Instability of a Tandem Network and its Propagation under RED

Richard J. La
University of Maryland, College Park
hyongla@eng.umd.edu.

***Abstract*— Random Early Detection (RED) mechanism has been proposed to control the average queue size at the bottlenecks inside the network. It has been shown that the interaction between a RED gateway and TCP connections can lead to a rich set of nonlinear phenomena in single bottleneck cases. In this paper we extend this model and study the interaction of TCP connections with RED gateways in a simple tandem network, using a nonlinear first-order discrete-time model. We demonstrate, using bifurcation diagrams, that the nonlinear behavior of TCP can result in both smooth and non-smooth bifurcations, leading to chaos. We show that the instabilities can be induced at both bottlenecks by changing the system parameters only at one of the bottlenecks while fixing the parameters at the other, thus demonstrating the propagation of instability. Moreover, we show that locally sufficient conditions for stability based on single node analysis are not sufficient for global network stability.**

I. INTRODUCTION

With the unprecedented growth and popularity of the Internet the problem of congestion control is emerging as a more crucial problem. Poor management of congestion can render one part of a network inaccessible to the rest and significantly degrade the performance of networking applications. Researchers have proposed various approaches to address this issue. One such approach is to control congestion level at each router through an Active Queue Management (AQM) mechanism.

The Random Early Detection (RED) mechanism, proposed by Floyd and Jacobson, attempts to control the congestion level at a bottleneck by monitoring the average queue size [2]. 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 poorly understood. Ranjan, Abed, and La have used a simple nonlinear model to investigate the behavior of a simple single link network with a RED gateway and TCP connections [10]. They have demonstrated that such a system leads to nonlinear phenomena, such as oscillations and chaos, if the system parameters are not selected carefully. 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 are caused by border collision bifurcations.

Although the RED was proposed almost a decade ago, it has not been widely deployed in practice, mainly due to a lack of agreement on parameter setting. This problem is further complicated by a lack of understanding of how congestion in one part of network affects another part of the network. In this paper we attempt to address these issues. We extend the model in [10] and study the interaction of the RED mechanism with TCP connections in a simple tandem network. This tandem network can be viewed as a network with two dominant bottleneck links. We show the existence of both smooth and non-smooth bifurcations, i.e., classical period doubling bifurcation and border collision bifurcation, as system parameters are varied, which lead to queue oscillation at the bottlenecks. We also demonstrate that these instabilities can be induced by varying the parameters only at one of the bottlenecks, while keeping the parameters at the other bottleneck fixed. This suggests that in some cases of general networks an instability induced by one node can spread to other nodes, making it difficult to isolate the source of the instability. Furthermore, the locally sufficient conditions for stability based on single node analysis are not sufficient to guarantee the global network stability.

The rest of the paper is organized as follows. Sections II and III present the nonlinear first-order discrete-time model that is used for our analysis. Section IV discusses the local stability of equilibrium points. Section V presents a numerical example based on our analytical model.

II. NONLINEAR FIRST-ORDER DISCRETE-TIME MODEL

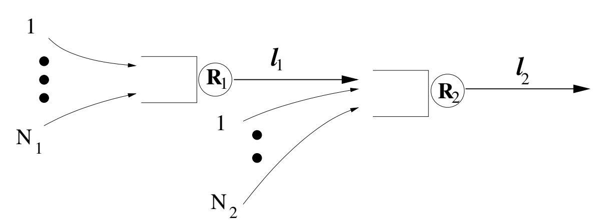

Fig. 1. Network model.

We consider a simple network of two links that are shared by many connections, as shown in Fig. 1. As mentioned before this network can be viewed as a network with two dominant bottleneck links but with other links that are not bottlenecks. We denote the set of connections that traverse both links l_1 and l_2 by $\mathcal{I}_1, \mathcal{I}_1 = \{1, \cdots, N_1\}$, and the connections that traverse only the second link l_2 by $\mathcal{I}_2, \mathcal{I}_2 = \{1, \cdots, N_2\}$. All connections are assumed to be TCP Reno connections. The capacity of links l_1 and l_2 are denoted by C_1 and C_2, respectively, and the buffer sizes at R_1 and R_2 by B_1 and B_2, respectively. The round-trip propagation delay (without any queuing delay) of connection i is given by d_i. We denote the rate or throughput of connection i by x_i, and the packet size by M. The parameter M can be interpreted as the average packet size of the connections. We assume that the RED queue management mechanism is implemented at nodes R_1 and R_2 in order to control the average queue size at the routers. 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 [2]:

