Modeling discrete-time analytical models based on random early detection: Exponential and linear

Hussein Abdel-Jaber*,§, Fadi Thabitah†,¶, and Mike Woodward‡,∥

*Department of Information Technology and Computing
Faculty of Computer Studies, Arab Open University
Kingdom of Saudi Arabia

†Ebusiness Department, Canadian University of Dubai, UAE

‡Department of Computing, University of Bradford
Bradford BD7 1DP, UK

§hadeljaber@arabou.edu.sa
¶fadi@cud.ac.ae
∥m.woodward@brad.ac.uk

Received 22 February 2013
Accepted 14 April 2015
Published 26 May 2015

Congestion control is among primary topics in computer network in which random early detection (RED) method is one of its common techniques. Nevertheless, RED suffers from drawbacks in particular when its “average queue length” is set below the buffer’s “minimum threshold” position which makes the router buffer quickly overflow. To deal with this issue, this paper proposes two discrete-time queue analytical models that aim to utilize an instant queue length parameter as a congestion measure. This assigns mean queue length (mql) and average queueing delay smaller values than those for RED and eventually reduces buffers overflow. A comparison between RED and the proposed analytical models was conducted to identify the model that offers better performance. The proposed models outperform the classic RED in regards to mql and average queueing delay measures when congestion exists. This work also compares one of the proposed models (RED-Linear) with another analytical model named threshold-based linear reduction of arrival rate (TLRAR). The results of the mql, average queueing delay and the probability of packet loss for TLRAR are deteriorated when heavy congestion occurs, whereas, the results of our RED-Linear were not impacted and this shows superiority of our model.

Keywords: Analytical modeling; congestion control; discrete-time queues; performance measures; simulation.

1. Introduction

Congestion is a situation that occurs at the network’s router buffer when the available network resources such as buffer rooms cannot accommodate the arriving
Congestion contributes in degrading the network performance through, (1) increasing the loss rate and the average queueing delay for packets, (2) reducing the throughput and (3) unbalanced fair share of resources among network sources. One solution for controlling congestion is using active queue management (AQM) techniques,\textsuperscript{3–9} such as RED,\textsuperscript{6} Gentle RED,\textsuperscript{8} Adaptive RED,\textsuperscript{7} dynamic random early drop (DRED),\textsuperscript{3} and others. An AQM method often detects congestion at a router buffer early before the router buffer overflows.\textsuperscript{3,4}

One of the known AQM techniques is RED, initially the performance of RED was satisfactory. Subsequently, the performance of RED became inadequate because of the variety of data traffic, i.e., voice, video, image, etc. This has caused problems primarily tuning the input parameters (min threshold, max threshold, etc.) to optimal values\textsuperscript{10} in order to obtain a satisfactory result. In addition, RED’s congestion measure (aql) may vary according to the congestion status. In particular, and for a short period of time, the packet arrival rate can unexpectedly increase which may cause RED’s router buffer to overflow. In this case, the aql value is indeed smaller than the min threshold position and it will take some time to increase. Therefore, the router buffer will not be able to drop packets even if the buffer has already overflown.

One possible solution of the second deficiency of RED is to utilize an instantaneous queue length as a congestion measure rather than the aql since the instantaneous queue length becomes larger than or equal to the min threshold position before the router buffer overflows, ensuring a probabilistic dropping of packets. Another way of controlling congestion is to analyze and model queueing network systems based on AQM techniques\textsuperscript{11–16} such as using discrete-time queues.\textsuperscript{17} A discrete-time queue uses a time unit called a slot.\textsuperscript{18} Relatively, little research works have conducted on building analytical models based on AQM methods, e.g., Refs. 14–17 and this is because of the complex nature in building analytical models and the variety of data traffic.

The main contribution of this paper is building two analytical models based on the RED called RED-Exponential and RED-Linear. These models aim to maintain their mean queue length (MQL) and average queueing delay at values smaller than those for RED to reduce buffers overflow in congestion situations. Moreover, the developed models rely upon the instantaneous queue length threshold as a congestion detector since this threshold passes over the min threshold position before the router buffer overflows as discussed earlier.

In the proposed RED-Exponential model, when congestion occurs at the router buffer the probability of packet arrival decreases exponentially and the packet dropping probability increases exponentially until the queue length reaches the max threshold position. When the instantaneous queue length is equal to or larger than the max threshold position, a heavy congestion occurs. To manage the heavy congestion, the packet arrival probability decreases to a constant value and drops packets at a maximum probability value.
Modeling discrete-time analytical models

On the other hand, and for the RED-Linear model, when the queue length is between min threshold and max threshold positions congestion occurs at the router buffer. The congestion is then controlled by decreasing the value of packet arrival probability linearly until the queue length reaches the maximum threshold position. Also, the packet dropping probability value increases linearly as long as the queue length increases from min threshold to the max threshold. In cases when the queue length is equal to or greater than max threshold, the RED-Linear performs like the RED-Exponential with reference to the packet arrival and dropping probability values.

This paper is organized as follows: Related work on discrete-time queues such as Threshold-based Linear Reduction of Arrival Rate (TLRAR) and the RED are presented in Sec. 2. The RED-Exponential and the RED-Linear analytical models are discussed in Sec. 3. Section 4 presents the performance evaluation of the proposed RED-Linear and TLRAR analytical model and Sec. 6 is devoted to the comparison of RED and the proposed analytical models with reference to different performance measures. Lastly, conclusions and future work are given in Sec. 5.

