restrict attention to the model of round-trip time (1). Two aspects of the model are investigated, the dependence on the loss probability and the dependence on the diffusion parameter σ . Measurements of a connection with source in San Jose and destination in Los Angeles were made during midday. From this data, the round-trip time model parameters and the conditional loss probability $\hat{g}(r)$ were determined. This loss probability was then scaled, i.e. $g(r) = scale \times \hat{g}(r)$. Figure 8 shows stationary density of the congestion window as the *scale* parameter varies. This figure also shows the relationship between the average loss probability, i.e., $\int g(r)p(r)dr$ vs. the average size of the congestion window. Note the apparent reciprocal relationship between the congestion window and the square-root of the average loss probability. This relationship has been extensively studied and is the basis for TCP friendly congestion control [6], [7].

Next, the *scale* parameter was set to one and the σ parameter was varied. σ controls how fast the transition probability converge to the stationary density. In other words, σ controls how fast round-trip time varies; a large σ yields faster variation. By examining (5), it is easy to see that the distribution of the round-trip time is independent of σ , that is $p(r)$ is independent of σ . Thus, *the average loss probability does not depend on σ* . Figure 9 shows a series of densities of the congestion window for different values of σ . Figure 9 also shows the relation between the average congestion window and the average loss probability. Note that these points must be on a vertical line since the average loss probability is not effected by σ . Figure 10 helps explain the effect of σ . This figure show the joint density of the congestion window and round-trip time for a small value of σ . This figure can be compared to the similar plot in Figure 7. Note that in Figure 10, the congestion window may take very large values when r is small. This is due to the fact that when round-trip time is small, the drop probability is small, so the congestion window is able to grow. If σ is small, then the condition of low drop probability is maintained for a longer time than if σ had been large. This effect is observed in real networks; during periods of light congestion the round-trip time and loss probability decreases allowing the congestion window to achieve a higher value.

Note that in Figure 9 that as σ increases, the density of the congestion window appears to converge. This convergence can be explained by recalling that for large σ , there is little correlation between different samples of round-trip time. Thus, in a very short amount of time, a flow would experience the full range of round-trip times and experience the average drop probability. Thus each flow, independent of its initial round-trip time, experience the same round-trip times and the same loss probability. In this case, there is little loss of generality when the round-trip time is replaced with its average value. This is essentially what was done in [9]. Thus, the work here

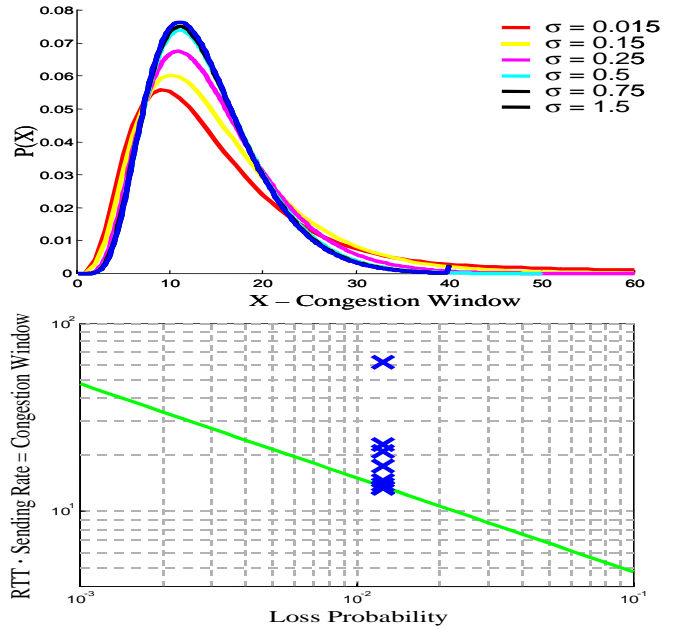


Fig. 9. Upper, Density of the Size of the Congestion Window for Different Values of σ ; lower, The Relationship Between the Average Loss Probability and the Average Size of the Congestion Window for Different σ . The average congestion window decreases monotonically as σ increases.

can be thought of an investigation into the case where σ is not necessarily large. The assumption that the average value of the round-trip time can replace the round-trip time is made by many researchers. Our network observation indicate that very small values of σ are observed. For example, $\sigma < 0.1$ is not uncommon.

