Reducing Queue-Fill Variability for Emergency Traffic in a Differentiated Services Network

Manali Joshi, Cory Beard
School of Computing and Engineering, University of Missouri-Kansas City
Kansas City, MO 64110, USA

Abstract—
Supporting emergency services applications using the Internet has gained considerable interest in the recent past. A simple Multiple Average Multiple Threshold (MAMT) Random Early Detection (RED) scheme is used to provide necessary Quality of Service for emergency traffic. We have studied the RED dropping function as one of the contributing factors towards the variation observed in the queue fill. This may cause interactive emergency applications to experience large delay and delay variation thus deteriorating the service provided to them. Hence we compare the performance of the linear dropping function in terms of average delay and variation in queue fill with that of three other dropping functions namely: concave, convex and step. With the help of an exhaustive simulation study conducted using the Network Simulator-2; we conclude that the concave function provides the best performance.

Keywords—simulations, active queue management, emergency services, quality of service.

I. INTRODUCTION

Random Early Detection (RED) [2] is a proactive congestion control mechanism that drops packets in a queue before it is full. It aims at maintaining a low average queue size while also avoiding global synchronization and bias against bursty traffic. It has the ability to maintain an upper bound on the average queue size even in the absence of cooperation from transport layer protocols. While the RED algorithm was primarily designed for responsive flows (TCP based applications) that can react to a packet loss and consequently reduce their sending rate, there are also a growing set of UDP based applications in the Internet like streaming video and voice etc. Since non-responsive UDP based applications do not throttle back in response to packet drops, RED can be used as a filter to drop packets and provide prioritized treatment to different classes of packets.

The RED algorithm is governed by 5 control parameters namely the maximum queue size, a minimum threshold ($min_{th}$), a maximum threshold ($max_{th}$), maximum dropping probability ($max_{prob}$) and the queue averaging parameter ($w_q$). The algorithm consists of two parts: Average Queue Estimation Process and the Randomized Dropping Mechanism. RED employs a linear dropping function. However, as will be discussed later in this paper, the linear function may not perform as well as other non-linear functions, especially when a quicker adjustment of dropping probability values is required in response to a changing queue fill. This is the subject of discussion here.

This paper is arranged as follows. Section II describes the concept of multi-level RED and its application to provide preferential treatment to emergency traffic in Diffserv networks. Section III describes the problem of large queue-fill variations observed with a RED queue, the contribution of the existing RED dropping function towards the same and hence the need to study non-linear dropping functions. Section IV details the simulation experiments and the corresponding observations. Section V provides a mathematical analysis validating the performance of non-linear dropping functions as observed during simulations and finally Section VI presents the conclusion.

II. MULTI-LEVEL RED FOR EMERGENCY SERVICES

Network traffic is highly diverse and each traffic type has unique requirements in terms of bandwidth, delay, loss and availability. The best-effort IP network introduces a variable and unpredictable amount of delay and also drops packets when the network is congested. Thus, it may not necessarily provide the required behavior for an application especially during times of heavy congestion. So Quality of Service (QoS) techniques need to be applied to IP networks in order to meet such different needs. The Differentiated Services (Diffserv) architecture has recently become the preferred method to address QoS issues in IP networks. Of the two standard Diffserv forwarding mechanisms, the AF PHB specifies four different traffic classes. Within a traffic class, a packet is assigned one of three levels of drop precedence (DP) which are also referred to in terms of colors as Green (DP0), Yellow (DP1) and Red (DP2) [5]. In case of congestion, packets of the highest DP (i.e. Red packets) are dropped in preference to packets of the lower DP (i.e. Yellow or Green packets).

Multi-level RED is a generic term used to describe a scheme where drop probability for packets of different DPs (or colors) needs to be calculated separately and this is achieved by maintaining multiple sets of RED thresholds—one for each DP [1]. When using multi-colors of packets, different ways of calculating average queue lengths have been proposed [6]. However, here we intend to focus on the application of multi-level RED to support emergency services. Hence we focus on using the Multiple Average Multiple Threshold (MAMT) approach as proposed in [3]. With this MAMT approach, each class has a separate set of RED thresholds; queue averages are computed per class in addition to using different dropping functions per class. However each function has the same form. It should be noted that many improvements to RED have been developed, such as PI, AVQ and ARED, but these only apply to a single class of traffic. For multiple classes, the current state of the art is based on multi-class implementations of RED, like MAMT studied here.

