Evaluating performance of Connection Division Scheme in WLAN

Hasani Mohammad Mehdi, Sanaz Pashmforoush and Ehsan Shafei

Islamic Azad University, Aiatolah Amoli Branch, Iran

Abstract: The transmission control protocol (TCP) is widely used in wired and wireless networks. It provides reliable transport services between end-to-end hosts. Since TCP performance affects the overall network performance, several analytical models were proposed to describe the steady-state throughput of bulk transfer TCP flows (i.e., a flow with large amount of data to send, such as FTP transfers). However, most TCP flows in the Internet world are short-lived to see few losses and they cannot reach the steady-state, consequently their performance is determined by the startup effects such as the connection establishment and the slow start mechanisms. Surprisingly, all of the previous models did not investigate the heterogeneity of wireless networks which is considered the most important issue in wired-cum-wireless networks. The heterogeneity is shown by different characteristics and different segment loss probability for various types of wireless networks such as IEEE 802.11 WLAN and 3G cellular network. Moreover, wireless TCP flows are much shorter than wired flows due to the time varying characteristics of wireless networks. In this research, a recursive and analytical model is developed and used to determine the performance of TCP in heterogeneous wired-cum-wireless networks in terms of average completion time for the short-lived TCP flows is proposed. The proposed model focuses on heterogeneous wireless networks. The model tackles one scheme, namely, Connection Division scheme that split the connection at base station aiming to distinguish different characteristics between different networks. The proposed model is based on the knowledge of average dropping probability, the average round trip time and the flow size both wired and wireless links. The analytical model has been validated by means of simulations and using NS-2 simulator. The performance metric of TCP is the average completion time, which is the time that the source takes to successfully transfer a given amount of data in addition to the connection establishment time. It is shown clearly that the proposed model reflects the behavior of TCP; especially in computing the average completion time that increases as the session size (i.e. the number of data segments in the flow) increases. The simulation results are very much the same as the values obtained from the analytical model. The percentage difference for the average completion time (in Sec) is less than 3% and 6% for the the Connection Division scheme. The model can be used by engineers to tune some of parameters to see its effect on the behaviour of TCP.

Key words: Wireless networks • TCP • Connection division • Performance • Analytical model

INTRODUCTION

Transmission Control Protocol (TCP) is the ubiquitous transport protocol used in the Internet world. The Transmission Control Protocol is a reliable, connection-oriented, full-duplex, byte-stream and end-to-end protocol that supports flow and congestion control [1]. Also, it is widely used to support applications like Telnet and FTP [2]. TCP’s performance influences Internet traffic behavior. Hence, many models of TCP latency and throughput have been proposed, trying to capture its characteristics [3, 4]. In most of these models, the TCP performance (latency and throughput) is described based on the network parameters such as TCP round trip time and packet loss rate. In today’s world, more and more people are moving towards using mobile and wireless technology for communication. This technology is very much required for people on the move and makes possible fast and easy installation in remote areas. Also, more and more people are using their mobile devices to access the Internet either for work or entertainment. In recent years, the modeling of the TCP behavior has received considerable attention and many analytical models have been proposed with the purpose of characterizing TCP performance [4-7]. The recent works show that most TCP flows are very short-lived flows with average sizes of
around 10 KB [4, 8]. Since TCP was designed for wired networks and because TCP is still being the only protocol used in the Internet for reliable transfers, it fails to meet the requirements of wireless networks without reducing the performance of TCP. The standard TCP error recovery algorithms decrease the throughput in wireless links because TCP assumes all losses are caused by congestion, while in fact errors and packet losses in wireless networks can result from random bit errors, fading, shadowing, mobility, low bandwidth, handovers, channel losses and link latency [9, 10]. Many of the assumptions made in the wired-domain networks are not valid in the wireless-domain networks. From the shared, open-air media, to the characteristics of the physical channels, to the radio signal propagation challenges, to supporting mobile devices, the transport control protocol (TCP) faces many challenges in responding to these emerging needs. Therefore, TCP flows in wireless networks are much shorter than those in wired networks and their performance is dominated by startup effects such as three-way handshake connection establishment and slow start mechanisms. So, short-lived flows do not reach the steady-state. Therefore, the previous proposed analytical models cannot be used for short-lived flows. Consequently, alternative analytical models for short-lived TCP flows were proposed in [3, 11]. All the previous models were focused on the TCP performance over wired networks. In [11], a recursive and analytical model for wired networks was proposed to determine the short-lived TCP performance in terms of connection setup time and completion time. On the other hand, Pack et al. [12] extended the previous recursive and analytical model in [11] and introduced a new model for short-lived TCP flows in wireless networks. The model did not take into account the heterogeneity of wireless networks that is demonstrated by different loss packet probability, different bandwidth and different round trip time. To date, different types of wireless networks are found, such as IEEE 802.11 WLAN, ad hoc wireless networks and 3G cellular networks. They have different characteristics and different packet error behaviors. These different characteristics were not investigated in [12]. Therefore, analytical modeling of the short-lived TCP flows over heterogeneous wireless networks provides a good method for characterizing the TCP performance in terms of average completion time (i.e. the time that the source takes to successfully transfer a given amount of data in addition to the connection establishment time). We propose a recursive and analytical model for heterogeneous wireless networks with integration of wired network based on the proposed model in [12] and we focus on the average completion time of the short-lived flow as TCP performance metric. We introduce this model for Connection Division scheme that breaks the TCP Connection into two connections at the base station. The purpose of this division is to distinguish between different characteristics of networks. The Connection Division scheme is different from the Split Connection scheme that has been proposed in [13-15] and is used to discriminate the corruption losses of wireless networks from the congestion [16] losses of wired networks. The NS-2 simulator was used to validate our proposed model. The simulation values were compared with the model values for this scheme. Based on the obtained results we found that the proposed models reflect the behavior of TCP well, especially in computing the average completion time for the short-lived TCP flows at the initial stage: connection setup time (i.e. the time that the source takes to successfully complete the three way handshake algorithm) and slow start phase. The rest of this paper is organized as follows: section2 presents the previous work that is similar to our work or has in some other way dealt with similar problems, section 3 presents the proposed model and the analysis of the analytical solution, section 5 demonstrates the results and gives a complete analysis of the generated plots. Finally, section 6 concludes the paper and highlights the major components of the proposed model.

