2005 American Control Conference
June 8-10, 2005. Portland, OR, USA

Stability Analysis and Filter Design for LRED Algorithm

Chengnian Long, Bin Zhao, Xinping Guan, and Bo Yang

Abstract— **Loss Ratio based Random Early Detection (LRED) scheme has been recently proposed for improving the response and robustness of Active Queue Management (AQM) in Internet routers. It measures the latest packet loss ratio periodically, and adjusts its packet loss probability around the measured value. Our goal is to identify the key factors that drive the effectiveness of this design. We present a stability analysis for the TCP/LRED delayed feedback system. Our analysis suggests that the stability of LRED scheme can be decoupled from TCP load N. Thus LRED scales well with a wide range of traffic conditions. Furthermore, based on the stability analysis, we propose a scalable algorithm both for TCP load N and for link capacity C. In the end, we study the packet loss ratio estimate scheme, which is a key factor of LRED, and present a full average loss (FAL) filter algorithm. Simulations validate that the proposed filter algorithm responds quickly to the change of traffic conditons and achieves a steady estimate in the presence of extremely short term flows.**

I. INTRODUCTION

Recent research efforts to design better Internet transport protocols combined with scalable Active Queue Management (AQM) [2] have led to significant advances in congestion control. The basic philosophy of AQM is to trigger packet dropping (or marking, if Explicit Congestion Notification (ECN) [23] is enabled) in advance while the onset of congestion is perceived. Thus AQM can mitigate the problems of long queuing delay and "TCP synchronization" in drop-tail queue management. AQM in Internet routers can improve application goodput and response time by detecting congestion early and improving fairness among flows. Most of promising approaches model AQM as a feedback controller on a time-delayed response system and apply control theory to design efficient controller for TCP traffic. Thus the explicit robust stability condition can be derived for the time-delayed feedback system [9],[15]. Unfortunately, those promising AQMs still face a critical deployment challenge since the network is highly dynamic. The design guidelines in existing approaches, such as PI [8],[9] and REM [1], are from the "worst-case" network parameters for robust stability. The conservative design leads to the slowly responsive and degraded performance, i.e., persistent buffer overflow or buffer emptiness. As a solution, the adaptive designs using real-time estimates of network conditions are emerging, such as Adaptive Virtual Queue (AVQ) [12], Adaptive RED (ARED) [5],[6], Stabilized RED (SRED) [20], self-configuring PI [27],[28], and Loss Ratio based RED (LRED) [26].

Due to the combination of distributed feedback and time delays, the stability of TCP/AQM interconnection system is a main concern in the study of AQM. Significant progress has been made in the past few years [9],[11],[14],[16],[18],[21],[22]. For rate-based congestion control framework proposed by Kelly [11], a local stability condition was given for homogeneous RTTs in [10] and, was generalized to handle heterogeneous RTTs in [18]. Vinnicombe [25] proposed an elegant lemma which relates the eigenloci of the product of a Hermitian matrix and a diagonal matrix to the product of the spectral radius of the Hermitian matrix and the convex hull of the entries of the diagonal matrix. Using this lemma and the generalized Nyquist criterion [4], he showed the correctness of the continuous-time analogue of the conjecture of Johari and Tan [10]. This result gives an interesting insight for distributed congestion-control algorithms, namely network stability can be guaranteed by simple, decentralized conditions on end sources and links. Thus each end system only needs knowledge of its own round-trip delay. Recently, Tian and Yang [24] proved out a more general stability criterion based on the clockwise property of parameterized curves and the general Nyquist criterion of stability. Further progress has been made in the research on the stability of primal-dual algorithm, where the overall network dynamics comprised of two decentralized dynamic systems interconnected by routing. These two dynamics are the TCP-like source algorithm and the AQM scheme, respectively. In [13], a stability condition was derived for AVQ with source dynamics.

However, the stability proofs for the interconnection dynamics between TCP dynamic model [19] and queue-based AQM are limited [7]. The stability analysis for RED and PI AQMs have been made in [8], [9] for a single congested link with homogeneous sources. In [17], the stability was generalized to the single congested link with heterogeneous sources. Recently, Han et al. [7] developed a local stability condition for queue-based TCP/AQM networks with arbitrary topologies and heterogeneous delays. In the paper, we focus on the stability analysis for LRED scheme based on the well-known TCP Reno model [19]. The motivation of our work is to provide an insight for its graceful properties. Both the single congested link network and the general topology network are considered. Interestingly, we present a more simple stability criterion for single congested link with heterogeneous RTTs than the original stability condition in [26]. Our results reveal that stability of TCP/LRED interconnection system is decoupled from the TCP load N. The analytical result provides a solid proof for the simulation finds that the LRED scheme is scalable for a wide range of traffic conditions. Furthermore, a scalable algorithm both for TCP load and link capacity can be

All authors are with the Department of Electrical Engineering, Yanshan University, Qinhuangdao 066004, P. R. China. dylcn@ysu.edu.cn

0-7803-9098-9/05/$25.00 ©2005 AACC

formulated from the stability condition. Because the update rule of LRED is to dynamic adjust around the dynamic equilibrium, we call our improved scheme as Dynamic Equilibrium Controller (DEC) throughout this paper.