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

where q_{min} and q_{max} are the lower and higher threshold values, and p_{max} is the selected marking/drop probability when $q^{ave} = q_{max}$. The average queue size is updated at the time of packet arrival according to the exponential averaging

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

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 and how fast the RED can react to time-varying load. On one hand, the averaging weight should be selected to be small enough so that transient, temporary congestion does not result in an oscillation of the packet marking/drop probability. On the other hand, the weight should be set large enough so that the RED can react to changes in load in a timely manner. These are conflicting goals, and the selection of the parameters affects the interaction of the RED mechanism with adaptive sources, such as TCP. In this paper, however, we show that the averaging weight cannot be set arbitrarily large in an attempt to improve the responsiveness of the RED, without causing an unstable behavior at the bottlenecks, which affects TCP performance.

The connections are assumed to be long-lived, 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 in $\mathcal{I}_1$ have the same round-trip propagation delay (RTPD) d_1 and all connections in $\mathcal{I}_2$ have the same RTPD d_2. Rather than interpreting this assumption as a requirement that the connections must have the same propagation delay, one should consider the delay d_j as the effective delay that represents the average propagation delay of the connections in $\mathcal{I}_j$ as shown in [10]. This allows us to reduce the problem with $N_1 + N_2$ connections to a two-connection system that represents the set of connections and study its behavior.

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

$$T(p, R) = \frac{MK}{\sqrt{p}R} + o\left(\frac{1}{\sqrt{p}}\right) \ , \quad (3)$$

where K is some constant in $\left[1, \sqrt{8/3}\right]$ [8], [9]. In this paper we assume $K = \sqrt{3/2}$ [9]. The exact value of K is not crucial to our qualitative analysis. We use the first term of this simple approximation, i.e., $MK/\sqrt{p}R$, for TCP throughput to facilitate our analysis. Although this may seem like a crude approximation, it provides a good enough approximation for our qualitative results when Explicit Congestion Notification (ECN) is adopted by TCP connections and the RED gateway breaks the synchronization among the connections as demonstrated in [11] and results in relatively uniform packet drops. Moreover, we are mainly interested in characterizing the stability conditions, and it has been observed that a system tends to be less stable when the load is light [7], [10], [5], in which case the packet marking/drop rate will be low and the $o(1/\sqrt{p})$ term will be relatively small compared to the first term [9]. Therefore, the first term in (3) will give us an accurate approximation in the region of our interest. This is validated through *ns*-2 simulation presented in [6]. Moreover, although we use this simple approximation for TCP throughput to facilitate our analysis, our qualitative results do not depend on this particular form of the approximation, and are consequences of rather benign nonlinear behavior of TCP and hold with more sophisticated models for TCP throughput.

TCP adjusts its transmission rate depending on whether it has detected a packet drop or not. Therefore, a network with an AQM mechanism can be viewed as a feedback system where TCP sources adjust their transmission rates based on the feedback from the AQM mechanism. This feedback is in the form of marked or dropped packets and is delayed by an RTT of the connection. This can be modeled as a stroboscopic map where the instant of observation is approximately one RTT. 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.

As mentioned before the exponential averaging weight should be chosen sufficiently small so that the average queue size q^{ave} does not fluctuate 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 RTT of the connections in order to avoid a fast oscillation in the packet drop probability. Therefore, with a large number of connections it is reasonable to assume that the TCP connections' dynamics operate at a faster time-scale than the evolution of the average queue size and that the aggregate throughput (or workload) presented by the connections sees a quasi-stationary behavior before the average queue size changes much.

We use a nonlinear dynamic first-order discrete-time model to analyze the interaction of the RED gateway with TCP connections, which was first proposed by Firoiu and Borden [1]. We define the control system as follows. The packet marking/drop probabilities $\underline{p}_k = (p^1_k, p^2_k)$ at period $k, k \geq 1$, where p^j_k is the marking/drop probability at node $R_j, j = 1, 2$, determine the throughput of the connections and the queue sizes $\underline{q}_{k+1} = (q^1_{k+1}, q^2_{k+1})$ at period $k+1$, based on the system constraints. The queue sizes at period $k+1$ are used to compute the average queue sizes $q^{j,ave}_{k+1}, j = 1, 2$, at period $k+1$ according to the exponential averaging rule in (2). Then, the average queue sizes $\underline{q}^{ave}_{k+1} = (q^{1,ave}_{k+1}, q^{2,ave}_{k+1})$ are used to calculate the packet marking/drop probabilities $\underline{p}_{k+1}$ at period $k+1$, which are the control variable of the AQM mechanism.