A. Overview of Emergency Services

Emergency services are mainly used for restoring community infrastructure to normal after occurrence of disasters like earthquakes, terrorist attacks etc. The commercial telecommunications infrastructure is rapidly
employing the MAMT scheme to support emergency services.

is indicated by the Blue color.

each of the AF queues to accommodate emergency traffic and

•

•

•

the network and dependability that its packets will receive the intended service. In order to achieve these goals, the scheme as proposed in [3] introduces an additional drop precedence in each of the AF queues to accommodate emergency traffic and is indicated by the Blue color.

There are certain objectives to be considered while employing the MAMT scheme to support emergency services.

- **Prevent emergency packet drops**: We maintain that the emergency packets should not be dropped before 100% of the non-emergency packets have been dropped. Thus, we stipulate a strict requirement that the dropping probability of a higher drop precedence must be almost equal to 1 before the dropping probability of a next lower drop precedence becomes non-zero.

- **Prevent Tail Drops**: Tail drops must be avoided because otherwise emergency packets would be dropped without any kind of preference over non-emergency packets.

- **Restrict instantaneous queue-fill variation**: This is necessary in order to provide better QoS in terms of delay and jitter for interactive emergency applications.

III. PROBLEM DEFINITION

A. Need for Restricting Queue-Fill Variations

There is a growing multitude of applications like IP telephony, interactive voice and video, etc. that have stringent delay requirements but may tolerate a little packet loss during times of congestion. Such applications are also jitter sensitive. The larger the instantaneous queue-fill variations, the higher is the average delay and delay variation.

RED was primarily designed as a proactive mechanism for congestion avoidance and providing fairness to TCP traffic, not as a filtering mechanism for non-adaptive traffic. But it is useful to be applied here to allow entry to a certain class of packets based on QoS requirements especially during times of severe congestion. One of the RED objectives and really the whole concept of queuing is to be able to absorb temporary bursts of input traffic and allow the router sufficient time for packet transmission. If the input traffic is bursty then it would be natural to expect some amount of variation in the queue fill. However, large queue-fill variations increase the possibility of Tail Drops. In a multi-level RED approach, occurrence of Tail Drops would also imply that higher priority traffic may be dropped without any preference over lower priority traffic. In order to prevent Tail Drops then, the queue limits would have to be unduly large which in turn would increase the queuing delay for applications. Also, oscillations in a RED queue are harmful because they cause periods of instantaneous queue fill drops to zero followed by periods of forced drops when the average queue size is greater than \( \text{max}_{ab} \). In a multi-level RED queue case, oscillations may cause poor discrimination among higher and lower priority traffic. This would be particularly true when dropping curves for two colors interact with each other at a certain load causing the instantaneous queue fill to constantly fluctuate between two different dropping curves [3]. This can cause higher priority traffic to be dropped before 100% of the lower priority traffic has been dropped, thus defeating one of the objectives stated earlier about having strict priorities for different traffic classes.

Thus, we can see the importance of restricting oscillations in the queue fill. The goal is then to achieve a state of restricted oscillations in the queue fill around a value between \( \text{min}_{ab} \) and \( \text{max}_{ab} \) so that the amplitude of oscillations of average queue size is significantly smaller and instantaneous queue fill always remains greater than zero but smaller than the maximum queue limit when average input load exceeds capacity. Thus we seek to simultaneously control average delay and delay variation by giving equal weight to both.

B. Effect of RED Dropping Function on Queue-Fill Variations

The slope of the loss probability curve is one of the important factors contributing towards oscillations observed in the queue fill. At any particular load, the slope determines how well and quickly the probability values adapt themselves in response to the increasing or decreasing queue fill. This in turn has an effect on how sooner or later the required number of packets is dropped and how the queue build-up and draining takes place. This variation decides how large or restricted the queue fill oscillations will be. The average delay experienced by the packets depends on the average queue fill. So, depending on the amount of variation, the average fill is correspondingly pushed up or down thus increasing or decreasing the average delay.

It is stated in [2] that for the linear RED function and TCP traffic, oscillations are discouraged when the packet marking probability changes fairly slowly with the average queue size. However, RED experiments with several different kinds of traffic show that this is not always the case. In fact, [3] shows how unresponsive ON-OFF type VBR traffic can produce very strong oscillations in the queue at even a moderate burstiness level. A lot of research work on RED has talked about RED parameter sensitivity which does not permit it to support a wide range of loads without causing strong oscillations in the queue fill or link under-utilization. This problem in itself has been attributed to the linearity of the RED dropping function which does not permit faster adjustment of dropping probability values in response to very low or high loads [4]. The optimality of the RED dropping function has not been studied especially in the context of its contribution to oscillations in the instantaneous queue fill and average delay. Hence performance of other dropping functions is studied.

