TCP throughput and timeout – steady state and time-varying dynamics

Stephan Bohacek
Dept. of Electrical and Computer Engineering
University of Delaware
bohacek@udel.edu

and

Khushboo Shah
Dept. of Electrical Engineering
University of Southern California
khushboo@usc.edu

Abstract— **While many models of the TCP's dynamics have been developed, few focus on the effects of timeout and high loss probability. Active queue management (AQM) is an important application of these dynamic models. However, recent work has shown that AQM provides little performance benefit over drop-tail queueing for HTTP traffic, except possibly at high utilizations. It is at these utilizations that the dynamic models of TCP are the least accurate. This paper presents a dynamic model of TCP that accurately models timeout. This model is also applicable to the static case. This paper also presents a model of the variance and the distribution of the congestion window. It is shown that, while the dynamics of the mean value of the congestion window are rather mild, the dynamics of timeout display large oscillations that take several seconds to decay. These oscillations cause the average bit-rate to also wildly oscillate. Finally, this paper includes results from several million simulations providing a detailed view of the dynamics of timeout.**

I. Introduction

Models of TCP are extensively used throughout networking research. For example, TCP models are used in fields such as AQM [1], [2], [3], [4], design of TCP friendly transport protocols [5], planning and provisioning networks [6], predicting file transfer time [7], [8], etc. In this way, TCP models are foundation for research. The accuracy of these models is critical to the accuracy and relevance of the work that rests upon them.

Models for TCP have evolved over time. Initially, a simple relationship between the sending rate and the loss probability was developed. The well-known ”1/sqrt(p)” formula, or more specifically $MSS \times \sqrt{3/2}/\left(RTT\sqrt{p}\right)$, was found and verified [9], [10]. While this formula is useful in many settings, it was found to not be accurate in situations where the packet loss probability is high. The reason for this inaccuracy is that this simple model does not account for TCP's timeout mechanism. Some studies have shown that timeout, as oppose to triple duplicate acknowledgement, is a significant way that TCP detects packet loss [11].

In order to include the impact of timeout, the simple ”square root of p” formula was extended to include the effect of timeout [11]. This formula was shown to be more accurate for large loss probabilities where timeout is significant. Sikdar et al. [12] developed a model for TCP throughput that includes time-out. While this model is different from [11], they report that it gives similar numerical results. However, we have performed extensive simulations, over 5 million in total, and have found that when applied to the ns-2 implementation of TCP-SACK and when loss is probabilistic, the model given in [11] results in errors as large as a factor of two. Considering the importance of the relationship between sending rate and loss probability, these differences are significant and require closer examination. This paper provides an accurate method to estimate TCP's sending rate even for large loss probabilities.

Another important extension to the TCP models was to make them dynamic. That is, the initial models assumed that the loss probability was fixed. In the case that the loss probability varied, dynamic models were required. The first model in this direction was given in [4], and several closely related models have followed [1], [3], [7]. The principle application of these models is for design and analysis of AQM. These models are extensively used to prove stability and to indicate the performance of AQM.

Since AQM was first introduced, there have been a large number of researchers active in this area. Recently, several groups have investigated what, if any, benefit AQM has on the performance of the network (see [13] and references therein). A recent paper along these lines showed that the AQM, be it drop tail or any more elaborate version, has little impact on the performance of the network for utilization below 90%. However, for utilization above 90%, some AQM techniques were shown to improve the performance [14].

This work is significant as it indicates the environment where AQM might have a positive impact, specifically, in networks with high utilization. On the other hand, this conclusion is troubling since the models of TCP used to design AQM neglect timeout. That is, in the environment where AQM might have the largest impact on performance, the analysis and design of AQM is least accurate. Because of this inaccuracy, the possible impact of AQM on network performance at high utilizations has never been carefully explored.

One reason that AQM design has not considered high utilization networks is that dynamics of TCP in such an environment are not known. By dynamics, we mean the variation of the throughput in the face of a time-varying loss probability. This paper provides a simple model of the dynamics of TCP's throughput that includes timeout. Perhaps one of the most significant contributions of this paper is that the dynamics of timeout is quite complicated and can result in wildly varying bit-rates that can take several seconds to decay. For example, the lower right of Figure 4 shows such oscillations. This dynamics is far more complicated than the

0-7803-8794-5/04/$20.00 © 2004 IEEE

dynamics of the mean congestion window used by most AQM design. As a result, this work calls into question the proof of stability of AQM at high loss probabilities.

As a by-product of this work, the variance and the distribution of the congestion window is also found. Specifically, it is shown that the window size is accurately modeled with a negative binomial distribution. And finally the huge number of simulations used to validate this model provide a detailed view of the behavior of the dynamics of TCP's timeout mechanism.