This can be mathematically written as follows:

$$\text{plant function:} \quad \underline{q}_{k+1} = G(\underline{p}_k) \tag{4}$$
$$\text{averaging function:} \quad \underline{q}^{ave}_{k+1} = A(\underline{q}^{ave}_k, \underline{q}_{k+1}) \tag{5}$$
$$\text{control function:} \quad \underline{p}_{k+1} = H(\underline{q}^{ave}_{k+1}) \ , \tag{6}$$

where $A(\underline{q}^{ave}_k, \underline{q}_{k+1}) = (1 - \underline{w}) \cdot \underline{q}^{ave}_k + \underline{w} \cdot \underline{q}_{k+1}$ as given in (2), $\underline{w} = (w_1, w_2)$, and the control function $H(\underline{q}^{ave}_{k+1}) = \underline{p}(\underline{q}^{ave}_{k+1})$ in (1). These multiplications are component-wise, and an underline is used to denote a vector. 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., and is described in the following section.

III. Interaction of RED gateways with TCP connections

We now describe the plant function $G(\cdot)$ that is used for analyzing the system. Let $p^3_k = 1 - (1 - p^1_k)(1 - p^2_k) = p^1_k + p^2_k - p^1_k \cdot p^2_k$. We define p^u to be

$$p^u = \inf \left\{ p \in [0, 1] \ \middle| \ \frac{MK}{\sqrt{p} d_1} \leq \frac{C_1}{N_1} \right\} = \left(\frac{N_1 MK}{C_1 d_1} \right)^2 \ , \tag{7}$$

which is the smallest probability such that all $p^3_k \geq p^u$ lead to $q^1_{k+1} = 0$ and is assumed to be smaller than one. Based on $\underline{p}_k$ we compute the queue sizes $\underline{q}_{k+1} = (q^1_{k+1}, q^2_{k+1})$ at period $k + 1$ as follows.

Case (1): $p^3_k \geq p^u$

We define

$$p^{2,u}(\underline{p}_k) = \inf \left\{ p \in \Re_+ \ \middle| \ \frac{N_2 MK}{\sqrt{p} d_2} \leq C_2 - \frac{N_1 MK}{\sqrt{p^3_k} d_1} \right\} \ . \tag{8}$$

This is the smallest probability that results in $q^2_{k+1} = 0$ given $p^3_k \geq p^u$.

Case (1a): $p^2_k \geq p^{2,u}(\underline{p}_k)$

In this case from (7) and (8) one can see that $\underline{q}_{k+1} = \underline{0}$.

Case (1b): $p^2_k < p^{2,u}(\underline{p}_k)$

In this case $q^1_{k+1} = 0$ and $q^2_{k+1} = \min(B_2, \tilde{q}^2(\underline{p}_k))$, where $\tilde{q}^2(\underline{p}_k)$ is the solution to

$$\frac{N_1 MK}{\sqrt{p^3_k} \left(d_1 + \frac{q^2 M}{C_2} \right)} + \frac{N_2 MK}{\sqrt{p^2_k} \left(d_2 + \frac{q^2 M}{C_2} \right)} = C_2 \ . \tag{9}$$

If we assume that $d_1 = d_2$, then $\tilde{q}^2(\underline{p}_k) = \frac{N_1 K}{\sqrt{p^3_k}} + \frac{N_2 K}{\sqrt{p^2_k}} - \frac{C_2 d_2}{M}$.