Here we intend to study the performance of three other dropping functions namely concave, convex and step as compared to the linear function in terms of the average delay and instantaneous queue-fill variation produced. The non-linear functions used here are quadratic adaptations of the existing linear function as shown in the Figure 1.

As will be shown in the simulation section, the non-linear
concave function is seen to give better performance than the existing linear function in terms of restricting instantaneous queue-fill variations and thus producing lower jitter and average delay.

C. Main Contributions of this Work:

- We have used MAMT RED scheme as the simplest mechanism to support special requirements of emergency services applications.

- Most of the previous work done with RED has been with respect to TCP traffic or interaction of TCP traffic with UDP traffic. However, we intend to study UDP traffic without the TCP response mechanism affecting the analysis. Moreover the MAMT approach used here proposes having homogeneous traffic types in each AF queue. So it is necessary to study the behavior of the dropping functions with respect to non-adaptive traffic.

- Another reason for considering UDP traffic for this study is because the current focus of emergency organizations is mainly on interactive applications that are delay and jitter sensitive.

- The MAMT AQM scheme is used mainly as a filter to drop packets based on QoS requirements.

- [4] proposes the use of non-linear convex function with TCP traffic in order to avoid the linear function drawbacks mentioned earlier. But the work does not deal with the issues of queue-fill variations or multi-level RED parameter configurations. The particular flat slope of the convex function in the initial part works well with the TCP feedback mechanism and helps in avoiding the drawbacks. However, the same is not true in case of unresponsive UDP traffic.

- Dynamic mechanisms employing control theory concepts, feedback mechanisms, etc. produce better results in terms of throughput when used in conjunction with the TCP feedback mechanism. But the same is not observed with unresponsive, interactive UDP type of applications where delay and jitter carries more importance.

- This work extends our previous work in [3] by taking the next logical step of considering the performance of non-linear dropping functions in controlling the queue-fill variations and average delay while maintaining configurations for other control parameters.

IV. SIMULATION RESULTS

This section presents the simulation results for the performance of the 4 dropping functions discussed above particularly with respect to the average delay and the variation produced in the instantaneous queue-fill. Experiments were performed with both single-level and multi-level RED using various types of representative UDP sources like CBR to bursty ON-OFF type VBR. However, due to space constraints, we present here only the MAMT RED results with VBR traffic. Note that the RED mechanism drops only newly arriving packets and hence all the delay values shown in our results are counted for packets already accepted in the queue. The simulation model consisting of five sources, 2 edge routers and a core router. The bottle-neck link is 10 Mbps. The size of each AF queue is assumed to be 80 packets. Exponential ON-OFF sources have average ON times of 4 ms and average OFF times of 1 ms. The following sub-sections present our findings.

A. Supporting a Wide Range of Loads over the Same RED Parameter Setting

Previous research work has shown that RED parameter \((\text{min}_h, \text{max}_h, \text{max}_p)\) sensitivity tends to restrict the range of loads that can be supported over the same parameter setting. Unknown load scenarios tend to produce large oscillations and hence forced drops as also periods of link under-utilization. During time of congestion, especially during emergency situations the network can be stressed and loads as high as three times the normal may be experienced. Thus we require that the same RED parameter setting be able to handle certain amounts of fluctuations in the load rather than having to configure them every time traffic dynamics change. A dynamic RED approach is not considered here due to the
complexity and stability issues of such algorithms that hinder actual deployment. Hence we try to make the best possible use of static and simple algorithms as proposed in [3]. It was shown in [3] that UDP traffic is not overly sensitive to RED parameter settings as is TCP traffic. With a sufficiently large queue size, it is possible to support a wide range of loads over the same RED parameter setting and the existing linear dropping function. Hence, the objective here is to study which of the dropping functions gives the best performance in terms of average delay, the instantaneous queue-fill variation and hence the overall service provided to interactive emergency traffic.

For each of the dropping functions, we ran a set of simulations by increasing the average VBR load from 11 Mbps to 25 Mbps. Figures 2 and 3 show the average delay and queue-fill variance graphs. We can observe that over the expected load range considered here, the concave function gives the best performance. If the dropping curve configurations are done properly by keeping appropriate space between the curves [3], we find that no higher priority packets are dropped before dropping 100% of the lower priority packets, regardless of the dropping function. Also no tail drops occur.

