A new TCP/AQM system analysis

Xu, Qin; Li, Fan; Sun, Jinsheng; Zukerman, Moshe

Published in:
Journal of Network and Computer Applications

Published: 01/11/2015

Document Version:
Post-print, also known as Accepted Author Manuscript, Peer-reviewed or Author Final version

License:
CC BY-NC-ND

Publication record in CityU Scholars:
Go to record

Published version (DOI):
10.1016/j.jnca.2015.07.004

Publication details:
Xu, Q., Li, F., Sun, J., & Zukerman, M. (2015). A new TCP/AQM system analysis. Journal of Network and Computer Applications, 57, 43-60. Advance online publication. https://doi.org/10.1016/j.jnca.2015.07.004

Citing this paper
Please note that where the full-text provided on CityU Scholars is the Post-print version (also known as Accepted Author Manuscript, Peer-reviewed or Author Final version), it may differ from the Final Published version. When citing, ensure that you check and use the publisher's definitive version for pagination and other details.

General rights
Copyright for the publications made accessible via the CityU Scholars portal is retained by the author(s) and/or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights. Users may not further distribute the material or use it for any profit-making activity or commercial gain.

Publisher permission
Permission for previously published items are in accordance with publisher's copyright policies sourced from the SHERPA RoMEO database. Links to full text versions (either Published or Post-print) are only available if corresponding publishers allow open access.

Take down policy
Contact lbscholars@cityu.edu.hk if you believe that this document breaches copyright and provide us with details. We will remove access to the work immediately and investigate your claim.

Download date: 03/02/2024
A New TCP/AQM System Analysis

Qin Xu\textsuperscript{a}, Fan Li\textsuperscript{b}, Jinsheng Sun\textsuperscript{c,}\textsuperscript{*}, Moshe Zukerman\textsuperscript{b}

\textsuperscript{a}School of Network and Communications Engineering, Jining Institute of Technology, Nanjing, Jiangsu Province, PRC
\textsuperscript{b}Department of Electronic Engineering, City University of Hong Kong, Hong Kong SAR, PRC
\textsuperscript{c}Department of Automation, Nanjing University of Science and Technology, Nanjing, Jiangsu Province, PRC

Abstract

The Misra, Gong and Towsley (MGT) fluid model for Transmission Control Protocol (TCP) has been used extensively to guide designs of Active Queue Management (AQM) schemes aiming to alleviate adverse effects of Internet congestion. In this paper, we provide a new analysis of a TCP/AQM system that aims to improve the accuracy of the MGT fluid model especially in heavy traffic conditions. The analysis is based on the consideration of two extreme congestion scenarios that leads to the derivation of upper and lower bounds for the queue length and marking probability dynamics and showing that they approach each other in steady state. Both discrete and continuous time models are provided. Simulation results demonstrate that the new model achieves a significantly higher level of accuracy than a simplified version of the MGT fluid model.

Keywords: MGT fluid model, control system, networking, congestion control, TCP/AQM

1. INTRODUCTION

For over three decades researchers have been seeking efficient and distributed means to control Internet traffic congestion. The common approach has been based on congestion indications generated by links and sent to their sources which then adapt their send rate. The scheme implemented at the links which decides on how and when to indicate congestion is called Active Queue Management (AQM) and the protocol that dictates how to adapt sources rate is mainly TCP (herein TCP means TCP Reno). The combined TCP/AQM system aims to achieve efficient resource utilization, acceptable packet loss and a stable and robust operation.

\*Corresponding author. Tel: +(86) 2584315467
Table 1: The Classification of AQM Schemes

| Congestion Detection | Dropping Probability Update | Event-Driven | Time-Driven |
|----------------------|-----------------------------|--------------|-------------|
| Queue Base           | RED [3] FPID [14]           | PI [4] PID [13] | MPAQM [15] |
| Rate Based           | Green [5] BLUE [6]          | AVQ [7]      |             |
| Queue and Rate Based | Yellow [9] REM [8] RaQ [10] |              |             |

Congestion indication can be done by either dropping or marking incoming packets. If Explicit Congestion Notifications (ECN) [1] is enabled, the packets will be marked instead of being discarded, which increases TCP goodput. Given this advantage of ECN, we focus in this paper on the AQM schemes with ECN enabled unless otherwise mentioned.

Many AQM schemes have been proposed, they can be classified into queue-based, rate-based and queue and rate-based combined according to their congestion detection methods [2]. RED [3] and PI [4] are queue-based. Green [5], BLUE [6] and AVQ [7] belong to the rate-based class. The combined queue and rate-based category includes REM [8], Yellow [9], and RaQ [10]. An alternative classification into event- and time-driven AQM schemes is based on the method of updating marking probability [11]. RED is a typical event-driven AQM scheme. The marking probability of RED is reset at packet arrival events. AVQ and E-AVQ [12] are also event-driven where resets take place at packet arrivals. BLUE [6] is event-driven as it updates marking probability when the buffer is empty or full. However, many AQM schemes are time-driven, including PI, PID [13], REM, RaQ. AQM schemes of this type update the marking probability when a certain timeout expires. Table 1 provides classifications of various AQM schemes.

Designing efficient, stable and robust TCP/AQM systems requires the understanding of their dynamics. To achieve such understanding, various models have been proposed [16–18]. However, the model of [16] is based on certain simplified assumptions which introduce significant inaccuracies, and the models proposed in [17] and [18] are not scalable. Misra, Gong and Towsley [19] proposed an analytical model for TCP/AQM systems which is called in [20] the MGT fluid model after the authors’ initials. A simplified version of the MGT fluid model was provided in [21] and has been widely used. It provided theoretical foundations for numerous analyses and syntheses of TCP/AQM systems and gave rise to many newly proposed AQM schemes, e.g. [22, 23].

In this paper, we use the model of [21] as a benchmark and we refer to it as the Simplified MGT model. It considers TCP/AQM as a feedback system where the number of packets allowed to be sent without acknowledgement by the source
(sender), designated as the congestion window, is adapted based on congestion indications generated by the AQM (either as lost or marked packets). However, it introduces certain inaccuracy because it assumes independence between the congestion window and the number of lost or marked packets. The inaccuracy tends to increase with increased load. Furthermore, the MGT model of [19] does not consider practical effects, such as the effect of ECN and modes of TCP operation, (e.g. slow start and congestion avoidance).

In our modeling of a TCP/AQM system we overcome the weaknesses of the Simplified MGT model by considering the above mentioned practical effects. An important distinction between our model and the Simplified MGT model is associated with the roles that TCP and AQM play in the feedback TCP/AQM system. The Simplified MGT model adopts the traditional approach, where TCP is the plant and the AQM is the controller that provides feedback to the TCP. In our model their roles are reversed where the TCP/AQM system is viewed as a system where AQM is the plant and TCP is the controller that provides feedback to the AQM.

We analyze the system dynamics by considering two congestion scenarios which give rise to upper and lower bounds of the queue length and marking probability processes, and show that these bounds are close to each other. Furthermore, we derive a continuous-time model based on differential equations that describe the TCP policy and the link queueing process. We also provide an equivalent discrete-time model for ease of implementation. For a given AQM scheme, our model can provide the transient behavior which include both queue length and marking probability. Both the Simplified MGT model and our model are directly applicable to time-driven AQM schemes and could be applicable to an event-driven AQM if it is approximated by a time-driven system. This paper focuses on time-driven AQM schemes and provides numerical results on PI, REM, RaQ which are all time driven.

The remainder of this paper is organized as follows. In Section 2, the Simplified MGT model is described. In Section 3, we provide an extensive analysis that leads to a new model of a TCP/AQM system which yields bounds for queue length and marking probability dynamics. Section 4 presents simulation results over a wide range of scenarios and parameter values that validate our new model, and illustrate the closeness of its bounds. Using the simulation results, we also demonstrate the improvement in accuracy achieved by our model versus that of the Simplified MGT model. Finally, we present our conclusions in Section 5.