Case (2): $p^3_k < p^u$

There exist a set of $(q^1, q^2) \geq \underline{0}$ such that

$$\frac{N_1 MK}{\sqrt{p^3_k} \left(d_1 + \frac{q^1 M}{C_1} + \frac{q^2 M}{C_2} \right)} = C_1 \ . \tag{10}$$

Let $q^{2,max}_k = \frac{N_1 K C_2}{\sqrt{p^3_k} C_1} - \frac{d_1 C_2}{M}$, i.e., q^2 that satisfies (10) with $q^1 = 0$, and $q^{2u}_k = \max(0, \frac{N_2 K C_2}{\sqrt{p^2_k}(C_2 - C_1)} - \frac{d_2 C_2}{M})$, where the second term is the solution to $\frac{N_2 MK}{\sqrt{p^2_k}\left(d_2 + \frac{q^2 M}{C_2}\right)} = C_2 - C_1$

Case (2a): $q^{2u}_k \leq \min(B_2, q^{2,max}_k)$

In this case the queue size $q^2_{k+1} = q^{2u}_k$ and $q^1_{k+1} = \min(B_1, \tilde{q}^1(q^2_{k+1}, \underline{p}_k))$, where $\tilde{q}^1(q^2_{k+1}, \underline{p}_k)$ is q^1 that satisfies (10) with $q^2 = q^2_{k+1} = q^{2u}$, i.e., $\tilde{q}^1(q^2_{k+1}, \underline{p}_k) = \frac{N_1 K}{\sqrt{p^3_k}} - \frac{d_1 C_1}{M} - \frac{q^2_{k+1} C_1}{C_2} = \frac{N_1 K}{\sqrt{p^3_k}} - \frac{d_1 C_1}{M} - \frac{q^{2u}_k C_1}{C_2}$.

Case (2b): $q^{2u}_k > \min(B_2, q^{2,max}_k)$

In this case the queue size $q^2_{k+1} = \min(B_2, \tilde{q}^2(\underline{p}_k))$ and $q^1_{k+1} = \max(0, \min(B_1, \tilde{q}^1(q^2_{k+1}, \underline{p}_k)))$, where $\tilde{q}^2(\underline{p}_k)$ is the solution to (9), and $\tilde{q}^1(q^2_{k+1}, \underline{p}_k) = \frac{N_1 K}{\sqrt{p^3_k}} - \frac{C_1 d_1}{M} - \frac{C_1 q^2_{k+1}}{C_2}$ as in the previous case.

From the above we can compute $\underline{q}_{k+1}$ as a function of $\underline{p}_k$, and hence we have the following plant function

$$G_1(\underline{p}_k) = q^1_{k+1} = \begin{cases} 0, & \text{if } p^3_k \geq p^u \\ \min(B_1, \frac{N_1 K}{\sqrt{p^3_k}} - \frac{d_1 C_1}{M} - \frac{q^{2u}_k C_1}{C_2}), & \text{if } p^3_k < p^u \text{ and } q^{2u}_k \leq \min(B_2, q^{2,max}_k) \\ \max(0, \min(B_1, \tilde{q}^1(\min(B_2, \tilde{q}^2(\underline{p}_k)), \underline{p}_k))), & \text{otherwise} \end{cases} \tag{11}$$

and

$$G_2(\underline{p}_k) = q^2_{k+1} = \begin{cases} 0, & \text{if } p^3_k \geq p^u \text{ and } p^2_k \geq p^{2,u}(\underline{p}_k) \\ \min(B_2, \tilde{q}^2(\underline{p}_k)), & \text{if } p^3_k \geq p^u \text{ and } p^2_k < p^{2,u}(\underline{p}_k) \\ q^{2u}_k, & \text{if } p^3_k < p^u \text{ and } q^{2u}_k \leq \min(B_2, q^{2,max}_k) \\ \min(B_2, \tilde{q}^2(\underline{p}_k)), & \text{otherwise} \end{cases} \tag{12}$$

to complete our discrete-time model described in Section II. From (4) through (6) and the plant function $G(\cdot)$ given in (11) and (12) we define a mapping

$$\begin{aligned} \underline{q}^{ave}_{k+1} &= (1 - \underline{w}) \cdot \underline{q}^{ave}_k + \underline{w} \cdot A(G(H(\underline{q}^{ave}_k))) \\ &:= g(\underline{q}^{ave}_k, \underline{\rho}) \end{aligned} \tag{13}$$

where $\underline{\rho}$ summarizes the system parameters, including the exponential averaging weights $\underline{w}$, and multiplications are component-wise. This maps the average queue sizes at period k to the average queue sizes at period $k + 1$.