2. Related Work

Discrete-time queueing is used to analyze and model network systems\textsuperscript{9,10,19} in order to measure their performance. The results of the analysis and modeling are, (1) the balance equations, (2) the equilibrium probabilities and (3) the performance measures numbers, e.g., MQL, throughput, average queueing delay, etc. Discrete-time queueing are also used to analyze queueing systems that employ AQM as of congestion control methods, such as DRED models,\textsuperscript{11,12,14,15} and GRED models.\textsuperscript{14} Hereunder we summarizes relevant research works and then examines two common approaches: RED and TLRAR since we compare with them in the experimental section.

In Ref. 20, an analytical model was presented based on the RED with the aim to eliminate the bias and bursty traffic in order to manage packets queueing delay. Woodward introduced a book related queueing systems within computer networks in order to obtain their performance. This book presented the way in which discrete-time queues approach is applied on computer and communication systems to improve their performance. References 3 and 21 presented other books to investigate queueing systems using discrete-time queues.

The authors of Refs. 13–17 developed discrete-time finite queue analytical models. For instance, in Ref. 16, a discrete-time queue analytical model based on DRED was proposed. The model used the instantaneous queue length as a congestion measure and a predetermined level equals to \(0.9 \times \text{half of the buffer capacity}\) as a congestion detector. Congestion happens when the queue length is equal to or larger than the predetermined level. This model controlled congestion by decreasing the value of packet arrival probability to another constant value. Abdeljaber \textit{et al.}\textsuperscript{13,16}
proposed a discrete time queue analytical model(s) based on BLUE\textsuperscript{12} to manage congestion. Congestion occurs when the queue length is larger than a predetermined level. Therefore, BLUE model of Abdeljaber \textit{et al.}\textsuperscript{16} controlled congestion when the queue length is larger than the predetermined level by decreasing the packet arrival probability value to another constant value. Abdeljaber \textit{et al.}\textsuperscript{13} model is similar to model Abdeljaber \textit{et al.},\textsuperscript{16} but they differ in decreasing their packet arrival probability values when congestion is occurred. Abdeljaber \textit{et al.}\textsuperscript{13} model controlled congestion linearly by decreasing the packet arrival probability value.

A discrete-time analytical model based on GRED was proposed in Ref. 14. This model relied upon the instantaneous queue length as a congestion measure and three thresholds (min threshold, max threshold, double max threshold) positions at the router buffer, where (min threshold < max threshold < double max threshold). When congestion happens it gets managed linearly by decreasing the packet dropping probability value. Also, if the queue length is equal to or greater than the double max threshold, a heavy congestion exists, and to control it, the probability of packet arrival gets reduced to another constant value.

Other discrete-time queueing network analytical models based on the DRED have been proposed in Refs. 11–12 and 15. For example, the proposed model in Ref. 15 is an extension of the DRED in Ref. 17. The new model in Ref. 15 was built by modeling two queue nodes where queue node 1 has a higher priority than the queue node 2 in receiving packets generated from outside the network. In Ref. 9, two discrete-time queues that can be used in packet switching network were developed. Packet arrivals to the queues were generated by different sources that can be correlated. Packet arrivals are assumed to occur in the end of a slot. Since queues are in tandem, the first queue result enters to the second queue, and the size of the two queues is infinite. The result from analyzing the two queues were given by the generating function of the steady state distribution of the two queue lengths, and the average queueing delay for packets, including packets currently in service.

A discrete-time queue analytical model called TLRAR was proposed in Ref. 19. The detail about this model is given in Sec. 2.2, and in Sec. 5 this model is compared with the proposed RED-Linear.

A discrete-time queue using a correlated packet arrivals and general service times was presented in Ref. 19. This queue node has an infinite queue size, and the arrivals have been modeled an on or off Bernoulli bursty source with geometrically distributed on and off periods. Zhou and Wang\textsuperscript{19} achieved a closed form equations for some performance measures, i.e., MQL, cell delay, unfinished work, etc. It has also examined some parameters’ effects on the performance measure results.

In Ref. 21, a discrete time-queue with batch service has been introduced in which infinite and finite buffer occupancies have been used. The packet arrivals are either in a single or batches in a slot. Packet inter-arrival times and service times are assumed to be i.i.d with geometrically distributed. The equilibrium probabilities and some performance measures were derived.\textsuperscript{21} Finally, the authors in Ref. 22 have developed a multi-time step model for a queue prediction based on a Discrete Time
Modeling discrete-time analytical models

Point Process (DTPP). This model has shown a fairly accurate result in queue prediction.

2.1. Random early detection model

RED is one of the known congestion control technique, which was primarily proposed to manage congestion before the buffer gets overflowed. The RED router buffer uses the aql as a congestion measure so for every arriving packet, the router buffer calculates the aql, then compares it with the min threshold and the max threshold. The max threshold value is at least double of the min threshold value aiming to maintain the throughput performance and to avoid an early dropping of packets. Now, when the aql is smaller than the min threshold position, no congestion will happen, and as a result no packets are dropped. Though, when the aql is between the min threshold and the max threshold positions, congestion occurs at the buffer, and the buffer alleviates congestion by dropping packets probabilistically, i.e., linearly. In cases when the aql $\geq$ max threshold, every arriving packet will be dropped ($D_p = 1$) to control the heavy congestion at the buffer. The RED pseudocode is given in Fig. 1.

