A Linear Packet Marking Probability Function for Fast Congestion Notification (FN)

Mohammed M. Kadhum
InterNetWorks Research Group
College of Arts and Sciences
Universiti Utara Malaysia
06010 UUM Sintok, MALAYSIA
kadhum@uum.edu.my

Suhaidi Hassan
InterNetWorks Research Group
College of Arts and Sciences
Universiti Utara Malaysia
06010 UUM Sintok, MALAYSIA
suhaidi@uum.edu.my

Summary

The efficiency of queue management mechanisms depends on how well their control decisions helps in satisfying their objectives regarding congestion avoidance and control. These decisions are implemented and compelled during the design of the packet mark probability and the mark activation functions. If a packet is dropped / marked rather than being allowed to the buffer, a congestion notification is delivered and the congestion avoidance and control carries on to be applied at the router. If the queue management mechanism realizes the need to apply more aggressive congestion control, the recently arrived packet should be dropped to provide early congestion notification. In this paper, we design a new packet drop probability function with a built-in drop activation function for Fast Congestion Notification (FN) mechanism. This design enables the two control decisions, packet admissions and congestion control directing, to be made along with each other. This permits sending congestion avoidance notification as early as required even if the queue is almost empty, and preventing congestion notification even if the queue is almost full but the arrival rate is controllable, thus the buffer is fully utilized and the congestion is detected properly.

Keywords:
Internet Congestion; Active Queue Management (AQM); Explicit Congestion Notification (ECN); Gateway buffers

1. Introduction

The TCP-based applications performance depends on the selection of queue management in network routers. Queue management is defined as the algorithms that manage the length of the packet queues by dropping packets when necessary or appropriate [1]. The efficiency of queue management mechanisms depends on how well their control decisions, on packet admission to the queue and congestion control directing, will help in satisfying their objectives regarding congestion avoidance and control. These decisions are implemented and compelled during the design of the packet mark probability and the mark activation functions. Packet admission and congestion control directing control decisions are dependent on each other. Based on the drop activation characteristic, the queue management policies can be classified into two categories. The first category is *reactive (passive) queue management policies*, which do not employ any preventive packet drop before the gateway buffer is flooded. The second category is *proactive* (*active*) *queue management policies(AQM)* which employ preventive packet drop before the gateway buffer gets full [2]. Drop-Tail, which is one of reactive queue management policies, is currently widely developed in the Internet routers. It introduces several problems, such as global synchronization, on the Internet. Active queue management policies, such as Random Early Detection (RED), are expected to eliminate global synchronization and improve Quality of Service (QoS) of the networks. The promised advantages of AQM are increase in throughput, reduce the delay, and avoid lock-out. AQM provides preventive measures to manage the router queue to eliminate the problems associated with passive queue management. AQM has the following attributes:

- Performing a preventive random packet drop before the queue is full.
- The probability of the preventive packet drop is proportional to congestion levels.

Preventive packet drop provides implicit feedback method to notify the traffic senders of the congestion onset [3]. As a reaction, senders reduce their transmission rate to moderate the congestion level. Arriving packets from the senders are dropped randomly, which prevents senders from backing off at the same time and thereby eliminate global synchronization [3].

2. Random Early Detection (RED)

RED is the default AQM mechanism that is recommended by IETF for the Internet routers [1], which was proposed by Floyd and Jacobson [4] in 1993. Fig. 1, Fig. 2, and Fig. 3, show the algorithm, gateway buffer, and packet drop function of RED, respectively. A router implementing RED detects early by computing the average buffer length (*avg*) and sets the two queue thresholds (Max_{th} and Min_{th}) for packet drop. The average buffer length at time t, is defined as

---

Manuscript received May 5, 2009
Manuscript revised May 20, 2009

$$avg\,(t) = (1 - w)\,avg\,(t - 1) + wq\,(t) \qquad (1)$$

and it is used as a control variable to perform active packet drop. The *avg* (*t*) is the new value of the average buffer length at time *t*, *q* (*t*) is instantaneous buffer length at time *t*, and *w*, which is normally less than one, is a weight parameter in calculating *avg*.

```
For each packet arrival
        Calculate the average queue size avg
        If Min_Threshold ≤ avg < Max_Threshold
            Calculate probability p
                With probability p:
                    Mark the arriving packet
        Else if Max_Threshold ≤ avg
                    Mark the arriving packet
```