2. THE SIMPLIFIED MGT MODEL

In a TCP/AQM system, TCP adjusts the send rate to avoid congestion according to the congestion indication received from the AQM. Fig. 1 shows the working mechanism of a TCP/AQM system. In the figure, the variable $w$ is the congestion window at the TCP end, and $p$ is the marking probability calculated by the AQM.
The original MGT fluid model [19] assumed that packet losses in a flow follow a non-homogeneous Poisson process. It considers the individual behaviors TCP sessions and their timeout mechanisms. In the Simplified MGT model [21], the authors ignored the TCP timeout mechanism and provided a set of differential equations that describe the dynamics of the TCP window at the sources and the queueing at the link for this case. Then, this Simplified MGT model is written as:

\[
\begin{align*}
\dot{W}(t) &= \frac{1}{R(t)} - \frac{W(t - R(t))W(t)}{2R(t)}p(t - R(t)), \\
\dot{q}(t) &= -C + N\frac{W(t)}{R(t)},
\end{align*}
\]

(1)

where \(N\) is the number of TCP sessions, \(W(t)\) is the expected congestion window size for a TCP session at time \(t \geq 0\), \(R(t)\) is the expected Round Trip Time (RTT) at time \(t \geq 0\), \(p(t)\) is the marking probability at time \(t \geq 0\), \(C\) is the link capacity, and \(q(t)\) is the expected link queue length at time \(t \geq 0\). The operating point \((W_0, q_0, p_0)\) is defined by \(\dot{W}(t) = 0\) and \(\dot{q}(t) = 0\), where \(q_0\) denotes the queue length at the operating point and

\[
p_0 = \frac{2N^2}{R_0^2C^2},
\]

(2)

\[
W_0 = \frac{R_0C}{N},
\]

(3)

where \(R_0\) is the RTT at the operating point.

The original MGT fluid model [19] and the Simplified MGT model [21] have gained popularity after they had been proposed since they provided analyses of TCP dynamics together with the stability region. Many publications adopted them as a stability guideline to design a stable TCP/AQM system. The authors of [24] reviewed fluid models of TCP traffic including [21], and proposed a nonnegative matrix method to model the dynamics of TCP flows that use AIMD congestion control algorithm and drop tail queues without considering
other AQM schemes. In [25], the author proposed a duality model of end-to-end congestion control model to better understand the equilibrium properties of TCP/AQM systems which adopted the Simplified MGT model as a local stability criterion. Other publications such as [26, 27] focus mainly on stability issues rather than the performance of model when TCP traffic is under heavy load.

The Simplified MGT model is known to be accurate for light traffic, but it is inaccurate under heavy traffic. This is partly because the model may yield \( p_0 > 1 \) in heavy traffic, and then if \( p_0 \) is truncated to 1, the model may not be able to converge to the target queue size, and if it is not truncated, there the marking probability is erroneous and the queue dynamic may exhibit much slower convergence than the real system (see Section 4). Our modeling approach overcomes the \( p_0 > 1 \) weakness as discussed below. For brevity, the Simplified MGT model hereafter designates the truncated version of the model. We will also discuss the untruncated version of the Simplified MGT model in Section 4.

3. A NEW TCP/AQM MODEL

We begin our study of the TCP/AQM system by modeling and analysis of the AQM marking. As mentioned, a time-driven AQM updates its marking probability every sampling interval. This enables the consideration of the sources reactions on the queue length in one sampling period. Next, two scenarios of TCP associated with two extreme congestion levels are considered, which lead to upper and lower bounds for the queue and marking probability dynamics. Both discrete- and continuous-time models are provided. Afterwards, we discuss the effect of ECN on the queue dynamics, followed by an analysis of the operating point where the two scenarios converge at the steady state under the assumption that the entire system is stable. To be able to test the accuracy of the model for a wide range of cases, three non-overlapping and exhaustive congestion levels are then defined and corresponding working scenario(s) are presented. This follows by a discussion on the closeness of the bounds, the system stability and on the setting of a key parameter.

3.1. Analysis of an AQM Marking Process

In most AQM schemes, packets will be marked or dropped according to marking probability and traffic conditions during a sampling interval. In our analysis, we assume that ECN is always ON and the buffers are sufficiently large so that packet dropping is excluded. Then for the marking process during a sample interval, we make the following assumptions:

1. For each packet, there are two mutually exclusive outcomes: marked or unmarked.
2. During each sampling internal, the marking of each packet is independent of the marking of other packets.
3. The marking probability remains constant within the duration of a sampling interval.
Accordingly, given the marking probability and the number of packets that arrive during a sampling interval, the number of marked packets within this interval is a Binomial random variable.

Let \( m(t) \) be the number of packets that arrive at the link during the time period from \( t \) to \( t + \tau \), where \( \tau > 0 \) is an infinitesimally small time period. Let \( p(t) \) be the marking probability. As the number of marked packets is Binomially distributed, its expectation is \( m(t)p(t) \), and the expected number of unmarked packets is \( m(t)(1 - p(t)) \). Let the sum of all TCP congestion windows \( W_s(t) \) increases by \( a(t) \) for every unmarked packet, and decreases by \( b(t) \) for every marked packet, at time \( t \geq 0 \). Let \( \Delta W_s(t + R(t)) \) be defined by:

\[
\Delta W_s(t + R(t)) = W_s(t + R(t) + \tau) - W_s(t + R(t)).
\]  

Then, we obtain

\[
\Delta W_s(t + R(t)) = a(t)m(t)(1 - p(t)) + b(t)m(t)p(t).
\]  

The function \( R(t) \) here is equivalent to the \( R(t) \) of the Simplified MGT model used in (1). It represents the average over all TCP sessions, that are active during the time period from \( t \) to \( t + \tau \), of their time delay from the moment a packet arrives at a link until its acknowledgement arrives at the sender. This delay is the RTT minus the time it takes the packet to reach the link from the moment it leaves the sender. However, for tractability, \( R(t) \) is approximated as the average RTT which is given by

\[
R(t) = T_p + \frac{q(t)}{C},
\]  

where \( T_p \) is the average propagation time, \( C \) is the link capacity, \( q(t) \) is the average queue length during the time period from \( t \) to \( t + \tau \), and \( q(t)/C \) is the mean queueing delay of a packet arriving in that time period. By the approximation we made in (6), our \( R(t) \) is the same as that of the MGT Fluid model which is also based on this approximation.

We assume that \( m(t) \) takes the form:

\[
m(t) = \frac{W_s(t)}{R(t)} \tau,
\]  

where \( W_s(t)/R(t) \) represents the packet arrival rate.

Marked packets may belong to different TCP sessions and therefore have different effects on TCP window sizes, for simplicity, we assume that the processes \( a(t) \) and \( b(t) \) are the same for all sessions. Appropriate values for them will be their expected values over all sessions which will be obtained through the analysis of TCP dynamics in the next subsection.

3.2. Two Working Scenarios of TCP Data Transmission

TCP is composed of four algorithms [28–30] which give rise to four data transfer phases: slow start, congestion avoidance, fast retransmit and fast recovery. There are two important variables for these algorithms: congestion window (\( cwnd \)) that limits the amount of data that TCP can send, and slow start
threshold size ($ssthresh$) that determines whether the slow start or congestion avoidance algorithm is used. Slow start (involving exponential growth of $cwnd$) continues until $cwnd$ reaches the threshold $ssthresh$ which has a lower bound of two packets. Then, congestion avoidance (involving linear $cwnd$ growth) is used. When congestion occurs, fast retransmit and fast recovery are used. A typical single-source single-link TCP $cwnd$ process is illustrated in Fig. 2.

To introduce our new model, it is convenient to consider the following two scenarios:

- **Scenario A:** In all the TCP sessions $cwnd < ssthresh$, so they are either in slow start, or in fast recovery.
- **Scenario B:** In all the TCP sessions $cwnd \geq ssthresh$, so they are either in congestion avoidance, or in fast recovery.

Scenarios A and B represent two extreme cases where all the TCP sessions behave in a similar way. For tractability, these scenarios exclude cases where in some sessions $cwnd < ssthresh$ and in others $cwnd \geq ssthresh$. Nevertheless, we demonstrate by simulations in Section 4 that our analysis of Scenario A and B is sufficient to lead to accurate evaluations for a wide range of system states even if the conditions of neither Scenario A nor B are satisfied.

In the example presented in Fig. 2, Scenario A takes place at time 0, and then Scenario B takes over at time 8. This example represents a light traffic case involving a single TCP session. As the number of sessions increases, and
more sessions need to share a limited capacity, their *cwnds* may be less than 2 packets and therefore less than *ssthresh*, then Scenario A dominates.

To derive the mathematical model of the two scenarios, consider a system where there is a single congested link with capacity $C$. Let $N$ be the number of TCP flows that use this link; they are labeled by $i = 1, 2, \ldots, N$. Then let $W_i(t)$ and $R_i(t)$ denote the window size and the RTT of flow TCP$_i$ ($i = 1, 2, \ldots, N$) at time $t > 0$, respectively. In the slow start phase, when an acknowledgement is received by the sender of an unmarked packet, the congestion window will increase by one packet. We obtain:

$$a(t) = \sum_{i=1}^{N} \frac{W_i(t)}{W_s(t)} \times 1 = \sum_{i=1}^{N} \frac{W_i(t)}{W_s(t)} = 1, \quad t > 0, \quad (8)$$

where $a(t)$ is defined in (5), which represents the increment of $W_s(t)$ for every unmarked packet, the ratio $W_i(t)/W_s(t)$ represents the proportion of total data packets that generated by TCP$_i$.

