A MODEL FOR CONGESTION CONTROL OF TRANSMISSION CONTROL PROTOCOL IN MOBILE WIRELESS AD HOC NETWORKS

Adib M. Monzer Habbal and Suhaiai Hassan
Internet Works Research Lab, School of Computing, Universiti Utara Malaysia, 06010 UUM Sintok, Kedah, Malaysia

Received 2012-06-02, Revised 2012-09-12; Accepted 2013-04-24

ABSTRACT

Transmission Control Protocol (TCP) is a fundamental protocol in the TCP/IP Protocol Suite. TCP was well designed and optimized to work over wired networks where most packet loss occurs due to network congestion. In theory, TCP should not care whether it is running over wired networks, WLANs, or Mobile Ad hoc Networks (MANETs). In practice, it does matter because most TCP deployments have been carefully designed based on the assumption that congestion is the main factor of network instability. However, MANETs have other dominating factors that cause network instability. Forgetting the impact of these factors violates some design principles of TCP congestion control and open questions for future research to address. This study aims to introduce a model that shows the impact of MANET factors on TCP congestion control. To achieve this aim, Design Research Methodology (DRM) proposed by BLESSING was used as a guide to present this model. The proposed model describes the existing situation of TCP congestion control. Furthermore, it points to the factors that are most suitable to be addressed by researchers in order to improve TCP performance. This research proposes a novel model to present the impact of MANET factors on TCP congestion control. The model is expected to serve as a benchmark for any intended improvement and enhancement of TCP congestion control over MANET.

Keywords: Transmission Control Protocol (TCP), Congestion Control, Wireless Ad Hoc Network, MANET, Design Research Methodology (DRM), Network Instability

1. INTRODUCTION

The escalating demand for mobile devices such as notebooks and PDAs has accompanied the growing interest in the use of Mobile Ad Hoc Network (MANET). MANET is a group of wireless mobile devices that connect to each other using a wireless channel (Boukerche, 2008), forming a temporary network without the aid of fixed infrastructure (Sarkar et al., 2007).

Recently, MANET has received considerable attention due to the attractive properties of building a network on-the-fly with ease, speed of deployment and decreased dependence on infrastructure. Moreover, MANET is suited for use in a large variety of applications, such as an emergency service at a disaster site, military operations in a remote area, business meetings held in venues without network infrastructure and sensor networks.

Transmission Control Protocol (TCP), meanwhile, is a de-facto standard protocol that provides reliable, end-to-end and connection-oriented data delivery over unreliable networks (Mohapatra and Krishnamurthy, 2005). It is the dominant protocol used on the Internet today, since it is responsible for the transmission of around 90% of Internet traffic (Mark and Aaron, 1999). TCP as an efficient transport protocol has several responsibilities, the main ones being to avoid congestion in the network and to adapt the transmission rate to available bandwidth. TCP congestion control is not just a service...
provided to the invoking application since it is a service for the Internet as a whole; it has been a critical factor in the robustness of the Internet (Floyd and Fall, 1999).

Using TCP over MANET will make network applications portable to MANET and will enable mobile users to interconnect with the wired network. These features and many others make TCP use over MANET a certainty.

So far, TCP was well designed to work over wired networks where most packet loss occurs due to network congestion. However, MANET suffers from different types of losses that are not related to congestion, thus leading to poor performance of TCP because TCP ignores the special characteristics of MANET, including high bit error rate, contention and mobility.

Many surveys have analyzed and compared the behavior of TCP over MANET (Floyd and Fall, 1999; Blessing and Chakrabarti, 2009; Wu and Wu, 2006; Andrei and Sally, 2004; Holland and Vaidya, 1999a; Ahmed et al., 2004), some of which have been conducted to improve and enhance the performance of TCP in this dynamic environment. However, this is not sufficient in providing a complete understanding of TCP current state over MANET. Moreover and based on our best knowledge, there is no study that presents one model that could be used as a benchmark for any intended improvement and enhancement of TCP over MANET. Thus, we have designed a model that aims to:

- Describe the existing situation of TCP
- Point at the factors that are most suitable to be addressed by researchers in order to improve TCP performance and
- Identify measurable factors that could be used to evaluate any proposed mechanism.

### 2. MATERIALS AND METHODS

Generally, a review of past research papers can be used as a starting point of any study, since it presents challenges and current issues related to the field of study. Furthermore, it helps researchers to identify the problem statement, objectives, scope. The same goes for TCP in MANET. However, we all know the proverb “a picture is worth a thousand words” and in this study, “a model is worth a thousand words” where the aim is to design a model to present the state-of-the-art TCP congestion control over MANET.