Another important factor of LRED is the packet loss ratio estimate. There is a clear trade-off between the responsiveness for sudden change in network condition and steady performance in the presence of noise. LRED measures the latest packet loss information and calculates it by an exponential moving average (EWMA) filter method. To track the current equilibrium quickly, the weight factor for history value is set to be small. However, the simple method may result in oscillation in the presence of short-lived TCP flows. To overcome the deficiency, we present a Full Average Loss (FAL) filter method.

Through an extensive simulation study, we evaluate DEC and compare it to alternative adaptive AQM mechanisms, including the ARED, SRED, AVQ, and LRED for a wide range of network and traffic conditions. Our simulation validates the robust stability and quick response of DEC, showing that DEC efficiently handles network congestion in all the tested traffic conditions.

Notation: $\lambda(X), \sigma_{\min}(X)$ and $\sigma_{\max}(X)$ denote the eigenvalues, smallest and largest singular values of square matrix X, respectively. $Co\{x_i\}$ denotes the convex hull of the set of points $\{x_1, x_2, \ldots\}$. $diag\{x_i\}$ denotes the matrix with the elements $x_1, x_2, \ldots$ on the leading diagonal and zeros elsewhere. Finally, $disc(r)$ denotes the closed disk $\{s \in \mathbb{C} : |s| \le r\}$.

II. LRED ALGORITHM

First, we describe the design principle of LRED scheme and present an initial stability condition for LRED in [26]. The packet loss ratio and queue length are employed to estimate the degree of link congestion in LRED. Thus the packet drop probability is calculated as follows

$$p(k) = \overline{l(k)} + \beta\sqrt{\overline{l(k)}}\left(b(k) - \hat{b}\right) \tag{1}$$

where $\overline{l(k)}$ is the measured packet loss ratio at period k, β is the gain parameter and $\hat{b}$ is the desired queue length. LRED calculates the loss ratio periodically for every small period. Let $l(k)$ be the packet loss ratio during the latest M measurement periods, then the measured packet loss ratio can be calculated as the following EWMA filter

$$\overline{l(k)} = mw * \overline{l(k-1)} + (1 - mw) * l(k) \tag{2}$$

where mw is the measured weight factor, which is set to be small in order to respond fast to the change. In [26], a stability condition for LRED is given with the same spirit of [12]. For comparison, we present the Theorem 4 of [26] as the follows.

Lemma 1: Given network parameters $\left(\hat{N}, \hat{C}, \hat{\tau}\right)$, and assume that $\hat{\beta}$ satisfies

$$\hat{\tau}\omega + \arctan\left(\frac{\omega}{K_{11}}\right) = \frac{\pi}{2}, \hat{\beta} > 0$$

where ω can be calculated as

$$\omega = \sqrt{0.5\left(\sqrt{K_{11}^4 + 4K_c^2 H_c^2} - K_{11}^2\right)}$$

and $K_{11} = \frac{2\hat{N}}{\hat{\tau}^2\hat{C}}, K_c = \frac{\hat{C}^2}{\eta\hat{N}}$, $H_c = \hat{\beta}\sqrt{p_0}$. If

$$\beta < \tilde{\beta} = \min\left(\hat{\beta}, \left(\sqrt{2\eta}\left(2\hat{N}\right)^4\right) \Big/ \left(\hat{C}^3\hat{\tau}^3\right)\right) \tag{3}$$

the TCP/LRED system remains stable for every values of $N \ge \hat{N}$ and $\tau \le \hat{\tau}$.

III. STABILITY ANALYSIS FOR TCP/DEC INTERCONNECTION SYSTEM

We adopt the system model as given in [19]. Consider a network with a set $\mathcal{L}$ of links and a set $\mathcal{I}$ of users. Let c_l denote the finite capacity of link $l \in \mathcal{L}$. Each user i has a fixed route $\mathcal{L}_i$, which is a non-empty subset of $\mathcal{L}$. The interconnections are described by the matrix $\mathcal{R}(s) = [R_{li}(s)]$ where

$$R_{li}(s) = \begin{cases} e^{-s\tau_{li}^f} & \text{if source } i \text{ traverses link } l \\ 0 & \text{otherwise} \end{cases}$$

and τ_{li}^f denotes the forward delay from source i to link l. The so-called routing matrix is $\mathcal{R}(0)$, which is assumed to be full row rank. Letting τ_{li}^b denote the return delay from link l to source i, we define the round-trip time (RTT) of source i by $\tau_i = \tau_{li}^f + \tau_{li}^b$. Associated with each link l is its dropping/marking probability $p_l(t)$ at time t, and with each source i its window $w_i(t)$ at time t. TCP Reno prescribes how congested window $w_i(t)$ is adjusted and AQM prescribes how $p_l(t)$ is updated. Together they form a delayed feedback system and can be interpreted as carrying out a distributed primal-dual algorithm to solve a welfare maximization problem over the Internet [16].