The paper proceeds as follows: The next section discusses the simulations used for verification. In Section III the distribution of the congestion window is found. Next, in Section IV, a model of the probability of a flow being in timeout is developed. The results from Sections III and IV yield a static model of TCP throughput. In Section V, this model is then extended to the dynamic case. The paper closes with some concluding remarks in Section VI.

II. The simulations

The goal of this effort is to closely examine the behavior of TCP. There is little doubt that in today's networks the sending rate of a TCP connection is complicated by interacting flows, router induced packet reordering, faulty load balancing induced packet reordering, server stalls, implementation idiosyncrasies, etc. While these effects are critical and a better understanding of them is necessary for a complete understanding of TCP, this paper focuses strictly on the behavior of TCP. The rationale for this is that in order to understand the behavior of TCP in the wild, it is necessary to thoroughly understand TCP in a controlled environment. Perhaps the most significant difference between these simulations and the environment found in the wild is that in these simulations packet losses are entirely random and never due to queue overflow. In the case of AQM, such simulations are appropriate as the goal of AQM is to drop packets in a controlled fashion, not when the queue fills.

The simulations presented in this paper are for a single flow over a single bottleneck topology where drops at the bottleneck are random. The ns-2 error model was used to produce the drops. The results presented here are for a fixed round-trip time of 30ms. Other round-trip times have been explored and yield the same results as the ones presented here. These simulations used ns-2 implementation of TCP-SACK in version 2.1b8a. Delayed acknowledgement and a maximum receiver window were not used.

Our investigation into TCP included over 5 million simulations. This large number of simulations is required to estimate the dynamics of timeout. These simulations were performed on the University of Southern California's Linux cluster [15] and the large simulation results were transferred via the Internet 2 to the University of Delaware for post-processing and storage. There is little doubt that without such high-performance computing, this work would not be possible.

III. The distribution of the congestion window

This section presents a simple model for the distribution of the congestion window. While the distribution may be of interest in its own right, we are particularly interested in this distribution since it is used to determine the probability of a flow entering timeout. We justify the model briefly with some analysis and then present simulations results.

Misra [4] introduced the idea of modeling TCP as a stochastic differential equation. Specifically, he suggested that the window size varies according to

$$dW_t = \frac{1}{RTT}dt - \frac{1}{2}W_t dN_t \qquad (1)$$

where N is a Cox process that counts the number of packet losses. This model is for the evolution of the congestion window when the flow is not in timeout. Thus, all probabilities should be conditioned on the flow not in timeout. We denote this condition as $\overline{TO}$. This model was further investigated in [7] where the partial differential equation of the window size was found. Specifically, if $p\left(w,t|\overline{TO}\right)$ is the probability density of the congestion window taking the value w at time t, then p satisfies

$$\frac{\partial p\left(w,t|\overline{TO}\right)}{\partial t} = \frac{1}{RTT}\left(-\frac{\partial p\left(w,t|\overline{TO}\right)}{\partial w}\right. \qquad (2)$$
$$\left.+w\times\delta\left(t-RTT\right)\left(4p\left(2w,t|\overline{TO}\right)-p\left(w,t|\overline{TO}\right)\right)\right),$$

where δ is the loss probability. If δ is constant, then in steady state, i.e., $\frac{\partial p\left(w,t|\overline{TO}\right)}{\partial t} = 0$, the distribution of the window solves

$$\frac{dp\left(w|\overline{TO}\right)}{dw} = w\times\delta\left(4p\left(2w|\overline{TO}\right)-p\left(w|\overline{TO}\right)\right). \qquad (3)$$

From (3) it is straightforward to show that

$$E\left(w^m|\overline{TO}\right) = \frac{c_m}{\delta^{m/2}},$$

for some constants c_m. If $m=1$, then this is the well-known "square root of p" formula where c_1 has been found to be between 1.1 and 1.3. Considering $m=2$, we find that the variance is of the form

$$Var\left(w|\overline{TO}\right) := E\left(w^2|\overline{TO}\right) - E\left(w|\overline{TO}\right)^2 = \frac{\gamma}{\delta}, \qquad (4)$$

where $\gamma := c_2 - c_1^2$. This relationship is borne out in simulations that indicates that $\gamma \approx 0.31$.

A closed form solution has not been found for (2), whereas a complicated closed form solution has been found for (3). However, simple yet accurate approximate solution to (3) is the negative binomial distribution, i.e.,

$$p\left(w\right) = \frac{\Gamma\left(N+w-1\right)}{\Gamma\left(N\right)\left(w-1\right)!}\left(1-q\right)^N q^{w-1}.$$

Since the mean of this negative binomial random variable is $Nq/\left(1-q\right)+1$ and the variance is $Nq/\left(1-q\right)^2$,

$$q = 1-\frac{E\left(w|\overline{TO}\right)-1}{Var\left(w|\overline{TO}\right)} \text{ and } N = \frac{1-q}{q}\left(E\left(w|\overline{TO}\right)-1\right) \qquad (5)$$

and from mean and variance of the w given above we get

$$q = 1-\frac{\delta}{\gamma}\left(\frac{c_1}{\sqrt{\delta}}+1\right) \text{ and } N = \frac{(1-q)}{q}\left(\frac{c_1}{\sqrt{\delta}}-1\right). \qquad (6)$$