In the congestion avoidance phase, the congestion window will increase by one packet in every RTT, then the increment of the congestion window can be expressed as:

$$a(t) = \sum_{i=1}^{N} \frac{W_i(t)}{W_s(t)} \frac{1}{W_i(t + R(t))} \approx \frac{N}{W_s(t + R(t))}, \quad t > 0. \quad (9)$$

The function $a(t)$ has two expressions for two different phases. When the system is in the slow start phase, the congestion window *cwnd* will be incremented by one packet for every unmarked packet. In this case, the congestion window *cwnd* will double (8). For other times $t$ during which the system is in the congestion avoidance phase, the congestion window *cwnd* will be incremented by one packet in every RTT (9).

Note that in (9) we make the following approximation:

$$\sum_{i=1}^{N} \frac{W_i(t)}{W_s(t)} \frac{1}{W_i(t + R(t))} \approx \sum_{i=1}^{N} \frac{W_i(t + R(t))}{W_s(t + R(t))} \frac{1}{W_i(t + R(t))}. \quad (10)$$

This approximation is made to simplify the derivation of the window increment $a(t)$, which has negligible effect as packets marking happens repeatedly and $R(t)$ is relatively small. This approximation is also applicable to the derivation of the window decrement $b(t)$.

In the fast recovery process, the congestion window will decrease by half its size when the sender receives a notification of a marked packet. We have

$$b(t) = -\sum_{i=1}^{N} \frac{W_i(t)}{W_s(t)} \frac{W_i(t + R(t))}{2} \approx -\sum_{i=1}^{N} \frac{W_i^2(t + R(t))}{2W_s(t + R(t))}, \quad (11)$$

where $b(t)$ is defined in (5), which represents the decrement of $W_s(t)$ for every marked packet. Notice that unlike the case of $a(t)$, the decrement $b(t)$ has the
same expression for both Scenarios A and B. Furthermore,

\[
\sum_{i=1}^{N} W_i^2(t + R(t)) \geq \frac{\left(\sum_{i=1}^{N} W_i(t + R(t))\right)^2}{W_s(t + R(t)) N},
\]

\[
\frac{\left(\sum_{i=1}^{N} W_i(t + R(t))\right)^2}{W_s(t + R(t)) N} = \frac{W_s(t + R(t))}{N}.
\]

Then, we write

\[
b(t) = -\frac{\rho W_s(t + R(t)) W_s(t)}{2N},
\]

where \(\rho\) is what we call the \textit{window distribution parameter}. It is a function of \(N\) and the \texttt{cwnd} values of all the sessions, and it satisfies \(1 \leq \rho \leq N\). For large \(N\), the extreme cases of \(\rho = 1\) and \(\rho = N\) are very rare. Having \(\rho = 1\) corresponds to the case where the congestion windows \(W_i(t + R(t))\) of all TCP sessions are the same and nonzero, and \(\rho = N\) corresponds to the case where there is always exactly one TCP session with \texttt{cwnd} > 0 and in all other sessions \texttt{cwnd} = 0. In practice, both \(N\) and the \texttt{cwnd} values of each session vary in time. Also, \(\rho\) depends on the particular scenario under consideration. In Section 3.5 below, we provide equations (27) and (28) by which one can calculate the unique value of \(\rho\) for each scenario. Nevertheless, we will show by simulations in Section 4 that if \(N\) is fixed and if a fixed value for the window distribution parameter \(\rho\) is correctly chosen, the analysis based on a fixed \(\rho\) leads to accurate results despite the fact that \texttt{cwnd} values vary. We also show that to model a case where \(N\) varies, if we can select the right \(\rho\) value for each value of \(N\), we still obtain accurate results. Obtaining the parameter \(\rho\) is a key issue in our modeling, and although we do not have an analytical way to derive it, we argue and demonstrate in the sequel that setting the window distribution parameter \(\rho = 1\) improves on the accuracy of the Simplified MGT model.

3.3. Modeling Window Dynamics for Scenarios A and B

Based on (5), (7) and the above results for \(a(t)\) and \(b(t)\), we have the following equation for Scenario A:

\[
\Delta W_s(t + R(t)) = \frac{W_s(t)}{R(t)} (1 - p(t)) \tau - \frac{\rho W_s(t + R(t)) W_s(t)}{2N R(t)} p(t) \tau.
\]

Similarly, for Scenario B, we obtain:

\[
\Delta W_s(t + R(t)) = \frac{NW_s(t)}{R(t) W_s(t + R(t))} (1 - p(t)) \tau - \frac{\rho W_s(t + R(t)) W_s(t)}{2N R(t)} p(t) \tau.
\]
The choice of the appropriate scenario to model the TCP/AQM system depends on the congestion level. In Section 3.6, we define a parameter named \textit{congestion level} used to determine such choice. Based on equations (14) and (15) which give the sum of all TCP congestion windows for the two Scenarios, we can present our new models in both discrete- and continuous-time as described in the following.

3.3.1. Discrete-time Model

Let time be divided into consecutive fixed-length time intervals each of size \( \Delta t \). Let \( k \) be an index for these intervals so that the \( k \)th interval represents the time between \( (k - 1)\Delta t \) and \( k\Delta t \). Recalling our definition of \( R(t) \) associated with the TCP sessions that are active between time \( t \) and time \( t + \Delta t \), let \( R(k) \) be the expected RTT of the TCP sessions that are active during the \( k \)th time interval. Let \( n(k) \) be defined as \( n(k) = \lfloor R(k)/\Delta t \rfloor \), and \( \Delta W_s(k + n(k)) \) be defined as \( \Delta W_s(k + n(k)) = W_s(k + n(k) + 1) - W_s(k + n(k)) \). We then obtain the following for Scenario A.

\[
\Delta W_s(k + n(k)) = \frac{W_s(k)}{R(k)}(1 - p(k))\Delta t - \frac{\rho W_s(k + n(k))W_s(k)}{2NR(k)}p(k)\Delta t. \quad (16)
\]

For Scenario B, we have:

\[
\Delta W_s(k + n(k)) = \frac{NW_s(k)}{R(k)W_s(k + n(k))}(1 - p(k))\Delta t - \frac{\rho W_s(k + n(k))W_s(k)}{2NR(k)}p(k)\Delta t. \quad (17)
\]

In the above two equations, the variables \( W_s(k) \) and \( p(k) \) are the expected values of the sum of all TCP congestion windows and marking probability at the \( k \)th time interval, respectively.

3.3.2. Continuous-time Model

According to the definition of time-derivative, we have:

\[
\dot{W}_s(t + R(t)) = \lim_{dt \to 0} \frac{W_s(t + R(t) + dt) - W_s(t + R(t))}{dt}. \quad (18)
\]

Recall the (14) and (15), we obtain the following form for Scenario A:

\[
\dot{W}_s(t + R(t)) = \frac{W_s(t)}{R(t)}(1 - p(t)) - \frac{\rho W_s(t + R(t))W_s(t)}{2NR(t)}p(t). \quad (19)
\]

For the Scenario B, we have:

\[
\dot{W}_s(t + R(t)) = \frac{NW_s(t)}{R(t)W_s(t + R(t))}(1 - p(t)) - \frac{\rho W_s(t + R(t))W_s(t)}{2NR(t)}p(t). \quad (20)
\]
The marking probability \( p(k) \) or \( p(t) \) depends on the specific AQM scheme, namely, different AQM schemes will update such marking probability based on different measures, for example, PI uses queue length error while REM uses both queue length error and sending rate error. Later in Section 4, we demonstrate the effect of marking probability of different AQM schemes on the dynamics of queue length and marking probability. As in our model, Scenarios A and B are used as upper and lower bounds, respectively, for the predicted values of queue lengths and marking probability, it is important to have these bounds close to each other. In Subsection 3.7 we will provide intuitive arguments and a rigorous proof for the closeness of the bounds, and in Section 4 we will provide a wide range of numerical results to provide further evidence of this closeness.