2.2. TLRAR analytical model

A discrete-time queue analytical model aimed to control the congestion of Internet traffic using a queue threshold scheme has been proposed in Ref. 18. Two thresholds ($L_1$ and $L_2$) were used to make the arrival rate reduces linearly between the two thresholds in this model. The single buffer with $L_1$ and $L_2$ of this model is given in Fig. 2. The model of Guan et al. assumes the following:

1. The probability of arrival in a slot before the number of packets in the system reaches $L_1$ and after reaches $L_2$ are $\alpha_1$ and $\alpha_2$, respectively.
2. The probability of a departure in a slot is $\beta$.

For every arriving packet at a RED router buffer, the router buffer does the following:

1. Calculate the aql.
2. Check the aql position with two threshold (min threshold and max threshold) positions at a RED router buffer, then the router performs its action as follows:

   - If $aql < \text{min\_threshold}$
     - No congestion is occurred
   - If $aql \geq \text{max\_threshold}$
     - A heavy congestion is existed. As a result, drops/marks every arriving packet with $D_p = 1$
   - If $\text{min\_threshold} \leq aql \&\& aql < \text{max\_threshold}$
     - Congestion is presented; therefore drops/marks every arriving packet randomly (probabilistically) with calculating its $D_p$ value

Fig. 1. The general pseudocode of RED technique.

1550028-5
Fig. 2. The single buffer with two thresholds (L₁ and L₂).

Fig. 3. The state transition diagram for the discrete-time finite queue with two thresholds (L₁ and L₂).

(3) When the number of packets in the system is between L₁ and L₂, arrival rate will be linearly reduced with a constant probability.

In the model of Guan et al., the packets dropping probability increases linearly from 0.0 to \((1 - \alpha_2/\alpha_1)\). The state transition diagram for the discrete-time finite queue with L₁ and L₂ is shown in Fig. 3.

State Transition Diagram
The state transition diagram of TLRAR model is shown in Fig. 3. Using Fig. 3, the arrival rate in part I is \(\alpha_1\). In part II, the arrival rate relies on the state, this means that every arrival rate is different with every state and will be linearly reduced by dropping packets.

3. The Proposed Models
3.1. RED's exponential
This section presents a new analytical model which is built by analyzing a single queue node (Fig. 4) and using discrete-time queue. The new model controls...
Modeling discrete-time analytical models

congestion at the router buffer by exponentially reducing the value of packet arrival probability. Further, it aims to alleviate its MQL and average queueing delay ($D$) at values smaller than those of RED to reduce buffer overflow in congestion scenarios. The congestion measure in the RED-Exponential is the instantaneous queue length since this parameter will pass the min threshold at the router buffer, and therefore the dropping of packets occurs before the router buffer overflows.

In the proposed model, the i.i.d Bernoulli process is used as the arrival process assuming $a_n$ is the number of arriving packets in a slot $n$, where $a_n \in \{0, 1\}$, $n = 0, 1, 2, \ldots, n$. First-come-first-served (FCFS) is the policy of queueing discipline in Fig. 4.

In Fig. 4, the probability of packet arrival in a slot when the queue length is less than the min threshold and equal to or larger than the max threshold at the router buffer is represented by $\alpha_1$ and $\alpha_2$, respectively. The packet arrival probability in a slot when max threshold $> \text{queue length} \geq \text{min threshold}$ is $\alpha_i$ ($\alpha_i$ is given later on in this section). The probability of packet departure from a queue node in a slot is denoted by $\beta$. The queue capacity at the router buffer is finite and equals to $K$, where $K$ is including packets currently in the service. Inter-arrival times of packets are geometrically distributed with means $\frac{1}{\alpha_1}$ (when queue length $< \text{min threshold}$), $\frac{1}{\alpha_2}$ (when queue length $\geq \text{max threshold}$). $\frac{1}{\alpha_i}$ (when max threshold $> \text{queue length} \geq \text{min threshold}$), and service time of packets is geometrically distributed with means $\frac{1}{\beta}$.

According to Fig. 4, when the queue length is smaller than the min threshold, no congestion will take place. Therefore, the probability of packet arrival is $\alpha_1$ and packet dropping probability ($D_p$) is 0, whereas if the queue length is between min threshold and max threshold, a congestion incident occurs, and to control it, the probability of packet arrival reduces exponentially from $\alpha_1$ to $\alpha_i$, where

$$\alpha_i = \alpha_1 - (\text{max threshold} - i + 2)^2 \cdot \frac{(\alpha_1 - \alpha_2) \cdot e^{-(\text{max threshold} - i + 2)}}{\text{(max threshold} - \text{min threshold})},$$

if min threshold $\leq i < \text{max threshold}$ and $i$ is the queue state.  

$$\alpha_1$$

$\alpha_2$

$\beta$

max threshold

min threshold

Exponentially decreasing in $\alpha_1$

Fig. 4. The single queue node system for RED-Exponential model.
Equation (1) is given aiming that RED-Exponential analytical model can offer more steady performance measure results than RED with reference to MQL and average queueing delay. This equation gives the value of the packet arrival probability based on the queue state, the queue state value is between the values of min threshold and max threshold. Moreover, as long as the queue state increases, the value of $\alpha_i$ decreases exponentially. When the queue length reaches the max threshold position, $\alpha_i$ decreases to $\alpha_2$ with an aim to control the heavy congestion. It is considered that the single queue node is in equilibrium and the queue length process is a Markov chain with finite state space, where the state space is: $\{0, \ldots, \text{min threshold}, \ldots, \text{max threshold}, \ldots, K\}$. Also, It is assumed that half $\alpha_1$ values (see Sec. 5 and Sec. 6.1) are satisfying $\beta > \alpha_1$ and the other half $\alpha_1$ values are satisfying $\beta < \alpha_1$, $\alpha_1 > \alpha_2$ and $\beta > \alpha_2$.