To accomplish this aim, we followed the Design Research Methodology (DRM) proposed by (Blessing and Chakrabarti, 2009). It is an approach and a set of supporting methods and guidelines that can be used as framework for performing design research. It helps making design research more rigorous, effective and efficient and its outcomes are academically and practically more worthwhile. Furthermore, Blessing advised that in presenting the current state of the studied phenomena, a reference model using special graphical symbols needs to be utilized. Here, we summarize the main characteristics of the graphical representation that we adopt from (Blessing and Chakrabarti, 2009) to present our model.

The oval nodes represent influencing factors. An influencing factor (or factor) is an aspect of the existing situation that influences other aspects of this situation, e.g., “the knowledge of wireless error loss” influences “the performance of congestion control”. An influencing factor is formulated as an attribute of an element that is considered relevant and can be observed, measured, or assessed. An example is “performance (attribute) of congestion control (element)”, as shown in Fig. 1.

Attributes have values, as explained in the following:

- Values can be qualitative, such as ‘high’ and ‘poor’, or quantitative, e.g., ‘20’, ‘larger than 20’, depending on the definition used in a particular study
- The existing or desired value of an attribute is attached to a link between factors by means of a ‘+’, ‘-‘, or ‘0’ sign. For example, a ‘-‘ next to the link and near the factor “performance of congestion control” indicates “low performance of congestion control”

Key Factors are those influencing factors that seem to be the most useful factors to be addressed in order to improve an existing situation. These are considered the core factors or root causes. The Key Factors are addressed directly by the proposed mechanism. In our example, the researchers might decide, on the basis of their investigations, that the “performance of congestion control” factor would be the most promising approach to improve “network resource utilization”.

![Fig. 1. Factor, attribute and element](image-url)
Success Criteria relate to the ultimate goal to which the research intends to achieve. These criteria usually reveal the purpose of the research and the eventual, expected contribution. In our model, the success criteria is to “increase efficient utilization of network resources”, where success criteria would relate to the Success Factors, plus its desired value.

The definition of success is still a topic of research, as many factors influence success. Sometimes, it could be difficult to measure or evaluate success factors within the research timeline or it could not be measured directly. What is needed in such cases are Measurable Success Criteria, i.e., criteria that are linked to the chosen success criteria and can be applied to evaluate the outcomes of the research, given the resources available. The factors where the desired values are taken as measurable success criteria are called the Measurable Success Factors. In our model, this would be to “increase the value of throughput”.

The link (arrow) between factors shows how the factors influence or desired to influence each other, i.e., it represents explicit statements about the existing or desired situation. The combination of ‘+’, ‘-’ and ‘0’ signs at the ends of a link describes how the value of the attribute of the factor at one end relates to the value of the attribute of the factor at the other end. Every link is labeled with the source of the statement it represents, using the following abbreviations:

[X]: the statement was published in the reference numbered X

Summarizing the above, a statement about the existing or desired situation can thus be modeled as two nodes representing the factors involved, connected by a line that is marked at either end with the values of the attributes of the factors to represent the details of their relationship. In the case of a causal link, the line becomes an arrow, pointing at the effect. The line is labeled with the source of the statement. Figure 2 shows our modeling terminology using the graphical representation of a statement from the literature source (Boukerche, 2008) stating that a “low knowledge of wireless error loss has a negative effect on the performance of congestion control”.

3. RESULTS

As stated in the previous section, this study introduces a novel model to present the state-of-the-art TCP congestion control over MANET. The model shows a roadmap to guide researchers’ steps towards enhancing TCP congestion control. It is called Reference Model as shown in Fig. 3. It consists of a collection of MANET factors that have impact on TCP and a group of measurable factors that could be used to study TCP performance. We identified factors and links that should be considered in our model based on intensive experimental studies conducted by us and other researchers. Furthermore, reviewing and analyzing the literature had helped us to verify and validate our model. Our discussion has been classified based on the following phenomena: high bit error rate, contention and mobility (Wu and Wu, 2006).

