Time-Scale Decomposition and Equivalent Rate-Based Marking

Yung Yi, *Student Member, IEEE*, Supratim Deb, and Sanjay Shakkottai, *Member, IEEE*

***Abstract*—Differential equation models for Internet congestion control algorithms have been widely used to understand network dynamics and the design of router algorithms. These models use a fluid approximation for user data traffic and describe the dynamics of the router queue and user adaptation through coupled differential equations. The interaction between the routers and flows occurs through marking, where routers indicate congestion by appropriately marking packets during congestion.**

In this paper, we show that the randomness due to short and unresponsive flows in the Internet is sufficient to decouple the dynamics of the router queues from those of the end controllers. This implies that a time-scale decomposition naturally occurs such that the dynamics of the router manifest only through their statistical steady-state behavior. We show that this time-scale decomposition implies that a queue-length based marking function (e.g., RED-like and REM-like algorithms, which have no queue averaging, but depend only on the instantaneous queue length) has an equivalent form which depends only on the data arrival rate from the end-systems and does not depend on the queue dynamics. This leads to much simpler dynamics of the differential equation models (there is no queueing dynamics to consider), which enables easier analysis and could be potentially used for low-complexity fast simulation.

Using packet-based simulations, we study queue-based marking schemes and their equivalent rate-based marking schemes for different types of controlled sources (i.e., proportional fair and TCP) and queue-based marking schemes. Our results indicate a good match in the rates observed at the intermediate router with the queue-based marking function and the corresponding rate-based approximation. Further, the window size distributions of a typical TCP flow with a queue-based marking function as well as the equivalent rate-based marking function match closely, indicating that replacing a queue-based marking function by its equivalent rate-based function does not statistically affect the end host's behavior.

***Index Terms*—Internet congestion control, marking functions, time-scale decomposition.**

I. INTRODUCTION

WE CONSIDER the problem of Internet congestion control when the network is accessed by a mixture of long-lived controlled flows, as well as short flows which do not react to congestion. The short flows model a mixture of real-time-based traffic (such as real-time multimedia) as well as web traffic (so-called web-mice), where the sessions are too short for the end systems to react to network congestion.

The transmission rate of the long-lived flows are controlled by the intermediate routers in the network. The task of these routers is to simply notify the end systems whenever they detect congestion in the network. Associated with each router is a marking function, which marks a fraction of the flow, and the fraction that is marked is a function of the arrival rate (rate-based marking) or the queue length (queue-based marking). In the Internet, marking is implemented via the Explicit Congestion Notification (ECN) mechanism [1], where packets have a bit in the header that can be set to "1" to indicate congestion. The end-host reacts to this information by suitably adapting its transmission rate, thus adapting to network congestion.

There has been extensive research on differential equation-based congestion control [2]–[7], where fluid models of a large number of flows were used to model the dynamics of the system based on a rate-based marking scheme. The source controllers are modeled by differential equations (i.e., a fluid model for data flow). These controllers adapt their transmission rate based on network feedback in the form of a fraction of fluid that is marked by the routers. In other words, with n flows in the network, the dynamics of the controller are described by

$$\dot{x}_n^i(t) = \kappa\left(w - U_i'(x_n^i(t))^{-1} p_r\left(\frac{1}{n}\sum_{j=1}^{n}(a_n^j(t) + x_n^j(t))\right)\right), \quad i = 1,\ldots,n \tag{1}$$

where w and κ are parameters of the controller that determine the equilibrium rate as well as the transient dynamics, $x_n^i(t)$ is the transmission rate of the controlled flow i at time t, $\sum_i a_n^i(t)$ represents the short-lived uncontrolled flows, and $U_i(x)$ is a concave utility function of the user i, when the transmission rate of the user i is x. Examples of such utility functions include $\log x$ (proportional fair controller) and $-1/x$ (TCP controller) [8]. The function $p_r(\cdot)$ is a rate-based marking function whose argument is the average arrival rate to the router (additional discussion is available later in this section). The marking function indicates the level of congestion at the router. Thus, $p_r(\cdot)$ is a monotone, increasing function with range [0, 1]. The larger the marking level is, the higher is the perceived congestion at the router. As seen in (1), the controller reacts to a congestion level by decreasing the transmission rate.

Alternately, instead of adapting based on the average arrival rate, the marking function at the router can adapt based on the queue length at the router. In other words, the router is associated with a queue-based marking function $p_q(\cdot)$. This is assumed to be a monotone increasing function over [0, 1] and Lip-

Manuscript received June 24, 2004; revised July 21, 2005; approved by IEEE/ACM TRANSACTIONS ON NETWORKING Editor F. Paganini. This work was supported by the National Science Foundation under Grant ACI-0305644, Grant CNS-0325788, and Grant CNS-0347400. A shorter version of this paper appeared in the Proceedings of the 38th Annual Conference on Information Sciences and Systems (CISS), Princeton, NJ, March 2004.

Y. Yi and S. Shakkottai are with the Department of Electrical and Computer Engineering, The University of Texas at Austin, Austin, TX 78712-0240 USA (e-mail: yi@ece.utexas.edu; shakkott@ece.utexas.edu).

S. Deb is with Bell-Labs Research, India Lucent Technologies, Bangalore, India (e-mail: supratim@lucent.com).

Digital Object Identifier 10.1109/TNET.2006.882862

schitz continuous with parameter L_q. The associated differential equation model for the end-system controller is given by

$$\dot{x}_n^i(t) = \kappa \left(w - U_i'(x_n^i(t))^{-1} p_q \left(\frac{1}{n} Q_n(t) \right) \right) \tag{2}$$

where

$$\dot{Q}_n(t) = \begin{cases} \sum_{j=1}^{n} \left(a_n^j(t) + x_n^j(t) \right) - nc, & \text{if } Q_n(t) > 0 \\ \left[\sum_{j=1}^{n} \left(a_n^j(t) + x_n^j(t) \right) - nc \right]^+, & \text{if } Q_n(t) = 0 \end{cases}$$

where $Q_n(t)$ is the queue length at the router, and nc is the capacity of the link. Examples of queue-based marking include random early discard (RED) [1], adaptive virtual queue (AVQ) [9], and random exponential marking (REM) [10].

It has been shown in [6] that the differential equation-based models described in (1) and (2) are valid models of in the Internet when there are a sufficiently large number of flows and the network capacity is large (scaled with the number of flows). In such a regime, the arguments of the marking functions are interpreted as the average arrival rate (averaging by the number of flows) or the scaled queue length (scaled by the number of flows), respectively. Physically, this scaling of the arguments corresponds to the fact that the arrival rates and capacity are large; see [6] for details.

In other words, for a network model with n flows and the capacity at the router being nc, the marking function at the router adapts either based on the average arrival rate $x(t) = 1/nX^n(t)$, where $X^n(t)$ is the total arrival rate to the router, or based on the average queue length $q(t) = (1/n)Q^n(t)$, where $Q^n(t)$ is the queue length at the router. In particular, this implies that, as the system size becomes larger, so does the associated queue length at the router. In other words, a finite nonzero queue length in the fluid differential equation model (2) indicates that the actual queue length in the router is large (of order n). Related work with a similar scaling (large buffer and capacity) for window based control is available in [11].

However, as link speeds in modern and future communication networks is becoming higher, high-speed memory buffer with high cost is required in the design of such networks. Therefore, it is questionable if the queue buffers at intermediate routers need to scale linearly with the number of flows [12]. In [12], [13], the authors have in fact shown that buffers need not scale with the link speed in order to achieve significant multiplexing gains.

In this paper, we focus on this regime where the queue length does not scale with the number of flows. Such a behavior occurs, for instance, if the queue-based marking function $p_q(\cdot)$ is invariant with the number of flows and is a function of the actual queue length, not the average queue length. Under such a regime, the queue dynamics occur on a much faster time-scale than that of the end system controller [14]. In this context, it is reasonable to expect that queueing dynamics are not visible to the end system controller. Instead, the queueing behavior at the router affects the end system controller only through the statistical behavior of the queue.

Recent related work includes [15], where the authors consider a discrete-time framework for congestion control. They have shown that, depending on the scaling, the limiting system could be a combination of queue- and rate-based marking, even if the unscaled system consists of only queue-based marking. In our paper, we consider a continuous-time framework, where we are primarily interested in a pure rate-based approximation to a queue-based marking function. Our focus is on deriving the equivalent rate-based marking function over a continuous-time framework and studying network dynamics by replacing a queue-based marking function (such as RED or REM) with an equivalent rate-based one. Further, the proof techniques employed are very different in the two approaches.

A. Main Contributions and Organization

The main contributions of this paper are the following.

1) This paper quantifies the heuristics based on time-scale separation by showing that, under suitable assumptions, queue-based marking (based on instantaneous queue length) and the associated queueing dynamics can be approximated by a rate-based marking function given by
$$p(x) \triangleq E_{\pi_\lambda^{c-x}}[p_q(Q)]$$
where π_λ^μ is the stationary queue-length distribution of an M/D/1 queue with Poisson arrival rate λ and capacity μ. The parameters x and λ are simply the average arrival rate from the controlled and the uncontrolled flows (averaging over flows, not time) to the router queue, respectively.
2) Using packet simulations by suitably modifying the *ns-2* [16] simulator, we compare queue-based marking schemes and their equivalent rate-based marking schemes for different types of controlled sources (proportional fair and TCP) and queue-based marking schemes (RED-like and REM-like algorithms without queue averaging). In addition, we show that the equivalent rate-based marking scheme behaves well even when drastic changes occur in the number of network connections. The simulation results indicate a good match between the queue-based marking and the equivalent rate-based marking in the steady states as well as the transient behaviors of end-sources.