### 3.1.1. State transition diagram and balance equations of RED-exponential

#### State transition diagram

The state transition diagram of the RED-Exponential is exhibited in Fig. 5 which is used as the basis to derive the balance equations below (Eqs. (2)–(10)).

#### Balance equations

The balance equations of the proposed RED-Exponential are defined as the equations that show the probability at each queue state ($i = 0, 1, \ldots, K$) equal to the flow probability for this queue state.

$$p_0 = (1 - \alpha_1)p_0 + [\beta(1 - \alpha_1)]p_1,$$

$$p_1 = \alpha_1p_0 + [\alpha_1\beta + (1 - \alpha_1)(1 - \beta)]p_1 + [\beta(1 - \alpha_1)]p_2.$$  \hspace{1cm} (2)

In general:

$$p_i = [\alpha_1(1 - \beta)]p_{i-1} + [\alpha_1\beta + (1 - \alpha_1)(1 - \beta)]p_i + [\beta(1 - \alpha_1)]p_{i+1},$$

where $i = 2, 3, 4, \ldots, \text{min threshold} - 2.$ \hspace{1cm} (4)

---

Fig. 5. The state transition diagram of the RED-Exponential and RED-Linear analytical model based on RED.
Modeling discrete-time analytical models

\[ P_{\text{min threshold} - 1} = [\alpha_1(1 - \beta)]P_{\text{min threshold} - 2} + [\alpha_1\beta + (1 - \alpha_1)(1 - \beta)]P_{\text{min threshold} - 1} + [\beta(\text{min threshold})]P_{\text{min threshold}}; \]
\]
\[ P_{\text{min threshold}} = [\alpha_1(1 - \beta)]P_{\text{min threshold} - 1} + [\alpha_{\text{min threshold}}\beta + (1 - \alpha_{\text{min threshold}})(1 - \beta)]P_{\text{min threshold}} + [\beta(\text{min threshold} + 1)]P_{\text{min threshold} + 1}, \]
\]
\[ p_i = \alpha_{i-1}(1 - \beta)p_{i-1} + [\alpha_i\beta + (1 - \alpha_i)(1 - \beta)]p_i + [\beta(1 - \alpha_{i+1})]p_{i+1}, \]
\]
where \( i = \text{min threshold} + 1, \text{min threshold} + 2, \ldots, \text{max threshold} - 2. \)
\]
\[ P_{\text{max threshold} - 1} = [\alpha_{\text{max threshold} - 2}(1 - \beta)]P_{\text{max threshold} - 2} + [\alpha_{\text{max threshold} - 1}\beta + (1 - \alpha_{\text{max threshold} - 1})(1 - \beta)]P_{\text{max threshold} - 1} + [\beta(\text{max threshold})]P_{\text{max threshold}}, \]
\]
\[ P_{\text{max threshold}} = [\alpha_{\text{max threshold} - 1}(1 - \beta)]P_{\text{max threshold} - 1} + [\alpha_2\beta + (1 - \alpha_2)(1 - \beta)]P_{\text{max threshold}} + [\beta(1 - \alpha_2)]P_{\text{max threshold} + 1}, \]
\]
\[ p_i = [\alpha_2(1 - \beta)]p_{i-1} + [\alpha_2\beta + (1 - \alpha_2)(1 - \beta)]p_i + [\beta(1 - \alpha_2)]p_{i+1}, \]
\]
where \( i = \text{max threshold} + 1, \text{max threshold} + 2, \ldots, K - 1. \)
\]

Lastly,
\[ p_K = [\alpha_2(1 - \beta)]p_{K-1} + [\alpha_2\beta + (1 - \beta)]p_K, \]
\]
where \( K = \text{max threshold} + i = \text{min threshold} + I + J, I = K - \text{min threshold} - J \) and \( J = K - \text{min threshold} - I. \)

3.1.2. Equilibrium probabilities

The equilibrium probabilities \((p_i, i = 0, 1, \ldots, K)\) of the RED-Exponential model are its steady state probabilities that help in calculating the performance measure results.

Consider \( \gamma_j = \frac{\alpha_j(1 - \beta)}{\beta(1 - \alpha_j)}, \) where \( j = 1, 2. \)
\]
In particular, \( j = 1 \) when queue length < min threshold (before congestion occurs) and \( j = 2 \) when queue length ≥ max threshold (after heavy congestion occurs), where \( j \) represents the router buffer state before and after congestion has occurred.
\]
Also it is assumed that \( \gamma_i = \frac{\alpha_i(1 - \beta)}{\beta(1 - \alpha_i)} \) when min threshold ≤ \( i < \) max threshold,
\]
where \( i \) represents the queue state when congestion has happened due to min threshold ≤ queue length < max threshold.
By applying Eqs. (12) and (13) in the balance equations, the equilibrium probabilities of the RED-Exponential model can be obtained. These generally are:

\[
p_i = \frac{\alpha_i^{\min \text{threshold}}(1 - \beta)^{i-1}}{\beta^i(1 - \alpha_i)^i},
\]

\[
p_0 = \frac{\gamma_i^{\min \text{threshold}}}{(1 - \beta)^{p_0}}, \text{ where } i = 1, 2, 3, \ldots, \min \text{threshold} - 1,
\]

\[
p_i = \frac{\alpha_i^{\min \text{threshold}}(1 - \beta)^{i-1} \prod_{j=\min \text{threshold}}^{i-1} (1 - \alpha_j)}{\beta^i(1 - \alpha_i)^i} p_0
\]

\[
= \frac{\gamma_i^{\min \text{threshold}}(1 - \alpha_i) \prod_{j=\min \text{threshold}}^{i-1} \gamma_j}{(1 - \alpha_i)(1 - \beta)} p_0,
\]

where \(i = \min \text{threshold}, \min \text{threshold} + 1, \ldots, \max \text{threshold} - 1\).

\[
p_{\max \text{threshold} + i} = \frac{\alpha_i^{\min \text{threshold}} \gamma_i^{\min \text{threshold}} 2^i (1 - \beta)^{\max \text{threshold} + i - 1}}{\prod_{j=\min \text{threshold}}^{i} (1 - \alpha_j)(1 - \beta)} p_0,
\]

The final probability which can be found is the probability that there is no packet in the single queue node, which is represented by \(p_0\). This can be calculated by substituting the above equilibrium probabilities in the following equation (17) and solving them recursively. The normalizing equation represents that the total of equilibrium probabilities for all queue states is equal to one.

\[
\sum_{i=0}^{\max \text{threshold} - 1} p_i = 1.
\]

Then the result of \(p_0\) is obtained as below:

\[
p_0 = \left[ \frac{1}{1 - \gamma_i^{\min \text{threshold}}(1 - \beta)(1 - \gamma_i)} + \frac{\gamma_i^{\min \text{threshold}}(1 - \alpha_i)}{(1 - \beta)} \right]^{-1}.
\]

3.1.3. Performance measures of RED-exponential

After calculating the equilibrium probabilities, the performance measures of the proposed RED-Exponential model are ready to be evaluated. These performance...
measures are: mean queue length \( mql \), throughput \( T \), average queueing delay \( D \) including packet in the service, overflow packet loss probability \( P_L \) and packet dropping probability \( D_p \). \( mql \) can be defined as the mean number of packets in the system. To compute the performance measures firstly, the generating function \( P(z) \) in Eq. (19) is used to evaluate the \( mql \). \( mql \) is derived by taking the first derivative of \( P(z) \) at \( z = 1 \), therefore the \( mql \) result is shown in Eq. (20).

\[
P(z) = \sum_{i=0}^{K} z^i p_i, \tag{19}
\]

\[
mql = P^{(1)}(1) = \frac{p_0}{(1 - \beta)} \left[ \frac{\gamma 1 - \gamma 1^{\min \text{threshold}}}{(1 - \gamma 1)^2} + \gamma 1^{\min \text{threshold}}(1 - \alpha 1) \right]
+ \gamma 1^{\min \text{threshold}} \left[ \sum_{i=\min \text{threshold}}^{\max \text{threshold}-1} \prod_{j=\min \text{threshold}}^{i-1} \left( \frac{\gamma_j}{(1 - \alpha_j)} \right) \right]
\left[ \frac{\max \text{threshold}(1 - \gamma 2)(1 - \gamma 2^{J+1}) + \gamma 2}{(1 - \alpha 2)(1 - \gamma 2)^2} \right]
+ \gamma 2^{J+1} \left[ \sum_{i=\min \text{threshold}}^{\max \text{threshold}-1} \prod_{l=\min \text{threshold}}^{i-1} \left( \frac{\gamma_l}{(1 - \alpha_l)} \right) \right]. \tag{20}
\]

Secondly, \( T \) is computed as the number of packets which have successfully passed through the queue node. Furthermore, \( T \) can be defined as a fraction of time that a router buffer is busy. The \( T \) can be derived as follows:

\[
T = \beta \sum_{i=1}^{K} p_i = \beta(1 - p_0) \text{packets/slot}. \tag{21}
\]

It is noted in Eq. (21) that \( T \) is equal to multiplying the probability of packet departure by the total of all equilibrium probabilities except \( p_0 \).

Thirdly, \( D \) is calculated as the average of time that packets are delayed at the router buffer. \( D \) can be evaluated relying on the results of \( mql \) and \( T \), in which they apply in Little’s law (Eq. (41)). Thus, the result of \( D \) is given as below:

\[
D = \frac{mql}{T \text{slots}} = \frac{P^{(1)}(1)}{T \text{slots}} = \frac{\sum_{i=0}^{K} i \times p_i}{T \text{slots}}. \tag{22}
\]

Equation (22) shows that the result of \( D \) is equal to dividing the result of \( mql \) by the result of throughput. Further, \( P_L \) is derived as the losing probability of packets due to overflowing the router buffer. This is:

\[
P_L = (1 - \beta) p_K. \tag{23}
\]

1550028-11
Equation (23) indicates that the result of $P_L$ is equal to the difference between 1.0 and the probability of packet departure time ($p_K$).