System Modeling and Analytical Solution

Model Objectives: The main goal of our proposed model is to estimate the average completion time (i.e. the time that the source takes to successfully transfer a given amount of data) for short-lived TCP flow in heterogeneous wireless networks integrated with wired networks.

A transfer is considered successful when the source receives the acknowledgment ACK for the last segment. This average completion time consists of two parts:

- The connection establishment time: the time it takes to successfully complete the three way handshake algorithm [2].
- The time to complete transferring the data segment.

Assumptions: Recent works have shown that the average size of most short-lived flows is around 10 KB [4, 8]. Assuming a unidirectional for data transfer with maximum transfer unit (MTU) of 1.5 KB, this will result in 10 KB / 1.5
KB = up to 8 segments maximum data transfer, which covers most of short-lived TCP flows. We consider the slow start algorithm implemented in [17]. Since all major TCP proposals such as Reno and New Reno implement the same slow start and the loss recovery algorithms.

We assume that the sender maintains only one timer for all in-flight segments [2, 17]. The timer is associated with the segment with the lowest sequence number that has not been acknowledged yet. When an ACK segment is received, the timer is reset and reassigned to the subsequent segment. Because of the small size of TCP flows, many of them never exit the slow start phase and the probability that the loss packet will trigger the fast recovery algorithm is very low. This is because the window size will be very small and the sender will not be able to send enough segments that generate the three duplicate acknowledgments that will trigger the fast recovery algorithm. Therefore, most losses will trigger the retransmission timeout (RTO) that will cause the TCP to reenter the slow start phase [12].

**Model Variables:** The proposed model must deal with two separate wired, in order to resolve the mismatch since the links are not homogeneous (i.e. determining the performance of TCP is not the purpose of this separation), to achieve the heterogeneity and to discriminate between different characteristics in different wireless networks, to make the system more practical such as finding a mobile host from a fixed host, to achieve the simplicity and tractability when splitting the connection from a fixed host to a mobile host at base station.

**Connection Division Scheme**

**Phase 1: Connection Establishment Time, C_{estt},**

**Derivation and Analysis:** In Connection Division scheme we divide a TCP connection at the base station into two separate TCP connections: wired and wired-wireless TCP connections. The main goal of Connection Division scheme is to distinguish between different characteristics of networks. Whereas, the Split Connection scheme is used to discriminate the congestion losses in wired networks from the corruption losses in wireless networks by splitting the connection at base station in to two connections: wired and wireless connections. We deal with wired-wireless link as End-to-End scheme.