Using the TCP fluid model, the congestion window for the ith source is approximated by the nonlinear differential equation

$$\dot{w}_i(t) = \frac{1}{\hat{\tau}_i} - \frac{w_i(t)^2}{\eta\hat{\tau}_i} q_i(t) \tag{4}$$

where $q_i(t)$ is the end-to-end probability

$$q_i(t) = \sum_l [R(0)]_{il}^T p_l\left(t - \tau_{li}^b\right) \tag{5}$$

$x_i(t)$ is the source i's rate at time t

$$x_i(t) = \frac{w_i(t)}{\hat{\tau}_i} \tag{6}$$

and $\hat{\tau}_i$ is the equilibrium for RTT. We model the lth congested link by

$$\dot{b}_l(t) = \sum_i [R_{li}(0)] \frac{w_i\left(t - \tau_{li}^f\right)}{\tau_i(t)} - c_l \tag{7}$$

where

$$\tau_i(t) = d_i + \sum_j [R(0)]_{ij}^T \frac{b_j(t)}{c_j}$$

and d_i is the propagation delay of source i. We use the equilibrium $\hat{\tau}_i$ to model the congestion window dynamics since the nonlinear differential equation (4) is a mean model for nonlinear stochastic differential equation in [19]. In queue dynamics (7), we consider the time-varying round-trip time incurred by queuing. But the delay is neglected in the $\tau_i(t)$ as in [7]. To model DEC, we assume that the measured packet loss ratio with FAL filter can approximate the recent equilibrium packet probability $\hat{p}_l$.

$$p_l(t) = \hat{p}_l + k_l\left(b_l(t) - \hat{b}_l\right) \tag{8}$$

Let $\hat{w}_i$, $\hat{p}_l$, $\hat{b}_l$ be the equilibrium congestion window, equilibrium packet loss ratio and the desired queue length respectively. It is apparent that there exist an unique equilibrium point for TCP/DEC interconnection system. Furthermore, we have the following relationship at the equilibrium point

$$\begin{cases} \hat{q} = R(0)^T \hat{p},\ R(0)\hat{x} = c,\ \hat{w}_i^2\hat{q}_i = \eta \\ \hat{\tau}_i = d_i + \sum_k [R(0)]^T_{ik}\frac{\hat{b}_k}{c_k},\ \hat{x}_i = \frac{\hat{w}_i}{\hat{\tau}_i} \end{cases} \tag{9}$$

We linearize the TCP/DEC Eqs. (4)-(8) to study its stability around equilibrium. Denote $\delta w_i(t) = w_i(t) - \hat{w}_i$, $\delta p_l(t) = p_l(t) - \hat{p}_l$. Then Linearizing Eq. (4) yields

$$\delta\dot{w}_i(t) = -\frac{1}{\hat{\tau}_i\hat{q}_i}\sum_l R_{li}\delta p_l\left(t - \tau^b_{li}\right) - \frac{\hat{q}_i\hat{w}_i}{\hat{\tau}_i}\delta w_i(t) \tag{10}$$

Around the equilibrium, the link dynamics can be expressed as

$$\begin{aligned} \delta\dot{b}_l(t) &= \sum_i [R_{li}(0)]\frac{\delta w_i\left(t - \tau^f_{li}\right)}{\hat{\tau}_i} \\ &\quad - \sum_i [R_{li}(0)]\frac{\hat{w}_i}{\hat{\tau}_i^2}\sum_j [R(0)]^T_{ij}\frac{\delta b_j}{c_j} \end{aligned} \tag{11}$$

and packet loss probability dynamics is

$$\delta p_l(t) = \frac{1}{c_l}d_l\delta q_l(t) \tag{12}$$

where $d_l = k_l c_l$. Let $\delta w_i(s)$, $\delta b_l(s)$ and $\delta p_l(s)$ denote the Laplace transforms of $\delta w_i(t)$, $\delta b_l(t)$ and $\delta p_l(t)$ respectively. Define $\delta w(s) = \left[\delta w_1(s), \delta w_2(s), \ldots \delta w_{|\mathcal{I}|}(s)\right]^T$, $\delta b(s) = \left[\delta b_1(s), \delta b_2(s), \ldots \delta b_{|\mathcal{L}|}(s)\right]^T$ and $\delta p(s) = \left[\delta p_1(s), \delta p_2(s), \ldots \delta p_{|\mathcal{L}|}(s)\right]^T$. It is straightforward to see that the Laplace transform of linearized version of (10), (11) and (12) can be rewritten as