The results of this paper potentially could enable low-complexity simulation and easier analysis for the following reasons. In the context of network simulation, several widely used discrete event-driven simulators are available [16]–[19]. However, with a large number of flows, the number of events at an intermediate router scales with the number of flows due to packet arrivals/departures and marking computation events. The equivalent rate-based model proposed in this paper uses a marking averaged by rates, along with the absence of queuing dynamics to enable a simulation complexity at intermediate routers that does not scale with the number of flows, leading to significant reduction of simulation complexity (see Section III-C for additional discussion).

Also, much of the work on stability analysis of the congestion control algorithm has been done based on the rate-based marking model at the intermediate routers. Thus, with our ap-

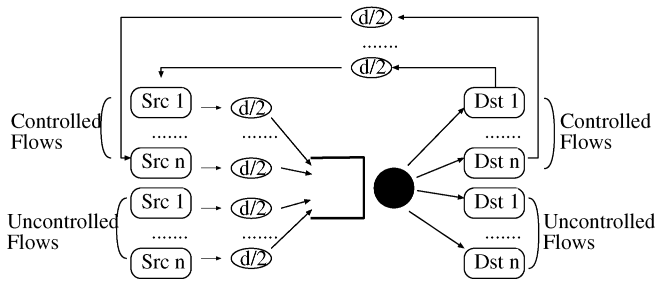


Fig. 1. System model.

proximate rate-based model, it seems easier to study queue-based marking systems using analytical tools developed for rate-based marking in literature [4], [6], [20].

In the remainder of this paper, we begin with a description of the system model in Section II. Next, in Section III, we show that there exists an equivalent rate-based marking function for a given queue-based marking scheme (under suitable conditions). Using these results, we derive expressions for the equivalent marking function with the RED-like and the REM-like controllers[1] in Section III-D. We finally present simulation results for RED and REM with proportional fair and TCP sources.

II. System Model

Consider the system shown in Fig. 1. We consider a single queue with the first-in-first-out (FIFO) scheduling discipline accessed by two types of flows: 1) controlled flows and 2) uncontrolled flows. We consider a sequence of systems indexed by n, the scaling parameter. In the nth system, the queue is fed by n independent, identically distributed uncontrolled flows and by n controlled flows determined by a congestion control algorithm.[2] The output capacity of the router queue scaled with n as nc pkts/s.

For the nth system, we model each uncontrolled flow by means of a point process $A_n^i(t)$ that represents the cumulative number of packets from flow i that arrive until time t. We assume that each $A_n^i(t)$ has the same distribution as a simple stationary point process A that satisfies the following assumptions [12], [21].

Assumption 2.1: A is a simple stationary point process satisfying the following three properties.

1) There exists $\lambda > 0$ such that $E[A(t)] = \lambda t$ for $t \in [0, \infty)$.
2) There exists $\theta_0 > 0$ and $K < \infty$ such that
$$\lim_{t \to 0^+} E[e^{\theta_0 A(t)} \mathbf{1}_{A(t)>K}] = 0$$
where $\mathbf{1}_E = 1$ if the predicate E is true, and 0 otherwise.
3) $\liminf_{t\to\infty} \frac{t\Lambda(x,t)}{\log t} > 0$, where $\Lambda(x,t) = \sup_{\theta \in \mathcal{R}}[\theta x - \frac{1}{t}\log E[e^{\theta A(t)}]]$.

Assumption 2.1 states that each uncontrolled arrival process satisfies the properties that: 1) multiple packets from a single uncontrolled source do not arrive at the same time; 2) all arriving packets are of the same size; and 3) the uncontrolled arrival process has a finite intensity (see [12] for further details).

From the controlled flows point of view, the system we have described above can be thought of as a closed-loop system with delay, and feedback control applied at the routers based on queue-based marking function is denoted by $p_q(\cdot)$. A popular modeling and analysis methodology for such closed-loop systems in the Internet context has been through functional differential equations based fluid models.

The generic model of such a system consists of a collection of user flows, a router modeled by marking functions which signal congestion by marking flows, and receivers which detect the marks and informs the respective flows to increase or decrease their transmission rate. We model flows by fluid processes. We denote the fluid rates of individual flows in the nth system by $\{x_n^i(t), i = 1, \ldots, n\}$, where $x_n^i(t)$ denotes the transmission rate of a controlled flow i at time t. The dynamics of the transmission rate for each user are governed by a differential equation-based controller as discussed in Section I. We comment that the controller in (2) is called a proportional-fair controller if $U(x) = \log(x)$ [14], as controllers of this form lead to a proportionally fair allocation of bandwidth across users. The results in this paper, however, apply to any differential equation-based congestion controller as long as $\dot{x}_n^i(\cdot)$ is bounded (i.e., the transmission rate is Lipschitz). In particular, suppose that the transmission rate $x_n^i(\cdot)$ is bounded by some constant L. This in-turn implies that $x_n^i(\cdot)$ is Lipschitz continuous with some parameter $M < \infty$ [22]. In the remainder of this paper, we assume that the transmission rate is Lipschitz continuous with parameter M.

Let $A_n(t) = \sum_i A_n^i(t)$ be the cumulative number of arrivals until time t due to uncontrolled flows, and $X_n(t) = \sum_i x_n^i(t)$ be the total arrival rate at time t due to controlled flows. From Assumption 2.1, $E(A_n(t)) = n\lambda t$.

For the controlled flows, let us denote the average arrival rate by
$$x_n(t) = \frac{1}{n} X_n(t).$$

Further, we define the total volume of arrivals (due to the controlled flows) until time t by $Y_n(t)$, where
$$Y_n(t) = \int_0^t X_n(z)dz = n\int_0^t x_n(z)dz.$$

Finally, we assume that the initial conditions satisfy
$$\begin{aligned} &x_n^i(0) \overset{n\to\infty}{\longrightarrow} x^i(0)\\ &x_n(0) \overset{n\to\infty}{\longrightarrow} x(0)\\ &Q_n(0) \overset{n\to\infty}{\longrightarrow} Q(0) < \infty\\ &x(0) + \lambda < c. \end{aligned} \tag{3}$$

Heuristically, these conditions correspond to the assumption that the initial condition is well defined and is a stable system.

III. Limiting Rate-Based Marking Function

Here, we will derive the equivalent rate-based marking function for a given queue-based marking function. In this paper, we

[1] Henceforth, for notional convenience we use the terms RED and REM to refer to queue-based RED and REM without queue averaging.

[2] For notational simplicity, we have assumed an equal number of controlled and uncontrolled sources. The results in this paper hold even if they are not the same, as long as the ratio of the number of controlled flows and the number of uncontrolled flows is finite.

focus on the instantaneous queue length process. Note that popular AQM algorithms such as RED and REM use (exponentially moving) average queue length to mark the incoming packets. We left the study on AQM algorithms with queue averaging as future work.

For a fixed $T > 0$, we are interested in studying the queue length process (which measures the volume of data at the router), denoted by $Q_n(t)$, over the time interval $\left[0, \frac{T}{n}\right]$. Thus, we are interested in the queueing behavior at the router over a short interval of time. Even over this small time interval, we will show that the queue reaches "steady-state" behavior. This occurs due to the fact that the capacity is very large (nc) and causes the queue to "regenerate" an arbitrarily large number of times over the interval $\left[0, \frac{T}{n}\right]$.

However, from a single end-system (the user) point of view, this corresponds to a very short interval of time. Thus, one can expect that the end-user will only perceive the statistical "steady-state" queueing behavior. The results in this section quantify the above heuristic.

For any $s \in \left[0, \frac{T}{n}\right]$, the queue length process is given by

$$\begin{aligned} Q_n(s) &= \sup_{r\in[0,s]} \Big[A_n(s) - A_n(r) + Y_n(s) - Y_n(r) \\ &\qquad - nc(s-r) + Q_n(r)\Big] \\ &= \sup_{r\in[0,s]} \Big[A_n(s) - A_n(r) + n\int_r^s x_n(z)\,dz \\ &\qquad - nc(s-r) + Q_n(r)\Big]. \end{aligned}$$

Now, let us study the processes (X_n, Y_n, A_n, Q_n) over a slowed-down time-scale. In other words, for $t \in [0,T]$, we define the processes

$$q_n(t) \triangleq Q_n\left(\frac{t}{n}\right),\ a_n(t) \triangleq A_n\left(\frac{t}{n}\right),\ y_n(t) \triangleq Y_n\left(\frac{t}{n}\right).$$

Thus, we have for any $t \in [0,T]$

$$\begin{aligned} q_n(t) &= Q_n\left(\frac{t}{n}\right) \\ &= \sup_{\frac{r}{n}\in[0,\frac{t}{n}]} \Big[A_n\left(\frac{t}{n}\right) - A_n\left(\frac{r}{n}\right) + Y_n\left(\frac{t}{n}\right) \\ &\qquad - Y_n\left(\frac{r}{n}\right) - \frac{nc(t-r)}{n} + Q_n\left(\frac{r}{n}\right)\Big] \\ &= \sup_{r\in[0,t]} [a_n(t) - a_n(r) + y_n(t) - y_n(r) \\ &\qquad - c(t-r) + q_n(r)]. \end{aligned} \tag{4}$$

By assumption, each individual data rate $(x_n^i(t))$ is Lipschitz continuous with some parameter $M < \infty$. This also implies that the average data rate $(x_n(r))$ is Lipschitz continuous with parameter M. Let us now define

$$\tilde{q}_n(t) \triangleq \sup_{r\in[0,t]} [a_n(t) - a_n(r) + (t-r)x(0) - c(t-r) + \tilde{q}_n(r)]. \tag{5}$$

A. Convergence of the Queue Length Trajectory

We now show that the queue length process over the slowed down time-scale converges weakly to the queue length process of a M/D/1 queue with service rate $c - x(0)$. In [12], the authors showed a similar result for the stationary distribution of the queue. In this paper, we are interested in the path properties of the queue because the marks received by the end-user depends on the integral of the marking function over the (unscaled) time interval $[0, T/n]$. Thus, it is not sufficient for us to consider only the stationary distribution. We show that the slowed down queue length process converges to the corresponding M/D/1 queueing process "uniformly" (to be precise, with respect to the Skorohod metric) over the time interval $[0,T]$.