Finally, the dropping probability of packets ($D_p$) is calculated as the probability of packets which dropped probabilistically prior the router buffer is full ($p_K$). This is:

$$D_p = \max_{i=\text{min threshold}}^{\max \text{ threshold} - 1} \sum_{i=\text{min threshold}}^{\alpha_1 - \alpha_i} \left( \frac{\alpha_1 - \alpha_i}{\alpha_1} \right) \times p_i + \left( \frac{\alpha_1 - \alpha_2}{\alpha_1} \right) \times \sum_{i=\max \text{ threshold}}^{K-1} p_i.$$  \hspace{1cm} (24)

Equation (24) represents the packet dropping probability of the RED-Exponential that is calculated when the queue length state is greater than or equal to min threshold position and less than $K$.

3.2. RED-linear analytical model

The RED-Linear model shown in Fig. 6 is similar to the RED-Exponential model with reference to goals. However, there are some differences between the two models, particularly, RED-Exponential decreases and increases the Packet arrival probability and the packet dropping probability exponentially when congestion control takes place. On the other hand, the RED-Linear decreases and increases the packet arrival probability and the packet dropping probability linearly. To elaborate on the differences, and when congestion occurs, the RED-Linear controls congestion by decreasing the probability of packet arrival linearly from $\alpha_1$ to $\alpha_i$,

$$\text{where } \alpha_i = \alpha_1 - (1 + i - \text{min threshold}) \frac{(\alpha_1 - \alpha_2)}{(1 + \text{max threshold} - \text{min threshold})}$$  \hspace{1cm} (25)

if $\text{min threshold} \leq i < \text{max threshold}$, where $i$ is the queue state. The value of $D_p$ increases linearly from 0 to $\left( \frac{\alpha_1 - \alpha_i}{\alpha_1} \right)$ as long as the queue length increases from min threshold to max threshold – 1.

---

Fig. 6. The single queue node system for RED-Linear model.
In the RED-Linear model, the decreasing of probability packet arrival equation relies on the state transition diagram (Fig. 5). This is since \( \min \text{threshold} \leq i < \max \text{threshold} \), where \( i \) is the queue state, and \( \beta > \alpha \). Figure 5 shows that \( \alpha_i \) as part of the “in” rate \((\alpha_i (1 - \beta))\) starting from \( \min \text{threshold} \) queue state. It also displays that the highest queue state shared using \( \alpha_i \) as part of the “in” flow rate is the \( \max \text{threshold} \), where the “in” flow rate is toward the \( \max \text{threshold} \) queue state and is coming from \( \max \text{threshold} - 1 \) queue state. This is represented by \( \alpha_i (1 - \beta) \). \((1 + i - \min \text{threshold})\) part of Eq. (25) which takes into account the “in” flow rate and not the “out” flow rate because they are linearly decreasing from \( \alpha_1 \) to \( \alpha_i \) starting when queue state \((i)\) is equal to \( \min \text{threshold} \). Thus, \((1 + i - \min \text{threshold})\)
depends on the state transition diagram of RED-Linear model.

Moreover, \((1 + \max \text{threshold} - \min \text{threshold})\) or \((\max \text{threshold} - (\min \text{threshold} - 1))\) part above depends on the state transition diagram of RED-Linear model. In this state transition diagram, \((1 + \max \text{threshold} - \min \text{threshold})\) or \((\max \text{threshold} - (\min \text{threshold} - 1))\) represents the difference between the highest queue state \((\max \text{threshold})\) and the lowest queue state \((\min \text{threshold} - 1)\) using \( \alpha_i \) as part of the “in” flow rate \((\alpha_i (1 - \beta))\) or the “out” flow rate \((\beta(1 - \alpha_i))\) since this difference can compute the sharing range using \( \alpha_i \).

Furthermore, \((\alpha_1 - \alpha_2)\) part of Eq. (25) depends on the state transition diagram of RED-Linear model since \((\alpha_1 - \alpha_2)\) represents the maximum dropping probability of packet arrival. As a result, \((1 + i - \min \text{threshold})\) \([(1+\max \text{threshold} - \min \text{threshold})\] part of \( \alpha_i \) relies on the state transition diagram of RED-Linear model, which is introduced based on RED algorithm. Finally, since \( \alpha_1 \) is linearly decreasing to \( \alpha_i \) when congestion occurs the amount of decreasing from \( \alpha_1 \) (before occurring of congestion) is equal to \((1 + i - \min \text{threshold})\) \([(1+\max \text{threshold} - \min \text{threshold})\] and this linearly decreasing amount is shown in the state transition diagram of RED-Linear.

From all the above description, it is clear that \( \alpha_i \) relies on the state transition diagram of RED-Linear model. Due to the probability of packet arrival value decreases linearly and the \( D_p \) value increases linearly, this model is called RED-Linear. Finally, the RED-Linear model is similar to the RED-Exponential model in terms of queueing discipline, balance equations, equilibrium probabilities and the performance measures.

4. Simulation Results
4.1. Simulation description

RED model was implemented based on simulation in which interarrival and departure times for packets are geometrically distributed with means \( \alpha^{-1} \) and \( \beta^{-1} \), respectively. The arrival process is identical, independently distributed (i.i.d) Bernoulli process, while the departure process is geometrical distribution. The capacity of RED’s router buffer is finite. In each slot, one packet may depart from and/or arrive to a single queue node. The packets are served in the manner of FCFS. RED is like the two proposed analytical models in considering that half \( \alpha_1 \)
values are convincing $\beta > \alpha_1$ and the other half $\alpha_1$ values are convincing $\beta < \alpha_1$, $\alpha_1 > \alpha_2$ and $\beta > \alpha_2$.

