

# Bifurcations of TCP and UDP Traffic Under RED

Priya Ranjan, Richard J. La, and Eyad H. Abed

*Abstract*—

Recently researchers have proposed active queue management (AQM) mechanisms as a means of better managing congestion at the bottlenecks inside the network. Random Early Detection (RED) mechanism has been proposed to control the average queue size at the congested routers. It has been shown that the interaction between an RED gateway and TCP connections can lead to period-doubling bifurcation and chaos. In this paper we extend this model and study the interaction of the RED gateway with TCP and UDP connections, using a discrete-time model.

We show that the presence of UDP traffic fundamentally changes the dynamics of the system. Second, with the help of bifurcation diagrams, we demonstrate the existence of nonlinear phenomena, such as oscillations and chaos, as the parameters of the RED mechanism are varied. We present numerical examples to validate our analysis.

## I. INTRODUCTION

With the unprecedented growth of the Internet both in its size and the number of subscribers the problem of congestion control is emerging as a more crucial problem in order to avoid a congestion collapse of the Internet as in 1980's. Researchers have proposed various approaches to address this issue. One such approach is to control the congestion level at each router through an active queue management (AQM) mechanism. There are various forms of AQM mechanisms that have been proposed in the past [7], [16], [12], [2], [11]. Although these AQM mechanisms are designed with different objectives in mind, one common objective is to monitor the congestion level and provide the information to the end hosts so that any incipient congestion can be avoided without a significant degradation of network performance.

The RED mechanism, proposed by Floyd and Jacobson, attempts to control the congestion level at a bottleneck by monitoring and updating the average queue size. Although the RED mechanism is conceptually very simple and easy to understand, its interaction with Transmission Control Protocol (TCP) connections has been shown to be rather complex and is somewhat poorly understood. Hollot *et al.* have developed a linearized model of a RED gateway with TCP connections and characterized the stability region of the system. Further, they have proposed Pro-

portional and Proportional-Integral controllers to improve the responsiveness and stability of the RED mechanism, using a similar linearized model [11], [10]. Ranjan *et al.* have used a simple nonlinear model to investigate the behavior of a simple network with a RED gateway with TCP connections [25]. They have demonstrated that such a system leads to nonlinear phenomena, such as oscillations and chaos, if the system parameters are not selected carefully [25]. It has been shown that the initial oscillation is a consequence of a smooth bifurcation of the stable equilibrium point of the system, while the latter chaos is caused by border collision bifurcations. Such phenomena were not captured in any of the previous studies.

In this paper we extend the model in [25] and study the interaction of the RED mechanism with both TCP and User Datagram Protocol (UDP) connections. This model is used to explain the instability, bifurcation, and routes to chaos of the system. First, we demonstrate that the presence of UDP traffic has a fundamental impact on the dynamics of the system. Second, we show the existence of both smooth and non-smooth bifurcations as control parameters are varied.

The rest of the paper is organized as follows. Section II presents the nonlinear discrete-time model that is used for our analysis. Section III discusses the methodology for our analysis, which is followed by the study of stability conditions of the system. Numerical examples are presented in section V.

## II. DISCRETE-TIME MODEL

We consider a simple network of single link, which is shared by many connections. Let  $\mathcal{I}, \mathcal{I} = \{1, \dots, N\}$ , denote the set of connections. Each connection is assumed to be either a TCP or UDP connection. Throughout this paper we assume that all TCP connections are TCP Reno connections. The capacity of the shared link is denoted by  $C$ , and the round-trip propagation delay (without any queuing delay) of connection  $i \in \mathcal{I}$  is given by  $d_i$ . We denote the rate or throughput of connection  $i$  by  $x_i$ , and the packet

size by  $M$ .<sup>1</sup> We assume that the Random Early Detection (RED) queue management mechanism is implemented at node  $r1$  in order to control the average queue size at the router. A RED gateway drops or marks a packet with a probability  $p$ , which is a function of the average queue size  $q^{ave}$  as follows [7]:<sup>2</sup>

$$p(q^{ave}) = \begin{cases} 0 & \text{if } q^{ave} < q_{min} \\ 1 & \text{if } q^{ave} > q_{max} \\ \frac{q^{ave} - q_{min}}{q_{max} - q_{min}} p_{max} & \text{otherwise} \end{cases} . \quad (1)$$

The average queue size is updated at the time of packet arrival according to the exponential averaging:

$$q_{new}^{ave} = (1 - w)q_{old}^{ave} + w \cdot q_{cur} , \quad (2)$$

where  $q_{cur}$  is the queue size at the time of arrival, and  $w$  is the exponential averaging weight, which determines the time constant of the averaging mechanism. Therefore, the control parameters of the RED mechanism are  $w$ ,  $q_{min}$ ,  $q_{max}$ , and  $p_{max}$ .

The network with an AQM mechanism can be modeled as a feedback system, where sources adjust their transmission rates based on the feedback from the AQM mechanism in the form of marked or dropped packets [8], [7]. For instance, if Explicit Congestion Notification (ECN) mechanism is implemented, the RED gateway marks the packet by setting the ECN bit in the IP header of the packet if the transport layer is ECN capable. This is indicated in the packet through ECN Capable Transport (ECT) bit in the IP header. If the source is not ECN capable, the RED gateway drops the packet [8].

The connections are assumed to be long-lived connections, and the set of connections remains fixed for the time period of interest. In order to have a tractable model we assume that all connections have the same round-trip propagation delay, i.e.,  $d_i = d$  for all  $i \in \mathcal{I}$ . Rather than interpreting this assumption as a requirement that the connections must have the same propagation delay, one should consider the delay  $d$  as the effective delay that represents the overall propagation delay of the connections as shown in [19]. This allows us to reduce the problem with  $N$  connections to a single connection system that represents the set of connections and study its behavior.

<sup>1</sup>For the simplicity of analysis we assume that all connections use the same packet size.

<sup>2</sup>In practice a RED gateway drops a packet with a modified probability in order to lead to a more uniform dropping pattern [7].

TCP adjusts its sending rate depending on whether it has detected a packet drop or not. Hence, this process can be modeled as a stroboscopic map where the instant of observation is approximately one round-trip time. This technique has been utilized before for different clocked systems in power electronics for modeling the dynamics of power converters [?]. Since the AQM mechanism should allow enough time for the connections to react to control actions before the control action changes significantly, it is natural to model the system as a discrete-time system. Although one would prefer that the sampling interval be regular, there are models where the dynamics is sampled at irregular intervals and the resulting maps are known as ‘‘impact maps’’ [5].