In order to determine the distribution of the window size, the parameters q and N must be determined. This can be done by plugging the mean and variance into (5). However, if only the loss probability is known, then (6) must be used, in which case the parameters c_1 and γ must be determined. We have found that if the objective is to use the distribution to determine the probability of timeout, the $c_1 = \sqrt{3/2}$ and $\gamma = 0.31$ is sufficient. On the other hand, if one is interested in the distribution of the congestion window (the objective of this section), then the selection of these parameters, especially c_1, requires some more analysis. However, to keep this paper focused, this issue is not discussed.

Once the parameters are determined, the distribution of the congestion window can be estimated. Figure 1 shows the distribution of the congestion window given the flow is not in the timeout state, i.e., $p\left(w|\overline{TO}\right)$. These figures shows the three types of curves, the observed distribution (histogram), the distribution with $c_1 = \sqrt{3/2}$, and the distribution where the parameters are determined by plugging in the observed mean and variance into (5). This last distribution coincides with the distribution given by (6) but with $c_1 = 1.27$ for $\delta = 10^{-4}$ and $c_1 = 1.14$ for $\delta = 0.05$. For $\delta = 0.01$, the value of c_1 plays less of a role, $c_1 = 1.14$, $c_1 = \sqrt{3/2}$, and $c_1 = 1.27$, or by using the observed mean and variance, all yield nearly the same distribution. In all cases, $\gamma = 0.31$.

IV. A MODEL OF TIMEOUT

In this section, we present a model of the probability of a flow being in the timeout state. We say that a flow is in the timeout state if the next packet will be sent only after the retransmission timer expires. To understand the development of this model, one should consider a large collection of TCP experiments all running in parallel on different networks with identical network characteristics (e.g., link speeds, propagation delays, etc.). Our goal is to determine the fraction of these flows that are in timeout.

A flow can enter the timeout state in the following three ways:

I If a flow experiences so many losses that triple duplicate acknowledgements are not received, i.e., if at least $\max(w-2,1)$ losses occur in one window.

II In the case of the ns-2 implementation of TCP-SACK, if more than $w/2$ packets are dropped[1].

III If a retransmitted packet is dropped.

Considering I and II, we see that a flow will timeout if a single drop is followed by at least $\max(\min(\lfloor w/2 \rfloor, w-3), 0)$ drops out of the next $w-1$ packets. Given that drops occur at a rate $\frac{w}{R}\delta$, drops that lead to timeout occur at a rate

$$\rho'(w,\delta) = \frac{w}{R}\delta \sum_{k=\max(\min(\lfloor w/2 \rfloor, w-3),0)}^{w-1} \binom{w-1}{k} \delta^k (1-\delta)^{w-1-k}.$$

Let $\lambda'(\delta)$ be the rate that a flow moves from congestion avoidance to timeout due to I and II. This rate is found taking the expected value of $\rho(w,\delta)$, i.e.,

$$\lambda'(\delta) = E\left(\rho'(w,\delta)\,|\overline{\text{TO}}\right) = \sum_w \rho'(w,\delta)\, p_\delta(w),$$

where $p_\delta(w)$ is given in Section III, however, here we have shown the dependence on δ. This dependence is made explicit by (6). Note that the approximation $E\left(\rho'(w,\delta)\,|\,\overline{\text{TO}}\right) \approx \rho'\left(E\left(w|\,\overline{\text{TO}}\right),\delta\right)$ will result in large errors. Consider the situation where $\delta = 0.09$ and, using $c_1 = \sqrt{3/2}$, the mean value of the congestion window is 4. In this case, $E\left(\rho'(w,\delta)\,|\,\overline{\text{TO}}\right) = 0.13 \times R$ while $\rho'\left(E\left(w|\,\overline{\text{TO}}\right),\delta\right) = 0.38 \times R$, a difference around a factor of three.

Next we determine λ'' (i.e., III), the rate that a flow moves from congestion avoidance to timeout because a retransmitted packet is dropped. The rate that a packet is first dropped is $\delta\frac{w}{R}$. The drop will lead to a retransmission only if triple duplicate acknowledgements are received, i.e., if less than $\max(\min(\lfloor w/2 \rfloor, w-3), 0)$ packets are dropped out of the next $w-1$ packets. In this case, the retransmission is dropped and timeout is entered with probability δ. Thus

$$\rho''(w,\delta) := \delta\frac{w}{R} \times \delta \times \left(1 - \sum_{k=\max(\min(\lfloor w/2 \rfloor, w-3),0)}^{w-1} \binom{w-1}{k} \delta^k (1-\delta)^{w-1-k}\right).$$

The average rate that a flow enters timeout due to dropped retransmissions is

$$\lambda''(\delta) := E\left(\rho''(w,\delta)\right)$$

Finally, the rate that a TCP flow moves into timeout is $\lambda' + \lambda''$.