Prior to presenting the main theorems, we first provide the following two lemmas.

Lemma 3.1: Given $\epsilon > 0$, we can find N such that $\forall n > N$

$$\|q_n(t) - \tilde{q}_n(t)\| < \epsilon \tag{6}$$

where $\|\cdot\|$ is the Skorohod metric [23] in the space $\mathcal{D}([0,T] : \mathcal{R}^+)$.

Proof: The proof is presented in the Appendix . ■

Lemma 3.2: Suppose that $a_n(t) \to a(t)$ in the space $\mathcal{D}([0,T] : \mathcal{R}^+)$. Then, given any $\epsilon > 0$, there exists N such that, $\forall n > N$, we have

$$\|q(t) - \tilde{q}_n(t)\| < \epsilon \tag{7}$$

where $q(t)$ is defined by

$$q(t) \triangleq \sup_{r\in[0,t]} [a(t) - a(r) + (t-r)x(0) - c(t-r) + q(r)] \tag{8}$$

and $a(t)$ is a Poisson process with arrival rate λ.

Proof: The proof is presented in the Appendix . ■

Theorem 3.1: As $n \to \infty$, we have

$$q_n(t) \xrightarrow{w} q(t), \quad t \in [0,T] \quad \text{over} \quad \mathcal{D}([0,T] : \mathcal{R}^+)$$

where $\xrightarrow{w}$ represents weak convergence, and $q(t)$ is the queue-length process of a single server M/D/1 queue, with deterministic service rate $c - x(0)$ and arrival process $a(t)$, which is a Poisson process of rate λ.

Proof: From the superposition theorem for point processes [21], we know that a_n converges weakly to a Poisson process with rate λ denoted by $a(t)$ in $\mathcal{D}([0,T] : \mathcal{R}^+)$. From the Skorohod representation theorem [23], we can find processes $a_n'(t)$ and $a'(t)$ in $\mathcal{D}([0,T] : \mathcal{R}^+)$ such that

$$a_n(t) \overset{\text{dist}}{=} a_n'(t)$$
$$a(t) \overset{\text{dist}}{=} a'(t)$$

where $\overset{\text{dist}}{=}$ means "equivalence in distribution" and

$$\|a_n'(t) - a'(t)\| \xrightarrow{n\to\infty} 0 \quad \text{in} \quad \mathcal{D}([0,T] : \mathcal{R}^+). \tag{9}$$

Corresponding to the arrival processes $a'(t)$ and $a_n'(t)$, let us define $q(t)$, $q_n'(t)$, and $\tilde{q}_n'(t)$ by (4), (5), and (8), respectively.

Then, it suffices to prove that, $\forall \epsilon > 0$, we can find N such that $\forall n > N, \|q'_n(t) - q'(t)\| < \epsilon$ in the space $\mathcal{D}([0,T] : \mathcal{R}^+)$. By the triangle inequality of Skorohod norm, we have

$$\|q'_n(t) - q'(t)\| \leq \|q'_n(t) - \tilde{q}'_n(t)\| + \|\tilde{q}'_n(t) - q'(t)\|.$$

By applying Lemma 3.1 to the first term of the right-hand side and Lemma 3.2 to the second term of the right-hand side, the result follows. ■

Using this result, we now show that the total volume of marks received over the (slowed down) time-interval $[0,T]$ converges to that given by an M/D/1 queue.

Theorem 3.2: Suppose that

$$q_n(t) \xrightarrow{w} q(t), \quad t \in [0,T] \quad \text{over} \quad \mathcal{D}([0,T] : \mathcal{R}^+)$$

where $q_n(t)$ and $q(t)$ is defined as (4) and (8). Then, we have

$$\int_0^T p_q(q_n(y))\, dy \xrightarrow{w} \int_0^T p_q(q(y))\, dy \tag{10}$$

$$\int_0^T x_n^i\left(\frac{y}{n}\right) p_q(q_n(y))\, dy \xrightarrow{w} \int_0^T x^i(0) p_q(q(y))\, dy. \tag{11}$$

Proof: From Theorem 3.1 and Skorohod representation theorem, we can find q'_n and q' in $\mathcal{D}([0,T] : \mathcal{R}^+)$ such that q'_n converges to q' in the Skorohod topology. By the definition of convergence in the Skorohod topology, we can find a strictly increasing, continuous function λ_n of $[0,T]$ onto itself and $N_1 > 0$ such that, for a given $\epsilon > 0$, and $\forall n > N_1$, we have

$$\sup_{t\in[0,T]} |q'_n(\lambda_n(t)) - q'(t)| < \epsilon$$
$$\sup_{t\in[0,T]} |\lambda_n(t) - t| < \epsilon. \tag{12}$$

By adding and subtracting a common term, we have

$$\begin{aligned}&\left|\int_0^T p_q(q'_n(y))\, dy - \int_0^T p_q(q'(y))\, dy\right| \\ &\quad \leq \left|\int_0^T p_q(q'_n(y))\, dy - \int_0^T p_q(q'_n(\lambda_n(y)))\, dy\right| \\ &\qquad + \left|\int_0^T p_q(q'_n(\lambda_n(y)))\, dy - \int_0^T p_q(q'(y))\, dy\right|.\end{aligned}$$

For the second term on the right-hand side, using the Lipschitz continuity assumption of p_q and the condition (12), we have

$$\begin{aligned}&\left|\int_0^T p_q(q'_n(\lambda_n(y)))\, dy - \int_0^T p_q(q'(y))\, dy\right| \\ &\qquad \leq \int_0^T L_q|q'_n(\lambda_n(y)) - q'(y)| < L_q T\epsilon \end{aligned} \tag{13}$$

where $0 < L_q < \infty$ is the Lipschitz constant of $p_q(\cdot)$.

Next, for the first term on the right-hand side, we know that $q'(s) \in \mathcal{D}([0,T] : \mathcal{R}^+)$ has a finite number of jumps denoted by $\mathcal{J}(q') < \infty$, since the arrival process is a Poisson process with a finite rate over the finite interval of time $[0,T]$. From the condition (12), we can find $N_2 > 0$ such that $\forall n > N_2,\ J \triangleq \mathcal{J}(q'_n) = \mathcal{J}(q')$. Let us denote the jump times of $q'_n(s)$ by $\{t_n^j, j = 1,\ldots,J\}$.

Now, we divide the entire interval $[0,T]$ into two sets of intervals A_1 and A_2, where $A_1 = \{I_j \triangleq [t_n^j - \epsilon, t_n^j + \epsilon], j = 1,\ldots,J\}$ and $A_2 = [0,T]\backslash A_1$. By taking $\epsilon < 0.5\min\{t_n^1, t_n^2 - t_n^1, \ldots, T - t_n^J\}$, this ensures that there is only one jump of the processes, $q'_n(\lambda_n(s))$ and $q'_n(s)$ in the interval I_j. From the Lipschitz continuity of p_q, $\forall s \in [0,T]$, we have

$$|p_q(q'_n(\lambda_n(s))) - p_q(q'_n(s))| \leq L_q|q'_n(\lambda_n(s)) - q'_n(s)|.$$

Let $N_{\max} = \max(N_1, N_2)$. Then, $\forall n > N_{\max}$, we have

$$|q'_n(\lambda_n(s)) - q'_n(s)| \leq \begin{cases} 1, & \text{if } s \in A_1 \\ \epsilon(c - x(0)), & \text{if } s \in A_2. \end{cases}$$

Thus

$$\begin{aligned}&\left|\int_0^T p_q(q'_n(y))\, dy - \int_0^T p_q(q'_n(\lambda_n(y)))\, dy\right| \\ &\leq \left|\int_{A_1} p_q(q'_n(y))\, dy - \int_{A_1} p_q(q'_n(\lambda_n(y)))\, dy\right| \\ &\quad + \left|\int_{A_2} p_q(q'_n(y))\, dy - \int_{A_2} p_q(q'_n(\lambda_n(y)))\, dy\right| \\ &\leq \int_{A_1} |p_q(q'_n(y)) - p_q(q'_n(\lambda_n(y)))|\, dy \\ &\quad + \int_{A_2} |p_q(q'_n(y)) - p_q(q'_n(\lambda_n(y)))|\, dy \\ &\leq 2L_q J\epsilon + L_q(c - x(0))\epsilon(T - 2J\epsilon) \\ &= \epsilon\left(2L_q J + L_q(c - x(0))(T - 2J\epsilon)\right).\end{aligned} \tag{14}$$

Since ϵ is arbitrary in (13) and (14), this completes the proof. The proof of (11) is analogous. ■

B. Equivalent Rate-Based Marking Function

This section defines an equivalent rate-based marking function based on Theorems 3.2 and 3.1. Let us consider the marks received over the time interval $\left[0, \frac{T}{n}\right]$ by some user i. By definition, the marked volume of data over this time interval is given by

$$\int_0^{\frac{T}{n}} x_n^i(y) p_q(Q_n(y))\, dy = \frac{1}{n}\int_0^T x_n^i(y/n)\, p_q(q_n(y))\, dy.$$

Thus, the time-averaged volume of marks received by user i over the time interval $\left[0, \frac{T}{n}\right]$ is

$$\frac{1}{T/n}\int_0^{\frac{T}{n}} x_n^i(y) p_q(Q_n(y)) dy = \frac{1}{T}\int_0^T x_n^i\left(\frac{y}{n}\right) p_q(q_n(y)) dy. \tag{15}$$

Thus, from Theorem 3.2, we have

$$1/\frac{T}{n}\int_0^{\frac{T}{n}} x_n^i(y) p_q(Q_n(y))\, dy \overset{n\to\infty}{\longrightarrow} x^i(0)\frac{1}{T}\int_0^T p_q(q(y))\, dy \tag{16}$$

where $q(y)$ is the queue-length process of an M/D/1 queue with Poisson arrival rate λ and capacity $c - x(0)$. Let us define

$$p_T(x(0)) \triangleq \frac{1}{T}\int_0^T p_q(q(y))\, dy. \tag{17}$$

For n sufficiently large, we see from (16) that the interaction between the router queuing process and the congestion controller at a fixed user occurs only through this function $p_T(\cdot)$.

