Formula-Based TCP Throughput Prediction with Available Bandwidth

Jae-Hyun Hwang and Chuck Yoo, *Member, IEEE*

***Abstract*—In the research area of TCP throughput prediction, the well-known throughput model, a function of loss rate, has some disadvantages such as difficulty of loss rate estimation and low accuracy under congestion. This letter proposes a new throughput model by a function of available bandwidth. Our model improves the accuracy in TCP throughput prediction by removing loss rate, which is usually hard to measure accurately, in the throughput function. Our formula-based predictor is validated by real experiments on LAN environments.**

***Index Terms*—TCP throughput prediction, formula-based predictor, available bandwidth.**

I. Introduction

As multi-path networking paradigms such as multi-homing, overlay, peer-to-peer network, and content delivery network are getting popular, the ability to determine which path is the best among available multiple paths has been identified as a key to maximize network performance. One possible solution is to estimate one-way delay or available bandwidth on network paths in an end-to-end manner [1], [2]. Many researchers however have pointed out that there is a significant difference between available bandwidth and application goodput, and this leads to research on predictability of TCP throughput to select a best path [3], [4], [5], [6].

There are two approaches for TCP throughput prediction – formula-based and history-based. The formula-based approach is to predict the expected throughput by modeling TCP behavior mathematically, which can be represented by a function of loss rate. On the other hand, the history-based approach predicts the on-going throughput from a time series of previous measurements on the same path, when such a history is available. Although the latter approach has more overheads than the former with regard to the history maintenance, the most related research works agree that the history-based predictor is superior because of its accuracy in the most cases [3], [4], [5]. There are some factors that make the accuracy of the formula-based approach low, but it is reported that the main reason is the variation of Round-Trip Time (RTT) and loss rate. That is, as network traffic increases, the average queuing delay and loss probability can be higher than those measured before a session begins [3]. Another main reason is that it is difficult to estimate loss rate only by end-to-end probing. To obtain the accurate loss rate, a scheme that polls SNMP MIB information from the bottleneck router has been proposed [6], but it cannot be a practical solution because finding a bottleneck router is another difficult problem. We notice that the formula-based approach has low accuracy because the existing TCP throughput models still do not address the prediction problem and simply use constant values for RTT and loss rate [7], [8]. If such weakness is overcome, the formula-based approach can be also an efficient alternative as a TCP throughput predictor.

In this letter, we propose an available bandwidth based TCP throughput model as a new formula-based predictor. The available bandwidth generally can be measured fast by end-to-end probing while loss rate requires longer time to estimate. From this, we can get the expected value of congestion window size without a random variable of loss rate, and can achieve better prediction accuracy.

II. TCP Throughput Model Based on Available Bandwidth

We derive an available bandwidth based throughput model from the Amherst model [7], a well-known early TCP throughput model. Like the Amherst model, our model is divided into a triple-duplicate (TD) acknowledgement (ACK) period and a timeout (TO) period. Due to the limited space, we assume that the reader is familiar with the Amherst model.

A. Triple-Duplicate(TD) Period

Modeling the TCP throughput starts from the sending rate. For the ith TD period, we denote by Y_i the number of transmitted packets in the period, and by A_i the duration of the period. Then, the sending rate B can be expressed by

$$B = \frac{E[Y]}{E[A]} \tag{1}$$

To derive $E[Y]$, we should obtain the expected value of W_i first, which is defined by the window size at the end of the ith TD period as described in the Amherst model. For the ith TD period, we newly define c_i (bps) to be the available bandwidth, r_i to be the available router buffer size, W_i' to be the window size at the end of the period when $r_i = 0$, and s (bits) to be packet size. Firstly, if $r_i = 0$, the sending rate can reach as much as the available bandwidth before the packet loss occurs. Then, we can get c_i and W_i' as follows.

$$\begin{aligned} c_i &= \frac{W_i' \times s}{RTT} \\ W_i' &= c_i \times \frac{RTT}{s} \end{aligned} \tag{2}$$

The packet queuing begins when the congestion window size, $cwnd$, is larger than W_i', and the congestion window can increase additionally by the buffer size r_i. Therefore, $cwnd$ at the end of the ith period can be expressed by

$$W_i = W_i' + r_i \tag{3}$$

Manuscript received November 25, 2009. The associate editor coordinating the review of this letter and approving it for publication was G. Lazarou.