Up to this point we considered the rate that a single flow moves into timeout. Now we will determine the fraction of flows entering timeout. Let $N * I_1(t)$, denote the rate that flows enter timeout at time t, N is the total number of flows. That is, in small time interval Δt, the probability that some flow enters timeout is $N * I_1(t)\,\Delta t$. Let RTO be the time in which flows remain in the timeout state upon the first timeout. If a flow exits timeout, but immediately experiences another drop, then this flow again enters timeout, but for $2 \times RTO$ seconds[2]. We denote the rate that flows enter timeout for this second time with $N * I_2(t)$. The fraction of flows in timeout are $\int_{t-RTO}^{t} I_1(\tau)\,d\tau + \int_{t-2RTO}^{t} I_2(\tau)\,d\tau$. The rate that flows enter timeout is the product of the number of flows not in timeout and the rate that such flows enter timeout, i.e.,

$$I_1(t) = \left(\lambda'(\delta(t)) + \lambda''(\delta(t))\right) \times \left(1 - \left(\int_{t-RTO}^{t} I_1(\tau)\,d\tau + \int_{t-2RTO}^{t} I_2(t)\,d\tau\right)\right). \tag{7}$$

[1] This way of entering time-out is discussed in [12]. RFC-3517 eliminates this way to enter time-out. It is straightforward to adjust the development below to reflect RFC-3517.

[2] We don't consider the case where a flow enters time-out a third time for a time interval $4 \times RTO$. While it is straightforward to extend the calculations, it does not appear to improve the accuracy of the model.

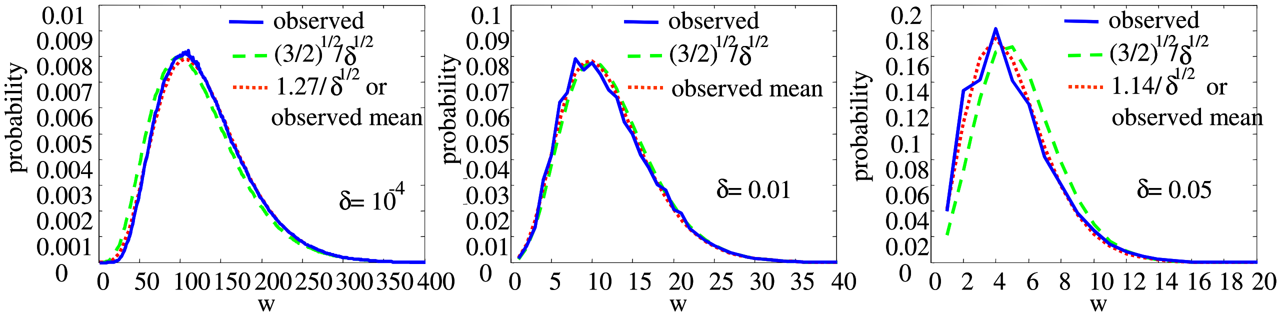

Figure 1. Distribution of the congestion window for loss probability of $\delta = 10^{-4}$ (left), $\delta = 0.01$ (middle), and $\delta = 0.05$ (right).

While the rate that flows exit timeout only to reenter timeout is

$$I_2(t) = \delta I_1(t - RTO). \tag{8}$$

In steady state, $I_1(t)$ and $I_2(t)$ are constant. So (7) and (8) reduce to

$$I_1 = (1 - I_1 \times RTO - 2I_2 \times RTO)(\lambda'(\delta) + \lambda''(\delta))$$
$$I_2 = \delta I_1.$$

This can be solved

$$I_1 = \frac{\lambda'(\delta) + \lambda''(\delta)}{1 + RTO(\lambda'(\delta) + \lambda''(\delta))(1 + 2\delta)}$$
$$I_2 = \frac{\delta(\lambda'(\delta) + \lambda''(\delta))}{1 + RTO(\lambda'(\delta) + \lambda''(\delta))(1 + 2\delta)}$$

Thus, the fraction of flows in timeout is

$$\begin{aligned} P(TO) &= \int_{t-RTO}^{t} I_1(\tau)\,d\tau + \int_{t-RTO}^{t} I_2(\tau)\,d\tau \quad (9)\\ &= \frac{(1+2\delta)RTO(\lambda'(\delta) + \lambda''(\delta))}{1 + (1+2\delta)RTO(\lambda'(\delta) + \lambda''(\delta))}. \end{aligned}$$

If the loss probability varies, then (7), (8), and the relationship $P(\text{in timeout at time } t) = \int_{t-RTO}^{t} I_1(\tau)\,d\tau + \int_{t-2RTO}^{t} I_2(\tau)\,d\tau$ must be used to determine the probability that a flow is in timeout.