A comparison between RED model and the proposed analytical models with reference to the following performance measures: mql, $T$, $D_L$, $P_L$ and $D_p$ have been conducted. This comparison mainly aims to determine which model offers better performance. The performance measure results of RED were obtained after the system reaches a steady state. The RED simulation program was run 10 times, where in each run the seed is changed. The performance measure results of RED are denoted by the mean results of the 10 runs. All models were implemented in Java on Intel Core i7 processor with 1.60 GHz and a 4 GB RAM.

4.2. The performance measure results based on variable $\alpha_1$

In this subsection, the performance measures are computed by setting the $\alpha_1$ parameter to variable values intending to,

1. Identify which model stabilizes the mql at a smaller value to avoid building up the queue at the router buffer.
2. Determine which model serves larger number of packets ($T$).
3. Specify which model offers lower average for packets that are delayed at the router buffer and in the service ($D$),
4. Identify the smallest generated result for overflow packet loss probability ($P_L$) among the compared models.
5. Discover which model drops the lowest amount of packets before the router buffer becomes full ($D_p$).

The parameters of the proposed models are set to the same values as those of RED. These parameter settings of the proposed models and original RED are aimed to create light congestion and congestion situations. For instance, $\alpha_2$ value and some values of $\alpha_1$ have been set smaller than $\beta$ since this will create a situation to prevent building up the queue at the router buffer aggressively, whereas some of $\alpha_1$ are set values higher than $\beta$ in order to build a congestion situation.

The performance measure results for all models are shown in Figs. 7–11 in which Figs. 7–9 illustrate the results of mql, $T$ and $D$ versus the values of $\alpha_1$, respectively. The results of $P_L$ and $D_p$ versus the $\alpha_1$ values are displayed in Figs. 10 and 11 respectively. It is clear in Figs. 7, 9 and 11 that RED and both RED-Linear and RED-Exponential models provide similar mql, $D$ and $D_p$ results when $\alpha_1 < \beta$ (either no congestion or light congestion situation).

RED and the proposed models are compared with regard to the performance measures in order to identify the one that offers more satisfactory performance results. A comparison was conducted by varying the value of packet arrival probability from 0.15 (low traffic state) to 0.75 (high traffic state). The results showed that the proposed analytical models offer more satisfactory results than those of
Modeling discrete-time analytical models

Fig. 7. \( m_q \) versus \( \alpha_1 \).

Fig. 8. \( T \) versus \( \alpha_1 \).

Fig. 9. \( D \) versus \( \alpha_1 \).
RED concerning the mql, average queueing delay and overflow packet loss probability when congestion occurs.

In particular, the RED-Linear model provides smaller mql and $D$ results than those for the RED and the RED-Exponential when congestion situation is existed, i.e., $\alpha_1 = 0.45$ (near $\beta$ value) or $\alpha_1 > \beta$. The smallest mql and $D$ results for the RED-Linear are obtained since it drops smaller amount of packets than both RED and RED-Exponential. Also at congestion situation, RED-Exponential derives smaller mql and $D$ results than RED since it drops fewer numbers of packets than RED.

Figure 8 indicates that RED, RED-Exponential and RED-Linear analytical models generate similar $T$ results whether $\alpha_1$ is smaller or greater than $\beta$ value. Moreover, the same figure shows that when congestion occurs ($\alpha_1 > \beta$), all compared models stabilize their $T$ results at $\beta = 0.5$. Furthermore, it is obvious in Fig. 10 that all models presents similar $P_L$ results when $\alpha_1 < \beta$ (the situation of either no congestion or light congestion). However, when congestion increases ($\alpha_1 > \beta$), both proposed analytical models lose fewer packets than RED due to buffer overflowing, and this is since both models have higher packets dropping
probability results than RED before router buffer becomes full. Moreover, Fig. 11 depicts that RED produces smaller $D_p$ results than RED-Exponential and RED-Linear when either $\alpha_1 = 0.45$ (near $\beta$) or $\alpha_1 > \beta$, and this is because the mql’s results for RED are greater than that of the proposed analytical models. Also, RED-Exponential offers smaller $D_p$ results than RED-Linear when $\alpha_1$ value is larger than $\beta$ but less than 0.75. However, both proposed models drop similar amount of packets before the router buffer is full in heavy congestion situations, i.e., $\alpha_1 = 0.75$.

4.3. The performance measure results based on variable min threshold

This subsection compares RED and the proposed models with reference to the performance measures mentioned in Sec. 4.1 by setting the min threshold parameter to different values. The parameters of RED and the proposed models have been set to values as in Sec. 6.1 with two exceptions, these are setting $\alpha_1$ to 0.78 (heavy congestion situations), and setting the min threshold to 4–8 since we need to examine the min threshold especially for values close to the max threshold.

The performance measure results for RED and the two analytical models versus the min threshold parameter are shown in Figs. 12–16 respectively. Specifically, Figs. 12–14 illustrates the results of mql, $T$ and $D$ versus $\alpha_1$, whereas the $P_L$ and $D_p$ results versus $\alpha_1$ are revealed in Figs. 15 and 16 respectively.