Given the round-trip time (RTT)  $R$  and marking/drop probability  $p$ , the stationary throughput of a TCP Reno connection can be approximated by

$$T(p, R) = \frac{MK}{\sqrt{p}R} , \quad (3)$$

where  $K$  is some constant in  $[1, \sqrt{\frac{8}{3}}]$  [25], [21], [9], [23]. In this paper we assume  $K = \sqrt{\frac{8}{3}}$ . The exact value of  $K$  is not crucial to our analysis. Although (3) may seem like a crude approximation, if the RED gateway works as intended and results in a relatively uniform packet drops, it yields a good enough approximation for our qualitative results. This is demonstrated through *ns-2* simulation results in [19]. We use this simple approximation for TCP throughput to facilitate our analysis. However, one will see that our qualitative results do not depend on this particular form of the approximation, and are consequences of rather benign nonlinear behavior of TCP [25].

We use a nonlinear dynamic first-order discrete-time model to analyze the interaction of the RED gateway with a mixture of TCP and UDP connections. We define the control system as follows. The packet marking/drop probability  $p_k$  at period  $k$ ,  $k \geq 1$ , determines the throughput of the connections and the queue size  $q_{k+1}$  at period  $k + 1$ , based on the system constraints. The queue size at period  $k + 1$  is used to compute the average queue size  $q_{k+1}^{ave}$  at period  $k + 1$  according to the exponential averaging rule in (2). Then, The average queue size  $q_{k+1}^{ave}$  is used to calculate the packet marking/drop probability  $p_{k+1}$  at period  $k + 1$ , which is the control variable of the AQM mechanism. This can be mathematical written as fol-

lows:

$$q_{k+1} = G(p_k) \quad (4)$$

$$q_{k+1}^{ave} = A(q_k^{ave}, q_{k+1}) \quad (5)$$

$$p_{k+1} = H(q_{k+1}^{ave}), \quad (6)$$

where  $A(q_k^{ave}, q_{k+1}) = (1-w)q_k^{ave} + w \cdot q_{k+1}$  as given in (2) and the control function  $H(q_{k+1}^{ave}) = p(q_{k+1}^{ave})$  in (1).

The exact form of the plant function  $G(\cdot)$  depends on the system parameters such as the number of connections, nature of connections, round-trip delays, etc. In the following section we briefly summarize the results shown for the case with only TCP connections, and in section III we investigate the case where TCP connections coexist with UDP connections.

### III. INTERACTION OF A RED GATEWAY WITH TCP AND UDP CONNECTIONS

In this section we study a single-link network with a mixture of TCP and UDP connections, using the discrete-time system model presented in the previous section. With the help of bifurcation diagrams, we show that the nonlinear phenomena exhibited by the network with TCP connections, such as oscillations and chaos, persist and that increasing load of UDP traffic tends to stabilize the system [25]. Throughout the rest of the paper, without loss of generality, we assume that TCP packets are dropped rather than marked, i.e., TCP connections are not ECN capable.

The idea behind the RED mechanism is to control the average queue size without reacting too sharply to transient congestion at the router. Hence, the exponential averaging weight should be chosen sufficiently small so that the average queue size  $q^{ave}$  does not fluctuate very much due to transient, temporary fluctuations in the arrival rate. This implies that the time constant determined by the exponential averaging weight should be at least the effective round trip time of the connections in order to avoid a fast oscillation in the packet drop probability. Tinnakornsriruphap and Makowski have shown that, with a very small feedback delay, the average queue size per connection approaches some deterministic process as the number of TCP connections increases [24]. This may be thought of as a consequence of being better able to control the system dynamics due to the finer granularity of feedback information the RED mechanism can provide with the increasing number of connections. Another example where finer granularity of feedback

improves the responsiveness and convergence of the system is given in [15]. Therefore, with a large number of connections it is reasonable to take a time-scale decomposition approach and assume that the TCP connections' dynamics operate at a much faster time-scale than the evolution of the average queue size and that the average throughput of the connections sees a quasi-stationary behavior before the average queue size changes much. Hence, a period in the discrete-time model described below may represent a round-trip time or a larger period so that the connections can see the quasi-stationary throughput. A similar time-scale decomposition approach is taken in [17].

Let  $\mathcal{I}_T$  and  $\mathcal{I}_U = \mathcal{I} \setminus \mathcal{I}_T$  denote the set of TCP and UDP connections, respectively. Each UDP connection generates a packet at a constant rate, i.e., a UDP sender is modeled as a constant-bit-rate (CBR) source. We denote the transmission rate of UDP connection  $j \in \mathcal{I}_U$  by  $\lambda_j$ . Let  $\lambda = \sum_{j \in \mathcal{I}_U} \lambda_j$ . For the stability of the system we assume that  $\lambda < C$ . The throughput of UDP connection  $j$  is given by  $\lambda_j(1-p)$ , where  $p$  is the packet drop probability.

We now describe the discrete-time model that is used for analyzing the system. Suppose that  $p_k$  denotes the packet drop probability at period  $k$ ,  $k \geq 1$ . Based on this packet drop probability  $p_k$  at period  $k$ , one can compute the queue size  $q_{k+1}$  at period  $k+1$  as follows. First, one can compute the steady state packet drop probability  $p_u$  such that

$$\begin{aligned} \sum_{i \in \mathcal{I}_T} T(p_u, d) + \lambda(1-p_u) \\ = N_T \cdot T(p_u, d) + \lambda(1-p_u) = C. \end{aligned} \quad (7)$$