**Derivation of Connection Establishment Time, C_{estt}:**

In this section, we will derive the connection establishment time depending on the variable distribution and using the geometric distribution.

\[ C_{estt} = C_{amp-wire} + C_{amp-wire-wire} \]

\[ = [R_{r} + (1 - p_{r}) O_{r} \sum_{i=1}^{n} 2^{2^{i-1}} T_{p} + \]

\[ R_{r}/2] + [R^{r}/2 + (1-p_{s}^{r}) O_{r} \sum_{i=1}^{n} (P_{s}^{r})^{2}]

\[ O_{r} \sum_{i=1}^{n} 2^{2^{i-1}} T_{p}]

\[ = 3R_{r}/2 + T_{p} (p_{s}^{r} (1-2p_{s}^{r})) + R^{r} + \]

\[ T_{p} (p_{s}^{r}) / (1-2p_{s}^{r})] \]

We derive \( p_{s}^{r} \) according to the properties of probabilities.

\[ p_{s}^{r} = \prod_{i=1}^{n} (p_{s}^{r}) + \prod_{i=1}^{n} \frac{1}{p_{s}^{r}} \times \frac{1}{p_{s}^{r}} \]

\[ R^{r} = (1/N) \times \sum_{i=1}^{n} [R_{r}(i) + R_{r}(i)] \]

\[ T_{p} = \text{Maximum} \{T_{p}(i), \text{worst case}\} \]

The base station sends the data to the destination node only after receiving all the data from the fixed host. The connection is started in wired-wireless links only after it is completed in the wired link, so an additional time (R/2), which is equal to half of the round trip time in the wired link is required. During the connection establishment time, if the segment is lost in wired link, the sender will retransmit the SYN segment after retransmission timeout (T_{p}) and double the SYN timeout value (T_{p} = 2 T_{p}). The terms R, R' is the average round trip time when there is no segment loss in the wired and wired-wireless links, the index (i) refers to the number of SYN segments dropped, the index (j) refers to the number of back-offs of the retransmission timer (T_{p}), \( p_{s}^{r} \) is the dropping probability for the SYN segment in wired link, \( (p_{s}^{r})_{\text{total}} \) is the overall dropping probability for SYN segment in wired-wireless links and \( (T_{p})_{\text{worst}} \) is the retransmission timer for the wired-wireless link.