Note that Figure 9 does not necessarily imply that the TCP-friendly formula is incorrect. In particular, for small σ , the round-trip time slowly varies. Hence, a short-term average of the round-trip time is perhaps very different from the actual, long-term average. Such a short-term average would yield a different relationship than the one shown in 9. This difference between the TCP friendly formula and model results presented here highlight the difficulty choosing of averaging windows or filters used to compute the averages required by the TCP friendly formula. In [27], experience has guided the choice of these parameter. However, the problem of window size and filtering methods appears to remain a difficult problem, e.g. see [28].

As mentioned in the introduction, the mean and median of the congestion window are not always the same. Figure 11 shows the mean and median as a function of σ . Note that for small σ , there is a substantial difference between mean and median. Figure 11 also shows the probability that the congestion window is less than the mean congestion window. Note that as σ decreases, flows that utilizes the mean congestion window would become more aggressive. It is interesting that in the results presented here the mean and median are nearly the same for large σ . Furthermore, the median remains fairly constant as σ varies. Hence, the mean

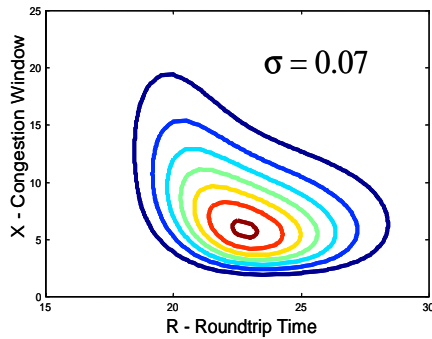


Fig. 10. Joint Density of the Congestion Window and the Round-trip Time for $\sigma = 0.07$.

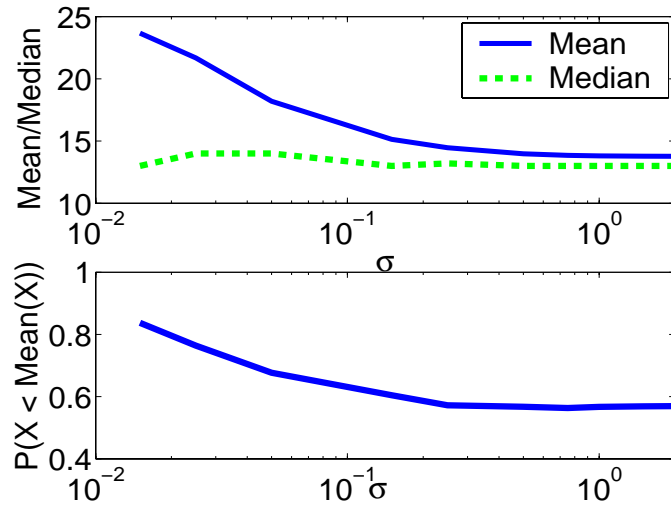


Fig. 11. *Upper*, Mean and the Median of the Congestion Window versus σ ; *lower*, The Probability that the Congestion Window is less than the Mean Value of the Congestion Window.

for large σ might be a reasonable estimate of the median for all σ . While more work is required, it is conceivable that the TCP-friendly formula is actually an estimate of the mean for large σ and, hence, a good estimate of the median for all σ .

IV. AN APPLICATION TO VIDEO SENDING RATE

Next the TCP model is employed to determine when if a unicast video transmission can be both fair and successful. We focus on previously recorded video (e.g., video on demand). However, with some modification, this work is applicable to live video as well.

The main difficulty with sending video data over networks without quality of service guarantees is that the compression scheme may require a sending rate that is too fast and not responsive to competing TCP flows. In this case, the video transmission may take an unfair proportion of the bandwidth. Thus, there is a need to match the video data rate to the sending rate that a TCP flow would send if it was in the same situation as the video flow, that is, the video sending rate must be TCP friendly [29]. However, TCP dictates a sending rate that

is rapidly oscillating according to congestion, whereas VBR compressed video bit-rate oscillates according to compression efficiency. One common approach has been to send data at TCP's average sending rate [2], [3]. Since the averaging leads to slow variations, the video is transmitted at this smoother rate. In parallel with this TCP smoothing, video smoothing techniques have been developed that lead to slowly varying video bit-rate [4], [5].