Fig 1. RED Algorithm

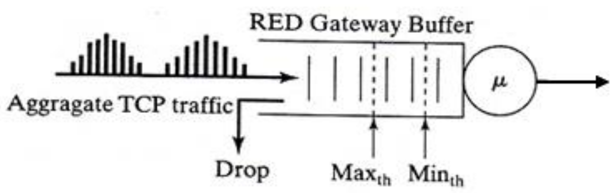

Fig 2. RED Gateway Buffer

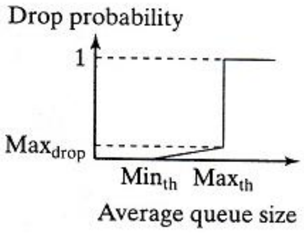

Fig 3. RED Packet Drop Function

2.1. Packet Drop Probability

Packet drop probability function determines the probability that the packet is dropped when the drop activation function imposes drop procedure initiation and the drop position function selects the specific packet to be dropped. For example in drop-tail, the chosen packet which is the packet at the tail of the queue is dropped with probability one.

In RED, the probability of dropping packet, *P*, is calculated by

$$P = Max_{drop}\,((avg - Min_{th}) / (Max_{th} - Min_{th})) \qquad (2)$$

The RED algorithm includes two computational parts: computation of the *average buffer length* and calculation of the *drop probability*.

The RED algorithm involves four parameters to regulate its performance. Min_{th} and Max_{th} are the queue thresholds to perform packet drop, Max_{drop} is the packet drop probability at Max_{th}, and *w* is the weight parameter to calculate the average buffer size from the instantaneous queue length. The average buffer length follows the instantaneous buffer length. However, because *w* is much less than one, *avg* changes much slower than *q*. Therefore, *avg* follows the long-term changes of *q*, reflecting persistent congestion in networks. By making the packet drop probability a function of the level of congestion, RED gateway has a low racket-drop probability during low congestion, while the drop probability increases the congestion level increases [3].

The packet drop probability of RED is small in the interval Min_{th} and Max_{th}. Furthermore, the packets to be dropped are chosen randomly from the arriving packets from different hosts. As a result, packets coming from different hosts are not dropped simultaneously. RED gateways, therefore, avoid global synchronization by randomly dropping packets.

The performance of RED significantly depends on the values of its four parameters [5], Max_{drop}, Min_{th}, Max_{th}, and *w*.

2.2. Calculating the Average Queue Length to Improve Response Time

RED uses four parameters and one state variable to regulate its performance. The state variable is the average buffer length, which is defined as

$$avg = (1\ \ w)\,avg + wq \qquad (3)$$

and works as a low pass filter (LPF) [4]. In the above expression, *w* is a weight parameter and *q* is the instantaneous queue size of gateway buffer. The average buffer length controls the active packet drop in the RED queue. The advantages of using average queue length to control active packet drop are:

- Accumulating short-term congestion
- Tracing long-term congestion.

Nevertheless, the low pass filter characteristic of average queue is also featured with slow-time response to the changes of long-term congestion in networks. This is harmful to the throughput and delay performance of RED gateway. For example, after a long-term congestion, the average buffer length stays high even if the instantaneous queue is back to normal or low; RED will, therefore, continue dropping packets even after the end of congestion [6] resulting in low throughput. The slow response of the average queue length will result in the throughput restoring slowly after heavy congestion [7]. A larger value of *w* can improve the response time, but at the expense of the RED queue tracing short-term congestion, which is against the AQM principle.

3. Fast Congestion Notification (FN)

The FN [8] queue management algorithm randomly marks (if ECN) / drops (if non-ECN) the arriving packets before the buffer overflows, to effectively control the:

- instantaneous queue length below a the optimal queue length to reduce the queuing delay and avoid the buffer overflows
- average traffic arrival rate of the queue in the proximity of the departing link capacity to enable the congestion and queue length control

FN integrates the instantaneous queue length and the average arriva1 rate of queue to compute the drop probability of the packet upon each arriving packet, as described in the following sections.