**Phase 2: Data Completion Time C (m, w), Derivation and Analysis:** In the following, we will derive the general form for average completion time by using a recursive manner. Let \( C (m, w) = C_{wired}(m, w) + C_{wired-wireless}(m, w) \) be the average completion time required to successfully send (m) data segments with an initial congestion window of size (w) in wired and wired-wireless links, respectively. We compute the average completion time for wired-wireless links exactly as it was computed in End-to-End scheme. Let \( C (m, 1) \) is the average completion time of a flow of size (m) segments with an initial congestion window of size one. \( C (m, 1) = C (1, 1) + C (m-1, 2) \) as discussed in End-to-End scheme. We will use up to eight segments as a maximum data transfer as discuss in section 3.2 to derive the average completion time \( C (m, w) \) which is defined by a general form as shown b
\[
\begin{align*}
C_{\text{send}}(1,1) & + C_{\text{send}}(m-1,2) + C_{\text{send-end}}(1,1) + C_{\text{send-end}}(m-1,2) \\
, & m \geq 1, w=1 \\
C_{\text{send}}(1,1) + C \rightarrow (1,1) \\
, & m \geq 1, w=0 \\
C_{\text{send}}(m,m) + C_{\text{send-end}}(m,m) \\
, & w \geq m \\
C_{\text{send}}(m,w) + C_{\text{send-end}}(m,w) \\
q^2(T_1) + q_1(T_1 + C_{\text{send}}(m-1,2)) + p^2(T_1 + C_{\text{send}}(m-1,1)) + p_1(T_1 + C_{\text{send}}(m,1)) + q^2_{\text{total}}(R_1 + C_{\text{send-end}}(m-1,2)) + p^2_{\text{total}}(T_2 + C_{\text{send-end}}(m,1)) + p^2_{\text{total}}(T_2 + C_{\text{send-end}}(m,1)) + q^2_{\text{total}}(R^* + C_{\text{send-end}}(1,1)) + p^2_{\text{total}}(R^* + C_{\text{send-end}}(1,1)) + q^2_{\text{total}}(R^* + C_{\text{send-end}}(3,1)) + p^2_{\text{total}}(T_2 + C_{\text{send-end}}(2,1)) + p^2_{\text{total}}(T_2 + C_{\text{send-end}}(4,1)) + q^2_{\text{total}}(R_1 + C_{\text{send}}(1,1)) + 2 q_1(T_1 + R_1 + C_{\text{send}}(2,2) + 2 R_1 + C_{\text{send}}(2,1)) + q_1(T_1 + C_{\text{send}}(4,2)) + p_1(T_1 + C_{\text{send}}(4,1)) + q^2_{\text{total}}(R^* + C_{\text{send-end}}(1,1)) + 2 p^2_{\text{total}}(T_2 + R^* + C_{\text{send-end}}(2,1)) + q^2_{\text{total}}(R^* + C_{\text{send-end}}(4,1)) + 2 p^2_{\text{total}}(T_2 + R^* + C_{\text{send-end}}(4,1)) + q^2_{\text{total}}(R^* + C_{\text{send-end}}(3,2)) + 3 p^2_{\text{total}}(T_2 + C_{\text{send-end}}(3,1)) + p^2_{\text{total}}(T_2 + C_{\text{send-end}}(5,1)) + q^2_{\text{total}}(R^* + C_{\text{send-end}}(1,1)) + 2 q^2_{\text{total}}(T_2 + R^* + C_{\text{send-end}}(2,1)) + 2 p^2_{\text{total}}(T_2 + R^* + C_{\text{send-end}}(4,1)) + q^2_{\text{total}}(R^* + C_{\text{send-end}}(3,2)) + 3 p^2_{\text{total}}(T_2 + C_{\text{send-end}}(3,1)) + p^2_{\text{total}}(T_2 + C_{\text{send-end}}(5,1)) \\
, & m = 4, w = 4 \\
C_{\text{send}}(m,w) = & q^2_{\text{total}}(R_1 + C_{\text{send}}(1,1)) + 2 q_1(T_1 + R_1 + C_{\text{send}}(2,2) + 2 R_1 + C_{\text{send}}(2,1)) + q_1(T_1 + C_{\text{send}}(4,2)) + 3 p_1(T_1 + C_{\text{send}}(4,1)) + q^2_{\text{total}}(R^* + C_{\text{send-end}}(1,1)) + 2 q^2_{\text{total}}(T_2 + R^* + C_{\text{send-end}}(2,1)) + 2 p^2_{\text{total}}(T_2 + R^* + C_{\text{send-end}}(4,1)) + q^2_{\text{total}}(R^* + C_{\text{send-end}}(3,2)) + 3 p^2_{\text{total}}(T_2 + C_{\text{send-end}}(3,1)) + p^2_{\text{total}}(T_2 + C_{\text{send-end}}(5,1)) \\
, & m = 5, w = 4 \\
\end{align*}
\]

R_\text{avg} = \text{the average round trip time in wired1 link}

T_\text{avg} = \text{the retransmission timeout for the data segment in wired1}

p_\text{avg} = \text{the dropping probability in wired1 link}

q_\text{avg} = \text{the success probability in wired1 link}

p^\text{total} = \text{the overall dropping probability for the data segment in wired-wireless link.}

q^\text{total} = \text{the overall success probability for the data segment in wired-wireless link.}

R^\text{avg} = \text{the average round trip time in wired-wireless links}

R^\text{avg} = \frac{1}{N^*} \sum_{i=10}^{N} [R_\text{avg}(i) + R_\text{avg}(0)]

T_\text{avg} = \text{the retransmission timeout for the data segment in the second connection at base station = Maximum \{T_\text{avg}(i), \text{worst case}\}}

The average completion time for data transmission in the Connection Division scheme is the sum of the completion times in the wired link and wired-wireless links. Each completion time is calculated separately using the same recursive model as in End-to-End scheme. The analysis of these equations is the same as discussed before for the End- to- End scheme, the main difference is in the use of variables in wired link and wired-wireless links.