To verify (9), a large number of ns-2 simulations were performed so that the probability of timeout could be determined even for small loss probabilities. Figure 2 shows the observed relationship of $P(\text{in timeout})$ and the loss probability as well as the estimate of this probability given by (9). These calculations used $c_1 = \sqrt{3/2}$. Note that the fit is quite good, the observed probability and the model are essentially the same.

Padhye et al. [11] provides an estimate of the throughput that accounts for timeout. This formula also provides an estimate of the fraction of time a flow spends in timeout, i.e., $P(TO)$. Specifically, [11] finds

$$P(\text{timeout}) \approx \frac{Q(W(\delta), \delta)\, G(\delta)\, T_o \frac{1}{1-\delta}}{\frac{R}{8}(W(\delta) + 1) + Q(W(\delta), \delta)\, G(\delta)\, T_o \frac{1}{1-\delta}}, \tag{10}$$

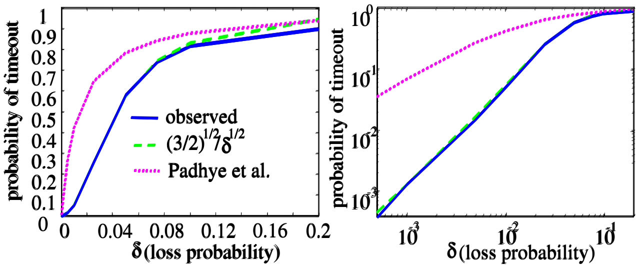

Figure 2. Probability of Timeout. The solid blue line represents the observed probability of a flow being in timeout. The dashed green line is the model's estimate of this probability. Note that these two lines nearly coincide. The third curve is the relationship derived from [11].

where

$$Q(w, \delta) = \min\left(1, \frac{\left(1 + (1-\delta)^3\right)\left(1 - (1-\delta)^{w-3}\right)}{\left(1 - (1-\delta)^w\right) / \left(1 - (1-\delta)^3\right)}\right)$$
$$W(\delta) = 1 + \sqrt{\frac{8(1-\delta)}{3\delta} + 1}$$
$$G(\delta) = 1 + \delta + 2\delta^2 + 4\delta^3 + 8\delta^4 + 16\delta^5 + 32\delta^6.$$

Figure 2 includes this estimate. We see that this model, while providing a qualitative fit in that the shape of the relationship is correct, it over estimates the probability of being in timeout. For example, for $\delta = 1\%$, the observed probability of being in timeout is 5%, while (10) gives an estimate of 43%. However, this comparison must be qualified. This paper focuses on TCP-SACK while the focus of [11] is on TCP-RENO. Furthermore, this paper examines how TCP behaves under random losses such as arises in AQM. In [11], the queueing discipline is drop tail. Furthermore, the loss probability in [11] was not exactly the probability of a packet being dropped, but more along the lines of the probability of a drop event occurring where a drop event leads to the rest of the packets in the "round" being dropped. Thus, the δ used in (10) is less than the packet loss probability. However, this model is useful as a benchmark as it was not known how accurate this model would be for TCP-SACK.

From the probability of timeout, the average through-put

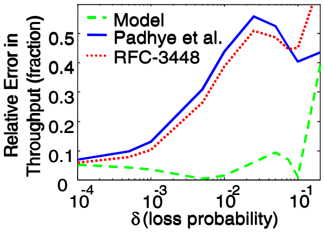

Figure 3. Relative Error of Predicted Throughput. Note that an error fraction of 0.5 implies a 50% error.

can be found via

$$T = \frac{c_1}{R\sqrt{\delta}} \left(1 - P\left(TO\right)\right) \times MSS.$$

Figure 3 shows the relative error of this model when compared to the observed throughput. Figure 3 also shows the relative error of the model in [11] as well as the model suggested in RFC-3448 for a TCP friendly sending rate. The figure shows only the relative error, an examination of the error shows that the model of [11] and RFC-3448 underestimate the sending rate (this is due to the over estimation of the timeout probability). As mentioned, these other models are for benchmarking purposes only.

Figure 2 shows that for even moderately large loss probabilities, the probability of being in timeout is quite high. Specifically, for $\delta > 1\%$, timeout plays an important role in the sending rate of TCP. For $\delta > 5\%$, a flow will spend 60% of the time in timeout. Thus, timeout dominates the sending rate of TCP. There are many implications of this. Recently there have been some proposed modifications to TCP such as limited transmit and ECN. These modifications significantly affect the behavior of timeout, specifically, they enter timeout much less frequently. Thus, for $\delta > 1\%$, one can expect that TCP implementations with these modifications will be more aggressive than implementations without these modifications.