![Figure 2: Effect of Dropping Functions on Average Delay](image)

![Figure 3: Effect of Dropping Functions on Instantaneous Queue-Fill Variation](image)

The reason why the concave function performs best can be explained as follows. As explained earlier, if the inter-dropping curve spacing is configured properly so as to minimize the curve interactions, the slope of the dropping curve is one of the major factors controlling the variance and the average delay observed at a particular load. The linearity of the existing RED dropping function results in a slower adaptation of the dropping probability values to the changing average queue-fill, especially at high loads. This results in a higher delay and variation as compared to a steeper sloped non-linear curve. The concave curve has a steeper slope in the initial part as compared to the linear curve. So any increase or decrease in the average queue fill is quickly counteracted by corresponding quick increases or decreases in the probability values in a non-linear fashion. This causes the required number of packets to be dropped in proper time providing appropriate adjustment to queue-fill changes. Thus the steeper portion of the concave curve helps in restricting large variations in the queue fill and also keeps the average fill low. While it is true that one could define a steeper linear curve, this also has limitations which are discussed later.

On the other hand, a convex curve has a flat slope in the initial part as compared to the linear curve. Thus when the average queue-fill increases or decreases, the corresponding probability value change is not so quick. When the average fill starts exceeding $\min_{\theta}$, it takes some time before the probability value actually reaches a value where sufficient number of packets would be dropped, so queue fill can be brought down to contain oscillations. This slow change consequently tends to push the average fill to a higher value as compared to a linear curve and gives greater possibility for larger variations of queue fill.

The step function is an extreme case of steep slope. Here we have only one threshold that is placed at $\max_{\theta}$. The average queue fill is allowed to reach till $\max_{\theta}$ before any single packet could be dropped. We note that the instantaneous queue fill values tend to oscillate very widely on either side of the average queue fill. They can dip down all the way to zero queue fill and also go much beyond the $\max_{\theta}$ value before the weighted average queue fill can actually exceed $\max_{\theta}$ for the packets to be dropped. Consequently, packets are not dropped soon enough resulting in high queue-fill variation. With multiple levels it becomes hard to contain oscillations within a small space around the $\max_{\theta}$ of a particular DP. Bringing the $\max_{\theta}$ of two DPs too close to one another can cause very large oscillations due to interactions between the adjacent DPs. In addition, this may actually result in drops of higher priority packets before 100 % of lower priority packets have been dropped, thus defeating one of our goals. Thus we see that in order to have a reasonable average delay using this function, the $\max_{\theta}$ value needs to be set carefully.

### B. Analysis of Linear and Non-linear Functions

From Figure 1 we can see that both the convex and concave functions exhibit a significant reversal in slope after a point i.e. concave curve becomes much flatter and convex curve becomes much steeper. So we expect their performance to reverse at high loads. However, over the entire wide range of loads considered in the previous section, the concave function gives the best performance in terms of both average delay and variance. In this section, we intend to explain this result and also observe at what point the reversal of performance takes place.

For this simulation experiment, we consider very high
average loads from 28 Mbps to 50 Mbps, once again, for a 10 Mbps link. We use the linear, concave and convex functions for these simulations. Figure 4 shows the variance plot obtained for multi-level RED. Even at such high loads as 50 Mbps, the concave function provides lower variance and the convex function provides more variance as compared to the linear function.

![Graph showing the variance of instantaneous queue fill against average VBR load for different functions.](image)

**Figure 4: Effect of Dropping Functions on Instantaneous Queue-Fill Variation at very High Loads**

In order to explain this result, we have plotted the actual calculated probability values during the simulation for each of the functions as shown in Figure 1. If we compare the concave and linear plots, we find that at a load of 50 Mbps, the ideal dropping percentage is still along the steeper portion of the concave curve in comparison with the linear curve. It naturally follows that the probability value adaptation in response to changing average queue fill is much quicker than the linear function as was explained earlier. Also we can observe from the plots that the concave curve maintains a steeper slope compared to the linear curve almost until it nears maxth at which point it starts to flatten out. We can make a similar observation for the convex curve. This point on the curve where we actually begin to see a significant change in slope of the concave or the convex curve corresponds to a very high dropping percentage (greater than 80%), which in turn means a very high load.