This probability is the smallest probability that results in a queue size of zero at the next period, and for all  $p > p_u$ , the queue size is zero at the next period. Hence, if  $p_k \geq p_u$ , we know that the throughput of the TCP connections is given by  $\frac{MK}{\sqrt{p_k d}}$  and the queue size at period  $k+1$  is zero, i.e.,  $q_{k+1} = 0$ . It is easy to see that if  $p_k < p_u$ , then the queue size will be strictly positive at the next period. The solution to (7) is the square of the positive, real solution of the following third-order polynomial:

$$\lambda dy^3 + d(C - \lambda)y - N_T MK = 0. \quad (8)$$

The positive, real solution to (8) is given by

$$y = \left( \frac{((108e_1^3 + 729e_2^2)^{1/2} - 27e_2)}{54} \right)^{\frac{1}{3}}$$

$$- \left( \frac{2e_1^3}{(108e_1^3 + 729e_2^2)^{1/2} - 27e_2} \right)^{\frac{1}{3}}, \quad (9)$$

where  $e_1 = \frac{C-\lambda}{\lambda}$  and  $e_2 = -\frac{N_T MK}{d\lambda}$ . One can find the corresponding queue size  $q_u^{ave}$  such that for any  $q_k \geq q_u^{ave}$ ,  $q_{k+1}$  is identically zero:

$$q_u^{ave} = \begin{cases} \frac{p_u(q_{max} - q_{min})}{p_{max}} + q_{min} & \text{if } p_{max} \geq p_u \\ q_{max} & \text{otherwise} \end{cases}.$$

Note that the introduction of UDP traffic does much more than simply reducing the capacity available for TCP connections by  $\lambda$ . This is clear from equation (8). In the absence of UDP traffic, i.e.,  $\lambda = 0$ , equation (8) reduces to a very simple first-order, linear function. This will be more clear shortly.

Suppose for the moment that the buffer size  $B$  is infinite. Then, given  $p_k < p_u$  one can compute the queue size  $q_{k+1}$  at period  $k+1$  as the solution of the following equation:

$$\frac{MK}{\sqrt{p_k}(d + \frac{q_{k+1}M}{C})} = \frac{C - \lambda(1 - p_k)}{N_T}. \quad (10)$$

The interpretation of (10) is as follows. Since the aggregate throughput of the UDP connections is  $\lambda(1 - p_k)$ , the remaining available capacity,  $C - \lambda(1 - p_k)$ , is equally divided among the TCP connections from the symmetry. Hence, the throughput of a TCP connection is given by  $T(p_k, R(q_{k+1})) = \frac{MK}{\sqrt{p_k}(d + \frac{q_{k+1}M}{C})} = \frac{C - \lambda(1 - p_k)}{N_T}$ . Hence, the queue size  $q_{k+1}$  is given by

$$q_{k+1} = \frac{C}{M} \left( \frac{MKN_T}{\sqrt{p_k}(C - \lambda(1 - p_k))} - d \right) \quad (11)$$

Now suppose that we have a finite buffer of size  $B$ . From (11) one can see that  $q_{k+1}$  is a strictly decreasing function of  $p_k$  and, hence, can compute the largest  $p_k$  such that the queue size  $q_{k+1}$  equals the buffer size  $B$ , i.e.,

$$p_l = \arg \sup_{p_k \in (0,1)} \left( \frac{C}{M} \left( \frac{MKN_T}{\sqrt{p_k}(C - \lambda(1 - p_k))} - d \right) \geq B \right) \quad (12)$$

One can show that this probability, which we denote by  $p_l$ , is the positive, real solution of the following equation.

$$\lambda p^{3/2} + (C - \lambda)p^{1/2} - \frac{CKN_TM}{BM + Cd} = 0. \quad (13)$$

The solution to (13) is given by the square of (9) with  $e_1 = \frac{C-\lambda}{\lambda}$  and  $e_2 = -\frac{CKN_TM}{(BM+Cd)\lambda}$ . Also, we can find the corresponding queue size  $q_l^{ave}$ :

$$q_l^{ave} = \frac{p_l(q_{max} - q_{min})}{p_{max}} + q_{min}.$$

It is obvious that for all  $p_k \leq p_l$ , i.e.,  $q_k \leq q_l^{ave}$ , we have  $q_{k+1} = B$ . From (7), (11), and (12) we have

$$G(p_k) = \begin{cases} 0, & \text{if } p_k > p_u \\ B, & \text{if } p_k < p_l \\ \frac{CN_T K}{\sqrt{p_k}(C - \lambda(1 - p_k))} - \frac{Cd}{M}, & \text{otherwise} \end{cases} \quad (14)$$

From (4) through (6) and (14) we define a mapping

$$\begin{aligned} q_{k+1}^{ave} &= (1 - w)q_k^{ave} + w \cdot A(G(H(q_k^{ave}))) \\ &= \begin{cases} (1 - w)q_k^{ave} & \text{if } q_k^{ave} \geq q_u^{ave} \\ (1 - w)q_k^{ave} + w \cdot B & \text{if } q_k^{ave} \leq q_l^{ave} \\ (1 - w)q_k^{ave} & \\ \quad + w \cdot \left( \frac{CN_T K}{\sqrt{p_k}(C - \lambda(1 - p_k))} - \frac{Cd}{M} \right) & \text{otherwise} \end{cases} \\ &:= g(q_k^{ave}, \rho), \end{aligned} \quad (15)$$

where  $\rho$  summarizes the system parameters, including the exponential averaging weight  $w$ , and  $p_k = \frac{q_k^{ave} - q_{min}}{q_{max} - q_{min}} p_{max}$  from (1). This maps the average queue size at period  $k$  to the average queue size at period  $k+1$ .

#### IV. STABILITY OF THE SYSTEM

A fixed point of the mapping  $g(\cdot)$  is an average queue size  $q^*$  such that  $q^* = g(q^*, \rho)$ . Solving (15) for such a fixed point  $q^*$ , if there exists any, leads to a fifth order polynomial, which does not depend on the exponential averaging weight  $w$  because neither the ‘‘queue law’’ nor the ‘‘feedback control law’’ is a function of  $w$ . The corresponding probability  $p^*$  of the fixed point  $q^*$  is given as the square of the positive, real solution of the following polynomial:

$$\begin{aligned} \frac{\lambda M}{\alpha} y^5 + \left( \frac{CM}{\alpha} + \lambda dC + \lambda M q_{min} - \frac{\lambda M}{\alpha} \right) y^3 \\ + (CM q_{min} - \lambda dC - \lambda M q_{min} + dC^2) y \\ - N_T MKC = 0, \end{aligned} \quad (16)$$

where  $\alpha = \frac{p_{max}}{q_{max} - q_{min}}$ . Further, one can show that when  $\lambda = 0$ , i.e., there are only TCP connections, and

the fifth order polynomial in (16) reduces to a third order polynomial given in [25], which has a closed form solution.