Further, we observe that $q(y)$ is a regenerative process when $\frac{\lambda}{c-x} < 1$ and $x < c$. Thus, from the ergodic theorem for a regenerative process [24] and Smith's theorem [25], for a given $\epsilon > 0$, $\exists T_0$ such that $\forall T > T_0$, we have

$$\left| \frac{1}{T} \int_0^T p_q(q(y))\, dy - E_{\pi_\lambda^{c-x}}[p_q(Q)] \right| < \epsilon \tag{18}$$

where π_λ^μ is the stationary distribution of an M/D/1 queue with arrival rate λ and capacity μ.

For T sufficiently large and by defining

$$p(x) = \begin{cases} E_{\pi_\lambda^{c-x}}[p_q(Q)], & \text{if } \frac{\lambda}{c-x} < 1 \text{ and } x < c \\ 1, & \text{if } x \geq c \text{ or } \frac{\lambda}{c-x} \geq 1. \end{cases} \tag{19}$$

we see from (17) and (18) that the congestion controller dynamics with a queue-based marking function $p_q(\cdot)$ can be well approximated by an equivalent system with only a rate-based controller $p(x)$ at the router, where x is simply the average arrival rate from the controlled flows (averaging over flows, not time) to the router queue.

Remark 3.1: We comment that, for each fixed T, as $n \to \infty$, the limiting approximate model (19) holds. Thus, for T sufficiently large (but finite), we can be within an ϵ bound of the time-averaged limit [see (18)]. Physically, this corresponds to the time-scale separation issue. In reality, we are interested in using this approximation for a finite n (number of flows). Thus, the time-scale T should be chosen such that the following two properties hold: 1) the time interval T is large enough such that the randomness due to uncontrolled flows enables the "law of large numbers" to hold (i.e, the M/D/1 queue time-average is close to the stationary distribution) and 2) the time-interval T/n is small enough such that the arrival rate from a user (controlled flow) does not significantly change. Thus, for any fixed $\epsilon > 0$, we choose T large enough such that the expected value of the (queue-based) marking function with an M/D/1 queue is "close" to the time-average. For this fixed T, we can apply the limit theorem (in n) to justify the rate-based approximation.

If there is insufficient randomness in the network, the value of T could be very large, thus requiring a large value of n (i.e., large number of flows and large capacity) for our analysis to hold. However, our simulations (see Section IV) indicate that, even with randomness generated due to short ON–OFF flows which occupy about 20%–30% of the link capacity, the value of $n = 100$ seems to be sufficient and leads to a match within 5% between a queue-based marking function and its equivalent rate-based model.

C. Application to Simulation Study

Over any fixed interval of time, the simulation complexity at intermediate routers of a queue-based simulation depends on the number of events to process. As the number of flows increases, the number of events (packet arrivals/departures, marking probability computation) increases, thus leading to increased simulation complexity. Also, the number of events scales linearly with the number of intermediate routers.

On the other hand, the equivalent rate-based model permits the following implementation. Fix a small time-step $\delta > 0$ such that the arrival rate from a controlled flow does not vary significantly over this time-step (i.e., δ is inversely proportional to the end-system congestion controller gain; see Fig. 2). For this fixed δ, suppose that the number of flows n is large enough such that there is a sufficient amount of randomness due to uncontrolled flows over this interval $[0, \delta]$ (i.e., the Poisson approximation holds for the chosen value of δ). Now at each equivalent rate-based router, a computation to determine the marking probability is performed only once in each time-step of size δ. Such a marking value is computed at each intermediate router, and the packets from end-systems are marked appropriately depending on the marking values. Thus, marking computation required scales as $\frac{1}{\delta} \times n_i$ (n_i is the number of intermediate routers) and is invariant with the number of flows n. Further, such a rate-based approximation will become increasingly accurate as the system scale (n) increases.

D. Examples: REM and RED

Here, we derive the closed form of equivalent rate-based marking functions for the simplified REM and RED controllers, which have no queue averaging, but depend only on the instantaneous queue length.

1) REM: The simplified version used in this paper has the following form of the queue-based marking function from [10]

$$p_q^{\text{rem}}(Q) = 1 - e^{-\alpha Q} \tag{20}$$

where α is a suitable constant predefined in the system, and Q is the queue length in the system.

First, from the P-K formula for stationary workload V of an M/D/1 queue [26], we have

$$E[e^{-sV}] = \frac{1-\rho}{1 - \frac{\lambda}{s}(1 - e^{-s/\mu})} \tag{21}$$

where μ is the service rate, λ is the arrival rate, and $\rho = \lambda/\mu$.

Further, for the fluid queueing system we consider, it follows from the definition of workload [26] that $V = Q/c$. Thus, we have

$$\begin{aligned} p(x) &= E_{\pi_\lambda^{c-x}}[p_q^{\text{rem}}(Q)] \\ &= E_{\pi_\lambda^{c-x}}[1 - e^{-\alpha c V}] \\ &= 1 - \frac{1-\rho}{1 - \frac{\lambda}{\alpha c}\left(1 - e^{-\alpha c/(c-x)}\right)} \end{aligned} \tag{22}$$

where $\rho = \frac{\lambda}{(c-x)}$. In the next section, we compare simulation results with queue-based marking with REM and compare that to numerical results using its equivalent rate-based marking function given by (22).

2) RED: The simplified queue-based marking function of RED controller is defined as

$$p_q^{\text{red}}(Q) = \max\left(\left(\frac{Q-a}{b}\right)^+, 1\right) \tag{23}$$

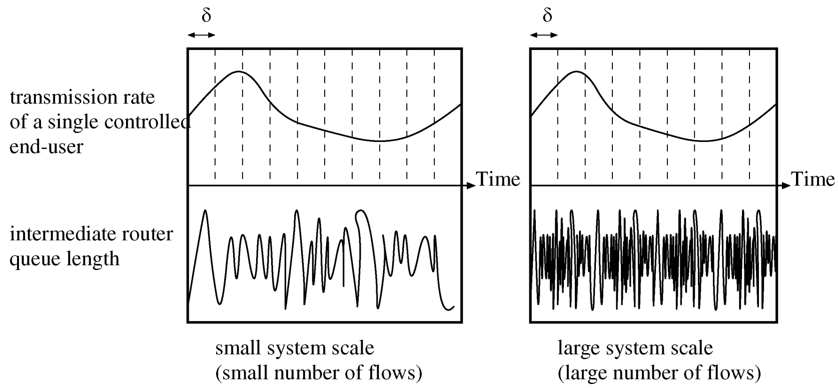


Fig. 2. System scale size and time-step size.

where a and b are suitable constants predefined at the intermediate router [1]. Thus, we have

$$\begin{aligned} p(x) &= E_{\pi_\lambda^{c-x}}[p_q^{\text{red}}(Q)] \\ &= E_{\pi_\lambda^{c-x}}\left[\max\left(\left(\frac{Q-a}{b}\right)^+, 1\right)\right] \\ &= \int_a^{a+b}\left(\frac{q-a}{b}\right) f_Q(q)\,dq + \Pr(Q > a+b) \quad (24) \end{aligned}$$

In order to evaluate (24), it suffices to determine the distribution of the random variable Q. We know that

$$\Pr(Q > y) = \Pr\left(V > \frac{y}{c}\right).$$

From [27] and [28], the unfinished work U for a Poisson process with arrival rate λ in the system with service rate μ has a steady-state distribution of the form

$$\Pr(U > x) = 1 - (1-\rho)e^{\rho x} Q_{\lfloor x \rfloor}(x - \lfloor x \rfloor) \quad (25)$$

where $\{Q_n(x), n = 0, 1, \ldots\}$ are polynomial functions (which can be calculated recursively as shown in [27] and [28]), and $\rho = \lambda/\mu$. From the definition of U and V, we have $U = \mu V$. Thus

$$\begin{aligned} \Pr(Q > y) &= \Pr\left(V > \frac{y}{c}\right) \\ &= \Pr\left(U > \frac{y(c-x)}{c}\right). \quad (26) \end{aligned}$$

From (25) and (26), we can evaluate (24).

IV. SIMULATION

In the previous section, we showed that the queue-based marking and the associated queueing dynamics can be approximated by a rate-based marking function under the fluid model. In this section, we use the *ns-2* [16] simulator to validate our results. The simulation results in this section show that both the steady-state behavior as well as the transients of the end sources have a good match between the queue-based marking and the equivalent rate-based marking.