Traffic entering at the edge router is subjected to traffic conditioning, which ensures that traffic confines to the traffic profiles as specified in the Traffic Conditioning Agreement. Moreover, we expect traffic engineering principles to be employed in the network so as to properly utilize the network resources. So a single path in the network cannot expect unreasonably high loads where 80% or even 60% of the load really needs to be discarded. Even in disaster or emergency situations, we expect the load in the network to go up to approximately 3 times the normal load which would be still imply a lower dropping percentage than 60%. Thus, loads as high as mentioned above are unreasonable to expect and the presence of such loads would be an indicator of bad network design and traffic engineering. So we can say that a concave function can provide a substantial advantage over a linear function in providing better delay and variance and it works well over all expected loads. Also the convex function is not useful over this range.

### C. Effects of Steep Linear Curve

In this section we study the effect of making the linear curve steeper and compare the average delay and variance results with those of the concave function with original parameter settings. The objective here is to see at what point the linear curve may give similar performance as the concave curve. This would happen when the linear curve is made steep enough that its slope equals that of the steeper portion of the concave curve. We have seen that a step function, which is an extreme case of steep slope, does not perform well with MAMT RED. So another objective is to observe what happens during the transition as we make the linear curve steeper till it eventually becomes a step function.

For this simulation experiment, we consider an average VBR load of 20 Mbps. We start with an initial parameter setting (minth, maxth and maxp) for all the DPs and note the average delay and variance for concave and linear functions. We know from previous results that the concave function will present better results. Now we reduce the maxth setting for all DPs for the linear curve while keeping concave settings the same. Thus, we have made the linear curve steeper and now we note the variance and average delay results. Table 1 shows the results obtained for MAMT RED case.

<table>
<thead>
<tr>
<th>RED</th>
<th>YELLOW</th>
<th>GREEN</th>
<th>BLUE</th>
<th>LIN Avg</th>
<th>CONC Avg</th>
<th>LIN Var</th>
<th>CONC Var</th>
</tr>
</thead>
<tbody>
<tr>
<td>7-12</td>
<td>20-30</td>
<td>35-45</td>
<td>50-60</td>
<td>34</td>
<td>129</td>
<td>115</td>
<td></td>
</tr>
<tr>
<td>7-12</td>
<td>20-25</td>
<td>35-40</td>
<td>50-55</td>
<td>34</td>
<td>115</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7-9</td>
<td>20-23</td>
<td>35-38</td>
<td>50-55</td>
<td>34</td>
<td>127</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Table 1. Effect on Variance with a Steep Linear Curve**

(All units in number of packets)

We observe that as we make the linear curve steeper by decreasing the maxth setting, the average delay as well as the variance reduces. This trend continues until the slopes of the concave and linear curve are almost the same. From this point, if the linear curve is made even steeper, the average delay and variance again starts to increase. Also we observe that the lower DP packets are dropped before 100% of the higher DP packets have been dropped. Thus, the throughput of the lower DP gets affected.

This trend can be explained as follows. As we make the linear curves of each DP steeper by reducing the maxth setting, we effectively provide a harsh treatment to the different color packets. Also the inter-dropping curve spacing increases. As mentioned in [3], this space needs to be properly configured because too large a gap can produce large oscillations in the queue when multiple curves start interacting at a particular load.

If both minth and maxth settings are pushed down to counter the oscillations and get better variance and average delay, we
tend to push the queue more towards zero queue fill which implies link under-utilization. Link under-utilization needs to be avoided because the buffer is not utilized when it could have absorbed burst in traffic. Configuring thresholds on a lower side may result in packet bursts not being accommodated when in fact sufficient space was available in the queue. This results in unnecessary packet drops and defeats the whole purpose of the queue which is to be able to absorb temporary packet bursts. On the other hand, bringing the curves closer in order to counter large queue-fill variations will result in higher priority packets being dropped before 100% of the lower priority packets have been dropped. Thus we maintain that with MAMT RED, making a linear curve steep can help achieve better performance only up to a certain point as the parameter configuration becomes more difficult. Also it would always be possible to find a concave curve that would fit in the changed settings that would always perform better than the linear curve in that setting.

D. Effects of the Characteristics of Input Traffic

In this section we study the effects of increase in burstiness of input traffic on variance of instantaneous queue fill for each of the dropping functions. We define traffic with a greater ON time as more bursty as compared to traffic with a lesser ON time. For all the simulations the ratio of ON time to OFF time is kept constant at 4:1. Figure 5 shows the queue-fill variance plot against the traffic burstiness for the four dropping functions.