3.4. Effect of ECN on the Queue Dynamics

ECN is enabled in some TCP sessions and it is not enabled in others. We will discuss now the different queue dynamics considering operation alternatives when the ECN is ON and OFF. The queue dynamics depends on the packet arrival rate which equals the sum of the send rate of the sources and the packet departure rate which is the link capacity \( C \). In the discrete-time model, the queue dynamics is described as:

\[
\Delta q(k) = \left( \frac{W_s(k)}{R(k)} - C \right) \Delta t,
\]

where \( \Delta q(k) = q(k+1) - q(k) \), and \( \frac{W_s(k)}{R(k)} \) represents the packet arrival rate.

Equivalently, for the continuous-time model, we obtain,

\[
\dot{q}(t) = \frac{W_s(t)}{R(t)} - C,
\]

where \( \dot{q}(t) \) is the time-derivative of \( q(t) \). The first and second term represent the packet arrival and departure rates, respectively.

When the ECN is off, the method of informing the sources of congestion is by dropping packets, instead of marking packets. The queue dynamics, which considers the dropped packets by disabling the ECN, can be expressed in both the discrete- and continuous-time models as:

\[
\Delta q(k) = \left( \frac{W_s(k)}{R(k)} - C - p(k) \frac{W_s(k)}{R(k)} \right) \Delta t,
\]

and

\[
\dot{q}(t) = \frac{W_s(t)}{R(t)} - C - p(t) \frac{W_s(t)}{R(t)},
\]

where the third term corresponds to the dropped packets, which is often ignored in the literature. However, it is important since the operating point can be affected by the dropped packets. This will be discussed next.
The Operating Point

The operating point \((W_{s0}, q_0, p_0)\) is defined by \(\dot{W}_s(t) = 0\) and \(\dot{q}(t) = 0\) according to the continuous-time model, so that

\[
\dot{q}(t) = 0 \Rightarrow W_{s0} = \begin{cases} 
R_0 C, & \text{if ECN is ON} \\
R_0 C / (1 - p_0), & \text{if ECN is OFF}
\end{cases}
\]

(25)

where \(W_{s0}, q_0\) and \(p_0\) are the sum of the congestion windows, the queue length, and the marking (dropping) probability when TCP/AQM system is in steady state, respectively. The value of \(W_{s0}\) as shown in (25) is dependent on whether ECN is ON or OFF.

In order to obtain \(p_0\) for Scenario A, we write

\[
W_s(t) = 0 \Rightarrow p_0 = 1 / \left(1 + \frac{\rho W_{s0}}{2N}\right) = \frac{2N}{2N^2 + \rho W_{s0}}.
\]

(27)

For Scenario B, we obtain

\[
W_s(t) = 0 \Rightarrow p_0 = 1 / \left(1 + \frac{\rho W_{s0}^2}{2N^2}\right) = \frac{2N^2}{2N^4 + \rho W_{s0}^2}.
\]

(28)

Note that the value of \(p_0\) in (27) and (28) cannot exceed 1, which is an important property that helps improve accuracy over the Simplified MGT model where \(p_0\) in (2) can take values higher than 1. Specifically, equation (27) and (28) correct such error by adding the value \(\rho W_{s0}\) or \(\rho W_{s0}^2\) to the denominator. Comparing the denominators of (2) and (28), and recalling that the window distribution parameter must satisfy \(1 \leq \rho \leq N\), we find that \((2N^2 + \rho W_{s0}^2) > W_{s0}^2\), so \(p_0\) of Scenario B will be always smaller than \(p_0\) of the Simplified MGT model. Furthermore, empirically, we have noticed that we examined that the real \(p_0\) was always lower than \(p_0\) based on the Scenario B model with \(\rho = 1\) in all the cases. If we take this empirical result as always true, it leads to an important conclusion that conservatively setting \(\rho = 1\), Scenario B always yields a better approximation than the Simplified MGT model.

Note that \(W_{s0}, q_0\) and \(p_0\), as well as \(\rho\) are steady state parameters, so their values will be the same for a set of AQM schemes which means that the corresponding TCP/AQM systems stabilize at the same operating point, while the transient states will distinguish between the dynamics of the different AQM schemes before they converge to steady-state. This is consistent with the results shown in Section 4.

Different Congestion Levels

As discussed, our model is based on Scenarios A and B which represent two extreme cases. These two scenarios are definitely not exhaustive as there are
Table 2: The Classification of Congestion Level

| Average cwnd & Scenario | Mild | Moderate | Severe |
|-------------------------|------|----------|--------|
| \( \bar{w} \)            | \( \bar{w} > 2 \) | \( 1 \leq \bar{w} \leq 2 \) | \( \bar{w} < 1 \) |
| Scenario                | B    | Both A and B | A      |

many cases where in some TCP sessions \( cwnd < ssthresh \) and in others \( cwnd \geq ssthresh \), and such cases are not covered by scenarios A and B. To demonstrate that our model is applicable in general, we will need to consider cases that are not included in Scenarios A and B, preferably cases that are very different from Scenarios A and B.

When the congestion level is mild, most TCP sessions are in the congestion avoidance phase, where Scenario B represents the system behavior. When the congestion level increases, the sizes of more and more congestion windows become less than two packets, which means that the \( ssthresh \) must be equal to 2 (since 2 is the lower bound of \( ssthresh \)). When most sessions are in the slow start phase, the congestion level is considered to be severe, where Scenario A represents the system behavior. Cases where neither Scenario A nor B can entirely represent the system behavior are characterized by many TCP sessions working in slow start phase and many other sessions working in the congestion avoidance phase.

Having discussed the relationship between Scenarios A and B and the congestion level, it is convenient to introduce a measure of the congestion level, denoted \( \bar{w} \), defined by \( \bar{w} = \frac{W_{s0}}{N} \). This measure is the average congestion window of all sessions at the operating point.

When \( \bar{w} < 1 \), most of the TCP sessions have small \( cwnd \) which is less than 2 packets. In such case, the \( cwnd \) is even lower than the lower bound of the \( ssthresh \) which indicate that the majority of TCP sessions work in slow start phase. The link is said to be severely congested in this condition. The condition that \( cwnd \) is greater than the 2 packet lower bound is the necessary condition for a TCP session to work in congestion avoidance phase. The larger \( \bar{w} \) is, the more TCP sessions work in congestion avoidance phase. Therefore, we use the condition \( \bar{w} > 2 \) to indicate a region where the link is said to be mildly congested where Scenario B accurately represents the system behavior. Having defined the congestion regions of mild and severe, we define the condition \( 1 \leq \bar{w} \leq 2 \) for the region where the link is moderately congested. Having the two thresholds of \( \bar{w} \), i.e, 1 and 2, we now have three non-overlapping and exhaustive regions for the congestion level. Table 2 shows the partition of the three congestion level regions of \( \bar{w} \) and Scenarios. As our Scenarios A and B do not cover the moderate congestion region where \( 1 \leq \bar{w} \leq 2 \), we will use Scenarios A and B as the lower and upper bounds to approximate the system behavior.
3.7. Closeness of the Bounds

As discussed, Scenarios A and B are applicable to the severe and mild congestion levels, respectively. However, for all the cases in-between these two extreme congestion levels, Scenario A and B models are used as bounds. It is therefore important to discuss their closeness to each other. Here we explain that their closeness can be achieved if we can choose an appropriate window distribution parameter \( \rho \) for each of the two scenarios and if the system is stable. Then, we show stability for a specific linearized PI AQM system.

Although Scenario A and B use different ways to increase \( cwnd \), the queue length and marking probability of the two models converge to the operating point under steady-state conditions if the system is stable. As mentioned, the operating point is defined by \( W_{s0} \), \( q_0 \) and \( p_0 \). Firstly, the value of \( q_0 \) depends on the AQM scheme. Hence, the queue length of our model can be stabilized at \( q_0 \) if the system is stable. Secondly, for a given set of network parameters, the values of \( W_{s0} \) of both scenarios are the same according to (25). Let \( \rho_A \) and \( \rho_B \) be the \( \rho \) values for Scenario A and B, respectively. Then, if the settings satisfy the following equation:

\[
\rho_A = \frac{W_{s0}}{N} \rho_B,
\]

the values of \( p_0 \) of both scenarios will be the same according to (27) and (28). Hence, for a given set of network parameters, Scenario A and B can have the same operating point if a suitable value of the window distribution parameter \( \rho \) is set for each scenario. The models of the two scenarios will converge to the same value under the assumption that the system is stable.

Although it is difficult to prove stability of every TCP/AQM system, we consider a particular system based on PI AQM under the settings used in Section 4 and in Appendix Appendix A we provide stability analysis and numerical verification for this TCP/AQM system for the case where the system is linearized at the operating point. Routh-Hurwitz stability criterion is then used to determine whether this TCP/AQM system is stable.