$$\begin{cases} \delta w(s) = -F(s)R(-s)^T\delta p(s) \\ \delta b(s) = (sI + \Omega)^{-1}R(s)\hat{T}^{-1}\delta p(s) \\ \delta p(s) = C^{-1}D(s)\delta b(s) \end{cases} \tag{13}$$

where $F(s) = diag\{f_i(s)\}$,

$$\begin{aligned} f_i(s) &= \frac{e^{-s\hat{\tau}_i}}{\left(s + \frac{2}{\hat{w}_i\hat{\tau}_i}\right)\hat{\tau}_i\hat{q}_i} \\ \Omega &= R(0)\hat{W}\hat{T}^{-2}R(0)^T C^{-1} \end{aligned}$$

the diagonal matrices are $\hat{T} = diag\{\hat{\tau}_i\}$, $\hat{W} = diag\{\hat{w}_i\}$, $C = diag\{c_l\}$, $D = diag\{d_l\}$. Thus, the return ratio for system (13) is

$$L(s) = F(s)R(-s)^T C^{-1}D(sI + \Omega)^{-1}R(s)\hat{T}^{-1} \tag{14}$$

Apparent, $L(s)$ is stable. To proof the stability of the TCP/DEC system (13), we should show that the eigenvalues of $L(j\omega)$ do not intersect $(-\infty, 1]$ for all $\omega \geq 0$ based on the Generalized Nyquist Theorem [4].

A. Single Congested Link with N Heterogenous Sources

For simplicity, we first derive the asymptotic stability condition for TCP/DEC system in the case of a single congested link with N heterogeneous sources. Dropping the link subscript l, the return ratio (14) is

$$\begin{aligned} L(s) &= diag\left\{\frac{e^{-s\hat{\tau}_i}}{\left(s + \frac{2}{\hat{w}_i\hat{\tau}_i}\right)\hat{\tau}_i\hat{p}}\right\}R(-s)^T \\ &\quad \times\frac{k}{s + \sum_{i=1}^N \frac{\hat{x}_i}{c\hat{\tau}_i}}R(s)\hat{T}^{-1} \end{aligned}$$

where $R(s) = \left[e^{-\tau^f_{11}s}, e^{-\tau^f_{12}s}, \ldots e^{-\tau^f_{1N}s}\right]$. To state this result, we introduce some addition notations. Let

$$z_i(j\omega) = \frac{e^{-j\omega\hat{\tau}_i}}{\left(j\omega\hat{\tau}_i + \frac{2}{\hat{w}_i}\right)\left(j\omega + \sum_{i=1}^N \frac{\hat{x}_i}{c\hat{\tau}_i}\right)}$$

and

$$H(j\omega) = \frac{kc}{\hat{w}_i\hat{p}}\sum_{i=1}^N \frac{\hat{x}_i}{c}z_i(j\omega) \tag{15}$$

Theorem 1: Consider the linearized network consisting of one congested link with N heterogenous sources. Then, the closed-loop system described by (13) is locally stable if the proportional gain k satisfies

$$k \leq \hat{p}\frac{\sum_{i=1}^N \frac{\hat{x}_i}{c}\frac{1}{\hat{\tau}_i}}{\sum_{i=1}^N \frac{1}{\hat{\tau}_i}} \tag{16}$$

The proof is given in the Appendix. Specially, the link dynamics of LRED AQM is

$$k = \beta\sqrt{\hat{p}}$$

Based on Theorem 1, we have the following result.

Theorem 2: Consider the linearized network consisting of one congested link with N heterogenous (homogenous) sources. Let the LRED parameter be chosen as: $\beta \in (0, \beta_{\max}]$ and assume that $\beta_{\max}$ satisfies

$$\beta_{\max} \leq \frac{\sqrt{\eta}}{c_{\max}\hat{\tau}_{\max}} \tag{17}$$

then the TCP/LRED interconnection system is locally stable for $c \leq c_{\max}, \hat{\tau}_i \leq \hat{\tau}_{\max}$.

The proof is simple by subtituting $\sqrt{\hat{p}} = \frac{\sqrt{\eta}}{\hat{w}} = \frac{\sqrt{\eta}}{c}\sum_{i=1}^N \frac{1}{\hat{\tau}_i}$ to equation (16).

Remark 1: Compared to the stability condition of Lemma 1, our result reveals that the stability of LRED is decoupled from the TCP load N. This result further validates the effectiveness of LRED under a wide range of traffic conditions.

So far, we have presented an explanation that LRED has a scalable property for TCP load. If the link capacity for best effort traffic can be measured [28], we can derive the following DEC controller to scale for link capacity

$$k_l = \frac{\gamma_l \sqrt{\hat{p}_l}}{c_l} \tag{18}$$

Theorem 3: Consider the linearized network consisting of one congested link with N heterogenous (homogenous) sources. Let the DEC controller parameter be chosen as: $\gamma \in (0, \gamma_{\max}]$ and assume that $\gamma_{\max}$ satisfies

$$\gamma_{\max} \leq \frac{\sqrt{\eta}}{\hat{\tau}_{\max}} \tag{19}$$

then the TCP/DEC interconnection system is locally stable for $\hat{\tau}_i \leq \hat{\tau}_{\max}$.