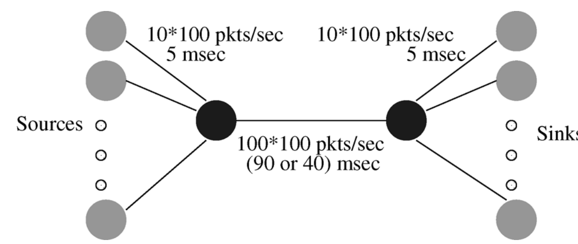


Fig. 3. Simulation topology.

A. Simulation Environment

The network topology used in the simulation is shown in Fig. 3. In Fig. 3, the bottleneck link is accessed by 100 controlled and 100 uncontrolled flows, and its bandwidth is set to be 100×100 pkts/s. We set the packet size to be 1000 bytes for controlled and uncontrolled sources in all simulations. The bandwidth and the propagation delay of each access link are set to be 10×100 pkts/s and 5 ms, respectively. We use 90 and 40 ms as the propagation delay of the bottleneck link (200 and 100 ms round-trip delay for the end sources). We assume that the bottleneck router has only marking functionality, i.e., there is no dropping of the packets due to the buffer overflow (see Section IV-B for additional discussion).

We use the equivalent marking functions for (simplified) RED and REM described in Section III-D as the AQM schemes, and we use two kinds of controlled sources, namely: 1) proportional fair controller [14] and 2) TCP Sack (the authors of [29] suggests the use of TCP Sack or TCP NewReno for network simulation and measurement). Proportional fair controlled source i is described by the following difference equation:

$$x^i[k+1] = x_n^i[k] + u\kappa\left(w - x^i[k-d]p_q\left(Q[k-d]\right)\right)$$

where u is the update interval, and d is the round-tip propagation delay. In our simulation, the update of the source rate is implemented by replacing $x_n^i[k-d]p_q\left(Q_n[k-d]\right)$ with N_k, which is the actual number of marks received over the update interval u. In our simulations, w and κ are set to be 5.5 and 1, respectively. In addition, we use a value of 200 ms as the update interval. All sources are started with small time differences in order to eliminate synchronization effects between the end-user systems. For notational convenience, we call the queue-based RED (REM) as

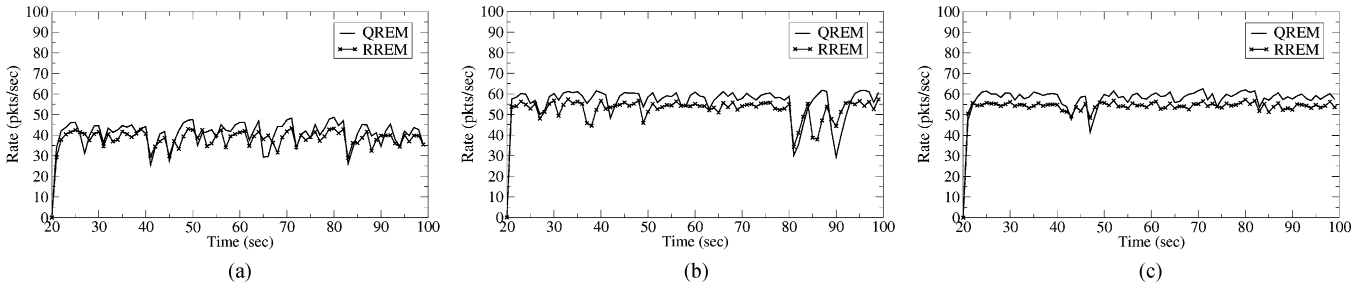


Fig. 4. Throughput of proportional fair source with REM. (a) $\lambda = 50$, ON-OFF(0.1), $\mathrm{rtt} = 200$ ms. (b) $\lambda = 35$, ON-OFF(0.2), $\mathrm{rtt} = 200$ ms. (c) $\lambda = 35$, ON-OFF(0.2), $\mathrm{rtt} = 100$ ms.

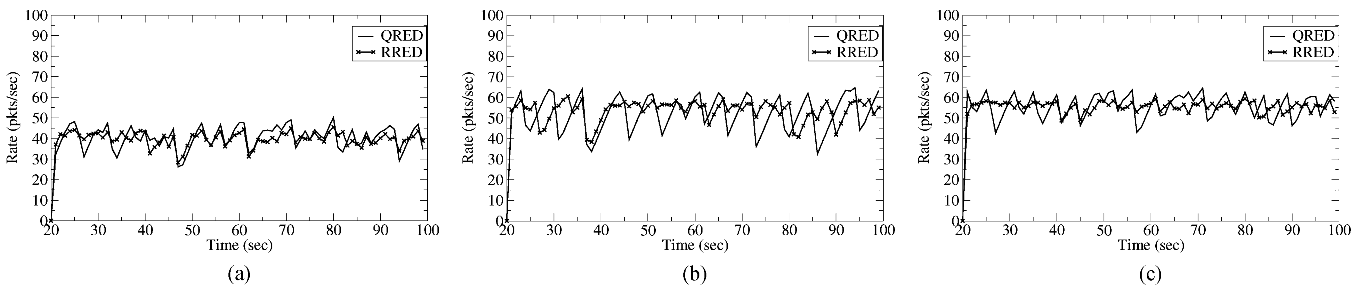


Fig. 5. Throughput of proportional fair source with RED. (a) $\lambda = 50$, ON-OFF(0.1), $\mathrm{rtt} = 200$ ms. (b) $\lambda = 35$, ON-OFF(0.2), $\mathrm{rtt} = 200$ ms. (c) $\lambda = 35$, ON-OFF(0.2), $\mathrm{rtt} = 100$ ms.

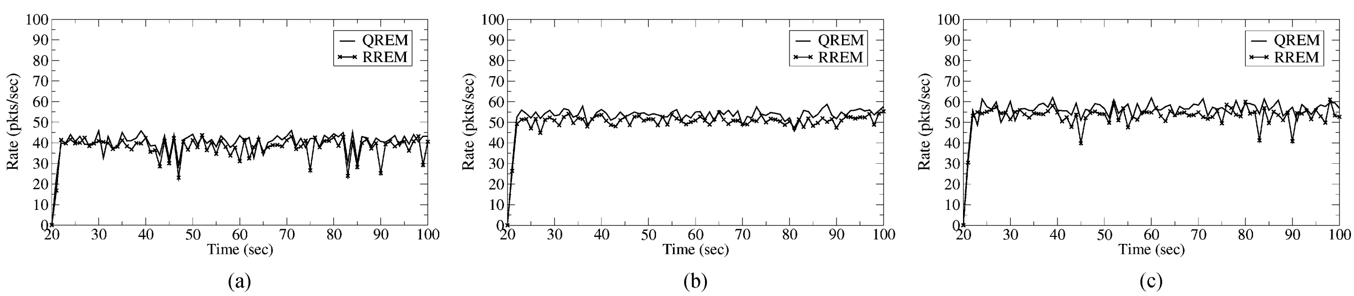


Fig. 6. Throughput of TCP with REM. (a) $\lambda = 50$, ON-OFF(0.1), $\mathrm{rtt} = 200$ ms. (b) $\lambda = 35$, ON-OFF(0.2), $\mathrm{rtt} = 200$ ms. (c) $\lambda = 35$, ON-OFF(0.2), $\mathrm{rtt} = 100$ ms.

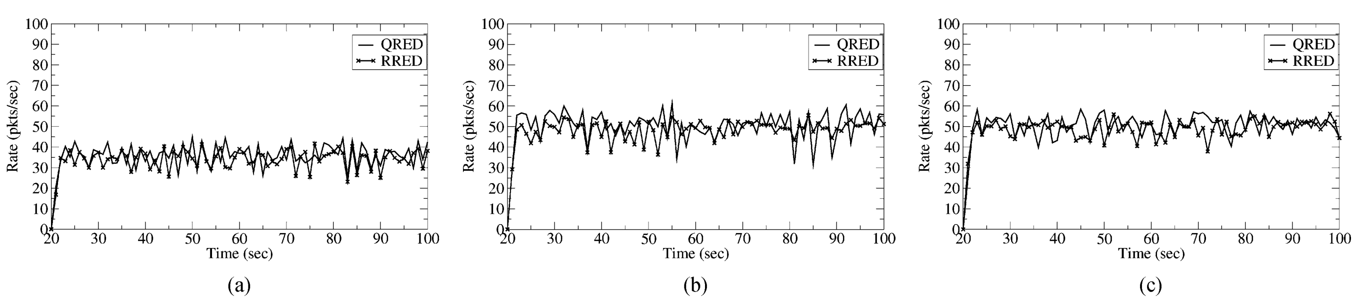


Fig. 7. Throughput of TCP with RED. (a) $\lambda = 50$, ON-OFF(0.1), $\mathrm{rtt} = 200$ ms. (b) $\lambda = 35$, ON-OFF(0.2), $\mathrm{rtt} = 200$ ms. (c) $\lambda = 35$, ON-OFF(0.2), $\mathrm{rtt} = 100$ ms.

QRED (QREM) and the equivalent rate-based marking scheme as RRED (RREM), respectively. In the simulations, we have used the following parameters for RED and REM: $\alpha = 0.02$ and $a = 10$, $b = 30$.

The uncontrolled flows are modeled by ON–OFF processes [30], where the ON and OFF periods are exponentially distributed with parameter 100 or 200 ms (denoted by ON-OFF(0.1) and ON-OFF(0.2)) and the packet transmission rate in the ON period is suitably set to be constant so that the total load due to the uncontrolled flows is a fixed fraction of the link capacity. We denote the average load due to a uncontrolled flow by λ pkts/s in the simulation results.