This work was supported by the Korea Research Foundation grant funded by the Korea government(MEST) (No. 2009-0076092).

The authors are with the Department of Computer Science and Engineering, Korea University, Seoul, Korea (e-mail: {jhhwang, hxy}@os.korea.ac.kr).

Digital Object Identifier 10.1109/LCOMM.2010.04.092309

Considering r_i to be i.i.d. random variable that is independent of the period number, it follows that

$$E[W] = E[W'] + E[r] \tag{4}$$

where $E[r] = \sum_{i=1}^{n} E[r_i]/n$ and n is the total number of the periods. To derive $E[W']$, we also consider the available bandwidth c_i to be i.i.d. random variable, which is assumed to be independent of the congestion window size, and thus independent of the period number, i. It follows that

$$E[W'] = E[c] \times \frac{RTT}{s} \tag{5}$$

where $E[c] = \sum_{i=1}^{n} E[c_i]/n$. Henceforth, we denote by $BW = E[c]$ the average value of the available bandwidth.

Next, from the Amherst model, Y_i is expressed by

$$Y_i = \sum_{k=0}^{\frac{X_i}{b}-1} \left(\frac{W_{i-1}}{2} + k\right) b + \beta_i, \tag{6}$$

where X_i is the number of rounds and β_i is the number of packets sent in the last round. We assume that the delayed ACK is not used for simplicity in this letter (i.e. $b = 1$). The congestion window size is increased by one for every round (RTT) from $E[W]/2$ to $E[W]$, and the total number of rounds in a TD period would be $E[W]/2 + 1$. $E[\beta]$ is replaced by $E[W]/2$ as the Amherst model does. From this, we can derive $E[Y]$, the expected value of transmitted packets, as follows.

$$\begin{aligned} E[Y] &= \sum_{k=0}^{E[W]/2} \left(\frac{E[W]}{2} + k\right) + \frac{E[W]}{2} \\ &= \left(\frac{E[W]}{2}\right)^2 + \frac{\frac{E[W]}{2}\left(\frac{E[W]}{2}+1\right)}{2} + E[W] \\ &= \frac{3}{8}E[W]^2 + \frac{5}{4}E[W] \end{aligned} \tag{7}$$

The next step is to derive $E[A]$. In the Amherst model, $E[A]$ is expressed by $(E[X]+1) \times RTT$, which is the product of mean RTT and the number of rounds. RTT basically is the sum of round-trip propagation delay and queuing delay. If the sending rate is not above bandwidth of a bottleneck link, RTT would be propagation delay largely. On the other hand, to include queuing delay by over-shooting, we should find the average value of marginal queuing delay, d, for a packet in the bottleneck buffer. With the assumption that bottleneck link and router are connected, we have

$$BW = \frac{s}{d} \rightarrow d = \frac{s}{BW} \tag{8}$$

And, RTT increases from round-trip propagation delay (RTT_{min}) to RTT_{max}, which can be expressed as

$$RTT_{max} = RTT_{min} + d \cdot E[r] \tag{9}$$

Thus, RTT_{max} means the average RTT at the last round of each period.

In the first round of a TD period, $cwnd$ starts from $E[W]/2$ which is less than $E[W']$, and RTT does not contain queuing delay (i.e. RTT_{min}). And then, it is increased by d as $cwnd$ is larger than $E[W']$. Thus, if $E[W]/2 \leq E[W'] < E[W]$, we can obtain $E[A]$ as follows.

$$\begin{aligned} E[A] &= \sum_{k=0}^{E[W]/2} RTT_{min} + \sum_{k=1}^{E[W]-E[W']} d \cdot k + RTT_{max} \\ &= RTT_{min}\left(\frac{E[W]}{2}+1\right) + \left(d \times \frac{E[r](E[r]+1)}{2}\right) \\ &\quad + (RTT_{min} + d \cdot E[r]) \end{aligned} \tag{10}$$