Here a different approach is taken in that smoothing is not the focus, rather determining compatibility is the objective. For a given network environment (round-trip time and loss probability), the distribution of the congestion window can be found and a fair sending rate can be determined. On the other hand, the dynamics of the video stream are also known, so the distribution of the video bit-rate can be determined. Hence, before the video is sent, it can be decided if the video transmission is compatible with the current network conditions. If they are not compatible, then either the video is not sent, the compression scheme is changed to a higher compression ratio (and therefore lower quality image), or, potentially, bandwidth could be purchased.

Most network video playback systems have a receiver buffer. This buffer is filled before playback begins. In this way, if video transmission rate must be momentarily decreased or the video stream momentarily requires more bandwidth, then the buffer can be momentarily emptied at a faster rate than it is filled. However, if the mismatch in rates persists for too long, the receiver buffer will underflow and result in missed frames, greatly degrading the quality. In order to avoid underflow, it is critical that the receiving buffer be prefilled to a level sufficient to accommodate any future mismatch between video bit-rate and TCP friendly sending rate.

Thus, before a video can be transmitted, there are two questions that must be answered. First, can the video be sent at all and second, if the video can be transmitted, how much must the receiving buffer be filled to avoid underflow. Before addressing these problems, a model of VBR compressed video is reviewed.

A. A Model for MPEG 1

There are a large number of models for VBR and a large number of compression schemes. For the approach developed here, the model and the compression scheme are not important. Since MPEG 1 video streams have been extensively studied, we consider these flows. The MPEG 1 standard calls for three classes of video frames; I, P and B frames. Of these, I frames are the least compressed and utilize no time correlation; they are compressed the same as a single JPEG image. B frames are the most compressed utilizing forward and backward time correlation. In terms of compression, P frames lie between B and I frames since they only utilize backward time correlations. A common implementation assumption is that the frames types follow a specific pattern. A single repetition of this pattern is called a group of pictures (GOP). A typical GOP is as follows I B B P B B P B B P B B . The video stream is made up to a stream of GOPs. It is not

uncommon that there is one more P,B,B pattern, making the GOP 15 frames long. Many models of VBR video have been developed. See [30] for a review of some models. We review a model presented in [31] that is specialized for typical full length motion pictures. Again, the approach here can easily be adapted to other video models.

[31] found that the distribution of the B and P frame sizes can be modeled as I.I.D. log-normal with parameters μ_B , σ_B , μ_P and σ_P . Furthermore, it was found that the size of the I frame is dependent on the scene. The size of the I frame is decomposed into $J + L$. The J component remains constants during a scene, while the L component is described by a low order AR process driven by a Normally distributed noise process. However, to slightly simplify the model, the distribution of L is approximated with zero mean normal distribution. The scene changes with probability q and the distribution of the size of the J component during a new scene log-normal and is independent of the size of all other frames.

Thus

$$J_{k+1} = \begin{cases} J_k & \text{with probability } q \\ Y_{k+1} & \text{with probability } 1 - q \end{cases}$$

where Y_{k+1} is independent of everything else and log-normally distributed with parameters μ_j and σ_j . In [31], a number of films were examined including The Wizard of OZ. The parameters of The Wizard of OZ are $\mu_J = 5.9$, $\sigma_J = 0.48$, $\sigma_L = 62$, $\mu_P = 4.8$, $\sigma_P = 0.64$, $\mu_B = 3.9$, $\sigma_B = 0.27$ and $q = 0.93$. These parameter sizes are for ATM cells, hence, must be converted into packet sizes. In the following, the packet size was taken to be 1500 bytes.

B. Video Transmission as a Queuing Problem

In order to ensure that the video does not send data too fast, a buffer is used. The rate at which the buffer empties varies and is chosen such that it's output rate is compatible with TCP. Specifically, the buffer output rate is set to the 50th percentile of TCP flows. Since the sending rate of TCP depends on the current round-trip time, the 50th percentile rate also depends on the round-trip time. Thus, we define the sending rate as $f_{50}(r)$ where r is the round-trip time. Clearly,

$$f_{50}(r) = \frac{1}{r} \left\{ X : \frac{1}{\int p(x,r) dx} \int_{\{x < X\}} p(x,r) dx = 0.5 \right\}, \quad (12)$$

where $p(x,r)$ is the stationary joint density function of the congestion window and the round-trip time. A typical sending rate is shown in the upper frame of Figure 12. Hence, the buffer is filled by the video stream that is modeled as in Subsection IV-A and the buffer empties at a rate f . Clearly, this scheme ensures that data will not be sent too fast, but it is not clear if the video flow is properly transmitted. In particular, it is conceivable that the occupancy of the buffer will grow without bound. On the other hand, the buffer might maintain an occupancy below level B with a probability near 1. In this case, the receiver must place a buffer on the received data and this buffer must be filled to a level of B

before playback begins. Thus, the probability distribution of the buffer occupancy must be determined.