3.1. High Bit Error Rate

It is a well-known fact that the wireless channel is weak, unreliable and unprotected from outside signals. Therefore, link layer technologies should deal with corrupted packets. However, sometimes corrupted packets could not be repaired in a timely fashion (Andrei and Sally, 2004). Holland and Vaidya (1999a) stated that the link layer retransmission technique can recover from channel error loss in low number of hops and can achieve rather low error rates, but when the number of hops increases the recovery becomes difficult. If the link layer abandons its attempts at link-level retransmission, TCP takes the loss as a congestion indicator. This misinterpretation of packet loss due to wireless channel error as congestion has badly affected TCP performance (Boukerche, 2008; Ahmed et al., 2004; Lochert et al., 2007). Therefore, there is a need for a detection mechanism to correctly interpret the loss as corruption instead of congestion (Habbal and Hassan, 2011).
Fig. 3. Reference model
From the previous discussion, we conclude with:

- Statement A-low knowledge of wireless error loss has negative impact on TCP performance (Boukerche, 2008; Lochert et al., 2007; Habbal and Hassan, 2010; 2011) (i.e., poor throughput); this statement is represented by Link A in Fig. 3.

3.2. Contention

Other researchers look from different perspectives since nodes in ad hoc networks use wireless channel as a shared medium to send and receive data. The shared channel allows a single sender to transmit at a time. Therefore, senders within a local neighborhood have to compete for wireless channel access before transmitting. In addition, the number of segments that can be in-flight concurrently is limited from a source to a destination. Fu et al. (2005) observed that packet drop due to link-layer contention dominates in ad hoc networks while buffer overflow-induced packet loss is rare.

We identified three factors that may contribute to increase the level of contention: (1) the large congestion window increases the contention among the TCP data packets in the forward path (Fu et al., 2005; Chen et al., 2003; Kim et al., 2005; Triantafyllidou et al., 2009) (2) generating ACK for every arrived packet will increase the contention between data and ACK packets in the forward and return path (Singh and Kankipati, 2004; Oliveira and Braun, 2007; Chen et al., 2008) and (3) hidden and exposed terminals (Chen et al., 2008; Cai et al., 2008). These factors cause an excessive number of medium accesses (Cai et al., 2008; Kim et al., 2006) that increases the level of contention. Consequently, the high level of contention may lead to one of the following: (1) degrade TCP performance; (2) increase the probability of route failure (Link E in Fig. 3) and again degrade TCP performance (Link F in Fig. 3); or (3) produce high delay (Boukerche, 2008), which may lead to unnecessary retransmission (Link I) and as a result, this would degrade TCP performance.

Previous discussion produces the following statements:

- Statement B-large congestion window size increases the level of contention (Fu et al., 2005; Chen et al., 2003; Triantafyllidou et al., 2009), which is represented by Link B in Fig. 3
- Statement C-Large number of generated ACKs increases the level of contention (Singh and Kankipati, 2004; Oliveira and Braun, 2007; Chen et al., 2008), which is represented by Link C in Fig. 3
- Statement D-Hidden and exposed terminal increases the level of contention (Fu et al., 2005; Cai et al., 2008), which is represented by Link D in Fig. 3
- Statement G-Large (high) contention level has negative impact on the TCP performance (Fu et al., 2005; Chen et al., 2008) (i.e., poor throughput and fairness and high delay and packet loss), which is represented by Link G in Fig. 3
- Statement E-Large (high) contention level increases the probability of route failure/change (Kim et al., 2006), which is represented by Link E in Fig. 3
- Statement F-Frequent route failure/change increases the delay (Boukerche, 2008), which is represented by Link F in Fig. 3
- Statement H-Exposed terminal increases the delay (Boukerche, 2008), which is represented by Link H in Fig. 3
- Statement Q-Large (high) contention level increases the delay (Boukerche, 2008), which is represented by Link Q in Fig. 3
- Statement I-Large (high) delay increases unnecessary retransmission (Kim et al., 2007; Fu and Atiquzzaman, 2005), which is represented by Link I in Fig. 3 and
- Statement J-Large (high) unnecessary retransmission has negative impact on the TCP performance (Kim et al., 2007; Fu and Atiquzzaman, 2005) (i.e., poor throughput), which is represented by Link J in Fig. 3.

3.3. Mobility

All devices are free to move randomly in MANET, so an ongoing connection should be kept alive to support mobility. However, once the host moves, the topology will change and this may lead to route failure/change or network partitioning.