Similarly, when $E[W'] < E[W]/2$, RTT already contains queuing delay as much as $E[W]/2 - E[W']$. With (4), an initial RTT in a TD period will be $(E[r] - E[W]/2)d + RTT_{min}$. In this case, it follows that

$$\begin{aligned} E[A] &= \sum_{k=0}^{E[W]/2} \left(\left(E[r] - \frac{E[W]}{2}\right) d + RTT_{min}\right) \\ &\quad + \sum_{k=1}^{E[W]/2} d \cdot k + RTT_{max} \\ &= \left(\frac{E[W]}{2}+1\right)\left(\left(E[r] - \frac{E[W]}{2}\right) d + RTT_{min}\right) \\ &\quad + \left(d \frac{\frac{E[W]}{2}\left(\frac{E[W]}{2}+1\right)}{2}\right) + (RTT_{min} + d \cdot E[r]) \end{aligned} \tag{11}$$

B. TimeOut (TO) Period

In the Amherst model, the sending rate B in a TO period is modeled by using the possibility Q, which means that TDP ends with timeout. That is, the amount of transmitted packets (R) and the time (Z^{TO}) in a TO period are multiplied by Q. The sending rate B that combines TD and TO period is expressed by

$$B = \frac{E[Y] + Q \times E[R]}{E[A] + Q \times E[Z^{TO}]} \tag{12}$$

In this step, we need packet loss rate p to obtain Q, $E[R]$, and $E[Z^{TO}]$. From the Amherst model, we bring the following expression,

$$E[Y] = \frac{1-p}{p} + E[W], \tag{13}$$

and from this, we have

$$p = \frac{1}{\frac{3}{8}E[W]^2 + \frac{1}{4}E[W] + 1} \tag{14}$$

By this loss rate p, we can get Q, $E[R]$, and $E[Z^{TO}]$, but dispense with the deriving process in this letter.

Finally, we should consider the limitation of $E[W]$ by advertisement window W_m from the receiver side (i.e. $W_m \leq E[W]$). In this case, the sending rate will be kept as much as W_m/RTT since it must have been lower than the network capacity. This RTT may have queuing delay in accordance with the difference between W_m and $E[W']$, and thus

$$RTT = RTT_{min} + \max(0, W_m - E[W']) \times d \tag{15}$$

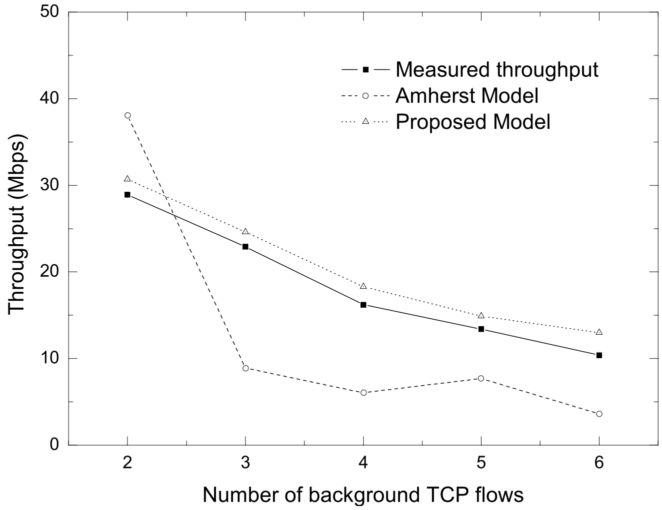

Fig. 1. Experiment results: throughputs of two models are predicted values.

Now, from (7), (10), (11), (15), we can express the TCP throughput model denoted by T, as a function of available bandwidth BW and router buffer size $E[r]$. To represent these parameters as packet unit, we explicitly use another capital letters, C and K for BW/s and $E[r]$ respectively. Then, we get the following for $T(C, K)$:

$$\begin{cases} \dfrac{\frac{1}{4}W\left(\frac{3}{2}W+5\right)+Q(W)}{RTT\left(\frac{W}{2}+2\right)+\frac{K}{C}\left(\frac{K+1}{2}+1\right)+QZ}, \\ \quad 0 < K \leq W/2, \quad W < W_m \\ \dfrac{\frac{1}{4}W\left(\frac{3}{2}W+5\right)+Q(W)}{RTT\left(\frac{W}{2}+2\right)-\frac{\left(\frac{W}{2}+1\right)\left(K-\frac{W}{4}\right)+K}{C}+QZ}, \\ \quad W/2 < K, \quad W < W_m \\ \dfrac{W_m}{RTT+\frac{\max(0, W_m - C \times RTT)}{C}}, \quad W_m \leq W \end{cases} \tag{16}$$

where QZ is $Q(W)E[Z^{TO}]$, and W is defined as a function of C and K from (4) and (5). It follows that

$$W(C, K) = C \times RTT + K \tag{17}$$

And also, we can derive $Q(W)$ and $E[Z^{TO}]$ from the Amherst model with (14) and (17). In (16) and (17), RTT is a minimum RTT, and TO is an initial RTO.