There has been extensive work in dynamic modeling of TCP, but most has neglected timeout. For example, many approaches to AQM utilize a model of TCP that neglects timeout. We see that such an approach is only reasonable if the loss probability is less than 1%. However, in the area of AQM, it is hoped that these models are applicable to higher loss probabilities. Next we develop a model for the dynamic behavior of TCP that includes timeout.

V. Dynamics of TCP

Here we extend the dynamics that is commonly used to model the dynamics of TCP. The commonly used model is for only the mean sending rate of TCP. Here we present a model for the variance. Once the variance is known, the distribution is found as in Section III. With the distribution, we can find the probability of being in timeout as was done in the previous section. Of course, the paper to this point has only considered a constant loss probability. Here we allow the loss probability to vary with time. Thus, this model is appropriate for AQM design.

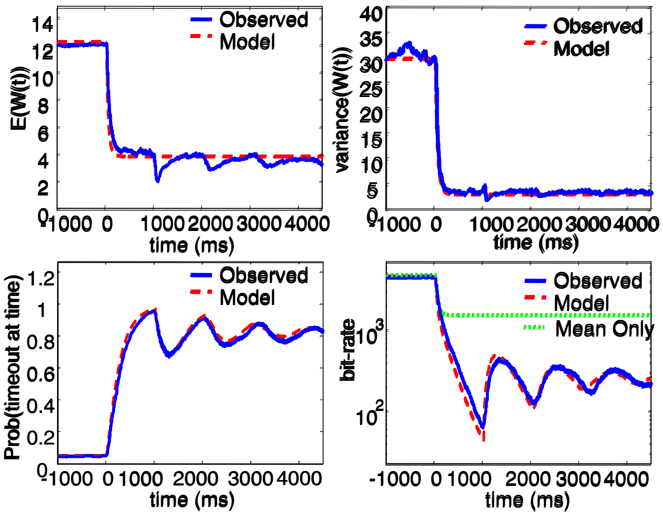

Figure 4. Simulation and modeling of a jump in loss probability from 1% to 10%. The jump occurred at $t = 0$. The solid blue curves show the observed and the dashed red curves shows the model. In the lower right, the dotted green indicates the bit-rate if only the mean size of the congestion window is used (i.e., the dashed red curve in the upper left). Note the units of the bit-rate is in 1000 bps.

Let $\bar{w}(t) = E\left(w(t) \mid \overline{\mathrm{TO}}\right)$ and $\overline{w^2}(t) = E\left(w^2(t) \mid \overline{\mathrm{TO}}\right)$. An approximation of the dynamics of the mean value of the congestion window is often given by

$$\frac{d}{dt}\bar{w}(t) = \frac{1}{R} - \frac{1}{c_1^2}\frac{1}{R}\delta(t-R)\,\bar{w}(t-R)\,\bar{w}(t). \qquad (11)$$

This is often approximated as

$$\frac{d}{dt}\bar{w}(t) = \frac{1}{R} - \frac{1}{c_1^2}\frac{1}{R}\delta(t-R)\,\bar{w}^2(t). \qquad (12)$$

Interestingly, this formula is often derived from (1). However, if proper stochastic calculus is applied, the correct dynamics for the mean are

$$\frac{d}{dt}\bar{w}(t) = \frac{1}{R} - \frac{1}{R}\delta(t-R)\,E\left(w(t-R)\,w(t)\right),$$

which can be approximated as

$$\frac{d}{dt}\bar{w}(t) = \frac{1}{R} - \frac{1}{R}\delta(t-R)\,\overline{w^2}(t). \qquad (13)$$

Note that in (13) the second moment is utilized, while (12) have $\frac{1}{c_1^2}$ multiplied by the first moment squared. Fortunately, we have found that (11) and (12) give good approximation to the dynamics of the mean. It seems that the mean varies slowly enough that this difference between $\overline{w^2}(t)$ and $\frac{1}{c_1^2}\overline{w(t)}^2$ is not significant.

From (1) it is possible to determine the dynamics of the second moment of the congestion window,

$$\frac{d}{dt}\overline{w^2}(t) = \frac{2}{R}\bar{w}(t) - \frac{3}{4}\frac{1}{R}\delta(t-R)\,E\left(w^3(t)\right).$$

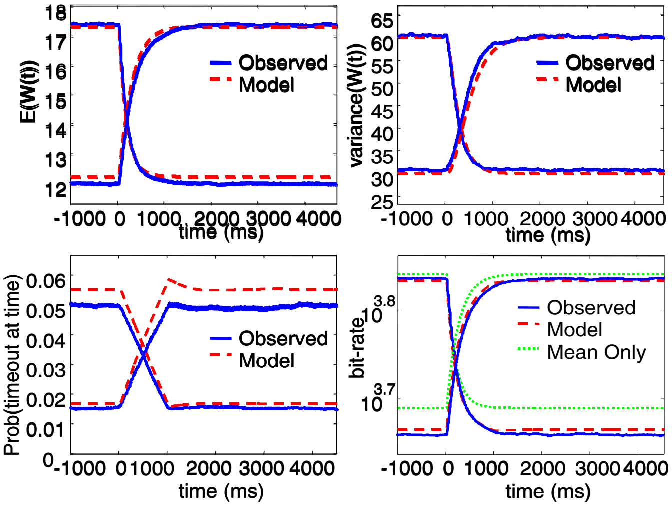

Figure 5. These figures are similar to the ones shown in Figure 4. However, here two sets of dynamics are shown, the loss probability jumping from 0.1% to 1% and falling from 1% to 0.1%. In both cases, the changes occur at $t = 0$.