The use of the instantaneous queue length in conjunction with the average queue speed (average arrival rate) can provide superior control decision criteria for an active queue management scheme [9].

3.1. Packet Drop/Mark Probability Function

The two control decisions of packet admissions and congestion control directing could be made in isolation at the gateway. Nevertheless, they are not independent of each other and affect each other. Whether a packet is allowed to the buffer or not will affect the direction taken by the congestion control process over time. If a packet is dropped (marked) rather than being allowed to the buffer, a congestion notification is delivered and the congestion avoidance and control carries on to be exercised at the gateway. On the other hand, the control exercised to guide the direction and aggressiveness of congestion control at the gateway over time affects whether a particular packet will be allowed to the buffer. If the queue management policy decides the need to exercise more aggressive congestion control, then the newly arrived packet has to be dropped to provide early congestion notification.

Particularly, the goal is to design a drop probability function with a built-in drop activation function which will enable the two control decisions of packet admissions and congestion control directing to be made along with each other, rather than in isolation. This is accomplished by computing a single packet marking (dropping) probability to enforce the two control decisions at the same time. The drop/mark probability function will determine the packet dropping (marking) probability based on both decision criteria of instantaneous queue length and average traffic rate. Every packet will be subject to a random drop/mark where the drop/mark probability will depend on the condition of the gateway at the time of the arrival of the packet in terms of the instantaneous queue length at the gateway and the average traffic rate. Packets can get dropped (marked) even when the queue is lightly occupied if the traffic rate is very high and conversely packets may not be dropped (marked) even if the queue is heavily occupied but the arriving traffic is moderate. This allows congestion avoidance notification to be delivered as early as necessary even if the queue is empty, and to be suppressed even if the queue is near-full but the arriving traffic is manageable. A linear dropping/marking probability function presumes that the average aggregate arrival rate remains constant over a period of time used for computing the packet drop probability. The period of time used for computing the packet drop/mark probability will be referred to as the *control time constant*. The designs presented have a built-in drop/mark activation function. In other words, the decision criteria of average buffer arrival rate and instantaneous queue length are only implicitly compared to the predetermined thresholds of a desired optimal queue length and the outgoing link capacity through the computation of the drop/mark probability. A zero probability will indicate that the traffic characteristics and buffer condition are such that the expected growth in the queue level over the next control time constant period will just match the available buffer capacity and therefore no packets are required to be dropped or marked. A negative drop/mark probability will indicate that the expected growth in the queue level over the next control time constant period will be less than the available buffer space at the gateway. In this case, the gateway will be able to hold the expected growth without having its buffer and link capacity resources flooded. In both cases the arriving packet is allowed to the buffer.

3.2. Linear Packet Marking Probability

The linear marking probability function is derived based on the assumption that the arrival traffic process remains unchanged over the control time constant period of length (T) seconds. In other words, it is supposed that immediately following the packet's arrival, the traffic continues to arrive at the fixed rate of (R) bits/sec, the estimated average arrival rate to the buffer computed upon the packet's arrival, for the period of the control time constant. The buffer has a capacity of (C) bits and is served by an outgoing link at a fixed rate of (μ) bits/sec. The packet drop/mark probability (P), is computed for, and applied to, every incoming packet, based on the above assumptions, with the goal of driving the instantaneous (current) queue length (Q_{cur}) to some desired optimal level (Q_{opt}) over the control time constant period (T). These are shown in Fig. 4.