B. General Routing Structure

In the section, we state a stability result for the general routing structure case. The return ratio (14) is equivalent to

$$L = F(s) R(-s)^T C^{-\frac{1}{2}} D (VG(s) V)^{-1} C^{-\frac{1}{2}} R(s) \hat{T}^{-1}$$

where V is unitary and

$$G(s) = diag\{g_l(s)\}, g_l(s) = s + \lambda_l$$

λ_l is an eigenvalue of $C^{-\frac{1}{2}} \Omega C^{-\frac{1}{2}}$. We also have

$$\lambda_{\min} \geq \frac{\min\left\{\hat{w}_i \hat{\tau}_i^{-2}\right\} \sigma_{\min}^2 [R(0)]}{c_{\max}} \tag{20}$$

Let

$$\begin{aligned} \hat{f}_i(s) &= \frac{m_i}{\hat{w}_i} f_i(s), \hat{F}(s) = diag\left\{\hat{f}_i(s)\right\} \\ \hat{R}(s) &= C^{-\frac{1}{2}} R(s) \left(\hat{W} M^{-1} \hat{T}^{-1}\right)^{\frac{1}{2}} \end{aligned}$$

where m_i is the number of congested link of source i and $M = diag\{m_i\}$. Using matrix commutation, we have

$$\lambda(L) = \lambda\left(\hat{R}(j\omega) \hat{F}(j\omega) \hat{R}(j\omega)^H D \left(VG(j\omega) V^H\right)^{-1}\right)$$

Now we can present local stability analysis for the system (13) using Generalize Nyquist Stability Criterion [4].

Theorem 4: Consider a multiple bottleneck TCP/DEC network described by (4-8) and (18). Let the DEC controller parameters be chosen as: $\gamma_l \in (0, \gamma_{\max}]$. Then, this interconnection system is locally stable around the equilibrium (9) if

$$\gamma_{\max} \leq \frac{\eta \sigma_{\min}^2 [R(0)]}{\max\left(m_i \hat{w}_i \hat{\tau}_i^2\right) c_{\max} \max\left(\sqrt{\hat{p}_l}\right)}, \quad \forall l, i : i \text{ uses } l \tag{21}$$

The proof is similar to that of [7] and is thus omitted here to conserve space.

IV. FULL AVERAGE LOSS FILTER

The calculation of the packet loss ratio is one of the most critical parts of DEC controller design. There is a clear trade-off between measuring the packet loss ratio over a short period of time and being able to respond rapidly to changes in the traffic conditions, versus measuring over a longer period of time and getting a signal that is much less noisy. LRED calculates the packet loss ratio by measuring the last four sample intervals and conducting an EWMA filter algorithm. To respond rapidly, the weight factor mw is set to a small value. The simple filter algorithm can track the current equilibrium quickly. However, the long-term steady performance may be degenerated with the short-lived TCP flow and bursty web flash flow. As a solution, we propose the FAL filter algorithm to achieve fast responsiveness and long-term steady performance simultaneously.

The FAL filter algorithm is shown in Fig. 1. DEC calculates the loss ratio periodically for each small sample period T_s, which is same as in the LRED. The FAL filter method differs from the EWMA in several ways. First, the FAL introduces a concept of computing unit to count the short-lived packet loss ratio. When a new sample interval is added, a new computing unit is formulated using the latest m sample intervals and all of the following computing units are correspondingly shifted down one. Let p_1 be the packet loss ratio during the latest m sample intervals, i.e. the ratio of dropped packets to total arrival packets during the latest m measurement periods. Second, the FAL filter method takes a weighted average of the last n computing units, with equal weights for the most recent $n/2$ computing units and smaller weights for older computing units. Thus the average packet loss ratio is calculated as follows:

$$\hat{p} = \frac{\sum_{i=1}^{n} w_i p_i}{\sum_{i=1}^{n} w_i}$$

for weights w_i

$$w_i = \begin{cases} 1, & 1 \leq i \leq n/2 \\ 1 - \frac{i - n/2}{n/2+1}, & n/2 < i \leq n \end{cases}$$

Because the weighted filter method averages a number of computing units, the naive FAL method measures reasonably the steady packet loss ratio. However, the method is slow to respond to a sudden increase/decrease in the loss ratio. For this reason we deploy *oblivious mechanism* as a component of the FAL method, to allow a more timely response to a sudden change in traffic conditions. *Oblivious mechanism* is used by the router after the identification of a particularly large or small packet loss ratio since the last sample interval. The details of the *oblivious mechanism* are as follows: If $p_1 \geq 2\hat{p}$ or $p_1 \leq \frac{1}{2}\hat{p}$, then the traffic conditions are changed suddenly in the most recent sample interval. To respond quickly to the change, it forgets the history value of computing units, i.e. $p_i = 0, i > 1$, and set the current measurement value to be the equilibrium packet loss ratio $\hat{p} = p_1$. Thus the router can estimate the new equilibrium using the above average filter method.