Throughout the rest of the paper we assume that  $p_{max} > p_l$ , which is the most interesting case from the application point of view. We normalize the queue sizes by the buffer size  $B$  so that the map is defined on the unit interval, i.e.,  $[0, 1] \mapsto [0, 1]$ . Normalization scheme reduces the number of parameters needed to completely characterize the dynamical behavior of the system. It also provides insight into how ratios of certain parameters affect the system rather than their absolute values. This will modify the system equations as follows:

$$\begin{aligned}
q_k^n &:= \frac{q_k^{ave}}{B} \\
\gamma &:= \frac{q_{max} - q_{min}}{p_{max} B} > 0 \\
q_{min}^n &:= \frac{q_{min}}{B} \\
b_2^n &:= \frac{q_l^{ave}}{B} \\
b_1^n &:= \frac{q_u^{ave}}{B} \\
q_{k+1}^n &= \begin{cases} (1-w)q_k^n & \text{if } q_k^n > b_1^n \\ (1-w)q_k^n + w & \text{if } q_k^n < b_2^n \\ (1-w)q_k^n + w \left( \frac{CN_T K}{B\sqrt{p_k}(C-\lambda(1-p_k))} \right) & \text{otherwise} \\ -\frac{Cd}{MB} & \text{otherwise} \end{cases} \quad (17)
\end{aligned}$$

where  $p_k = \frac{q_k^n - q_{min}^n}{\gamma}$ . Equation (17) maps unit interval into itself. Using partial fraction (17) can be rewritten as

$$\begin{aligned}
q_{k+1}^n &= \begin{cases} (1-w)q_k^n & \text{if } q_k^n > b_1^n \\ (1-w)q_k^n + w & \text{if } q_k^n < b_2^n \\ (1-w)q_k^n + w \left( \frac{CN_T K}{B(C-\lambda)} \left( \frac{1}{\sqrt{p_k}} - \frac{\lambda\sqrt{p_k}}{(C-\lambda(1-p_k))} \right) - \frac{Cd}{MB} \right) & \text{otherwise} \end{cases} \\
&:= f(q_k^n, \rho) \quad (18)
\end{aligned}$$

Stability of a map is determined by its eigenvalue. For (18) it is given as:

$$\begin{aligned}
\frac{\partial f(q_k^n, \rho)}{\partial q_k^n} \Big|_{q_k^n = q^{n*}} &= \frac{\partial f(q_k^n, \rho)}{\partial p_k} \frac{\partial p_k}{\partial q_k^n} \Big|_{q_k^n = q^{n*}} \\
&= 1 - w + \frac{wCN_T K}{\gamma B(C-\lambda)} \\
&\quad \left( -\frac{1}{p_k^{3/2}} - \frac{\lambda(C-\lambda(1+p_k))}{2\sqrt{p_k}(C-\lambda(1-p_k))^2} \right)
\end{aligned}$$

$$\begin{aligned}
&:= 1 - w + \frac{wCN_T K}{\gamma B(C-\lambda)} (I_1 + I_2) \\
&:= \mu(\rho) \quad (19)
\end{aligned}$$

where  $I_1 = -\frac{1}{p_k^{3/2}}$  and  $I_2 = -\frac{\lambda(C-\lambda(1+p_k))}{2\sqrt{p_k}(C-\lambda(1-p_k))^2}$ . These expressions will later be used to prove the piecewise monotonicity of Mixed-RED which in turn provides information about the possible bifurcation behavior of the system. Now we can formulate the linear stability criterion ( $|\lambda(\rho)| < 1$ ) [14] as follows:

$$\left| 1 - w - \frac{wCN_T K}{\gamma B(C-\lambda)} \times \left( \frac{1}{p_k^{3/2}} + \frac{\lambda(C-\lambda(1+p_k))}{2\sqrt{p_k}(C-\lambda(1-p_k))^2} \right) \right| < 1 \quad (20)$$

We look for parameter values for which eigenvalue given by (19) becomes  $-1$  and gives rise to a Period Doubling bifurcation (PDB) leading to first oscillatory behavior in the system. Further, we compute the nature of emerging period doubling bifurcation in terms of second and the third derivatives of the normalized map. A function  $S = \left( \frac{1}{2} \left( \frac{\partial^2 f}{\partial q_k^{n2}} \right)^2 + \frac{1}{3} \left( \frac{\partial^3 f}{\partial q_k^{n3}} \right) \right)$  of the second and the third derivatives of a map are used to determine the nature of a period doubling bifurcation (see [13], pp.158). Positive  $S$  implies that the bifurcation is supercritical, and a negative  $S$  implies a subcritical bifurcation, which can be potentially dangerous due to the discontinuous appearance of the period two orbit. We avoid the explicit expression for  $S$  but it is something which should be kept in mind while designing a Mixed-RED system to avoid any unexpected oscillations in router queues. We have shown earlier that in case of TCP-RED this can be an important concern [26].

## V. NUMERICAL EXAMPLES

In this section we take a more careful look at the map  $f(\cdot)$  given by eq. 18. As the eigenvalue in (19) gets closer to the unit circle when a control parameter, e.g. exponential averaging weight  $w$ , is varied, the fixed point becomes unstable. Depending on the nature of bifurcation there can be new fixed points with a larger period or chaos. If the fixed point collides with either border  $b_1^n$  or  $b_2^n$ , it may lead to border collision bifurcation [25].

A bifurcation diagram shows the qualitative changes in the nature or the number of fixed points

of a dynamical system with varying parameters. In this section we vary the exponential averaging weight and study the stability of the system with varying UDP traffic load. However, similar results can be obtained by varying another control parameter such as  $q_{min}$ . The  $x$ -axis is the control parameter that is being varied, and the  $y$ -axis plots the set of fixed solutions (with a period of one or higher) corresponding to the value of the control parameter. For generating the bifurcation diagrams, in each run we randomly select four random initial average queue sizes,  $q_1^{ave}(0)$ ,  $q_2^{ave}(0)$ ,  $q_3^{ave}(0)$ , and  $q_4^{ave}(0)$ , and these average queue sizes evolve according to the map  $g(\cdot)$  in (15), i.e.,

$$q_i^{ave}(k) = g(q_i^{ave}(k-1), \rho), \text{ for } k = 1, \dots, 1,000$$

and  $i = 1, 2, 3, \text{ and } 4$ .