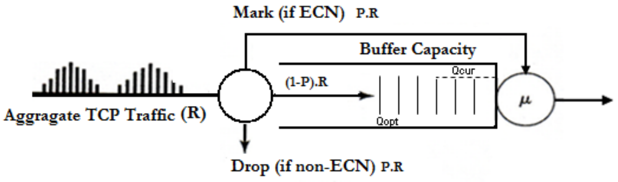

Fig 4. FN Queuing Mode1

Before deriving the drop/mark probability function, we specify some necessary quantities related to the queue, as follows:

- Growth in the queue due to packet arrivals (Q^+): Packets arriving at the buffer at a constant rate of (R) bits/sec over a period of time of length (T) seconds cause the queue size to grow by:

$$Q^{+} = R.T \quad \text{bits.} \tag{4}$$

- Drain in the queue due to packet transmissions (Q_{μ}^{-}): Packets are transmitted from the buffer at a rate of (μ) bits/sec. Over a period of time of length (T) seconds, a total of ($\mu.T$) bits are drained from the buffer and the queue level will be reduced by just as much:

$$Q_{\mu}^{-} = \mu.\ T \quad \text{bits.} \tag{5}$$

- Drain in the queue due to random packet dropping/marking (Q_{P}^{-}): If packets are dropped/marked with a probability of (P) over the period of length (T) seconds, then (P) fraction of the total queue growth that would have resulted in the absence of random packet dropping/marking, ($R.T$), will be converted to queue drain:

$$Q_{P}^{-} = P\ .(R.\ T\ ) \quad \text{bits.} \tag{6}$$

- Total decrease in the queue due to departures & random marking (Q^{-}).

- Total decrease in the queue due to packet transmissions and random packet marking is the sum of (Q_{μ}^{-}) and (Q_{P}^{-}):

$$Q^{-} = Q_{\mu}^{-} + Q_{P}^{-} = (\mu.\ T) + (P.\ (R.\ T\ )) = (\mu + P.\ R).\ T \quad \text{bits.} \tag{7}$$

- Change in the queue level due to arrivals & departures (ΔQ_{μ}): In the absence of random packet drops/marks, the queue growth and drain would be governed only by packet arrivals and departures. The change in the queue level would equal the difference of queue growth and queue drain:

$$\Delta Q_{\mu} = Q^{+} - Q_{\mu}^{-} = (R.\ T) - (\mu.\ T) = (R - \mu).\ T \quad \text{bits.} \tag{8}$$

A positive value indicates growth, a negative value indicates drain, and a zero value indicates the total growth and drain due to arrivals and departures to be equal, neutralizing the effect of each other.

- Total actual growth/drain in the queue due to arrivals, departures, & random marking (ΔQ_{a}).

- The total change in the instantaneous queue length is the difference of the growth due to packet arrivals and the drain due to packet transmissions and random packet drops/marks:

$$\begin{aligned} \Delta Q_{a} &= Q^{+} - Q^{-} \\ &= (R.T) - ((\mu + P.R).T\ )) \\ &= ((R - \mu).T\ ) - ((P.R).T)) \quad \text{bits.} \end{aligned} \tag{9}$$

A positive value indicates a growth and a negative value indicates a drain while zero indicates a steady queue level.

- Total allowed-growth/desired-drain in the queue (ΔQ_{d}): Depending on the relative sizes of the desired optimal queue length (Q_{opt}) and the instantaneous queue length (Q), their difference specifies the total allowed growth in the queue over the control time constant period, if the instantaneous queue length is smaller than the desired optimal queue length ($Q < Q_{opt}$), and it specifies the total desired drain, if the instantaneous queue length is larger than the desired optimal queue length ($Q > Q_{opt}$). If instantaneous queue length is identical to the desired optimal queue length, ΔQ_{d} computes to zero indicating the need to avoid queue growth, not requiring but also not forbidding queue drain: $\Delta Q_{d} = Q_{opt} - Q$ bits.

FN mechanism tries to direct the current instantaneous queue length (Q_{cur}) to the desired optimal queue length (Q_{opt}). The design and dynamics of the packet dropping/marking probability function determines the path that the instantaneous queue length follows in transiting from its current value to the desired optimal value over the control time constant period. Even though a variety of curves can be used as the path in designing the probability function, a good candidate is a straight line. Hence, upon the arrival of every packet, the drop/mark probability will be computed assuming that it is desired to direct the instantaneous queue length from its current value (Q_{cur}) to its desired value (Q_{opt}) along a straight line over a time interval of (T) seconds immediately following the arrival of each packet. The computed drop/mark probability for the newly arrived packet will be the rate at which packets will have to be dropped over the control time constant period to direct the instantaneous queue length from its current position to the desired optimal level if the traffic conditions would continue unchanged over that period. This is illustrated in Fig. 5.