Let U_t be the buffer occupancy at time t . Then the buffer dynamics are

$$\begin{aligned} dU_t &= -1_{\{U_t > 0\}} f(R_t) dt + \sum_k \left(\delta_{\{t=kG\}} (J_k + L_k) \right. \\ &\quad \left. + \sum_i \delta_{\{t=kG+\tau_i^B\}} B_k^i + \sum_i \delta_{\{t=kG+\tau_i^P\}} P_k^i \right) \\ dR_t &= \mu(R_t) dt + \sigma(R) dB_t \\ J_{k+1} &= \begin{cases} J_k & \text{with probability } q \\ Y_k & \text{with probability } 1 - q \end{cases}, \end{aligned}$$

where the δ is such that if $U_0 = 0$ and $dU_t = \delta_{\{t=1\}}$, then $U_t = 0$ for $t < 1$ and $U_t = 1$ for $t \geq 1$. Here G is the duration of a GOP. Thus, the k^{th} GOP begins at time kG . Hence, at time kG the I frame is dumped into the buffer. Then, at the appropriate times, the P and B frames are dumped into the buffer. For example, the i^{th} B frame is dumped τ_i^B after the start of the GOP. All the while, the buffer is being emptied at a rate $f(R_t)$. The distribution of the buffer occupancy is given by

$$\begin{aligned} &\frac{\partial p(u,r,j,t)}{\partial t} \\ &= f(r) \frac{\partial p(u,r,j,t)}{\partial u} - \frac{\partial}{\partial r} (\mu(r) p(u,r,j,t)) \\ &\quad + \frac{1}{2} \frac{\partial^2}{\partial r^2} (\sigma^2(r) p(u,r,j,t)) \end{aligned} \quad (13)$$

for $t \neq kG, kG + \tau_i^B, kG + \tau_i^P$, and

$$\begin{aligned} &p(u,r,j,kG) \\ &= \int p(u-v,r,v,kG^-) p_I(J_k = j | J_{k-1} = v) dv \\ &p(u,r,j,kG + \tau_i^B) \\ &= \int p(u-v,r,j,(kG + \tau_i^B)^-) dP_B(B_k^i = v) \\ &p(u,r,j,kG + \tau_i^P) \\ &= \int p(u-v,r,j,(kG + \tau_i^P)^-) dP(P_k^i = v), \end{aligned}$$

otherwise, where p_I is the transition probability density function for the size of the I frame, $dP_B = \frac{d}{dv} P(B_k^i < v)$, and $dP_P = \frac{d}{dv} P(P_k^i < v)$. Letting $t \rightarrow \infty$, the stationary density of the buffer occupancy can be determined.

Figure 12 shows the results of The Wizard of OZ movie parameters along with the parameters estimated from a connection with source in San Jose and destination in Los Angeles. The upper plot in Figure 12 shows TCP's 50th percentile sending rate. Figure 12 also shows the expected value of the buffer size over the course of one GOP and the 99 percentile of the buffer size. The arrival of the frames is clearly evident. The average data rate of the 50 percentile ($\int p(r) f(r) dr$) was found to be 1.22Mbps while the average data rate required by the video flow is 1.19Mbps. Clearly, the buffer occupancy will not grow without bound. However, in order to accommodate