We plot  $q_i^{ave}(k)$ ,  $k = 991, \dots, 1,000$  and  $i = 1, 2, 3, \text{ and } 4$ . Hence, if there is a single stable fixed point or attractor  $q^*$  of the system at some value of the control parameter, all  $q_i^{ave}(k)$  will converge to  $q^*$  and there will be only one point along the vertical link at the value of the control parameters. However, if there are two stable fixed points,  $\tilde{q}_1^{ave}$  and  $\tilde{q}_2^{ave}$ , with a period of two, i.e.,  $g(\tilde{q}_i^{ave}, \rho) \neq \tilde{q}_i^{ave}$  and  $g(g(\tilde{q}_i^{ave}, \rho)) = \tilde{q}_i^{ave}$ ,  $i = 1, 2$ , then there will be two points along the vertical lines and the average queue size will alternate between  $\tilde{q}_1^{ave}$  and  $\tilde{q}_2^{ave}$ . The top and bottom dotted straight lines represent  $q_u^{ave}$  and  $q_l^{ave}$ , respectively.

The system parameters are as follows:

$$C = 30.0 \text{ Mbps}, K = \sqrt{\frac{8}{3}}, B = 1,500 \text{ packets},$$

$$d = 0.1 \text{ sec}, q_{min} = 250, q_{max} = 1,000,$$

$$p_{max} = \frac{1}{3}, N_T = 100, M = 500 \text{ bytes}$$

Figures 1 through 2 plot the set of fixed points and the eigenvalues as a function of the exponential averaging weight with different UDP traffic loads. One can clearly see that given the UDP traffic load the system has a single stable point until the eigenvalue becomes -1 and the initial bifurcation occurs. Figures 1 through 2 also indicate that a border collision bifurcation takes place when the larger stable fixed point with a period of two hits the boundary corresponding to  $q_u^{ave}$ , which leads to a chaotic behavior. A border collision in our context indicates a change in the operating regime. This is discussed in more details in

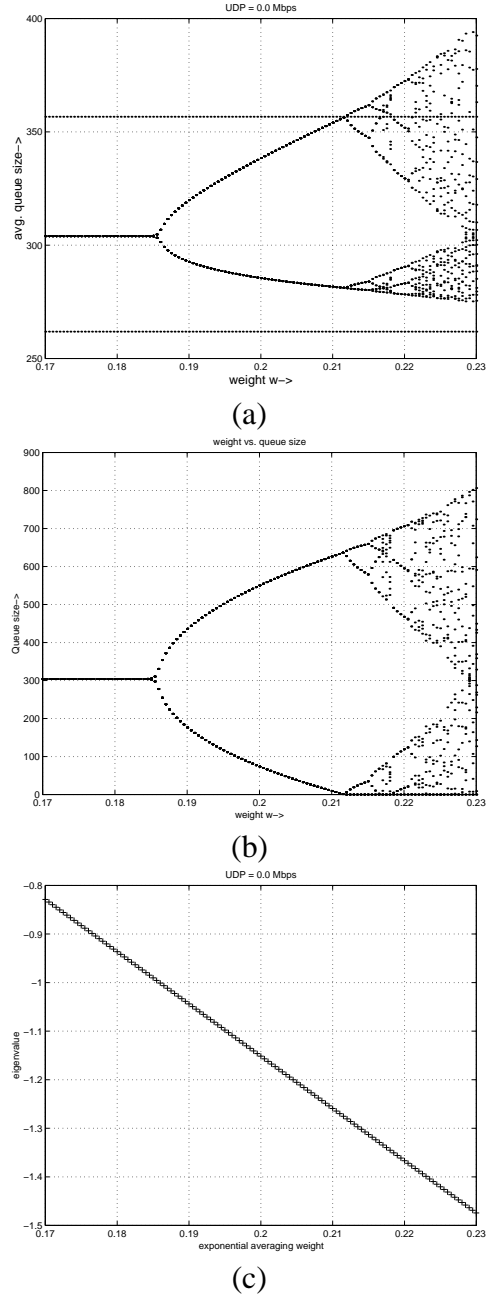


Fig. 1. Bifurcation diagram, instantaneous queue size, and eigenvalue (UDP = 0.0 Mbps).

section VI. Figure 1(b) plots the instantaneous queue size  $q_i(k)$ ,  $k = 991, \dots, 1,000$  and  $i = 1, 2, 3, \text{ and } 4$ , i.e.,  $G(H(q_i^{ave}(k-1)))$ . Although the amplitude of oscillation in the average queue size may not seem large, one can easily see from Figure 1(b) that the amplitude of the oscillation of the actual queue size is large and the queue may empty out quite often.

Figure 3 shows the first-return and second-return maps, i.e.,  $g(\cdot)$  and  $g^2(\cdot) = g(g(\cdot))$ . The dotted line is the 45 degree line. The fixed points of the maps,  $g(\cdot)$  and  $g^2(\cdot)$ , are the points where the maps cross the

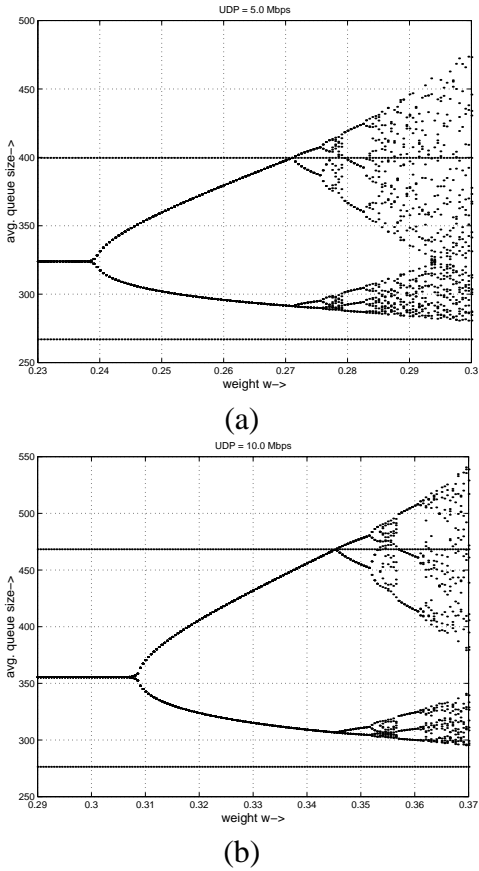


Fig. 2. Bifurcation diagrams. (a) UDP = 5.0 Mbps, (b) UDP = 10.0 Mbps.