3.8. The Window Distribution Parameter

We have already discussed the importance of setting the window distribution parameter \( \rho \) to achieve accuracy and closeness of the bounds. As discussed, we do not have an analytical method to obtain it. Nevertheless, we know that it is bounded within \( 1 \leq \rho \leq N \) and that simply and conservatively setting \( \rho = 1 \) gives improved results over the Simplified MGT model, which will be demonstrated through a large sample of cases over a wide range of parameters in the next section. To appreciate the difficulty in obtaining \( \rho \), notice that \( \rho \) could have been obtained by (25), (26), (27), and (28) if \( p_0 \) is available analytically. However, obtaining \( p_0 \) analytically is a well known open problem. By the same equations, if \( \rho \) is analytically available then \( p_0 \) can be obtained. This explains that obtaining \( \rho \) is equivalent to the known difficult problem of finding \( p_0 \).

One way is to obtain \( p_0 \) (and \( \rho \)) is by simulation. Such values will be used in many simulations in the next section to demonstrate how accurate our model
could be if we were able to obtain $\rho$ analytically. As discussed, the results for the cases of $\rho = 1$ will also be discussed and compared with. As mentioned, $\rho$ is a steady-state parameter, therefore if $\rho$ is correct for one AQM scheme, it can be reused for the other AQM schemes who has the same operating point. This property will be illustrated in the next section.

4. NUMERICAL RESULTS

In this section, we show the improvement of the proposed model compared to the Simplified MGT model for two settings of the window distribution parameter $\rho$: one is the conservative case of $\rho = 1$ and the other is the case where $\rho$ is obtained based on the NS2 simulation results.

We first introduce the network topology used for the NS2 simulation, followed by information on the setting of the default parameters of three AQM schemes: PI, REM and RaQ. Then, we conduct a series of simulations to evaluate the new model by comparing the analytical results with simulation results. All the time-dependent analytical results are based on the discrete-time model of Scenarios A and B and a discrete time version of the Simplified MGT model. In all the figures in this section that show comparative simulation results, for clarity, we use the label Simplified MGT to represent the results obtained for the Simplified MGT model, we use the labels Scenario A, Scenario B, with $\rho = 1$ to represent the analytical results obtained for Scenarios A and B setting $\rho = 1$, and equivalently we use the labels Scenario A, Scenario B together with the specific $\rho$ value obtained from the NS2 simulations to represent an ideal analytical results for Scenarios A and B indicating the limit of possible improvement based on our model. We observe in all the numerical results that all Scenario A and B curves (including the cases of $\rho = 1$) converge to the target queue length. This is consistent with our discussions in Subsections 3.7. Strong evidence for the accuracy of this model has been achieved for these cases where simulation results are obtainable. Validation for realistic parameters values (e.g. much larger link capacity) is achievable by experiments and this will be done in future study.

4.1. Experiment Topology and Default Parameters

In our simulations, we use a dumb-bell network topology shown in Fig. 3. The link between router B and router C is the bottleneck link. Unless mentioned otherwise, the following parameters are set as default: the mean packet size is 1000 bytes, the propagation time is 100 ms, the bottleneck link capacity is 45 Mb/s, the buffer size is 1125 packets, the target queue length is 500 packets, and the default number of TCP sessions is set at $N = 500$. TCP/Reno is adopted as the TCP plant and ECN is set ON unless otherwise stated.

The default parameters of the AQM schemes are set as follows: For PI, $a = 0.00001822$ and $b = 0.00001816$. For REM, $\gamma = 0.001$ and $\Phi = 1.001$. For RaQ, $q_{kp} = 0.0077$, $q_{ki} = 0.0005$, and $r_{kp} = 0.0095$. The above PI [4] and REM [8] parameters are the NS2 defaults and the RaQ parameters are as in [10].

The default sampling period $\Delta t$ used for the analytical results for Scenarios A and B and the Simplified MGT model is 0.0005 s, and the default sampling
period $T$ used for the three AQM schemes PI, REM and RaQ is 0.005 s. Note that $\Delta t$ should be smaller than or equal to $T$. Further discussion of these two sampling periods will be given in Subsection 4.7 below.

4.2. Performance Under Different Congestion Levels

As discussed, our analytical model that are based on Scenario A and B should be tested for a wide range of congestion levels. To this end, we consider five cases, three of which represent the three congestion levels: mild, severe and moderate, and the other two represent borderline cases mild/moderate and moderate/severe that approximately correspond to the cases $\overline{w} = 2$ and $\overline{w} = 1$, respectively. These five cases are described in Table 3. The number of TCP sessions $N$ is adjusted to achieve different congestion levels. In particular, we use the $N$ values of 200, 500, 800, 1100 and 2000. For example, in the first case, of $N = 200$, $\overline{w} = CR_0/N = 5625 \times 0.1889/200 = 5.3128$, which is more than 2 packets and therefore, this case is within the mild congestion level region.

| $N$ | $\overline{w}$ | Congestion level       |
|-----|----------------|------------------------|
| 200 | 5.3128         | Mild                   |
| 500 | 2.1251         | Mild/Moderate          |
| 800 | 1.3282         | Moderate               |
| 1100| 0.9660         | Moderate/Severe        |
| 2000| 0.5313         | Severe                 |

4.2.1. Mild Congestion

In this case, we consider PI as the AQM scheme and $N = 200$. As this case falls within the mild congestion region, it is well modeled by Scenario B. Based on NS2 simulations, we obtain $p_0 = 0.0442$, which gives by (15), $\rho = 1.5318$. By (2) and (28), the $p_0$ values of the Simplified MGT model and the Scenario B model (with $\rho = 1$) are given in Table 4. We reuse the value $\rho = 1.5318$ in the other two AQM schemes REM and RaQ. Results for all three AQM schemes are
presented in Fig. 4, Fig. 5 and Fig. 6, for PI, REM and RaQ, respectively. We observe that the results of all analytical models for all three AQM schemes are fairly accurate in both in steady state and during the transient periods before steady state. For example, all analytical results for all three AQM schemes exhibit convergence to the queue length target of 500 packets. In addition, we notice for all three AQM schemes that the marking probability dynamics based on Scenario B with $\rho = 1$ is somewhat more accurate than the results obtained by the Simplified MGT model which is consistent with the steady state results shown in Table 4.

| N  | Simplified MGT | Scenario B ($\rho = 1$) | NS2 |
|----|----------------|-------------------------|-----|
| 200| 0.0708         | 0.0662                  | 0.0442 |

Table 4: Values of $p_0$ for $N = 200$.

![Figure 4: Queue length and marking probability versus time for PI AQM with $N = 200$.](image)

Our analytical model, as the Simplified MGT model, is based on simplifications that rely on expected values of parameters that affect the transient queue length results. Therefore, it does not exhibit the fluctuations obtained by the NS2 simulations. Nevertheless, our results are closer to the NS2 simulation results than those of the Simplified MGT model. Similar observations will apply to many of the numerical results presented in this section.

4.2.2. Severe Congestion

In this case, we again first consider PI as the AQM scheme and we set $N = 2000$. This case falls within the severe congestion region, so it is modeled by Scenario A. Based on NS2 simulations, we obtain $p_0 = 0.4879$, which gives by (14), $\rho = 3.9516$. By (2) and (27), the $p_0$ values of the Simplified MGT model and Scenario A (with $\rho = 1$) are given in Table 5. We notice that the
$p_0 = 0.7901$ predicted by Scenario A (with $\rho = 1$) improves the $p_0 = 7.0827$ of the Simplified MGT model, while still it is not close to the NS2 $p_0 = 0.4879$. We reuse the value $\rho = 3.9516$ in the other two AQM schemes REM and RaQ. Results for all three AQM schemes are presented in Fig. 8, Fig. 9 and Fig. 10, for PI, REM and RaQ, respectively.

We observe that the results of Scenario A (with $\rho = 1$) for all three AQM schemes improve the accuracy both in steady state and during the transient periods. We observe in Fig. 8 (a), that although the curve of Scenario A ($\rho = 1$) does not very accurately predict the transient period, it converges to the target queue length of 500 packets, despite the fact that the $p_0$ is inaccurate. By comparison, the Simplified MGT model predicts that the queue length saturates the buffer. As for the marking probability dynamics, the curve of Scenario A
(with $\rho = 1$) converges to the value 0.7901 after 100 second, while the Simplified MGT model converges to 1 only because it is truncated to 1. In the other two figures associated with REM and RaQ AQM, Scenario A (with $\rho = 1$) again converges to the target queue length and provides more accurate marking probability than the Simplified MGT model. The results also demonstrate that Scenario A (with $\rho = 3.9516$) provides curves closest to the curves of the NS2 simulations by comparison to the Simplified MGT model and Scenario A (with $\rho = 1$) for both queue length dynamics and marking probability.