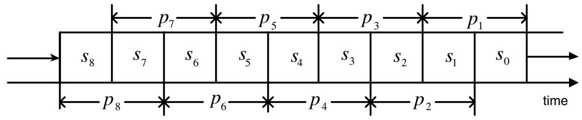

Fig. 1. Full average loss filter diagram (m = 2,n=8)

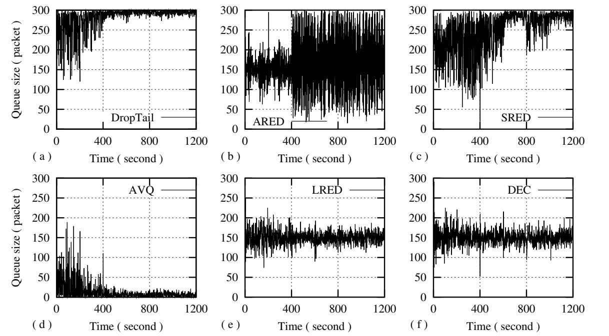

Fig. 2. Queue evolution over a range of TCP traffic loads

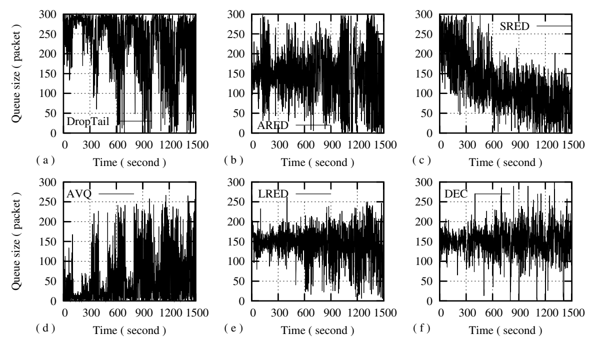

Fig. 3. Queue evolution for different AQMs over a range of link capacity

V. PERFORMANCE EVALUATION

This section compares the performance among several adaptive AQMs, such as ARED, SRED, AVQ, LRED and our proposed DEC by detailed simulations. Unless otherwise noted, a dumbbell network topology is used with a bottleneck link capacity of 10 Mbps and a maximum packet size of 500 bytes. Round-trip link delays are randomly uniformly distributed over the range [120:220]ms. The physical queue limit is set to be 300 packets. The desired queue length is set to be 150 packets. The settings of the parameters for various AQMs are based on their authors' recommendations. For LRED, we choose the β according to Lemma 1 in the paper. In all simulations, we use the drop mechanism unless otherwise stated. Each simulation has 300 background Web-like sessions that start evenly distributed during the first 30 seconds. Each Web session requests pages with 1 object drawn from a Pareto II distribution with a shape parameter of 1.2 and an average size of 10 Kbytes. The Web sessions have an exponentially distributed think time with a mean of 7 seconds.

A. Long-Lived TCP Traffic Load Variation

This experiment compares the performance of DEC, ARED, SRED, LRED, AVQ and Drop-Tail over a range of traffic loads with long-lived TCP flows. Each simulation begins with 50 FTP flows at initial time. After 200 seconds, an additional 50 FTP flows are added. The additional 300 FTP flows and 600 FTP are added at time 400s and 600s, respectively, resulting in 400 FTP flows at time 400s and 1000 FTP flows at time 600s. At time 800s, 400 FTP flows are stopped and 200 FTP flows among them are restarted at time 1000s. All the traffic are terminated at time 1200s. Fig. 2 depicts the queue dynamics for different AQMs. Drop-tail exhibits the expected large queue oscillations when there are few flows and stable, but large, queue sizes when there are many flows. ARED and SRED don't efficiently control the queue under the dynamic environment. Furthermore, ARED exhibits large oscillation when there are 400 FTP flows or more FTP flows. AVQ exhibits a little oscillation when there are few flows but stabilizes with relatively low queue once there are 400 or more FTP flows. Both LRED and DEC can regulate queue length to the desired value with small overshoot. In summary, LRED and DEC are both robust and adaptive quickly to the variance of long-lived TCP traffic.

B. Congested Link Capacity Variation

For the next set of simulations, the bottleneck link capacity is initially set to 10Mbps. It is increased by 10 Mbps every 300 simulation seconds up to 50 Mbps. In each 300 second, the number of FTP flows is as follows: 50 FTP flows are started at the first 100s, an additional 250 flows are added at the second 100s, and 100 FTP flows are stopped at the end of the second 100s. Except for 300 Web traffic started at the initial stage, the short-lived TCP traffic is added randomly throughout the simulation. The results of this simulation are presented in Fig. 3. From Fig. 3(e) and Fig. 3(f), DEC has a better scale for link capacity than that of LRED. It contrasts the main result of Theorem 3.