45 degree line. One can see in Figure 3 that the number of fixed points of the second-return maps increase from one to three when the eigenvalue becomes  $-1$ , and the fixed point of the map  $g(\cdot)$  becomes unstable, yielding two stable fixed points with a period of two, that are plotted in Figures 1 through 2. This can be seen from that any small perturbation about the unstable fixed point pushes the system even further away from the fixed point because the slope of the mapping  $g(\cdot)$  is larger than one around the neighborhood of the fixed point. However, the other two fixed points of the second-return map are stable as shown in the figure.

## VI. BORDER COLLISION BIFURCATION (BCB)

In this section, we provide some analytical insights into the bifurcations in the model given by (18). As evident from the bifurcation diagrams in Figure 1 and in Figure 2 the system without and with UDP traffic respectively shows a period doubling sequence of bifurcations leading to chaos. A careful look at the bifurcation diagram with respect to exponential averaging weight  $w$  reveals that second period doubling

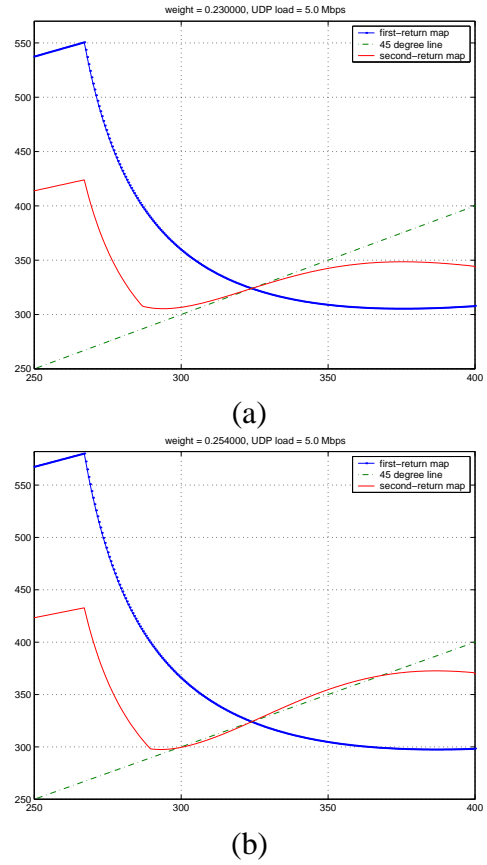


Fig. 3. Plot of first-return and second-return maps ( $\lambda = 5.0$  Mbps).

bifurcation is not a classical period doubling bifurcation ( $eigenvalue = -1$ ), but rather a period doubling bifurcation due to Border Collision [4]. We will prove that under certain general conditions Border Collision type bifurcations are inevitable in our model. We will also provide numerical evidence for this claim.

### A. Background of Border Collision Bifurcation

If the fixed point collides with the border(s) with a change in the parameters, there is a discontinuous change in the derivative of  $\frac{\partial f}{\partial x}$  and the resulting phenomenon is called *border collision bifurcation*. This kind of bifurcation has been reported widely in economics [22], mechanical systems, power electronic models [3], [6], [4].

Border collision is a local bifurcation and hence it can be studied by characterizing the local properties of a map in the neighborhood of the colliding border. It is shown in [4] that a normal form, which is an affine approximation of  $f$  in the border neighborhood, is sufficient to quantify the border collision bi-

furcations. This normal form is described as follows.

$$G(x, \mu) = \begin{cases} ax + \mu, & \text{if } x \leq 0 \\ bx + \mu, & \text{if } x \geq 0 \end{cases} \quad (21)$$

where

$$a = \lim_{x \rightarrow x_b^-} \frac{\partial}{\partial x} f(x, \mu^*)$$

$$b = \lim_{x \rightarrow x_b^+} \frac{\partial}{\partial x} f(x, \mu^*)$$

and  $\mu^*$  is the parameter for which border collision happens. It can be assumed to be 0 without any loss of generality.

There are various types of bifurcation scenarios possible depending on the different values for coefficients  $a$  and  $b$  in the normal form given by (22). For the sake of simplicity, we will discuss only the case relevant to the observed phenomena in our model and provide a numerical proof by computing the one sided coefficients (eigenvalues) for the same.

For our model, the case of interest is when

$$b < -1 < a < 0 \text{ and } ab < 1 \quad (22)$$

This is mentioned as case 7 in [4]. It is shown that in this case a fixed point can bifurcate into a period two attractor as  $\mu$  is varied from less than zero to greater than zero. This is the exact phenomena we are observing for our model when bifurcation parameter  $w$  is varied and a stable period two orbit goes through a period doubling bifurcation again.

Now, we begin with defining a monotone map and state a few assumptions that will be used in the proofs of the lemma which rules out a smooth period doubling route to chaos in the map given by eq. 18.

*Definition 1:* A map  $f(x, \mu)$  is piecewise monotone if  $\frac{\partial f}{\partial x}$  changes sign only at the borders. For an one dimensional map border is a point at which map changes its description.

*Assumption 1:* Assume that derivative of the normalized map in (17) is negative for the  $q_k^n = b_1^n$ .

All assumption 1 says is that system parameters are chosen such that RED control action on mixed TCP and UDP flows is working.

*Assumption 2:* Assume that  $\lambda < \frac{C}{1+p_{max}}$  or in the worst case UDP traffic is not taking more than half of the total bandwidth.

Assumption 2 can be further relaxed by analyzing the precise nature of  $I_2$  as a function of  $\lambda$ , i.e., it is a sufficient condition, but not a necessary condition.

*Lemma 1:* Map described by (17) is piecewise monotone.

**Proof** First, note that map is continuous by construction. The only two points where continuity needs to be checked is  $b_1^n$  and  $b_2^n$ . For  $0 \leq q_k^n < b_2^n$  and  $B \geq q_k^n \geq b_1^n$ , piecewise monotonicity can be shown by differentiating the map. Its differential for all  $q_k^n > b_1^n$  or  $q_k^n < b_2^n$  is  $1 - w$  which implies that map is strictly increasing in these two regions of unit interval.