We also presented the result of MGT model without truncation (for both marking probability and queue length) in Fig. 7. It is shown that the queue length for RaQ or PI can converge to the target queue length of 500 packets, while the marking probability is stabilized to more than 7. However, the marking probability of REM does not exceed 1 because its calculation is based on the function $1 - \Phi^{-\rho l}$, but the consequence is that the queue length cannot converge to the target queue length. We also present the queue length dynamics of the Simplified MGT model without truncation in Fig. 10. We observe that although some improvement in the modeling of queue length dynamic is achieved without truncation, relative to the truncated counterpart, the accuracy under severe congestion is not satisfactory. Note that the convergence of the untruncated Simplified MGT for PI takes 2000 seconds for the queue size and 4000 for the marking probability, so the results are not included in Fig. 8. We also exclude the untruncated MGT results from Fig. 9 because they do not improved on their truncated counterparts.

| $N$   | Simplified MGT | Scenario A ($\rho = 1$) | NS2   |
|-------|----------------|-------------------------|-------|
| 2000  | 7.0827         | 0.7901                  | 0.4879|

### 4.2.3. Mild/Moderate, Moderate and Moderate/Severe Congestion

We now consider the congestion levels associate with the cases $N = 500, 800, 1100$. For each case, we obtain $p_0$ from the NS2 simulations of PI AQM, and derive the $\rho$ values from (27) and (28) provided in Table 7. We also obtain $p_0$ of the Simplified MGT model and Scenarios A and B for $\rho = 1$ by (2), (27) and (28), respectively. The values are presented in Table 6. Fig. 11, Fig. 12 and Fig. 13 show the numerical results of the NS2 simulation and of the analytical models. Comparing to the results of NS2 simulations, it can be found that the convergence time of queue length and marking probability at the steady state increase when $N$ increases. For the cases that $\rho$ are set to the values based on NS2 simulation (Table 7), the results show that, as expected, Scenario B performs better than Scenario A near the bound of Mild/Moderate, while Scenario A performs better near the bound of Moderate/Severe. In the moderate congestion range, the curves of NS2 simulations are between the Scenario B and Scenario A curves. Therefore, these results confirm that the milder the congestion level
is, the better Scenario B model performs, and vice versa. This conclusion also applies to the cases where $\rho = 1$. In addition, Scenario A and B curves are approaching each other in the cases where the $\rho$ values are set based on the NS2 simulations. This is consistent with the discussion in Subsection 3.7. Also Scenario A and B curves with $\rho = 1$ are reasonable close to each other for queue dynamic results of the three AQM schemes.

4.3. Turning OFF ECN

In the case where ECN is turned off, by NS2 simulations, we obtain $p_0 = 0.1416$. The congestion level index is $\bar{\pi} = CR_0/(N - NP_0) = 2.4756$. By
Figure 9: Queue length and marking probability versus time for REM AQM with $N = 2000$.

Figure 10: Queue length and marking probability versus time for RaQ AQM with $N = 2000$.

Table 6: Values of $p_0$ for $N = 500, 800, 1100$

| $N$  | Simplified MGT | Scenario B ($\rho = 1$) | Scenario A ($\rho = 1$) | NS2 |
|------|----------------|--------------------------|--------------------------|-----|
| 500  | 0.4429         | 0.3069                   | 0.4848                   | 0.2004 |
| 800  | 1.1337         | 0.5313                   | 0.6009                   | 0.3504 |
| 1100 | 2.1434         | 0.6819                   | 0.6743                   | 0.4212 |

Table 2, such congestion level falls into the mildly congested region, so the model of Scenario B is adopted. Then from (25) and (27), we obtain

$$\rho = \frac{(1 - p_0)^3 \cdot 2N^2}{C^2 R_0^2 p_0} = 1.9789.$$  (30)
Table 7: Parameter $\rho$ Of New Model For $N = 500, 800, 1100$ Based On NS2 Simulations

| $N$ | $\rho$ for Scenario B | $\rho$ for Scenario A |
|-----|------------------------|------------------------|
| 500 | 1.7670                 | 3.7551                 |
| 800 | 2.1022                 | 2.7921                 |
| 1100| 2.9450                 | 2.8448                 |

Since the Simplified MGT model does not specifically consider ECN, the previous case of $N = 500$, which yields $p_0(\text{MGT}) = 0.4429$ is used here. While the $p_0$ of Scenario B model ($\rho = 1$) when turning off ECN is obtained by (25) and (28):
Figure 13: Queue length and marking probability versus time for PI AQM with $N = 1100$.

\begin{equation}
p_0(\text{Scenario B}, \text{ECN OFF}) = \frac{2N^2}{2N^2 + R_0^2C^2} \frac{R_0^2C^2}{(1 - p_0)^2}.
\end{equation}

Isolating $p_0$ and solving this cubic equation, we obtain $p_0(\text{Scenario B}, \rho = 1) = 0.2146$. The simulation and analytical results are shown in Fig. 14, Fig. 15 and Fig. 16.

Figure 14: Queue length and marking probability versus time for PI AQM with $N = 500$ when ECN is OFF.

The results exhibit that the curves of Scenario B (with $\rho = 1$) are closer to the curves of NS2 simulations in both Fig. 14 and Fig. 15 compared to the
Figure 15: Queue length and marking probability versus time for REM AQM with $N = 500$ when ECN is OFF.

Figure 16: Queue length and marking probability versus time for RaQ AQM with $N = 500$ when ECN is OFF.

Simplified MGT models results. The marking probabilities are consistent with the theoretical value calculated by (31).

Note that the target queue length is set as 500 packets in NS2 simulations. In Fig. 15, where REM AQM is adopted, the stable queue length of the NS2 simulation is around 450 packets, which is lower than the target queue length (500 packets). In this case, Scenario B ($\rho = 1.9789$) can still captures this discrepancy, while the queue length of the Simplified MGT model converges to 500 packets. This fact demonstrates that the correct parameter $\rho$ is reusable under the stability assumption, regardless of which AQM scheme is used. For the other two AQM schemes, Scenario B model ($\rho = 1.9789$) provides accurate
match for both the queue length and dropping probability dynamic.

### 4.4. Varying Number of TCP Sessions

This set of simulations tests the ability of tracking a link where the number of TCP sessions is varied. Specifically, 300 TCP sessions start at the time 0, 200 additional sessions join at 65 s, and 200 TCP sessions stop at 130 s. According to (25), (26) and (28) and based on NS2 simulation, we have the values of $\rho = 1.6575$ and $\rho = 1.7670$ when $N = 300$ and $N = 500$, respectively. By (2) and (28), the theoretical $p_0$ of Scenario B ($\rho = 1$) and the Simplified MGT model for $N = 300$ and $N = 500$ are given in Table 8.

| $N$   | Simplified MGT | Scenario B ($\rho = 1$) | NS2      |
|-------|----------------|-------------------------|----------|
| 300   | 0.1594         | 0.1375                  | 0.0877   |
| 500   | 0.4429         | 0.3069                  | 0.2004   |

The results are shown in Fig. 17, Fig. 18 and Fig. 19. In Fig. 17 and Fig. 18, the marking probability of PI and REM AQM schemes have not reached the theoretical values presented in the above table. This is because the TCP/AQM system has not reached steady state when $N$ changes at times 65 s and 130 s. Nevertheless, the results in Fig. 19 confirm the theoretical values due to the fast convergence of the RaQ scheme. As expected, in such Mild and Mild/Moderate congestion region when $N$ changes, the curves of Scenario B ($\rho = 1$) are closer to the NS2 simulation curves in both queue dynamic and marking probability compared to the curves of the Simplified MGT model.
4.5. Different RTT

Now simulation and analytical results of two different RTTs are presented. The propagation time are set as 0.05 s and 0.15 s, then the corresponding RTTs (26) are $0.05 + 500/5625 = 0.1389$ and $0.15 + 500/5625 = 0.2389$, respectively. The marking probabilities are 0.2894 and 0.1402 from NS2 simulation, then the corresponding $\rho = 2.0107$ and $\rho = 1.6984$ can be derived. By (2) and (28), the $p_0$ of Simplified MGT model and of Scenario B ($\rho = 1$) under different RTTs are provided in Table 9.