![Figure 5: Effect of Dropping Functions on Instantaneous Queue-Fill Variation with Increasing Burstiness](image)

We observe that as the burstiness of input traffic increases, variation in the instantaneous queue fill increases and oscillations become stronger [3]. The concave function produces the least variance. The reason for this is the same as explained in part A of simulation results. Now, if burstiness of traffic goes on increasing, the large variations in queue fill can ultimately lead to Tail Drops. Thus from the results shown above, we can say that the concave dropping function helps in controlling variation produced by an increase in the burstiness of traffic in a much better way than the other dropping functions. Thus the probability of getting Tail Drops also reduces. Apart from that there is not much difference in the variance trend exhibited by the different dropping functions in response to increasing traffic burstiness.

V. CONCAVE VERSUS LINEAR CURVE SLOPES

In this section, we show the mathematical analysis for the concave dropping function that indicates the point after which the slope of this function becomes flatter as compared to the linear function. After this point, we would observe a reversal of performance for the concave function.

For the purpose of this analysis, we assume that \( \min_{th} \) is set at 0. From figure 1, the slope of the linear function can be written as:

\[
\frac{\max_p}{\max_{th}}
\]

Similarly slope of the concave function can be written as:

\[\frac{\max_p}{\max_{th}} \left( 1 - \frac{q_{avg}}{\max_{th}} \right) \left[ 1 - \frac{1}{\max_{th}} \right] \]

For the slope of the concave curve to be steeper in the initial part as compared to the linear curve,

\[\frac{\max_p}{\max_{th}} \left( 1 - \frac{q_{avg}}{\max_{th}} \right) \left[ 1 - \frac{1}{\max_{th}} \right] > \frac{\max_p}{\max_{th}} \]

Therefore, \( q_{avg} \geq \frac{1}{2} \max_{th} \)

Thus we observe that when the weighted queue average is halfway between \( \min_{th} \) and \( \max_{th} \) or less, there is a change in the slope of the concave curve as compared to the linear curve and its performance would reverse. Considering \( q_{avg} = \frac{1}{2} \left( \max_{th} \right) \), we can find the exact point on the concave dropping curve as:

\[\frac{\max_p}{\max_{th}} \left( 1 - \frac{q_{avg}}{\max_{th}} \right) \left[ 1 - \frac{1}{\max_{th}} \right] = 0.75 \max_p \]

If \( \max_p = 1 \), the point on the concave dropping curve where the slope change takes place is 0.75 which corresponds to a dropping percentage of 75%. This implies that the total load on the link would be greater than 40 Mbps. This analysis substantiates our simulation findings as explained in the simulation results section B of this paper. Since loads as high as 40 Mbps are unreasonable to expect along any single path in the network, we can say that the concave function performs the best over all expected load ranges.

VI. CONCLUSION

We have focused on the application of multi-level RED to provide the necessary QoS to emergency service applications. Of the important QoS metrics, the RED mechanism itself handles the issue of achieving the maximum possible throughput at a particular load by its proactive approach. We handle the issue of loss of packets by employing a multi-class approach with a different set of RED parameters for each class and thus prioritizing loss as per requirements. The next logical step is to improve performance for parameters like average delay and delay variation as has been considered in this work. We have focused on the linear RED dropping function as one of the contributing factors towards variations in the instantaneous queue-fill. From the exhaustive simulation study
conducted using the four different dropping functions (linear, convex, concave and step), we find that all of them offer the same performance in terms of protection of emergency traffic and prevention of tail drops if the MAMT RED parameter settings ($min_{th}$, $max_{th}$, $max_{p}$) are done as per recommendations made in [3].

Importantly we observe that over all expected ranges of loads considered here, the concave dropping function gives the best performance in terms of average delay as well as the variance observed in the instantaneous queue fill. The presence of a combination of steep and flat slopes in the same function does not restrict the range of loads that this function can support because the curve tends to flatten out around the $max_{th}$ level which corresponds to extremely high loads. The performance of this function is consistent over both single level as well as multi-level RED queues. Since, it produces the least variation in the instantaneous queue fill, it gives the least delay variation which is important for interactive or playback type of applications. We also maintain that the concave function is a definite candidate for further analytical study and actual implementation. A multi-level RED mechanism employing a concave dropping function can definitely be used for AF queues carrying interactive UDP type of traffic to achieve desired QoS.

VII. REFERENCES