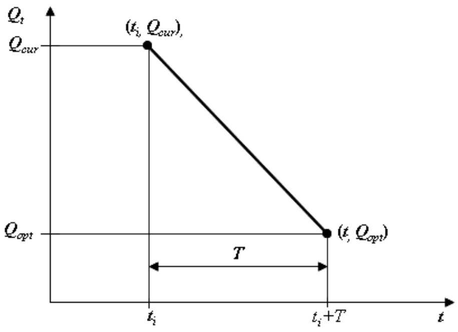

Fig 5. Queue Directing Control Dynamics

The packet dropping/marking probability is updated upon every packet arrival and depends on the difference between the traffic arrival rate and the outgoing link capacity and the difference between the current queue size and the optimal queue length. Upon the arrival of the i^{th} packet (Pkt_i), FN presumes that traffic will continue to arrive at the fixed rate of (R) over the next (T) seconds and computes the drop probability (P) as the fraction (Q_P^- / Q^+) of queue growth due to traffic arrival ($Q^+ = R_i . T$) over this period that has to be discarded by dropping/marking packets randomly. Packets are randomly dropped/marked to assist the outgoing link capacity in lowering the effective traffic rate to the buffer $((1 - P^i).R_i)$, to below the link capacity (μ) and in directing the instantaneous queue length (Q_{cur}) to the desired optimal level (Q_{opt}), over the period of length (T) seconds. Consequently, the drop/mark probability should be chosen such that the desired $(\{\Delta Q_d\}^{(i)})$ and actual $(\{\Delta Q_a\}^{(i)})$ changes in the queue over the period of time of length (T) will become equal:

$$\{\Delta Q_d\}^{(i)} = \{\Delta Q_a\}^{(i)} \tag{10}$$

$$Q_{opt} - Q_{cur} = \{Q^+\}^{(i)} - \{Q^-\}^{(i)} \tag{11}$$

$$Q_{opt} - Q_{cur} = \{Q^+\}^{(i)} - \{Q_\mu^-\}^{(i)} + \{Q_P^-\}^{(i)} = \Delta Q_u - \{Q_P^-\}^{(i)} \tag{12}$$

$$Q_{opt} - Q_{cur} = (R_i .T) - ((\mu .T) + (P^{(i)} .R_i).T)) \tag{13}$$

$$Q_{opt} - Q_{cur} = ((R_i - \mu).T) - (P^{(i)}(.R_i .T)) \tag{14}$$

The equation above can be solved for (P) to attain:

$$P^{(i)} = \frac{((R_i - \mu).T) - (Q_{opt} - Q_{cur})}{R_i .T} = \frac{\Delta Q_u - \Delta Q_d}{Q^+} \tag{15}$$

To show how the designed drop probability function allows the two decisions regarding average arrival rate control and queue length control interact with each other, the drop probability function can be written as a sum of two components :

$$P^{(i)} = \frac{((R_i - \mu).T) - (Q_{opt} - Q_{cur})}{R_i .T} \tag{16}$$

$$P^{(i)} = \frac{((R_i - \mu).T)}{R_i .T} + \frac{(Q_{cur} - Q_{opt})}{R_i .T} \tag{17}$$

$$P^{(i)} = P_R^{(i)} + P_Q^{(i)} \tag{18}$$

The (P_R) component expresses the average arrival rate effects and the (P_Q) component form the instantaneous queue length effects. The relative influence of each decision control on the other is specified by the relative sizes of the two components of the drop probability function, (P_R) and (P_Q). This depends on the degree to which the arrival rate differs from the outgoing link capacity and the instantaneous queue length from the optimal desired queue length.

4. Conclusion

We have presented the design and derivation of the linear version of the FN drop probability function based on the assumptions of constant rate. This permits sending congestion avoidance notification as early as required even if the queue is almost empty, and preventing congestion notification even if the queue is almost full but the arrival rate is controllable, thus the buffer is fully utilized and the congestion is detected properly.