The simulation results are presented in Fig. 20 and Fig. 21, it is shown that Scenario B ($\rho = 1$) improves the accuracy of queue length and marking probability in both cases relative to the Simplified MGT model. Note that Scenario B ($\rho = 1.6984$) in Fig. 21 matches NS2 simulations better than Scenario B.
Table 9: Values of \( p_0 \) for Different RTTs

| RTTs (s) | Simplified MGT | Scenario B \((\rho = 1)\) | NS2 |
|----------|----------------|--------------------------|-----|
| 0.1389   | 0.8191         | 0.4503                   | 0.2894 |
| 0.2389   | 0.2769         | 0.2168                   | 0.1402 |

Figure 20: Queue length and marking probability versus time for PI AQM with propagation time is 0.05 s.

\((\rho = 2.0107)\) in Fig. 20. This can be explained by the congestion level. According to (25) and \( \overline{w} = W_{s0}/N \), we obtain \( \overline{w} = 1.5626 \) and \( \overline{w} = 2.6876 \) for \( RTT = 0.1389 \) and \( RTT = 0.2389 \), respectively. As shown in Table 2, the
second case falls into the mildly congested region while the first one falls into
the moderately congested region, hence the Scenario B model performs better
in the second case.

4.6. Different Link Capacities

We now consider two cases where the link capacities are 15Mb/s and 95
Mb/s. By (2) and (28), the \( p_0 \) of Simplified MGT and of Scenario B model
\((\rho = 1)\) for the two cases are given in Table 10. From the NS2 simulations,
the marking probabilities for the cases of \( C = 15 \) Mb/s and \( C = 95 \) Mb/s are
0.3426 and 0.0973, respectively. Then the corresponding \( \rho \) values are 2.0297 and
1.6286, respectively. The results are shown in Fig. 22 and Fig. 23. In the two
figures, the Scenario B curves with the \( \rho \) values of 2.0297 and 1.6286 match the
NS2 simulation results very well, and the curves generated by Scenario B \((\rho = 1)\)
are closer to the NS2 simulation results than the curves of the Simplified MGT
model.

| Link Capacities (Mb) | Simplified MGT | Scenario B \((\rho = 1)\) | NS2    |
|---------------------|---------------|--------------------------|-------|
| 15                  | 1.0577        | 0.5140                   | 0.3426|
| 95                  | 0.1756        | 0.1494                   | 0.0973|

Figure 22: Queue length and marking probability versus time for PI AQM with link capacity C is 15 Mb/s.

4.7. Time Interval \( \Delta t \)

In practice the marking probability may change during the time interval \( \Delta t \),
while we assume that it stays constant in our model (and also the discrete-time
implementation of the Simplified MGT model). Therefore, the longer \( \Delta t \) is, the larger the error it introduces. It is required that \( \Delta t \) is bounded above by the AQM sampling period denoted by \( T \), to allow sufficient time to update the marking probability before the next sampling period. Here we increase the value of \( \Delta t \) from its previous value of 0.0005 s to 0.2 s and the AQM sampling period from 0.005 s to 0.2 s, and we examine the effect on the accuracy of our model and of the Simplified MGT model. Notice that the \( \Delta t \) and AQM sampling period are now longer that the RTT which is 0.1889 s by (26). Because of RaQ's short convergence time, we only present results here for RaQ AQM. The parameter \( \rho = 1.7670 \) by (28) is used, and Scenario B is chosen of this case. The result is shown in Fig. 24, which demonstrates that even if \( \Delta t \) is greater than the RTT, the Scenario B model (with \( \rho = 1.7670 \)) still has the capability to track the simulation results well. The performance of the Simplified MGT model is also quite good under the same settings, while it is not as good as that of Scenario B (with \( \rho = 1 \)) for both the queue length and marking probability dynamics.

5. CONCLUSIONS

In this paper, we have proposed a new model for a time-driven TCP/AQM system. This new model rigorously considers the four algorithms embedded in TCP, and it contributes to performance studies of AQM schemes. We have provided both discrete and continuous-time models. Our proposed model is based on bounds associated with heavy and light congestion, but we have shown that the bounds are tight in steady state and the model is applicable to a wide range of congestion levels. In addition, the model has been validated using extensive simulation results for a wide range of parameters and various AQM schemes and it was demonstrated that it is more accurate than the Simplified
Figure 24: Queue length and marking probability versus time for RaQ AQM with step time \( \Delta t \) is 0.2 s.

MGT model. In fact, the accuracy of our model was consistently demonstrated in all cases studied.

Appendix A. Stability under Linearization

Recalling Scenario A described by (19), we define:

\[
f(W_s, W_{sR}, p_R, q_R) \triangleq W_s(t) = \frac{W_{sR}(t)}{T_p + \frac{T_p}{C}} (1 - p_R) - \frac{\rho W_{sR} p_R}{2N(T_p + \frac{T_p}{C})} \quad (A.1)
\]

where \( W_{sR}(t) \triangleq W_s(t - R) \), \( p_R(t) \triangleq p(t - R) \), \( q_R(t) \triangleq q(t - R) \).

Taking partial derivatives at the operating point \((W_{s0}, q_0, p_0)\) of this model yields:

\[
\frac{\partial f}{\partial W_s} = -\frac{\rho W_{s0} p_0}{2NR_0} \quad (A.2)
\]

\[
\frac{\partial f}{\partial W_{sR}} = \frac{1 - p_0}{R_0} - \frac{\rho W_{s0} p_0}{2NR_0} \quad (A.3)
\]

\[
\frac{\partial f}{\partial p_R} = -\frac{W_{s0}}{R_0} - \frac{\rho W_{s0}^2}{2NR_0} \quad (A.4)
\]

\[
\frac{\partial f}{\partial q_R} = -\frac{W_{s0}(1 - p_0)}{R_0 C} + \frac{\rho W_{s0}^2 p_0}{2N R_0^2 C}. \quad (A.5)
\]
Hence Scenario A linearized form is expressed as:

\[
\delta \dot{W}_s(t) = \frac{\partial f}{\partial W_s} \delta W_s(t) + \frac{\partial f}{\partial W_{sR}} \delta W_{sR}(t - R) + \frac{\partial f}{\partial p_R} \delta p(t - R) + \frac{\partial f}{\partial q_R} \delta q(t - R),
\]

(A.6)

where

\[
\begin{aligned}
    \delta W_s(t) & \doteq W_s(t) - W_{s0} \\
    \delta q(t) & \doteq q(t) - q_0 \\
    \delta p(t) & \doteq p(t) - p_0.
\end{aligned}
\]

(A.7)

We linearize Scenario B model in a similar way to (A.6), where the partial derivatives are replaced by:

\[
\begin{aligned}
    \frac{\partial f}{\partial W_s} &= -\frac{N}{N(1 - p_0)} R_0 W_{s0} - \frac{\rho W_{s0} p_0}{2N R_0 C} R_0^2 \\
    \frac{\partial f}{\partial W_{sR}} &= \frac{N}{N(1 - p_0)} R_0 W_{s0} - \frac{\rho W_{s0} p_0}{2N R_0 C} R_0^2 \\
    \frac{\partial f}{\partial p_R} &= \frac{N}{N(1 - p_0)} R_0 - \frac{2N R_0}{2N R_0 C} R_0^2 \\
    \frac{\partial f}{\partial q_R} &= \frac{N}{N(1 - p_0)} R_0 + \frac{\rho W_{s0} p_0}{2N R_0 C} R_0^2.
\end{aligned}
\]

(A.8)

Linearizing (22) when ENC is enabled gives:

\[
\delta \dot{q}(t) = \frac{1}{R_0} \delta W_s(t) - \frac{W_{s0}}{R_0^2 C} \delta q(t).
\]

(A.9)

We obtain the block diagram presented in Fig. A.25. We choose PI as the AQM algorithm and we obtain a closed-loop system. In Laplace area, the PI controller has the form \( k_p + \frac{k_i}{s} \), where \( k_p \) and \( k_i \) are the proportion and integral coefficient, respectively. Then the diagram of Fig. A.25 is transformed to the diagram of Fig. A.26. For simplicity, we approximate the time delay as the first-order lag \[ e^{sR_0} \approx \frac{1}{1 + sR_0} \]. Thus, we obtain the characteristic equation:

\[
s^4 + \alpha_1 s^3 + \alpha_2 s^2 + \alpha_3 s + \alpha_4 = 0,
\]