**Performance Evaluations and Results**

**Simulation Methodology:** NS-2 [18] was used as a simulation tool to validate our proposed model and to evaluate its performance in terms of average completion time for short-lived TCP flow.
Average Completion Time in End-to-End Scheme with N = 6

Fig. 1: Average Completion Time in Connection Division Scheme with N = 2

Average Completion Time in Connection Division Scheme with N = 3

Fig. 2: Average Completion Time in Connection Division Scheme with N = 3

Average Completion Time in Connection Division Scheme with N = 6

Fig. 3: Average Completion Time in Connection Division Scheme with N = 6
Table 1: Difference between Simulation Values and Model Values for Connection Division Scheme

<table>
<thead>
<tr>
<th>Number of wireless networks (N)</th>
<th>Average differences for the average completion time (ms)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>20.702</td>
</tr>
<tr>
<td>3</td>
<td>23.75</td>
</tr>
<tr>
<td>6</td>
<td>25.36</td>
</tr>
</tbody>
</table>

All values used in our simulations are the same as those used previously in other analytical models in [11,12]. The number of wireless networks (N) is a simulation parameter that varies. The bandwidth and round-trip delay of the wired links are 10 Mbps and 20 ms, respectively. On the other hand, the wireless link is assumed to follow the specification of IEEE 802.11 with omni antenna, the bandwidth, round-trip delay are 1 Mbps and 20 ms, respectively. Initial retransmission timers (RTO) in wired and wireless links are set to 1500 and 500 ms, respectively. The dropping probabilities in all wired links are 0.1%. Whereas the dropping probabilities in wireless links are different from one wireless network to another in order to achieve the heterogeneity which is due to the differences in their characteristics such as low bandwidth, random bit errors, fading, shadowing, channel losses and link latency. In terms of packet dropping, the uniform distribution is used.

Our simulations are based on the TCP-Reno, which is the most popular implementation in the Internet world. In all simulations, TCP-Reno sender (Fixed Host) transfers data segments to the receivers (Mobile Host). In the following sections, we discuss the results obtained for End-to-End scheme and Connection Division scheme.

Results and Performance Analysis: In this section we run the simulators using the same simulation methodology mentioned in section 3.1. The difference here is that we have two different configurations; one from fixed host to the base station and the other from base station to the mobile host. We used two, three and six wireless networks in these simulations. The dropping probabilities for these wireless networks are 1%, 2%, 6%, 4%, 8% and 2% respectively. Figures 10, 11 and 12, show that the average completion time increases as the session size (in segment) increases. From the above plots (10, 11 and 12), it is observed that the NS-2 simulation and the model result obtained from equations for the Connection Division scheme is relatively close to each other as shown in Table 1.

These plots give a conclusion that the percentage of difference between simulation and model value for average completion time (in Sec) is less than 6%. A comparison of the results shows that the difference between the simulation and the model values of the Connection Division scheme is larger than that of the End-to-End scheme. This is because the proposed model for Connection Division scheme calculates the total average completion time as a simple sum of the average completion times in wired and wired-wireless links. Also, the purpose of Connection Division scheme is different from the Split Connection scheme. The Connection Division scheme is used to distinguish between different characteristics of networks. The connection is started in wired-wireless links only after it is completed in the wired link. This will result in addition time (R)/2. Whereas, the Split Connection scheme is used to discriminate the congestion losses in wired networks from the corruption losses in wireless networks.

CONCLUSIONS

The Transmission Control Protocol (TCP) is widely used in the Internet world and several analytical models were proposed to improve the performance of TCP in terms of throughput. However, these models cannot be used for short-lived TCP flow which is considered as the most important characteristic in the Internet.

In this paper, we proposed a recursive and analytical model for short-lived TCP flows and we focused on heterogeneous wireless networks that have different packet dropping probability. We proposed this model for Connection Division scheme. The average completion time in the proposed wireless TCP model for heterogeneous wireless networks is calculated to determine the TCP performance using the average round trip time, retransmission timeout, flow size and the segment dropping probability. We used the NS-2 simulator to verify the accuracy of our model and we compared the simulated result for heterogeneous wireless topology with the calculated values that will be obtained from our proposed model.

Based on the obtained results, we found that the proposed model for Connection Division scheme reflects TCP behavior well, specifically the session completion time at the initial stage (i.e. slow start).

The results obtained show that the simulation and the model values are relatively close to each other with the percentage of difference for the average completion time (in Sec) being less than 3% and 6% for Connection Division scheme.
REFERENCES