IV. Stability of the System

A fixed point of the mapping $g(\cdot)$ is a vector of average queue sizes $\underline{q}^*$ such that $\underline{q}^* = g(\underline{q}^*, \underline{\rho})$. Given the system parameters one can find such a fixed point $\underline{q}^*$, if there is any, by solving (13) with the given $A(\cdot), G(\cdot)$, and $H(\cdot)$. The local stability of the fixed point $\underline{q}^*$ can be studied by looking at the eigenvalues of the Jacobian matrix $J = [J_{ij}, \ i, j = 1, \ 2]$ of the mapping $g(\cdot, \cdot)$ in (13), where $J_{ij} = \frac{\partial g_i}{\partial q^j}$. Depending on where the fixed point $\underline{q}^* = (q^{1*}, q^{2*})$ of the system lies, there could be several different cases. In this section we only consider the case where $\underline{q}^*$ satisfies case (2a) in Section III, which is the most interesting case from the operational point of view. In this case, $G_1(\underline{p})$ and $G_2(\underline{p})$, where $\underline{p} = H(\underline{q}^*)$, are given by

$$G_1(\underline{p}) = \min(B_1, \frac{N_1 K}{\sqrt{p^3_k}} - \frac{N_2 K C_1}{\sqrt{p^2_k}(C_2 - C_1)} + \frac{(d_2 - d_1) C_1}{M})$$
$$G_2(\underline{p}) = \max(0, \frac{N_2 K C_2}{\sqrt{p^2_k}(C_2 - C_1)} - \frac{d_2 C_2}{M}) \tag{14}$$

where $p^1 = \frac{q^{1*} - q^1_{min}}{q^1_{max} - q^1_{min}} p^1_{max}$, $p^2 = \frac{q^{2*} - q^2_{min}}{q^2_{max} - q^2_{min}} p^2_{max}$, and $p^3 = p^1 + p^2 - p^1 \cdot p^2$. We assume that the fixed point is an interior point, i.e., $\underline{0} < \underline{q}^* < \underline{B}$.

From (14) one can see that $g_2(\cdot, \underline{\rho})$ does not depend on q^{1*} and, hence, $J_{21} = 0$. Therefore, the eigenvalues of the Jacobian matrix are given by J_{11} and J_{22}, where

$$J_{11} = 1 - w_1 - w_1 \cdot \frac{N_1 K (1 - p^2) \alpha^1}{2(p^1 + p^2 - p^1 \cdot p^2)^{1.5}} \ ,$$
$$J_{22} = 1 - w_2 - w_2 \cdot \frac{N_2 K C_2}{2\sqrt{p^2}(q^{2*} - q^2_{min})(C_2 - C_1)} \ ,$$

and $\alpha^j = p^j_{max}/(q^j_{max} - q^j_{min}), j = 1, 2$. We investigate how these eigenvalues behave as the exponential averaging weights vary in the following section.

V. Numerical Results

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 only vary the exponential averaging weights and study the stability of the system. However, similar results can be obtained with any of the system parameters [10], [5]. The x-axis is the system 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 system parameter. For generating the bifurcation diagrams, in each run we randomly select four random initial average queue sizes, $\underline{q}^{ave}_1(0), \underline{q}^{ave}_2(0), \underline{q}^{ave}_3(0)$, and $\underline{q}^{ave}_4(0)$, and these average queue sizes evolve according to the map $g(\cdot,\cdot)$ in (13), i.e.,

$$\underline{q}^{ave}_i(k) = g(\underline{q}^{ave}_i(k-1), \underline{\rho}) \ , \qquad \text{for } k = 1, \cdots, \ 1{,}000$$
$$\text{and } i = 1, \ 2, \ 3, \ \text{and } 4 \ .$$

We plot $\underline{q}^{ave}_i(k), k = 991, \cdots, 1{,}000$ and $i = 1, 2, 3,$ and 4. Hence, if there is a single stable fixed point or attractor $\underline{q}^*$ of the system at some value of the system parameter, all $\underline{q}^{ave}_i(k)$ will converge to $\underline{q}^*$ and there will be only one point along the vertical line at the value of the system parameter. However, if there are two stable fixed points, $\tilde{q}^{ave}_1$ and $\tilde{q}^{ave}_2$, with a period of two, i.e., $g(\tilde{q}^{ave}_i, \underline{\rho}) \neq \tilde{q}^{ave}_i$ and $g(g(\tilde{q}^{ave}_i, \underline{\rho})) = \tilde{q}^{ave}_i, i = 1, 2$, then there will be two points along the vertical line and the average queue sizes will alternate between $\tilde{q}^{ave}_1$ and $\tilde{q}^{ave}_2$.