C. Web-mixed Traffic

we stress-test the AQMs with a Web-mixed flow environment in this experiment. Recent Internet measurements show that 75% of the flows and bytes are Web [3]. For this simulation, the number of FTP flows is fixed at 25 while the Web-traffic load is varied between 40 and 80 percent of the bottleneck bandwidth. Fig. 4 depicts the queue dynamics with different AQMs. Similar to the former simulations, LRED and DEC outperform ARED and SRED in terms of queue dynamics.

VI. CONCLUSION

We have demonstrated the scalable property of LRED algorithm. Based on the stability condition, we propose a scalable algorithm both for TCP load and link capacity. Thus it implies that TCP/DEC becomes unstable only when the network scales up in delay. To avoid the effect of short-lived TCP traffic, we have presented a FAL filter method. Finally, simulation results show that DEC achieves robustness and adaptiveness better than other adaptive AQMs. As

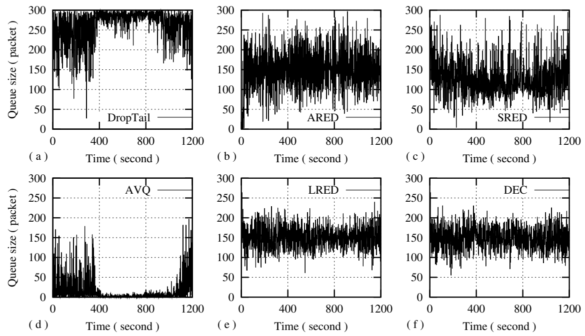

Fig. 4. Queue evolution over a range of short-lived Web traffic loads

future work, we are evaluating their response performance with the lived Web traffic.

VII. ACKNOWLEDGMENTS

We are indebted to Dr. Chonggang Wang for providing the simulation scripts. The work was supported in part by the NSFC under Grants 60174010, 60404022 and the Key Scientific Research Project of the Education Ministry of China under Grant 204014.

REFERENCES

[1] S. Athuraliya, S. H. Low, V. H. Li, and Q. Yin, "REM: active queue management," *IEEE Network*, vol. 15, May/June 2001, pp. 48–53.

[2] B. Braden, D. Clark, J. Crowcroft, et al., "Recommendations on Queue Management and Congestion Avoidance in the Internet," *IETF RFC2309*, Apr. 1998.

[3] K. Claffy, G. Miller, and K. Thomposn, "The nature of the beast: recent traffic measurements form an internet backbone," in *Proceedings of ISOC INET*, July 1998.

[4] C. A. Deseor and Y. T. Yang, "On the generalized Nyquist Stability Criterion," *IEEE Trans. Automat. Contr.*, vol 25, 1980, pp. 187-196.

[5] W. Feng, D. Kandlur, D. Saha, and K. Shin. A Self-Configuring RED Gateway, in *Proc. of IEEE INFOCOM*, Mar 1999.

[6] S. Floyd, R. Gummadi, and S. Shenker, "Adaptive RED: An algorithm for increasing the robustness of RED's active queue management," August 2001. Available:http://www.icir.org/floyd/papers.html.

[7] H. Han, C. V. Hollot, Y. Chait, and V. Misra, "TCP networks stabilized by buffer-based AQMs," in *Proceedings of IEEE INFOCOM*, Hongkong, 2004.

[8] C. V. Hollot, V. Misra, D. Towsley, W. Gong, "On designing improved controllers for AQM routers supporting TCP flows," in *Proceedings of IEEE INFOCOM*, vol. 3, 2001, pp. 1726–1734.

[9] C. V. Hollot, V. Misra, D. Towsley and W. B. Gong, "Analysis and design of controllers for AQM routers supporting TCP flows," *IEEE Trans. Automat. Contr.*, Vol. 47, no. 6, Jun. 2002, pp. 945–959.

[10] R. Johari, D. Tan, "End-to-end congestion control for the internet: Delays and stability," *IEEE/ACM Trans. Networking*, vol. 9, no. 6, Dec. 2001, pp. 818–832.

[11] F. P. Kelly, A. Maulloo, and D. Tan, "Rate control in communication networks:shadow prices, proportional fairness and stability," *Journal of the Operational Research Society*, vol. 49, no. 3, 1998, pp. 237–252.

[12] S. Kunniyur and R. Srikant, "Analysis and design of an adaptive virtual queue (AVQ) algorithm for active queue management," in *Proc. of ACM SIGCOMM*, San Diego, USA, August 2001, pp. 123-134.

[13] S. Kunniyur and R. Srikant, "Stable, scalable, fair congestion control and AQM schemes that achieve high utilization in the Internet," *IEEE Trans. Automat. Contr.*, vol. 48, no.11, Nov. 2003, pp. 2024-2029.