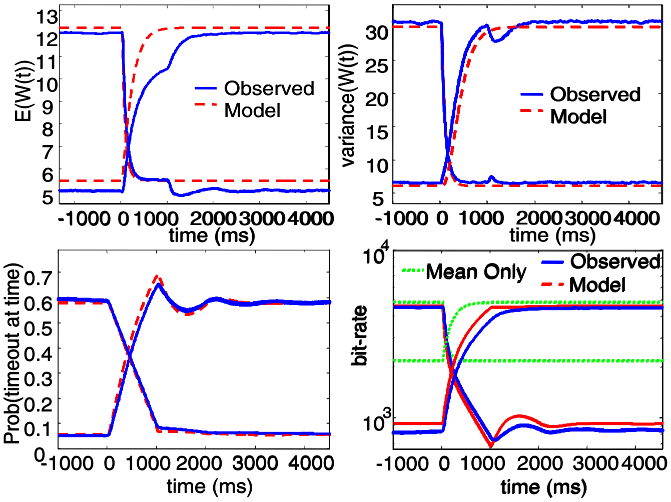

Figure 6. Validation of the Model. These figures are near identical to Figure 5, however, here the loss probability jumps between 1% and 5%.

Using the same idea as above, we approximate $E\left(w^3(t)\right) \approx \frac{8}{3}\frac{\beta}{\left(c_1^2-0.31\right)^{3/2}}\left(\overline{w^2}(t)\right)^{3/2}$ and arrive at

$$\begin{aligned}\frac{d}{dt}\overline{w^2}(t) &= \frac{2}{R}\overline{w}(t) \qquad (14)\\ &-\frac{3}{4}\frac{8}{3}\frac{\beta}{\left(c_1^2-0.31\right)^{3/2}}\frac{1}{R}\delta(t-R)\left(\overline{w^2}(t)\right)^{3/2}.\end{aligned}$$

As in the case of (12), this approximation is exact in steady state.

Equations (12) and (14) form a system of ordinary differential equations (ODEs). With the mean and second moment, the distribution of the congestion window can be found as discussed in Section III. With this approximation of the distribution, the probability of being in timeout can be found as in Section IV, i.e., $P(TO \text{ at time } t) = \int_{t-RTO}^{t} I_1(\tau)\,d\tau + \int_{t-2RTO}^{t} I_2(\tau)\,d\tau$, where $I_1(\tau)$ is given by (7) and $I_2(\tau)$ is given by (8).

Figures 4-6 provide some validation of this dynamic model of TCP. Through extensive simulation, $E\left(\overline{w}(t)\,|\overline{TO}\right)$, $E\left(\overline{w^2}(t)\,|\overline{TO}\right)$, and $P(TO \text{ at time } t)$ were found under different time variations of the loss probability. Specifically, in each simulation, the loss probability took two values δ_0 and δ_1. The simulation began and ran for 20 seconds with loss probability set to δ_0. In these simulations, 20 seconds was enough to ensure that the system was in steady state (i.e., $E\left(\overline{w}(t)\,|\overline{TO}\right)$ was constant). Then, 20 seconds after the beginning of the simulation, the loss probability was switched to δ_1. In the figures, this moment is labeled as $t = 0$. This simulation was carried out 200,000 times. With these simulations it is possible to determine the value of the statistics $E\left(\overline{w}(t)\,|\overline{TO}\right)$, $E\left(\overline{w^2}(t)\,|\overline{TO}\right)$, and $P(TO \text{ at time } t)$ at each time point. For example, to determine $P(TO \text{ at time } 500\ ms)$, we found the fraction of the 200,000 flows that were in the timeout state at time $t = 500\ ms$. This large number of simulations allows for high confidence in the observations and allows the details of the dynamics to be observed.

Figures 4-6 indicate that the dynamic model is quite accurate. They also show, as expected, that using the dynamics of the mean congestion window given by (12) is not adequate for large loss probability, but is adequate for small loss probability. We also see that the dynamics for the mean and the variance of the congestion window perform rather poorly for very large loss probability. However, in this case, the behavior of TCP is completely dominated by the timeout, which is well modeled. As a result, the bit-rate is accurately modeled.