III. Evaluation

We perform real experiments for TCP throughput prediction over 100Mbps LAN environments with a simple dumpbell topology. The evaluation methodology is similar to [3]. An experiment starts with available bandwidth measurement for 20 to 60 seconds using Spruce [9], followed by a 60-sec measurement of RTT and loss rate using a ping utility that generates a 60-byte probing packet every 100ms, followed by a 60-sec TCP transfer (target flow) using IPerf [10]. To obtain the router buffer size, we use the router buffer estimation scheme, which is $(RTT - RTT_{min}) \times C$ [11].

To simulate network congestion situations in WAN, we increase the number of background TCP flows. TCP throughputs are predicted by both Amherst model and ours before the target TCP transfer starts. Then, we compare the predicted throughputs with the measured one as shown in Fig. 1. Through these experiments, we found that RTT increases linearly as TCP flows increase, but loss rate is fluctuated even by one or two packet losses because the number of probing packets is insufficient to estimate exactly (about 600 packets for 60 seconds). For instance, loss rate for one packet drop is about 0.17% and two-packet drop is about 0.33%. However, we also found that more frequent and invasive probing does not improve the accuracy of loss rate. In addition, the pre-measured RTT is different with experienced one since the target flow also affects the queuing delay. These factors lead the prediction by the Amherst model to under or over-estimate. On the other hand, our model overcomes such weakness by replacing the function parameter from loss rate to available bandwidth, and we confirm that the proposed scheme predicts the real throughput well.

IV. Conclusion

This letter proposes a modified TCP throughput model to improve the weakness of existing model. While the existing throughput model is a function of loss rate that is difficult to obtain with accuracy, our model is based on available bandwidth. We show that the accuracy of the formula-based predictor can be improved significantly based on our model. As future work, various wide-area network experiments need to be performed.

References

[1] J.-H. Choi and C. Yoo, "Analytical derivation of one-way delay," *IEE Electron. Lett.*, vol. 39, no. 25, pp. 1871–1872, Dec. 2003.
[2] http://www.icir.org/models/tools.html.
[3] Q. He, C. Dovrolis, and M. Ammar, "On the predictability of large transfer TCP throughput," in *Proc. ACM SIGCOMM*, 2005.
[4] D. Lu, Y. Qiao, P. Dinda, and F. Bustamante, "Characterizing and predicting TCP throughput on the wide area network," in *Proc. IEEE ICDCS*, 2005.
[5] M. Mirza, J. Sommers, P. Barford, and X. Zhu, "A machine learning approach to TCP throughput prediction," in *Proc. ACM SIGMETRICS*, 2007.
[6] M. Goyal, R. Guerin, and R. Rajan, "Predicting TCP throughput from non-invasive network sampling," in *Proc. IEEE INFOCOM*, 2002.
[7] J. Padhye, V. Firoiu, D. F. Towsley, and J. F. Kurose, "Modeling TCP Reno performance: a simple model and its empirical validation," *IEEE/ACM Trans. Networking*, vol. 8, no. 2, pp. 133–145, Apr. 2000.
[8] N. Parvez, A. Mahanti, and C. Williamson, "An analytic throughput model for TCP NewReno," to appear in *IEEE/ACM Trans. Networking*, 2009.
[9] J. Strauss, D. Katabi, and F. Kaashoek, "A measurement study of available bandwidth estimation tools," in *Proc. ACM SIGCOMM Conference on Internet Measurement*, 2003.
[10] http://dast.nlanr.net/Projects/Iperf/.
[11] K. Kaneko and J. Katto, "Reno friendly TCP Westwood based on router buffer estimation," in *Proc. Joint ICAS and ICNS*, 2005.