It is indicated in Figs. 12–14 that the mql, $T$ and $D$ results for the RED are not impacted by the min threshold parameter. Therefore, the min threshold parameter of RED cannot affect the results of mql, $T$ and $D$ even under heavy congestion. Furthermore, in Figs. 15 and 16 the min threshold parameter of RED can influence the results of $P_L$ and $D_p$. In particular, the most satisfactory $P_L$ result has been obtained when the congestion measure (aql) of RED reaches the min threshold, the router buffer starts dropping packets.
Fig. 13. $T$ versus min threshold.

Fig. 14. $D$ versus min threshold.

Fig. 15. $P_T$ versus min threshold.
packets before it becomes full until it overflows, then it starts losing packets. Therefore, as long as the min threshold value increases the $P_L$ results increase and the $D_p$ results decrease (see Fig. 16). So, the most optimal min threshold value for RED in order to produce satisfactory $D_p$ result, is the nearest value to the max threshold, e.g., 8.

After analyzing Figs. 12 and 14, the best mql and $D$ results for the RED-Linear model are generated when the min threshold is set to the farthest value from the max threshold, which is 4. This is since when the congestion measure (instantaneous queue length) of RED-Linear model reaches 4, the router buffer drops packets earlier than the other min threshold values, e.g., [5–9]. Further, Figs. 13, 15 and 16 reveal that $T$, $P_L$ and $D_p$ results of the RED-Linear model are not affected by the min threshold, and thereby this model is insensitive to the min threshold with reference to the $T$, $P_L$ and $D_p$ results. Lastly, it is clear in Figs. 12–16 that the RED-Exponential model is insensitive to the min threshold since its performance measure results were not impacted by the min threshold setting values.

4.4. RED-linear versus TLRAR model results

This section presents the performance measure results of the two analytical models: the proposed RED-Linear and TLRAR model. TLRAR model results are generated using the simulation of the single queue node presented previously in Fig. 2. We aim to evaluate the pros and cons of our model comparing with TLRAR in the presence of light congestion and congestion situations. The reason for comparing RED-Linear with TLRAR is that both models use the linear way in decreasing the packet arrival probability and increasing the probability of dropping packet. Both models were implemented in Java. The parameters of the RED-Linear analytical model and TLRAR are set as follows: $\alpha_1$, $\alpha_2$, $\beta$, $K$, min threshold (or $L_1$), max threshold (or $L_2$), have been set to 0.15–0.75, 0.1, 0.5, 20, 3 and 9, respectively. $\alpha_2$ value and some values of $\alpha_1$ have been set smaller than $\beta$ since this will create a situation to prevent building up the queue at the router buffer aggressively. $K$ has been set to 20 in order to measure the performance at small buffer size.
Tables 1 and 2 provide different numerical results for RED-Linear and TLRAR with respect to different evaluation measures including mean queue length, throughput, average queuing delay, the probability of packet loss, overflow packet loss probability and packet dropping probability. From Tables 1 and 2, it is noted that the proposed RED-Linear offers more satisfactory performance measure results than those of the analytical model of Guan et al.\textsuperscript{18} when the value of probability of packet arrival is less than or equal to 0.48 (light congestion or noncongestion situation) and these due to the results of mql and average queuing delay of the proposed linear model are less than those of TLRAR model since the router buffer of RED-Linear is built up with content of packets less than that of TLRAR. Also, the throughput results of RED-Linear are not exactly the same as those of TLRAR, and therefore, the RED-Linear provides higher throughput performance than those of TLRAR. The results of overflow packet loss probability or the packet dropping probability of our model are less than those of TLRAR since the RED-Linear gives smaller mql results than those of the analytical model of Guan et al.\textsuperscript{18}

5. Conclusions and Future Work

Two analytical models based on RED called RED-Exponential and RED-Linear were proposed and built by analyzing a single queue node using discrete-time
Modeling discrete-time analytical models

queues. Both proposed models employed the instantaneous queue length as a congestion measure to prevent early buffers overflow. In RED-Exponential and when congestion exists, the probability of packet arrival decreases exponentially and the packet dropping probability increases exponentially, until the queue length reaches the maximum threshold position. Whereas, in RED-Linear the congestion is controlled by decreasing the value of packet arrival probability linearly until the queue length reaches the maximum threshold position. The proposed models have been compared with RED and TLRAR models with reference to different performance measures. After experimentations and results analysis, we concluded the following:

- RED and the proposed models offered similar mql, $D$ and $D_p$ results when $\alpha_1 < \beta$ (either no congestion or light congestion situation is occurred). Further, the Linear model outperformed RED and the Exponential models in regard to mql and $D$ results in congestion situation.
- RED and the proposed models offered similar $T$ results with or without congestion and all analytical models produced similar $P_L$ results in light congestions. However, when congestion increases ($\alpha_1 > \beta$), the proposed models generate better $P_L$ results than RED.
- The mql and $D$ results for the Linear model are sensitive to the min threshold in which the most satisfactory results were accomplished when the min threshold was set to the farthest set value from the max threshold. Moreover, the results of $T$, $P_L$ and $D_p$ for the Linear model are insensitive to the min threshold parameter.
- The proposed RED-Linear offers more satisfactory performance measure results than those of TLRAR analytical model in light congestion or no congestion situations.
- When the packet arrival probability value is greater than $\beta$ (>0.63), the performance measure results produced TLRAR were deteriorated when compared to our analytical model.
- When the value of packet arrival probability is between 0.48 and 0.63, TLRAR model’s results in regards to probability of packet loss are decreased.

In near future, it is intended to apply the two analytical models as congestion control methods in internet and cellular networks.

References