References

[1] B. Braden, D. Clark, J. Crowcroft, B. Davie, S. Deering, D. Estrin, S. Floyd, V. Jacobson, G. Minshall, C. Partridge, L. Peterson, K. Ramakrishnan, S. Shenker, J. Wroclawski, and L. Zhang, *Recommendations on Queue Management and Congestion Avoidance in the Internet*: RFC Editor, 1998.

[2] S. Leonardo, P. Adriano, and M. Wagner, Jr., "Reactivity-based Scheduling Approaches For Internet Services," in *Proceedings of the Fourth Latin American Web Congress*: IEEE Computer Society, 2006.

[3] M. Hassan and R. Jain, *High Performance TCP/IP Networking: Concepts, Issues, and Solutions*: Pearson Prentice Hall, 2004.

[4] S. Floyd and V. Jacobson, "Random early detection gateways for congestion avoidance," *Networking, IEEE/ACM Transactions on,* vol. 1, pp. 397-413, 1993.

[5] F. Wu-Chang and D. K. Dilip, "Adaptive packet marking for maintaining end-to-end throughput in a differentiated-services internet," *IEEE/ACM Trans. Netw.*, vol. 7, pp. 685-697, 1999.

[6] M. May, C. Diot, B. Lyles, and J. Bolot, "Influence of Active Queue Management Parameters on Aggregate Traffic Performance," Institut National De Recherche En Informatique Et En Automatique, 2000.

[7] M. Christiansen, K. Jeffay, D. Ott, and F. D. Smith, "Tuning RED for Web traffic," *Networking, IEEE/ACM Transactions on,* vol. 9, pp. 249-264, 2001.

[8] M. M. Kadhum and S. Hassan, "Fast Congestion Notification mechanism for ECN-capable routers," in

Information Technology, 2008. ITSim 2008. International Symposium on, 2008, pp. 1-6.

[9] M. M. Kadhum and S. Hassan, "The Design Motivations and Objectives for Fast Congestion Notification (FN)," in *Proceedings of the APAN Network Research Workshop* Malaysia, 2009.

Mohammed M. Kadhum is a Lecturer in the Department of Computer Science for Postgraduate Studies, Northern University of Malaysia (Universiti Utara Malaysia). He is currently pursuing his PhD in computer networking. His current research interest is on Internet Congestion. He has been awarded with several medals for his outstanding projects. His professional activity includes being positioned as Technical Program Chair for International Conference on Network Applications, Protocols and Services 2008 (NetApps2008), which has been held successfully in the Northern University of Malaysia (Universiti Utara Malaysia). So far he has published various papers including on well-known and influential international journal.

Associate Professor Dr. Suhaidi Hassan is currently the Assistant Vice Chancellor of the College of Arts and Sciences, Universiti Utara Malaysia (UUM). He is an associate professor in Computer Systems and Communication Networks and the former Dean of the Faculty of Information Technology, Universiti Utara Malaysia.

Dr. Suhaidi Hassan received his BSc degree in Computer Science from Binghamton University, New York (USA) and his MS degree in Information Science (specializing in Telecommunications and Networks) from the University of Pittsburgh, Pennsylvania (USA). He received his PhD degree in Computing (specializing in Networks Performance Engineering) from the University of Leeds in the United Kingdom.

In 2006, he established the ITU-UUM Asia Pacific Centre of Excellence (ASP CoE) for Rural ICT Development, a human resource development initiative of the Geneva-based International Telecommunication Union (ITU) which serves as the focal point for all rural ICT development initiatives across Asia Pacific region by providing executive training programs, knowledge repositories, R&D and consultancy activities.

Dr. Suhaidi Hassan has been elected as a senior member of the Institute of Electrical and Electronic Engineers (IEEE) since March 2008 which he actively involved in both the IEEE Communications and IEEE Computer societies. He served as the Vice Chair (2003-2007) of the IEEE Malaysia Computer Society. He also serves as a technical committee for the Malaysian Research and Educational Network (MYREN) and as a Council Member of the Cisco Malaysia Network Academy.