The system parameters are as follows:

$$\underline{q}_{max} = (600, \ 1200), \ \underline{q}_{min} = (200, \ 400), \ p_{max} = 1/8$$
$$\underline{C} = (12, \ 30) \text{ Mbps}, \ K = \sqrt{3/2}, \ \underline{B} = (1000, \ 2000),$$
$$d_1 = d_2 = 0.1 \text{ sec}, \ M = 4{,}000 \text{ bits}, \ N_1 = N_2 = 100 \ .$$

A. Instability of the tandem network

Fig. 2(a) and 2(b) plot the set of stable fixed points as a function of exponential averaging weights $\underline{w} = w \cdot (1, \ 1)$, where w is the value along the x-axis. The y-axis plots the average queue sizes per flow, i.e., $q^{1,ave}/N_1$ and $q^{2,ave}/(N_1 + N_2)$. Fig. 2(c) and 2(d) show the actual queue sizes per flow. One can see that there is a unique fixed point of the system $\underline{q}^* = (2.12 \cdot N_1, 2.78 \cdot (N_1 + N_2))$ for $w < 0.3862$. At $w = 0.3862$ the initial period doubling bifurcation occurs, and the

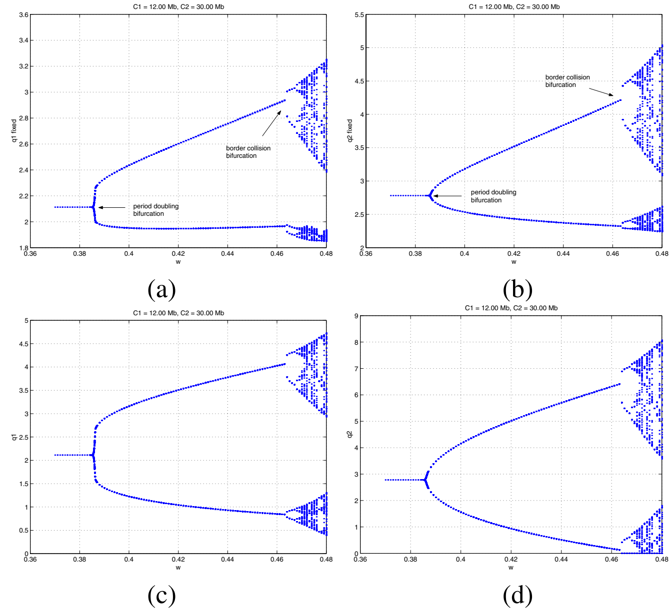

Fig. 2. Bifurcation diagrams. (a) $q^{1,ave}$, (b) $q^{2,ave}$, (c) q^1, and (d) q^2.

queue size begins to oscillate. This can be also verified by computing the eigenvalue J_{22} and showing that it hits the unit circle at -1 [4].

A careful investigation of the bifurcation diagrams reveals that the second period doubling bifurcation at $w = 0.464$ is not a classical period doubling bifurcation, but rather a period doubling bifurcation due to a border collision. Here a border is a region at which the behavior of queue sizes changes from one case to another in section III. If a fixed point collides with a border, there is a discontinuous change in the Jacobian matrix of the mapping. This border collision bifurcation leads to a cascade of bifurcations, resulting in chaos as shown in Fig. 2.

This border collision bifurcation illustrates the impact of an instability on the system throughput. Note that the distance between the initial period doubling bifurcation point and the border collision bifurcation point is relatively short. Hence, once the system becomes unstable, the queue sizes quickly start oscillating widely, often leading to empty queues. Since empty queues mean that there are no packets to be scheduled in the buffer, it indicates a waste of resources and a decrease in TCP throughput. Further, an oscillatory behavior of the queue size leads to larger RTT variance, and as a consequence it takes TCP connections longer to detect packet losses, for instance, due to changes in load, through retransmission time-outs when there are not enough duplicate acknowledgments.