Authors in (Al-Hanbali et al., 2005; Fu et al., 2002; Ahuja et al., 2000; Abdullah-Al-Mamun et al., 2006; Holland and Vaidya, 1999b) observed that mobility-induced frequent route failures have a significant impact on the performance of TCP, due to TCP’s inability to recognize the difference between route failure/change loss and congestion loss (Zaini et al., 2012). This leads to the following statements:

- Statement K-mobility increases the number of route failure/change (Al-Hanbali et al., 2005; Fu et al., 2002; Ahuja et al., 2000; Abdullah-Al-Mamun et al., 2006; Holland and Vaidya, 1999b; Premalatha and
Natarajan, 2011), which is represented by Link K in Fig. 3; and

- Statement F-frequent route failure/change has negative impact on TCP performance (Al-Hanbali et al., 2005; Fu et al., 2002; Ahuja et al., 2000; Abdullah-Al-Mamun et al., 2006; Holland and Vaidya, 1999b; Premalatha and Natarajan, 2011) (i.e., poor throughput, high loss and delay), which is represented by Link F in Fig. 3

A network partition, meanwhile, occurs when a node in MANET moves away from the network, thereby causing an isolation of some part of the network by breaking it into two parts. These isolated parts are called partitions. During this period, the TCP sender cannot receive any ACKs transmitted by the TCP receiver. If the disconnectivity persists for a duration greater than the Retransmission Time Out (RTO) of the sender node, the TCP agent will trigger the exponential back-off algorithm, which consists of doubling the RTO whenever the timeout expires. The RTO back-off mechanism exaggerates the effect of network partition, hence, when network partition occurs, a series of expired RTOs will follow (Link M) and the RTO value will double (Link N), leading to a large RTO value which will reduce the responsiveness of TCP (Link O). In other words, Consecutive timeout is the key factor that affects TCP performance (Touati et al., 2009; Dyer and Boppana, 2001; Lin et al., 2006).

Previous discussion produces the following statements:

- Statement L-mobility increases the probabilities of network partition (Habbal and Hassan, 2010; Touati et al., 2009; Lin et al., 2006; Choi et al., 2003), which is represented by Link L in Fig. 3

- Statement M-long network partition produces a series of timeouts (Touati et al., 2009; Lin et al., 2006; Choi et al., 2003), which is represented by Link M in Fig. 3

- Statement N-series of timeouts increases the value of RTO (Touati et al., 2009; Lin et al., 2006; Choi et al., 2003), which is represented by Link N in Fig. 3

- Statement O-a large RTO value has negative impact on TCP performance (Touati et al., 2009; Lin et al., 2006; Choi et al., 2003), which is represented by Link O in Fig. 3

4. DISCUSSION

All derived statements from A to O have been combined to design our reference model, as shown in Fig. 3. Based on our model, researchers can easily determine the most suitable factor (key factor) to be addressed in their study to improve TCP performance in MANET. For instance, a researcher may want to focus on reducing and eliminating the influence of wireless error loss on TCP congestion control performance. Accordingly, he or she would try to come up with a sort of mechanism that changes the “low knowledge of wireless error loss” to “high knowledge of wireless error loss”. As a result, this would reflect on the relation between the knowledge of wireless error loss and the performance of TCP congestion control. In addition, the researcher will be able to design a conceptual model to describe the desired situation that is expected, as a consequence of addressing TCP congestion control using the proposed detection mechanism.

5. CONCLUSION

Enhancing TCP to perform well in MANET is such a challenging task in the presence of various factors, including but not limited to high bit error rate, contention, mobility and limited power.

Thus, setting up challenging task in the presence of various factors, including but not limited to high bit error rate, contention, mobility a graphical representation of the existing situation in the form of a reference model helps to structure research findings and acts as an important basis for future research. In this study, we presented the design of a reference model, which serves as a benchmark for any intended improvements and enhancements of TCP over MANET. Accordingly, the resulting model will help the community to:

- Obtain a better understanding of the existing situation of TCP by identifying and clarifying the factors that influence its performance
- Illustrate and clarify the line of argumentation for developing a specific mechanism
- Provide a list of measurable success factors that could be used to evaluate the proposed mechanism (e.g., throughput, packet loss ratio and delay)

6. ACKNOWLEDGMENT

The researchers would like to thank Mr. Faisal bin Zulhumadi, MS. Haniza Binti Nahar and Mr. Khuzairi Zaini et al. (2012) for their useful comments on this study. Furthermore, this work is funded by Ministry of...
Higher Education, Malaysia, under the ERGS scheme S/O code: 12474.

7. REFERENCES