(A.10)

where

\[
\begin{aligned}
    \alpha_1 \doteq & \frac{1}{R_0} - \frac{\partial f}{\partial W_s} + \frac{W_{s0}}{R_0^2 C} \\
    \alpha_2 \doteq & -\frac{\partial f}{\partial W_s} + \frac{\partial f}{\partial W_{sR}} + \frac{W_{s0}(1 - \frac{\partial f}{\partial W_s} R_0)}{R_0^2 C} \\
    \alpha_3 \doteq & -\frac{W_{s0}(\frac{\partial f}{\partial W_s} + \frac{\partial f}{\partial W_{sR}})}{R_0^2 C} - \frac{\partial f}{\partial q_R} + \frac{\partial f}{\partial p_R} k_p \\
    \alpha_4 \doteq & -\frac{\partial f}{\partial p_R} k_i \\
\end{aligned}
\]

(A.11)
According to the Routh stability criterion, the system is stable if and only if

$$\begin{align*}
\alpha_1 &> 0 \\
\beta_1 &= \alpha_1 \alpha_2 - \alpha_3 > 0 \\
\beta_2 &= \alpha_3 (\alpha_1 \alpha_2 - \alpha_3) - \alpha_1^2 \alpha_4 > 0 \\
\alpha_4 &> 0.
\end{align*}$$

\tag{A.12}

As an example we demonstrate that the settings of the Subsection 4.2 leads to a stable system. Computation based on these settings are given in Table A.11 and A.12, where PI is chosen as the AQM scheme. The values shown in the two tables demonstrate that they all satisfy the conditions, so the system is stable.

**ACKNOWLEDGEMENT**

This paper was supported by a grant from City University of Hong Kong (Project No. 9380044), and two grants from National Natural Science Foundation of China (Project No. 60974129) and (Project No. 70931002).
Table A.11: The Values for Scenario A

| $N$  | $\alpha_1$ | $\beta_1$ | $\beta_2$ | $\alpha_4$ |
|------|-------------|------------|------------|------------|
| 500  | 14.8205     | 946.6351   | 1.2581e+3  | 0.0472     |
| 800  | 14.0721     | 799.4057   | 8.3581e+4  | 0.0270     |
| 1100 | 13.6513     | 732.6899   | 6.7882e+4  | 0.0225     |
| 2000 | 13.2988     | 672.6120   | 5.5029e+4  | 0.0194     |

Table A.12: The Values for Scenario B

| $N$  | $\alpha_1$ | $\beta_1$ | $\beta_2$ | $\alpha_4$ |
|------|-------------|------------|------------|------------|
| 200  | 12.4921     | 536.3968   | 3.5160e+4  | 0.0403     |
| 500  | 45.8829     | 2.6924e+4  | 2.6975e+7  | 0.0422     |
| 800  | 15.7664     | 1.1551e+3  | 1.7474e+5  | 0.0203     |
| 1100 | 16.9312     | 1.4268e+3  | 2.6368e+5  | 0.0232     |

References

[1] S. Floyd, TCP and explicit congestion notification, SIGCOMM Comput. Commun. Rev. 24 (1994) 8–23.

[2] G. Chatranan, M. A. Labrador, S. Banerjee, A survey of TCP-friendly router-based AQM schemes, Comput. Commun. 27 (15) (2004) 1424C–1440.

[3] S. Floyd, V. Jacobson, Random early detection gateways for congestion avoidance, IEEE/ACM Trans. Netw. 1 (4) (1993) 397–413.

[4] C. V. Hollot, V. Misra, D. Towsley, W.-B. Gong, On designing improved controllers for AQM routers supporting TCP flows, in: Proc. IEEE INFOCOM 2001, Vol. 3, Anchorag, AK, USA, 2001, pp. 1726–1734.

[5] B. Wydrowski, M. Zukerman, GREEN: An active queue management algorithm for a self managed internet, in: Proc. IEEE ICC 2002, Vol. 4, New York, United States, 2002, pp. 2368–2372.

[6] W.-C. Feng, K. G. Shin, D. D. Kandlur, D. Saha, The BLUE active queue management algorithms, IEEE/ACM Trans. Netw. 10 (4) (2002) 513–528.

[7] S. Kunniyur, R. Srikant, Analysis and design of an adaptive virtual queue (AVQ) algorithm for active queue management, SIGCOMM Comput. Commun. Rev. 31 (4) (2001) 123–134.

[8] S. Athuraliya, S. H. Low, V. H. Li, Q. Yin, REM: Active queue management, IEEE Netw. 15 (3) (2001) 48–53.
[9] C. Long, B. Zhao, X. Guan, J. Yang, The Yellow active queue management algorithm, Comput. Netw. 47 (4) (2005) 525–550.

[10] J. Sun, M. Zukerman, RaQ: A robust active queue management scheme based on rate and queue length, Comput. Commun. 30 (8) (2007) 1731–1741.

[11] M. H. Suzer, K.-D. Kang, C. Basaran, Active queue management via event-driven feedback control, Comput. Commun. 35 (4) (2012) 517–529.

[12] H. Wang, C. Liao, Z. Tian, Effective adaptive virtual queue: a stabilising active queue management algorithm for improving responsiveness and robustness, IET Communications 5 (1) (2011) 99–109.

[13] Y. Fan, F. Ren, C. Lin, Design a PID controller for active queue management, in: Proc. ISCC 2003, Vol. 2, Kemer - Antalya, Turkey, 2003, pp. 985–990.

[14] J. V. Chen, F.-C. Chen, J. M. Tarn, D. C. Yen, Improving network congestion: A RED-based fuzzy PID approach, Comp. Stand. Inter. 34 (5) (2012) 426C–438.

[15] P. Wang, H. Chen, X. Yang, Y. Ma, Design and analysis of a model predictive controller for active queue management, ISA transactions 51 (1) (2011) 120–131.

[16] M. Mathis, J. Semke, J. Mahdavi, T. Ott, The macroscopic behavior of the tcp congestion avoidance algorithm, SIGCOMM Comput. Commun. Rev. 27 (3) (1997) 67–82.

[17] F. P. Kelly, A. K. Maulloo, D. K. H. Tan, Rate control for communication networks: shadow prices, proportional fairness and stability, J. Oper. Res. Soc. 49 (3) (1998) 237–252.

[18] S. H. Low, A duality model of TCP and queue management algorithms, IEEE/ACM Trans. Netw. 11 (4) (2003) 525–536.

[19] V. Misra, W.-B. Gong, D. Towsley, Fluid-based analysis of a network of AQM routers supporting TCP flows with an application to RED, SIGCOMM Comput. Commun. Rev. 30 (4) (2000) 151–160.

[20] M. Ajmone Marsan, M. Garetto, P. Giaccone, E. Leonardi, E. Schiattarella, A. Tarello, Using partial differential equations to model TCP mice and elephants in large IP networks, IEEE/ACM Trans. Netw. 13 (6) (2005) 1289–1301.

[21] C. V. Hollot, V. Misra, D. Towsley, W.-B. Gong, A control theoretic analysis of RED, in: Proc. IEEE INFOCOM 2001, Vol. 3, Anchorage, AK, USA, 2001, pp. 1510–1519.
[22] L. Yu, M. Ma, W. Hu, Z. Shi, Y. Shu, Design of parameter tunable robust controller for active queue management based on $h_{\infty}$ control theory, J. Netw. Comput. Appl. 34 (2) (2011) 750–764.

[23] J. Wang, L. Rong, Y. Liu, Design of a stabilizing AQM controller for large-delay networks based on internal model control, Comput. Commun. 31 (10) (2008) 1911–1918.

[24] R. Shorten, F. Wirth, D. Leith, A positive systems model of TCP-like congestion control: asymptotic results, IEEE/ACM Trans. Netw. 14 (3) (2006) 616–629.

[25] S. Low, A duality model of tcp and queue management algorithms, IEEE/ACM Trans. Netw. 11 (4) (2003) 525–536.

[26] V. Firoiu, J.-Y. Le Boudec, D. Towsley, Z.-L. Zhang, Theories and models for Internet quality of service, Proc. IEEE 90 (9) (2002) 1565–1591.

[27] P. Ranjan, E. Abed, R. La, Nonlinear instabilities in TCP-RED, IEEE/ACM Trans. Netw. 12 (6) (2004) 1079–1092.

[28] W. Stevens, RFC2001: TCP Slow Start, Congestion Avoidance, Fast Retransmit, and Fast Recovery Algorithms (Jan. 1997).

[29] W. Stevens, RFC2581: TCP Congestion Control (Apr. 1999).

[30] E. Blanton, RFC5681: TCP Congestion Control (Sep. 2009).

[31] G. F. Franklin, J. D. Powell, A. Emami-Naeini, Feedback Control of Dynamic Systems, Prentice Hall, 2006.