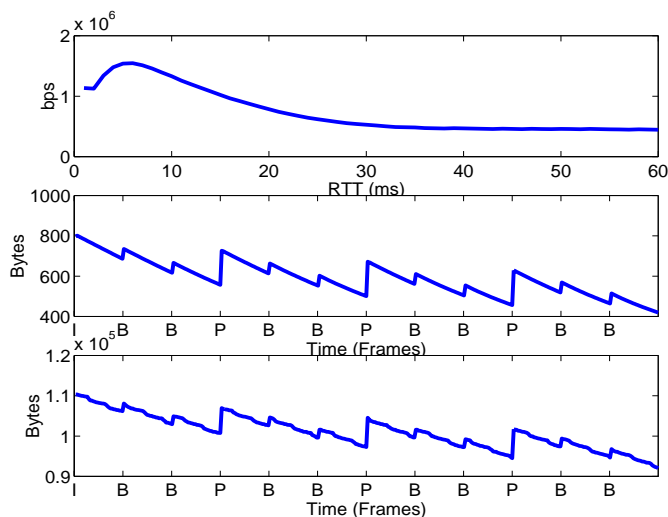


Fig. 12. Upper, TCP's 50th Percentile Sending Rate; middle, The Expected Value of the Buffer Size Versus Time in MPEG 1 Frames; lower, 99th Percentile of the Buffer Size Versus Time in MPEG 1 Frames.

99% of the realizations, it is necessary to permit the buffer to hold up to 110KB bytes. Furthermore, the buffer on the receiving end will need to initially receive 110KB bytes before playback can begin. Note that as the sending rate gets closer to the video rate, the 99th percentile of the maximum buffer occupancy rapidly increases.

Remark 3: The great difference between the average buffer occupancy and the 99th percentile occupancy shows the danger of only considering mean behavior.

Remark 4: One approach not investigated here is the possibility of work-ahead transmission. That is, the above method dictates that no data is sent if the buffer is empty. However, for prerecorded video, it is possible to send data faster than the video bit-rate. An investigation of this approach will be left for future investigation.

V. CONCLUSION

The TCP model presented is able to give detailed information about TCP's performance and provide a basis for deciding whether a video may be fairly transmitted. Furthermore, this modeling approach provides a deeper insight as to how TCP reacts to different network conditions. For example, the dependence of the congestion window on the rate that the round-trip time varies is a new insight. Furthermore, the difference between the mean and the median TCP sending rate has never been investigated. More application of this model approach remains. For example, through this type of modeling, it is straight forward to determine if an alternative approach to TCP (e.g. GAIMD) is fair to TCP.

A drawback of this modeling approach is that it appears complex. For example, the PDE's (10), (11) and (13) look formidable. However, this is not the case. Since the domain of these equations are rectangles, these linear parabolic equations can be efficiently solved with spectral methods. Furthermore, preliminary work indicates that, given the round-trip time, the

distribution of the congestion window is well modeled by a lognormal distribution. Indeed, it may be possible to find a close form approximation of distributions.

A more serious drawback is that the models contain parameters that must be estimated. However there has been extensive work in estimation that can easily be applied to this problem. Furthermore, previous work indicates that end hosts can be clustered in such a way that a connection to one member of the cluster has similar dynamics to other members of the cluster. Thus, it appears that it is not necessary to estimate each connection independently. Finally, while more measurements are required, our data indicates that large parts of the network behave similarly. Hence, prior measurements can be used as an initial estimate of the parameters and as data is sent over the connection, the parameters estimates can be further refined.

A more interesting effect of the parameters is that they may vary over time. It must be emphasized that the parameters typically change slowly and sometimes remain stable for long periods of time. However, the effect of parameter variation should not be neglected. The methods developed in this paper have been extended to the case of time vary parameters. The details of this extension will be presented elsewhere.