From Assumption 1, map in (17) has a negative differential as given by (19) evaluated at  $q_k^n = b_1^n$ . Now, for any  $q_k^n \in [b_2^n, b_1^n]$ ,  $q_k^n - q_{min}^n$  will be smaller than  $b_1^n - q_{min}^n$  or  $p_k$  becomes smaller as move towards  $b_2^n$ . This implies that  $I_1$  in the expression of the differential of the map in (19) will be getting more negative as  $q_k^n$  moves towards the left of  $b_1^n$  until  $b_2^n$ . Now if we can prove that  $I_2$  also gets more negative as  $q_k^n$  moves towards the left of  $b_1^n$  until  $b_2^n$  then we are done.

Assumption 2 implies that the numerator of  $I_2$ ,  $-\lambda(C - \lambda(1 + p_k))$  remains negative since  $p_k < 1$ . Since, the denominator of  $I_2$  always remains positive  $I_2$  will stay negative for all  $q_k^n \in [b_2^n, b_1^n]$ . This implies that  $\frac{\partial f}{\partial q_k^n}$  stays negative for any  $q_k^n \in [b_2^n, b_1^n]$ . Hence, we have proved that the map is strictly decreasing in the interval  $[b_2^n, b_1^n]$  and the map given by (17) is piecewise monotone. •

*Lemma 2:* Map given by (17) depends smoothly on  $w$  but not on  $\gamma$ .

**Proof** The differential of this map with respect to the parameter  $\gamma$  in the three regions where it is piecewise monotone with respect to  $q_k^n$ , is  $\frac{\partial f(q_k^n, w, \gamma)}{\partial \gamma} = 0$  for all  $q_k^n \geq b_1^n$  and  $q_k^n \leq b_2^n$  but for  $b_1^n \geq q_k^n \geq b_2^n$ ,  $\frac{\partial f(q_k^n, w, \gamma)}{\partial \gamma} \neq 0$ . Hence, it does not depend smoothly on parameter  $\gamma$ .

$$\frac{\partial f(q_k^n, w, \gamma)}{\partial w} = \begin{cases} -q_k^n & \text{if } q_k^n > b_1^n \\ 1 - q_k^n & \text{if } q_k^n > b_2^n \\ -q_k^n + \left( \frac{CN_T K}{B\sqrt{p_k}(C - \lambda(1 - p_k))} - \frac{Cd}{MB} \right) & \text{otherwise} \end{cases} \quad (23)$$

Similarly, evaluating the partial differential with re-

spect to  $w$  as given by eq. 23 from the both sides near the borders  $b_1^n$  and  $b_2^n$ , it can be easily shown that their values agree. Hence, this map is smooth in  $w$  •

*Lemma 3:* If a fixed point of map given by (17) undergoes a smooth (*eigenvalue* =  $-1$ ) period doubling bifurcation at  $w_1$  and the resulting period two orbit also goes through a smooth period doubling for  $w_2 > w_1$  then the periodic orbit must collide with the border for some  $w \in [w_1 w_2]$ .

**Proof** Proof is a simple application of lemma given in [4] for one dimensional piecewise monotone maps •

Border collision has implications for network buffer. As shown in Figure 2, collision with border  $b_1^n$  implies that buffer will now be getting empty either periodically or in a chaotic manner. Lemma 3 is interesting in the sense that it rules out a very well known mechanism of period doubling to chaos which says that chaos sets in smoothly as a bifurcation parameter is varied. Presence of Border Collision bifurcations indicate the possibility of chaos of large periodicities setting in quickly and hence making the system very sensitive to parameters.

Now we are ready to use the *border collision bifurcation* theory [3], [6], [22], [4] to analyze the bifurcations due to the variation of parameter  $w$  numerically. However, the same theory can not be used to analyze the bifurcations with respect to  $\gamma$  as the map does not depend smoothly on  $\gamma$  (see Lemma 2). As shown in the plots 1 and 2, first bifurcation is a smooth period doubling one and it appears for both the parameters  $w$  and  $\gamma$ .

To provide evidence to our claim, we compute the eigenvalue of period two orbit of the map numerically and show that indeed one sided coefficients (eigenvalues) follow the condition given in (22). This computation is done for the set of parameters corresponding to Figure 2(a).

In Table I, first and second rows shows the two consecutive states (normalized size of the exponentially averaged queue at the router) corresponding to the parameter  $w$  just before the period doubling bifurcation. It also shows the eigenvalues corresponding to those fixed points. It is clear that the eigenvalues are moving out of the unit circle along the negative real axis as  $w$  increases which leads to a period doubling bifurcation.

The third row depicts the same data just before a Border Collision bifurcation. Here, one of the eigenvalues does not lie in the unit circle but eigenvalue corresponding to second iterate has magnitude less than one, which is why system has a stable period two orbit. The fourth row shows the same data after a *border collision* bifurcation from period one to period two for the second iterate. Comparing the states with the border reveals that these two consecutive states lie on the different side of the border. The eigenvalues corresponding to these states confirms the condition given in (22). It is shown that  $b$  in (21) is less than  $-1$ . The multiplication of eigenvalues in third row corresponds to  $a$  in (21) which is just before the *border collision* bifurcation and its magnitude is between zero and one. Their product  $ab = -0.0161 * -3.059 = 0.0492499$  is also less than one as required by the nonsmooth period doubling criterion given by (22). Also, note that the eigenvalues are changing discontinuously as  $w$  is varied. This is a hallmark of *border collision* type bifurcations. This supports our contention that there is a *border collision* bifurcation in through which system exhibits a non smooth

## VII. CONCLUSIONS

In this paper we have presented a simple, nonlinear, discrete-time model that is used to capture the dynamics of the interaction of a RED gateway with both TCP and UDP connections. A normalization scheme is proposed which shows that system need not be studied by varying all the parameters individually and a parameter reduction can be achieved to characterize the dynamical behavior of the system. We also provide a detail study of bifurcations as system parameters are varied using this model. Existence of both classical (Period Doubling) relatively new type of rich bifurcation phenomena (Border Collision) is demonstrated. We also provide some analytical results regarding the bifurcation behavior of the system.

The existence of bifurcations is significant for various reasons. It provides insight into the actual system parametric sensitivity. Understanding the specific bifurcation sequence and its route to chaotic behavior provides a basis for the design of control schemes to improve the dynamical behavior of the network.

## REFERENCES

- [1] M. Allman, V. Paxson, and W. Stevens, "TCP Congestion Control", RFC 2581.