We have two kinds of figures (throughput for both sources and additionally congestion window size for TCP sources) to validate the equivalent rate-based marking function proposed in this paper. In the figures titled "throughput" (see Figs. 4–7 and 11), we measure and plot the aggregate instantaneous throughput (over all controlled sources) every 0.5 s, average them over

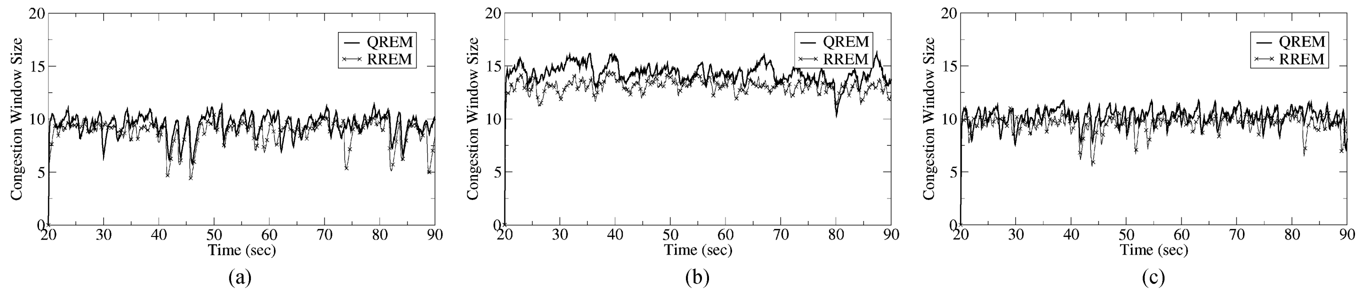


Fig. 8. Congestion window size of TCP with REM. (a) $\lambda = 50$, ON-OFF(0.1), $\mathrm{rtt} = 200$ ms. (b) $\lambda = 35$, ON-OFF(0.2), $\mathrm{rtt} = 200$ ms. (c) $\lambda = 35$, ON-OFF(0.2), $\mathrm{rtt} = 100$ ms.

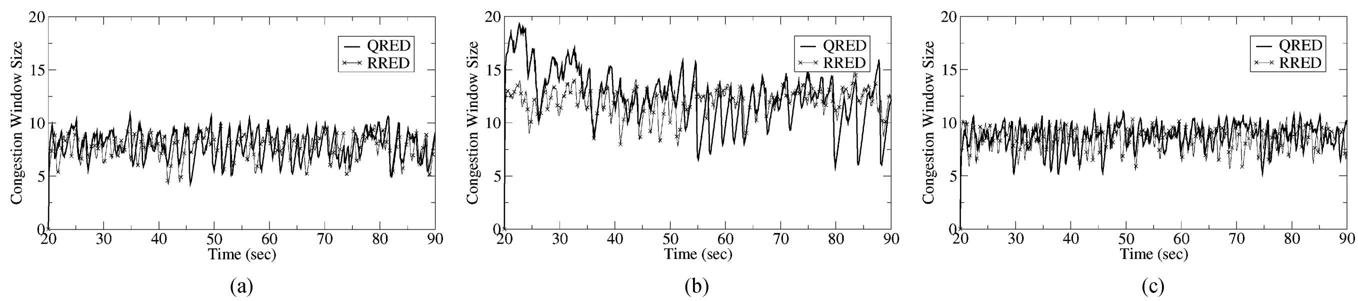


Fig. 9. Congestion window size of TCP with RED. (a) $\lambda = 50$, ON-OFF(0.1), $\mathrm{rtt} = 200$ ms. (b) $\lambda = 35$, ON-OFF(0.2), $\mathrm{rtt} = 200$ ms. (c) $\lambda = 35$, ON-OFF(0.2), $\mathrm{rtt} = 100$ ms.

the number of controlled sources, and plot the samples at one second intervals. In addition, we also plot the average rate over (horizontal line in the figures). In the figures for the congestion window size, we trace the instantaneous cwnd (congestion window size) value of all TCP sources, average them across the sources, (sampled every 1 s) and plot this as a time series (see Figs. 8 and 9). Finally, we also present the complementary distribution function of the congestion window size (cwnd) for a typical flow under different network conditions (see Fig. 10).

B. Implementation Issues

Prior to presenting the simulation results, we describe a few implementation issues. In the simulations in this section, we use a sliding window of time-step size to estimate the arrival rate from controlled and uncontrolled sources at the router. Time-step size is set to be 5 ms and the time-interval of the sliding window is chosen to be the round-trip time of the sources [31], i.e., the number of sliding window slots is equal to $\mathrm{rtt}/0.005$. We estimate the arrival rate by computing the instantaneous arrival rate over the time-step size and averaging the sliding window length of current and previous instantaneous arrival rates.

Further, the actual number of flows is not needed at the router to compute the marking probability based on the equivalent rate-based marking function. Computing the total controlled and uncontrolled rates (as opposed to the average arrival rate) is sufficient as the equivalent marking function is automatically "normalized."

In the simulations, the buffer size (also called queue limit) is set to be sufficiently large such that only marking functionality affects the transmission rate of controlled flows, i.e., physical dropping of the arriving packets does not occur due to queue overflow.

In practice, however, packet drops could occur due to finite buffers at routers. The equivalent rate-based model in this paper can be extended to such a finite queue length system by adding an equivalent rate-based dropping function (thus the intermediate router has a pair of probability (p_m, p_d), where p_m and p_d are marking and dropping probability computed based on the equivalent rate-based model), since the dropping function for a finite size of queue in a queue-based system is a step function. Further, this model can be extended to more complicated queue-based dropping function (e.g., RED) than a step function. However, in this paper, we restrict to an intermediate router with only marking functionality.

C. Experiment 1: Proportional Fair Controller

Figs. 4 and 5 show the average (over flows) instantaneous throughput as well as the long-term average over the entire simulation time (straight horizontal line in the figures) for the REM and RED controllers. As probabilistic marking is employed at the router (with both queue-based and equivalent rate-based marking), the starting times of flows and transmission pattern of uncontrolled flows are randomized, and the instantaneous rates will not be identical in a path-wise sense. However, a good match between the two schemes implies that the statistical behavior should be close to each other.

Both figures include the results for different network parameters such as the round-trip time, the load of uncontrolled flows, and the burstiness of uncontrolled flows. Through these results, we can compare the long-term behavior of a queue-based scheme and a rate-based scheme. The results show that there is about less then 5% difference between the rate-based and queue-based schemes. Further, the instantaneous rate show similar statistical path behavior.

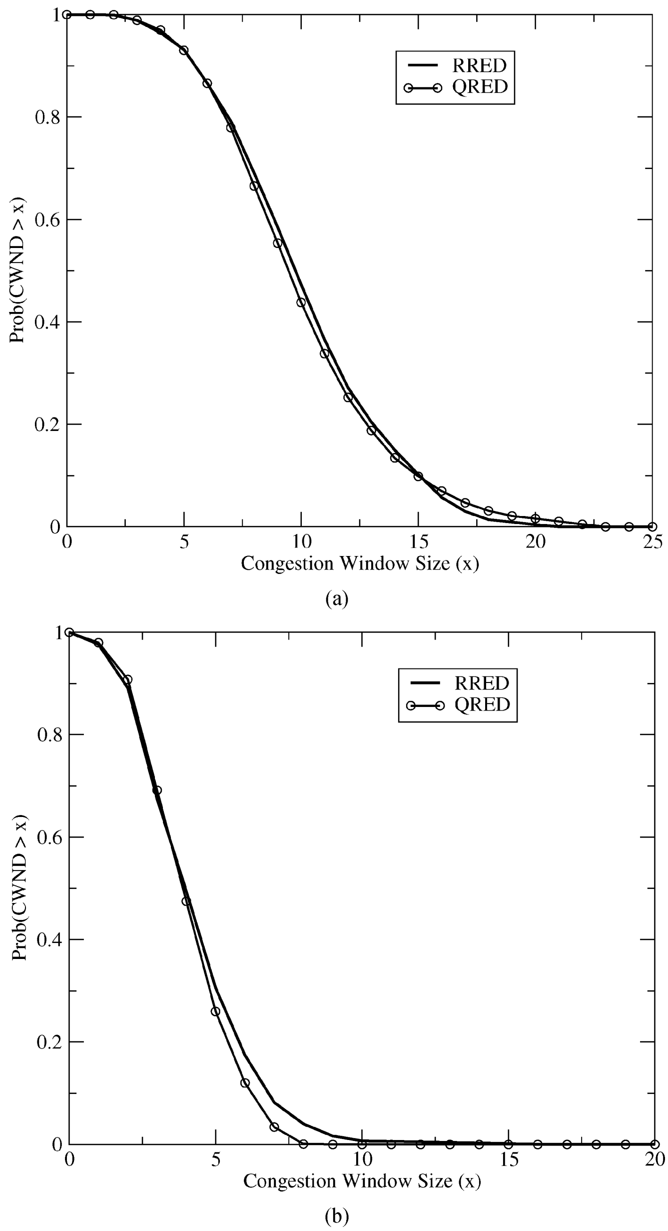


Fig. 10. Congestion window size distribution of a typical TCP source with RED: ON-OFF(0.1) and $\mathtt{rtt} = 200$ ms. (a) $\lambda = 35$. (b) $\lambda = 50$.