Perhaps the most striking aspect of Figures 4-6 is the dynamics of the bit-rate at large loss probabilities. Most significantly, we see that the bit-rate experiences wild oscillations that can take several seconds to decay. In all cases, the model for the mean congestion window (12) does not produce any oscillations. To the best of the authors' knowledge, this behavior has not been previously observed, much less modeled. Oscillating step-responses are often a sign of instability, thus it is likely that these dynamics will impact the stability results of previous AQM research.

Figure 6 shows some unexpected behavior of the mean and the variance of the congestion window. Specifically, we see that the mean has two distinct phases of growth with a transition at $t = 1000\ ms$ while the variance shows non-monotonic growth. The ODEs (12) and (14) do not model this behavior. This behavior is due to the feedback of the fraction of flows in timeout to the mean and the variance of the congestion window. Notice that the non-monotonic growth of $\overline{w^2}$ occurs at RTO, just as flow exit timeout. However, since the probability of a flow being in the timeout state is so high, a slight error in the mean congestion window has a minor impact on the average bit-rate.

VI. Conclusions

We have developed a new model for the sending rate of TCP that includes the effect of timeout. Previous models have included the effect of timeout, but we have found these models to greatly over estimate the probability of being in timeout. The model presented is easily extended to a dynamic model of timeout. This dynamic model was shown to agree with simulations quite well. Furthermore, it was shown that timeout can cause the bit-rate to oscillate substantially for several seconds after a change in the loss probability. It was also shown that for large loss probabilities, the behavior of TCP is dominated by the behavior of timeout.

This model is significantly different from those used in previous design of AQM. Future work will examine the possibility of designing an AQM that utilizes these models for TCP. Perhaps, more substantial performance gains will be achieved for large loss probability.

References

[1] S. Athuraliya, V. H. Li, S. H. Low, and Q. Yin, "REM: Active queue management," *IEEE Network*, vol. 15, pp. 48–53, 2001.

[2] C. V. Hollot, V. Misra, D. F. Towsley, and W. Gong, "On designing improved controllers for AQM routers supporting TCP flows," in *INFOCOM*, 2001.

[3] S. Kunniyur and R. Srikant, "Analysis and design of an adaptive virtual queue (AVQ) algorithm for active queue management," in *Proceedings of ACM SIGCOMM*, 2001.

[4] V. Misra, W. Gong, and D. Towsley, "Stochastic differential equation modeling and analysis of tcp-windowsize behavior," in *Proceedings of PERFORMANCE99*, 1999.

[5] M. Handley, S. Floyd, J. Padhye, and J. Widmer, "RFC 3448 - TCP friendly rate control (TFRC): Protocol specification," 2003. http://www.faqs.org/rfcs/rfc3448.html.

[6] J. Beckers, I. Hendrawan, R. Kooij, and R. V. der Mei, "Generalized processor sharing performance models for internet access lines," in *9th IFIP CONFERENCE on PERFORMANCE MODELLING AND EVALUATION OF ATM and IP NETWORKS*, (Budapest), 2001.

[7] S. Bohacek, "A stochastic model of tcp and fair video transmission," in *Infocom*, 2003.

[8] N. Cardwell, S. Savage, and T. Anderson, "Modeling TCP latency," in *INFOCOM (3)*, pp. 1742–1751, 2000.

[9] T. Ott, J. H. B. Kemperman, and M. Mathis, "Window size behavior in TCP/IP with constant loss probability," in *Proc. Of the DIMACS Workshop on Performance of Realtime Applications on the Internet*, 1996.

[10] M. Mathis, J. Semke, J. Mahdavi, and T. Ott, "The macroscopic behavior of the TCP congestion avoidance algorithm," *ACM Comput. Comm. Review*, vol. 27, no. 3, 1997.

[11] J. Padhye, V. Firoiu, D. Towsley, and J. Krusoe, "Modeling TCP throughput: A simple model and its empirical validation," *Proceedings of the ACM SIGCOMM '98 conference on Applications, technologies, architectures, and protocols for computer communication*, pp. 303–314, 1998.

[12] K. S. V. Biplab Sikdar, Shivkumar Kalyanaraman, "Analytic models for the latency and steady-state throughput of TCP tahoe, reno, and SACK," *IEEE/ACM Transactions on Networking*, vol. 11, no. 6, pp. 959–971, 2003.

[13] S. Bohacek, K. Shah, G. Arce, and M. Davis, "A review of central issues - signal processing challenges in active queue management," *IEEE Signal Processing Magazine*, vol. 21, pp. 69– 79, 2004.

[14] L. Le, J. Aikat, K. Jeffay, and F. D. Smith, "The effects of active queue management on web performance," in *SIGCOMM*, 2003.

[15] "http://www.usc.edu/hpcc/hpc.html." http://www.usc.edu/hpcc/hpc.html.