TABLE I

EIGENVALUES COMPUTED FOR DIFFERENT VALUES OF  $\mu$ . QUEUE VALUES GIVEN ARE MULTIPLES OF BUFFER SIZE  $B$ 

No.	$w$	$b_1^n$	$q_k^n$	$q_{k-1}^n$	$\lambda(q_k^n, w)$	$\lambda(q_{k-1}^n, w)$	$\lambda(q_k^n, w)\lambda(q_{k-1}^n, w)$	Legend
1	0.235	399.595/B	323.937	323.937	-0.9640	-0.9640	0.9292	Close to PDB
2	0.240	399.595	317.3197	331.4380	-1.2713	-0.7685	0.9770	After PDB
3	0.268	399.595	394.2552	292.3771	0.0042	-3.7999	-0.0161	Close to BCB
4	0.273	399.595	403.6452	290.526	0.7270	-4.2085	-3.059	After BCB

- [2] S. Athuraliya, S. Low, V. H. Li, and Q. Yin, "REM: Active Queue Management", *IEEE Network*, pp. 48-53, May/June 2001.
- [3] S. Banerjee, P. Ranjan and C. Grebogi, "Bifurcations in two-dimensional piecewise smooth maps - Theory and applications in switching circuits," *IEEE Trans. on Circuits and Systems-I: Fundamental Theory and Applications*, Vol. 47, no. 5, pp. 633-643, 2000.
- [4] S. Banerjee, M. S. Karthik, G. Yuan, J. A. Yorke, "Bifurcations in one-dimensional piecewise smooth maps - Theory and applications in switching circuits," *IEEE Trans. on Circuits and Systems-I: Fundamental Theory and Applications*, Vol. 47, no. 3, pp. 389-394, 2000.
- [5] M. di Bernardo, F. Garofalo, L. Glielmo and F. Vasca, "Analysis of Chaotic Buck, Boost and Buck-Boost Converters through Impact Maps," *Proceedings PESC97 (IEEE Power Electronics Specialist Conf.)*, St. Louis, USA, 1997.
- [6] M. D. Bernardo, M. I. Feigin, S. J. Hogan and M. E. Homer, "Local analysis of  $C$ -bifurcations in  $n$ -dimensional piecewise smooth dynamical systems," *Chaos, Solitons and Fractals*, vol. 10, no. 11, pp. 1881-1908, 1999
- [7] S. Floyd and V. Jacobson, "Random early detection gateways for congestion avoidance", *IEEE/ACM Transactions on Network*, August 1993.
- [8] S. Floyd, "TCP and explicit congestion notification", *ACM Compute Communication Review*, vol. 24, pp. 10-23, October 1994.
- [9] J. P. Hespanha, S. Bohacek, K. Obraczka, and J. Lee, "Hybrid modeling of TCP congestion control", *Lecture notes in computer science*, no. 2034, pp. 291-304, 2001.
- [10] C. V. Hollot, V. Misra, D. Towsley, and W. Gong, "A control theoretic analysis of RED", *Proc. of IEEE INFOCOM 2001*, Vol. 3, pp. 1510-1519, Anchorage, AK, 2001.
- [11] C. V. Hollot, V. Misra, D. Towsley, and W. Gong, "On designing improved controllers for AQM routers supporting TCP flows", *Proc. of IEEE INFOCOM 2001*, Vol. 3, pp. 1510-1519, Anchorage, AK, 2001.
- [12] R. J. Gibbens and F. Kelly, "Resource pricing and the evolution of congestion control", available at <http://www.statslab.cam.ac.uk/~frank>, June 1998.
- [13] J. Guckenheimer and P. Holmes, *Nonlinear oscillations, dynamical systems, and bifurcations of vector fields*, Springer-Verlag, NY, First Edition, 1983.
- [14] E. A. Jackson, *Perspectives of nonlinear dynamics*, Cambridge University Press, NY, First Edition, Vol. 1, pp-150, 1991.
- [15] K. Kar, S. Sarkar, and L. Tassiulas, "A simple rate control algorithm for maximizing total user utility", *Proc. of IEEE INFOCOM*, Anchorage, AK, April, 2001.
- [16] S. Kunniyur and R. Srikant, "Analysis and design of an adaptive virtual queue algorithm for active queue management", *Proc. of ACM SIGCOMM 2001*, San Diego, August 2001.
- [17] S. Kunniyur and R. Srikant, "A time scale decomposition approach to adaptive ECN marking", *Proc. of IEEE INFOCOM 2001*, Anchorage, AK, April 2001.
- [18] R. J. La and V. Anantharam, "Window-based congestion control with heterogeneous users", *Proc. of IEEE INFOCOM 2001*, Anchorage, AK, April 2001.
- [19] R. J. La, P. Ranjan, and E. H. Abed, "Nonlinear dynamics of mixed TCP and UDP under RED", *Submitted to SIGCOMM 2002 and available at <http://www.ece.eng.umd.edu/hyongla>*. February 2002.
- [20] S. Low and D. Lapsley, "Optimization flow control", *IEEE/ACM Transactions on Networking*, Vol 7(6), pp. 861-74, June 1997.
- [21] M. Mathis, J. Semke, J. Mahdavi, and T. Ott, "The macroscopic behavior of the TCP congestion avoidance algorithm", *Computer Communications Review*, Vol. 27(3), 1997.
- [22] H. E. Nusse and J. A. Yorke, "Border collision bifurcations for piecewise smooth one dimensional maps," *Int. J. Bifurcat. Chaos*, vol. 3, pp. 1573-1579, 1993.
- [23] J. Padhye, V. Firoiu, D. Towsley, and J. Kurose, "Modeling TCP Reno performance: a simple model and its empirical validation", *IEEE/ACM Transactions on Networking*, Vol. 8(2), pp. 133-145, April 2000.
- [24] P. Tinnakornsrisuphap and A. Makowski, "Queue dynamics of RED gateways under large number of TCP flows", *Proc. of 2001 Globecom*, San Antonio, TX, December 2001.
- [25] P. Ranjan, E. H. Abed, and R. J. La, "Nonlinear instabilities in TCP-RED", *To appear in the proc. of IEEE INFOCOM 2002*.
- [26] P. Ranjan, E. H. Abed, "Nonlinear analysis and control of TCP-RED in a simple network model", *To appear in the proc. of 2002 American Control Conference*.