In this simulation, the arrival rate could exceed the service capacity. In this situation, while the equivalent rate-based marking marks all the packet, not all packets are marked in the practical packet systems. In addition to the Poisson approximation, we posit that one of the reasons for performance difference between the rate-based model and the queue-based model is due to this reason.

However, if the system scale is large enough, this effect becomes small for the following reason: when the arrival rate exceeds the capacity, the queue-length increases, leading to an unstable queue. Thus, as the system scale increases (indexed by n), the queue length is of order n, and thus an increasingly large fraction of the incoming packets will be marked (as the marking function is unscaled). This is supported by the simulation results that even 100 controlled and uncontrolled flows are enough to decrease the performance error to less than 5%.

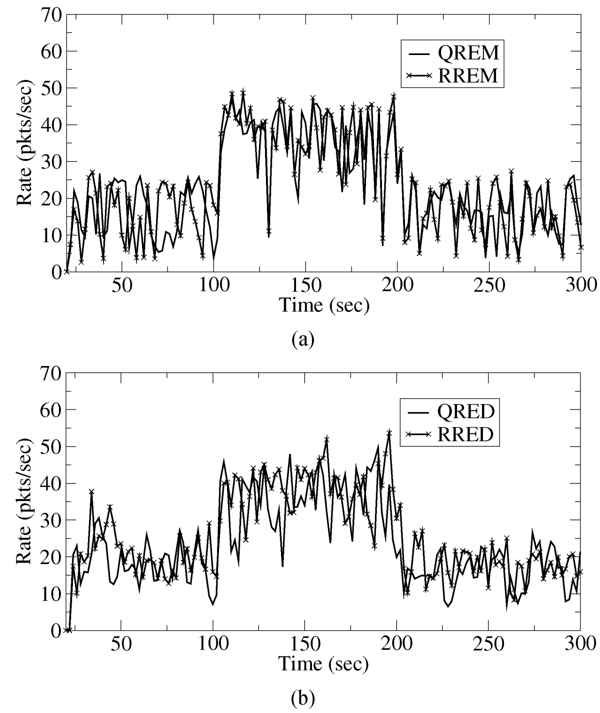


Fig. 11. Sensitivity to change of network connections. (a) Proportional fair, REM. (b) TCP, RED.

D. *Experiment 2: TCP Controller*

Figs. 6 and 7 plot the throughput with the REM and RED controllers respectively, and Figs. 8 and 9 plot the congestion windows (averaged over flows) as a time series. Further, Fig. 10 shows the complementary distribution function of the congestion window for a typical TCP flow, for two different loads and with the RED controller at the router. The plots indicate that the window sizes are statistically very similar for the queue-based marking and the equivalent rate-based marking. Thus, this indicates that replacing a queue-based marking function by its equivalent rate-based function does not statistically affect the end host's behavior.

E. *Experiment 3: Sensitivity to Change of Network Connections*

In this simulation, we study the sensitivity of the equivalent rate-based marking function to the change of network connections. The simulation starts with 100 uncontrolled and 200 controlled sources. After 100 s, the number of controlled sources is reduced to 100. Then from 200 s, the number of uncontrolled sources increases to 150. The bottleneck bandwidth maintained at 100×100 pkts/s and the mean transmission rate of an uncontrolled source is sustained, which means that the uncontrolled load after 200 s is changed to 75%. In Fig. 11, we can see the robust behavior of the proposed rate-based marking function even with the change of network connections.

V. Conclusion

In this paper, we have shown that the randomness due to short and unresponsive flows in the Internet is sufficient to decouple the dynamics of the router queues from those of the end controllers. This implies that a time-scale decomposition naturally occurs such that the dynamics of the router manifest only through their statistical steady-state behavior. This time-scale decomposition implies that a queue-length based marking function such as RED or REM have an equivalent form which depend only on the data arrival rate from the end-systems and do not depend on the queue dynamics. This leads to much simpler dynamics of the differential equation models (there is no queueing dynamics to consider), which enables low-complexity simulation and easier analysis.

Using packet-based simulations, we have studied queue-based marking schemes and their equivalent rate-based marking schemes for different types of controlled sources (proportional fair and TCP) and marking schemes (RED and REM). Our results indicate a good match in the rates observed at the intermediate router with the queue-based marking function and the corresponding rate-based approximation. Additionally, the simulation complexity is reduced due to the absence of queueing dynamics in the network simulation. Further, the window size distributions of a typical TCP flow with a queue-based marking function as well as the equivalent rate-based marking function match closely, indicating that replacing a queue-based marking function by its equivalent rate-based function does not statistically affect the end host's behavior.

Appendix I
Proof of Lemma 3.1

Proof: Let $J \triangleq J(a_n)$ be the number of jumps (number of arrivals) of the trajectory $a_n(t),\ t \in [0,T]$. It can be shown that, for n sufficiently large, $J < \infty$ almost surely, since $a_n(\cdot)$ converges to $a(.)$, which is a finite rate Poisson process, in the Skorohod topology (thus, for n sufficiently large, and over a compact time interval, the number of arrivals in $a_n(\cdot)$ is the same as the number of arrivals in $a(\cdot)$; see also the proof of Lemma 3.2). Let $\{t_n^j, \ , \ t_n^j \le T,\ j = 1, 2, \ldots, J\}$ be the jump times, respectively.

Next, let $\xi_n(r) = y_n(r) - x(0)r$. Then, we have

$$\begin{aligned} |\xi_n(t) - \xi_n(r)| &= \left| n \int_{r/n}^{t/n} (x_n(z) - x_n(0))\, dz \right| \\ &\le \left| n \int_{r/n}^{t/n} Mz\, dz \right| \\ &= \frac{nM}{2}\left(\left(\frac{t}{n}\right)^2 - \left(\frac{r}{n}\right)^2\right) \\ &= \frac{M}{n}(t^2 - r^2). \end{aligned} \quad (27)$$

Then, from the definition of $q_n(t)$ and $\tilde{q}_n(t)$, we have

$$\begin{aligned} q_n(t) = \sup_{r\in[0,t]} &[a_n(t) - a_n(r) + x(0)(t-r) + \xi_n(t) \\ &- \xi_n(r) - c(t-r) + q_n(r)] \\ \tilde{q}_n(t) = \sup_{r\in[0,t]} &[a_n(t) - a_n(r) + (t-r)x(0) \\ &- c(t-r) + \tilde{q}_n(r)]. \end{aligned}$$

Let $\Delta\tilde{q}_n(s) = |q_n(s) - \tilde{q}_n(s)|,\ s \in [0,T]$. Then, it suffices to show that for given $\epsilon > 0$, there exists N such that $\forall n > N$, we have

$$\sup_{s\in[0,T]} \Delta\tilde{q}_n(s) < \epsilon$$

since convergence with respect to uniform topology implies convergence with respect to Skorohod topology.

Consider any two jump times t_n^j and t_n^{j+1}. Then, we have

$$\begin{aligned} \Delta\tilde{q}_n(t_n^{j+1}) &\le \Delta\tilde{q}_n(t_n^j) + |\xi_n(t_n^{j+1}) - \xi_n(t_n^j)| \\ &\le \Delta\tilde{q}_n(t_n^j) + \frac{M}{n}(t_n^{j+1} + t_n^j)(t_n^{j+1} - t_n^j) \\ &\le \Delta\tilde{q}_n(t_n^j) + \frac{M}{n}(2T)(t_n^{j+1} - t_n^j). \end{aligned}$$

Further, for any $z_n^j \in [t_n^j, t_n^{j+1}]$, we have

$$\begin{aligned} \Delta\tilde{q}_n(z_n^j) &\le \Delta\tilde{q}_n(t_n^j) + \frac{M}{n}(2T)(z_n^j - t_n^j) \\ &\le \tilde{q}_n(t_n^j) + \frac{M}{n}(2T)(t_n^{j+1} - t_n^j). \end{aligned} \quad (28)$$

Thus, it is enough to check $\Delta\tilde{q}_n(s)$ when $s \in \{t_n^j, , t_n^j \le T, j = 1, 2, \ldots, J\}$. Noting that $\Delta\tilde{q}_n(0) = 0$ since $\tilde{q}_n(0) = q_n(0)$, we have

$$\begin{aligned} \Delta\tilde{q}_n(t_n^1) &\le \frac{M}{n}2T(t_n^1 - 0) \\ \Delta\tilde{q}_n(t_n^2) &\le \frac{M}{n}2T(t_n^1 - 0) + \frac{M}{n}2T(t_n^2 - t_n^1) \\ &= \frac{M}{n}2T(t_n^2) \\ &\ldots. \end{aligned}$$

By induction, we have

$$\Delta\tilde{q}_n(t_n^j) \le \frac{M}{n}2T(t_n^j) \le \frac{M}{n}2T^2. \quad (29)$$