REFERENCES

- [1] A. Misra and T. J. Ott, "The window distribution of idealized TCP congestion avoidance with variable packet loss," in *INFOCOM*, 1999.
- [2] S. Floyd, M. Handley, J. Padhye, and J. Widmer, "Equation-based congestion control for unicast applications," in *SIGCOMM 2000*, Stockholm, Sweden, 2000.
- [3] J. Padhye, J. F. Kurose, D. F. Towsley, and R. Koodli, "A TCP-friendly rate adjustment protocol for continuous media flows over best effort networks," in *Measurement and Modeling of Computer Systems*, 1999, pp. 220–221.
- [4] Z. Jiang and L. Kleinrock, "A general optimal video smoothing algorithm," in *INFOCOM*, 1998, pp. 676–684.
- [5] Z. Zhang, J. Kurose, J. Salehi, and D. Towsley, "Smoothing, statistical multiplexing and call admission control for stored video," *Journal on Special Areas in Communications*, 1997.
- [6] S. Floyd, "Connections with multiple congested gateways in packet-switched networks part 1: One-way traffic," *Computer Communication Review*, vol. 21, pp. 30–47, 1991.
- [7] M. Mathis, J. Semke, J. Mahdavi, and T. Ott, "The macroscopic behavior of the TCP congestive avoidance algorithm," *Comput. Commun. Rev.*, vol. 27, 1997.
- [8] J. Padhye, V. Firoiu, D. Towsley, and J. Kurose, "Modeling TCP throughput: A simple model and its empirical validation," in *ACM Sigcomm*, Sept. 1998.
- [9] V. Misra, W. Gong, and D. Towsley, "Stochastic differential equation modeling and analysis of TCP-window size behavior," in *Performance '99*, Istanbul, Turkey, 1999.
- [10] F. Baccelli and D. Hong, "TCP is max-plus linear," in *SIGCOMM 2000*.
- [11] A. Kumar, "Comparative performance analysis of versions of TCP in a local network with a lossy link," *IEEE ACM Transactions on Networking*, vol. 6, pp. 485–498, 1998.
- [12] E. Altman, K. Avrachenkov, and C. Barakat, "A stochastic model of TCP/IP with stationary random losses," in *ACM Sigcomm*, Sep. 2000.
- [13] N. Cardwell, S. Savage, and T. Anderson, "Modeling TCP latency," in *INFOCOM 2000*, 2000.
- [14] J. Pointek, F. Shull, R. Tesoriero, and A. Agrawala, "Netdyn revisited: A replicated study of network dynamics," Dept. of Computer Science, University of Maryland, Tech. Rep. CS-TR-3696, 1996.
- [15] A. Mukherjee, "On the dynamics and significance of low frequency components of Internet load," *Internetworking: Research and Experience*, vol. 5, no. 4, pp. 163–205, 1994.
- [16] A. Acharya and J. Saltz, "A study of internet round-trip delay," Department of Computer Science, University of Maryland, Tech. Rep. CS-TR-3736, 1996.

- [17] M. S. Borella and G. B. Brewster, "Measurement and analysis of long-range dependent behavior of internet packet delay," in *INFOCOM*, 1998.
- [18] S. Bohacek and B. Rozovskii, "A diffusion model of roundtrip time," *Computational Statistics and Data Analysis*, Accepted.
- [19] J. C. Cox, J. E. Ingersoll, and S. A. Ross, "A theory of the term structure of interest rates," *Econometrica*, vol. 53, pp. 385–407, 1985.
- [20] L. Kleinrock, *Queueing Systems*. New York: Wiley, 1975.
- [21] M. Mathis, J. Semke, J. Mahdavi, and T. Ott, "The macroscopic behavior of the TCP congestion avoidance algorithm," *Computer Communication Review*, vol. 27, 1997.
- [22] E. Altman, K. Avrachenkov, and C. Barakat, "TCP in the presence of bursty losses," in *ACM SIGMETRICS*, June 2000.
- [23] S. Savari and E. Telatar, "The behavior of certain stochastic processes arising in window protocols," in *IEEE Globecom*, Dec. 1999.
- [24] M. Yajnik, S. B. Moon, J. F. Kurose, and D. F. Towsley, "Measurement and modeling of the temporal dependence in packet loss," in *INFOCOM*, 1999, pp. 345–352.
- [25] R. L. Eubank., *Spline Smoothing and Nonparametric Regression*, R. L. Eubank., Ed. New York: M. Dekker, 1988.
- [26] B. Oksendal, *Stochastic Differential Equations : An Introduction with Applications*. New York: Springer, 1998.
- [27] J. Widmer and M. Handley, "Extending equation-based congestion control to multicast applications," in *Proc of ACM SIGCOMM 2001*, 2001.
- [28] M. Vojnovic and J. L. Boudec, "Some observations on equation-based rate control," in *Proceedings of ITC-17, Seventeenth International Teletraffic Congress*, Salvador da Bahia, Brazil, September 2001.
- [29] J. Mahdavi and S. Floyd, "TCP-friendly unicast rate-based flow control," Note sent to the end2end-interest mailing list, 1997.
- [30] M. R. Izquierdo and D. S. Reeves, "A survey of statistical source models for variable-bit-rate compressed video," *Multimedia Systems*, vol. 7, pp. 199–213, 1999.
- [31] M. Krunz and S. K. Tripathi, "On the characterization of VBR MPEG streams," in *Proc. Of ACM SIGMETRICS '97*, 1997.