[14] C. N. Long, J. Wu and X. P. Guan, "Local stability of REM algorithm with time-varying delays," *IEEE Commun. Lett.,* vol. 7, no. 3, March 2003, pp. 142-144.

[15] C. N. Long, B. Zhao, X. P. Guan, and J. Yang, "The yellow active queue management algorithm," *Computer Networks*, 2004 (to appear).

[16] S. H. Low, F. Paganini, J. C. Doyle, "Internet congestion control," *IEEE Control Systems Magazine*, vol. 22, 2002, pp. 28-43.

[17] S. H. Low, F. Paganini, J. Wang, and J. C. Doyle, "Linear stability of TCP/RED and a scalable control," *Comput. Networks*, vol. 34, 2003, pp. 633-647.

[18] L. Massoulié, "Stability of distributed congestion control with heterogeneous feedback delays," *IEEE Trans. Automat. Contr.*, vol. 47, no. 6, Jun. 2002, pp. 895–902.

[19] V. Misra, W. Gong, and D. Towsley, "A fluid-based analysis of a network of AQM router supporting TCP flows with an application to RED," in *Proceedings of ACM SIGCOMM*, Stockholm, Sweden, Sep. 2000.

[20] T. J. Ott, T. V. Lakshman, and L. Wong, "SRED: stabilized RED," in *Proc. IEEE INFOCOM*, pp. 1346-1355, March 1999.

[21] F. Paganini, "A global stability result in network flow control," *Systems & Control Letters,* vol. 46, 2002, pp. 165-172.

[22] F. Paganini, Z. Wang, S. H. Low, and J. C. Doyle, "A new TCP/AQM for stability and performance in fast networks," in *Proc. of IEEE Infocom*, San Francisco, USA, 2003.

[23] K. Ramakrishnan and S. Floyd, "A Proposal to Add Explicit Congestion Notification (ECN) to IP," *IETF RFC2481*, Jan. 1999.

[24] Y. Tian, H. Yang, "Stability of the Internet congestion control with diverse delays," *Automatica*, vol. 40, 2004, pp. 1533–1541.

[25] G. Vinnicombe, "On the stability of networks operating TCP like congestion control," in *Proceedings of the IFAC World Congress*, Barcelona, 2002.

[26] C. Wang, B. Li, Y. Thomas Hou, K. Sohraby, and Y. Lin, "LRED: A robust active queue management scheme based on packet loss ratio," in *Proceedings of IEEE INFOCOM*, Hongkong, 2004.

[27] W. Wu, Y. Ren, and X. Shan, "A Self-configuring PI Controller for Active Queue Management," in *Asia-Pacific Conference on Communications (APCC)*, Session T53, Tokyo, Japan, 2001.

[28] Honggang Zhang, C.V.Hollot, Don Towsley, Vishal Misra, "A Self-Tuning Structure for Adaptation in TCP/AQM Networks," in *Proceedings of Globecom*, pp. 3641-3646, 2003.

APPENDIX

Proof: [Proof of Theorem 1] Because eigenvalues are invariant under matrix commutation, we have

$$\lambda\left(L\left(j\omega\right)\right)=\lambda\left(H\left(j\omega\right)\right)$$

Because $\sum_{i=1}^{N}\frac{\hat{x}_i}{c}=1$, thus

$$H\left(j\omega\right)\in\frac{kc}{\hat{w}_i\hat{p}}\cdot Co\left\{z_i\left(j\omega\right),i=1,\dots N\right\}$$

i.e. $H\left(j\omega\right)$ lies in the convex hull defined by the N points $z_i\left(j\omega\right)$ in the complex plane. Then we need prove that this convex hull is bounded away from $(-1,0)$, i.e. every $z_i\left(j\omega\right)$ does not enclose $\frac{\hat{w}_i\hat{p}}{kc}$. We can write the eigenvalue λ of $H\left(j\omega\right)$

$$\lambda\in\frac{Co\left\{\frac{e^{-j\omega\hat{\tau}_i}}{j\omega\hat{\tau}_i+\frac{2}{\hat{w}_i}}\right\}}{Co\left\{j\omega\epsilon+\epsilon\sum_{i=1}^{N}\frac{\hat{x}_i}{c\hat{\tau}_i}\right\}}$$

where $\epsilon=\frac{\hat{w}_i\hat{p}}{kc}$. Now, for $\frac{2}{\hat{w}_i}>0$, $\Re\left(\frac{e^{-j\omega\hat{\tau}_i}}{j\omega\hat{\tau}_i+\frac{2}{\hat{w}_i}}\right)>-1$. Consequently, the eignloci cannot cross the real axis at or to the left of the point $(0,-1)$ if

$$\epsilon\sum_{i=1}^{N}\frac{\hat{x}_i}{c\hat{\tau}_i}\geq 1 \tag{22}$$

Substituting $\hat{w}_i,\hat{p}$ in (9) yields the theorem. ■