From (28) and (29), the result follows. ■

Appendix II
Proof of Lemma 3.2

Proof: By definition of convergence in Skorohod topology, for a given $0 < \delta < 1$ and sufficiently large n, we can find a strictly increasing, continuous function λ_n of $[0,T]$ onto itself such that

$$\begin{aligned} \sup_{s\in[0,T]} |a_n(\lambda_n(s)) - a(s)| &< \delta \\ \sup_{s\in[0,T]} |\lambda_n(s) - s| &< \delta \end{aligned} \quad (30)$$

and, for n sufficiently large, we have

$$J \triangleq \mathcal{J}(a) = \mathcal{J}(a_n) \quad (31)$$

where $\mathcal{J}(\cdot)$ is the number of jumps over the space $\mathcal{D}([0,T] : \mathcal{R}^+)$. The fact that $a(s), s \in [0,T]$ is a Poisson process with a finite rate ensures that we have a finite number of jumps almost surely over the finite interval of time. In addition, as any "extra jump" would lead to a distance of 1 which contradicts condition (30), (31) follows. Let us denote the arrival times of a and a_n by $\{t^j, j = 1, \ldots, J\}$ and $\{t_n^j, j = 1, \ldots, J\}$, respectively. We know that the arrival times of $a(t)$ ($a_n(t)$) are equivalent to the jump times of $q(t)$ ($q_n(t)$). Also, we know that

$$\sup_{1 \le i \le J} |t_n^i - t^i| < \delta$$

for sufficiently large n from (30). For a function $\lambda_n(s)$ satisfying (30), we choose a piecewise linear function such that $\lambda_n(t^i) = t_n^i, i = 1, \ldots, J$. This construction implies that $\lambda_n(s)$ is a continuous and strictly increasing function over the interval $[0,T]$. To complete the proof, it suffices to show that $\sup_{s \in [0,T]} |q_n(\lambda_n(s)) - q(s)|$ is arbitrarily small. Let

$$\Delta q_n(s) = |q_n(\lambda_n(s)) - q(s)|.$$

Then, we have the following recurrence relation:

$$\Delta q_n(t^{j+1}) \le \Delta q_n(t^j) + \delta(c - x(0))$$

since $q_n(\lambda_n(s))$ and $q(s)$ has only one jump at time t^{j+1} in the interval $(t^j, t^{j+1}]$. Thus, the queue size difference is only that due to the amount of service with rate $c - x(0)$ over the time difference $|\lambda_n(s) - s|$. Further, for any $r \in [t^j, t^{j+1}]$, we have

$$\Delta q_n(r) = \max[\Delta q_n(t^j), \Delta q_n(t^{j+1})]$$

as $q_n(\cdot)$ and $q(\cdot)$ are piecewise linear between jumps. Thus, it is enough to check only the jump times of the process $q(s), s \in [0,T]$. Thus, defining $t^0 = 0$, and choosing n sufficiently large such that $|q_n(0) - q(0)| < \delta$, we have

$$\begin{aligned}\sup_{s \in [0,T]} |q_n(\lambda_n(s)) - q(s)| &\le \max_{0 \le j \le J} \Delta q_n(t^j) \\ &\le \delta \max\{1, J(c - x(0))\}.\end{aligned}$$

Choosing δ small enough, we are done. ∎

REFERENCES

[1] S. Floyd and V. Jacobson, "Random early detection gateways for congestion avoidance," *IEEE/ACM Trans. Netw.*, vol. 1, no. 4, pp. 397–413, Aug. 1993.

[2] F. Kelly, "Mathematical modeling of the Internet," in *Mathematics Unlimited – 2001 and Beyond*, B. Engquist and W. Schmid, Eds. Berlin, Germany: Springer-Verlag, 2001, pp. 685–702.

[3] S. H. Low, F. Paganini, and J. C. Doyle, "Internet congestion control," *IEEE Control Syst. Mag.*, vol. 22, no. 1, pp. 28–43, Feb. 2002.

[4] S. Deb, S. Shakkottai, and R. Srikant, "Stability and convergence of TCP-like congestion controllers in a many-flows," in *Proc. INFOCOM*, San Francisco, CA, 2003, pp. 884–894.

[5] S. Low, F. Paganini, J. Wang, S. Adlakha, and J. C. Doyle, "Dynamics of TCP/RED and a scalable control," in *Proc. IEEE INFOCOM*, New York, 2002, vol. 1, pp. 239–248.

[6] S. Shakkottai and R. Srikant, "How good are deterministic fluid models of Internet congestion control?," in *Proc. IEEE INFOCOM*, New York, Jun. 2002, vol. 2, pp. 497–505.

[7] V. Misra, W.-B. Gong, and D. Towsley, "Fluid-based analysis of a network of AQM routers supporting TCP flows with an application to RED," in *Proc. ACM SIGCOMM*, 2000.

[8] S. Kunniyur and R. Srikant, "End-to-end congestion control: Utility functions, random losses and ECN marks," in *Proc. IEEE INFOCOM*, Tel Aviv, Israel, Mar. 2000, vol. 3, pp. 1323–1332.

[9] ——, "Analysis and design of an adaptive virtual queue algorithm for active queue management," in *Proc. ACM SIGCOMM*, San Diego, CA, Aug. 2001, pp. 123–134.

[10] S. Athuraliya, V. H. Li, S. H. Low, and Q. Yin, "REM: Active queue management," *IEEE Network*, vol. 15, no. 3, pp. 48–53, May/Jun. 2001.

[11] P. Tinnakornsrisuphap and A. Makowski, "Limit behavior of ECN/RED gateways under a large number of TCP flows," in *Proc. IEEE INFOCOM*, San Francisco, CA, Apr. 2003, pp. 873–883.

[12] J. Cao and K. Ramanan, "A poisson limit for buffer overflow probabilities," in *Proc. IEEE INFOCOM*, New York, Jun. 2002, vol. 2, pp. 994–1003.

[13] M. Mandjes and J. H. Kim, "Large deviations for small buffers: An insensitivity result," *Queueing Syst.*, vol. 37, pp. 349–362, 2001.

[14] F. P. Kelly, "Models for a self-managed Internet," *Philos. Trans. Roy. Soc. A*, vol. 358, pp. 2335–2348, 2000.

[15] S. Deb and R. Srikant, "Rate-based versus queue-based models of congestion control," in *Proc. ACM SIGMETRICS*, 2004.

[16] *Ns-2*, [Online]. Available: http://www.isi.edu/nsnam/ns/

[17] Parallel and Distributed NS PDNS [Online]. Available: http://www.cc.gatech.edu/computing/compass/pdns/

[18] *GloMoSim*, [Online]. Available: http://pcl.cs.ucla.edu/projects/glomosim/

[19] *QualNet*, [Online]. Available: http://www.scalable-networks.com

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

[21] D. Daley and D. Vere-Jones, *An Introduction to the Theory of Point Processes*. New York: Springer-Verlag, 1988.

[22] S. Shakkottai and R. Srikant, "Mean FDE models for Internet congestion control under a many-flows regime," *IEEE Trans. Inf. Theory*, vol. 50, no. 6, pp. 1050–1072, Jun. 2004.

[23] P. Billingsley, *Convergence of Probability Measures*. New York: Wiley-Interscience, 1999.

[24] W. L. Smith, "Regenerative stochastic processes," in *Proc. Roy. Soc. London A*, 1955, vol. 232, pp. 6–31.

[25] S. I. Resnick, *Adventures in Stochastic Processes*. Boston, MA: Birkhauser, 1992.

[26] R. W. Wolff, *Stochastic Modeling and the Theory of Queues*. Englewood Cliffs, NJ: Prentice-Hall, 1989.

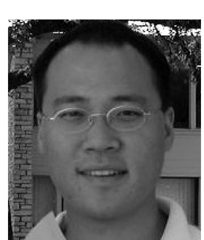
[27] U. M. J. Roberts and J. Virtamo, Broadband Network Teletraffic, Final Report of Action COST 242 Birkhauser, Boston, MA, 1992.

[28] J. Virtamo, "Numerical evaluation of the distribution of unfinished work in an M/D/1 system," *Electron. Lett.*, vol. 31, no. 7, pp. 531–532, 1995.

[29] S. Floyd, "Thoughts on the evolution of TCP in the Internet," in *Proc. 2nd Int. Workshop Protocols for Fast Long-Distance Networks*, 2004.

[30] C. V. Hollot, Y. Liu, V. Misra, and D. Towsley, "Unresponsive flows and AQM performance," in *Proc. INFOCOM*, San Francisco, CA, Apr. 2003, vol. 1, pp. 85–95.

[31] H. Kim and J. C. Hou, "Network calculus based simulation for TCP congestion control: Theorems, implementation and evaluation," in *Proc. IEEE INFOCOM*, Mar. 2004, vol. 4, pp. 2844–2855.



Yung Yi (S'03) received the B.S.E. and M.S.E. degrees from the School of Computer Science and Engineering, Seoul National University, Seoul, Korea, in 1997 and 1999, respectively. He is currently working toward the Ph.D. degree at the Department of Electrical and Computer Engineering, The University of Texas at Austin, Austin.

His current research interests include scheduling and QoS for wireless and ad hoc networks, congestion control in the Internet, design and performance analysis of sensor networks, and resource allocation for heterogeneous networks.

Supratim Deb received the B.E. degree in electrical and electronics engineering from the Birla Institute of Technology and Science, Pilani, India, in 1997, the M.E. degree in telecommunication from the Indian Institute of Science, Bangalore, India, in 1999, and the Ph.D. degree in electrical engineering from the University of Illinois at Urbana-Champaign in 2003.

He was with the IBM India Research Lab during 1999–2000. Following his doctoral work, he was a Post-Doctoral Associate with the Massacusetts Institute of Technology, Cambridge, before joining Bell-Labs Research India, Bangalore, in February 2005. His research interests include Internet congestion control, resource allocation in wireless and wireline networks, network coding, and distributed systems.

Sanjay Shakkottai (M'01) received the Ph.D. degree from the University of Illinois at Urbana-Champaign in 2002.

He is currently with The University of Texas at Austin, where he is an Assistant Professor with the Department of Electrical and Computer Engineering. He was the finance chair of the 2002 IEEE Computer Communications Workshop in Santa Fe, NM. His research interests include wireless and sensor networks, stochastic processes and queueing theory.

Dr. Shakkottai was the recipient of the National Science Foundation CAREER Award in 2004.