B. Propagation of the instability

In this subsection we only vary the exponential averaging weight of the second bottleneck link. Similarly as in Fig. 2, Fig. 3(a) and 3(b) plot the average queue sizes per flow, i.e., $q^{1,ave}/N_1$ and $q^{2,ave}/(N_1 + N_2)$, as a function of $\underline{w} = (0.1, w)$, and Fig. 3(c) and 3(d) show the actual queue sizes. One can see that, as in Fig. 2, there is a unique fixed point of the system $\underline{q}^* = (2.12 \cdot N_1, 2.78 \cdot (N_1 + N_2))$ for $w < 0.3862$. At

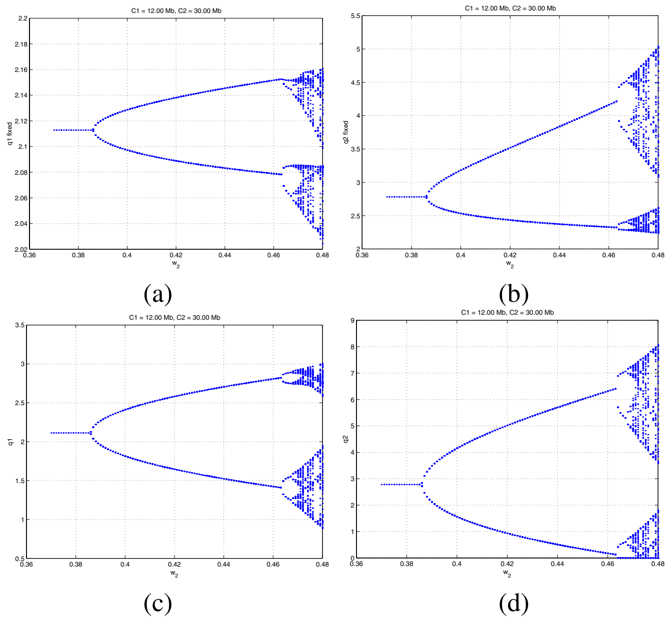

Fig. 3. Bifurcation diagrams. (a) $q^{1,ave}$, (b) $q^{2,ave}$, (c) q^1, and (d) q^2.

$w = 0.3862$ the initial period doubling bifurcation takes place as in the previous subsection. Hence, although w_1 is much smaller in this case (w_1 is fixed at 0.1), the initial bifurcation occurs at the same value of w as in Fig. 2. In fact the border collision bifurcation, which leads to chaos, occurs at the same value as well, as shown in Fig. 2 and 3. This can be explained as follows. In this example the fixed point $\underline{q}^*$ satisfies case (2a), and the eigenvalue of the Jacobian matrix with larger magnitude for the given range of w in this example is J_{22}, which does not depend on q^{1*} because $g_2(\cdot, \underline{\rho})$ is a function only of q^2 as shown in Section IV.

This example demonstrates that in a general network an instability caused by one bottleneck may propagate to other parts of the network. Hence, even if the majority of the routers are properly configured, if a handful of routers are misconfigured, then a large portion of the network may experience an instability, which may make it difficult to isolate the source of instability and correct it.

This numerical example also illustrates the fact that locally sufficient conditions for stability are not enough to guarantee the global network stability as will be shown here. Note from [10] that the initial period doubling bifurcation point in a single bottleneck case is given by

$$w^* = \frac{2}{1 + \frac{NK}{2\sqrt{\alpha}(q^* - q_{min})^{1.5}}}, \tag{15}$$

where $\alpha = p_{max}/(q_{max} - q_{min})$, and N is the number of connections. The fixed point q^* can be computed as the positive, real solution to the following third order polynomial

$$(q^* - q_{min}) \left(q^* + \frac{dC}{M} \right)^2 - \frac{(NK)^2}{\alpha} = 0 \ .$$

Suppose that we isolate each bottleneck link and assume that the other link is not a bottleneck. In other words, we take one bottleneck link in the tandem network at a time and remove the other bottleneck and the connections not traversing the selected bottleneck link. This allows us to study how congestion at one bottleneck affects the queue behavior at another and the stability of the overall network. Eq. (15) yields the bifurcation points of 0.5674 and 0.4692 for the first and second bottleneck links, respectively, when they are considered separately. Therefore, from Fig. 2 and 3 one can see that the locally sufficient conditions that $w_1 < 0.5674$ and $w_2 < 0.4692$ based on the single node analysis, are not sufficient to guarantee the global network stability. The *ns*-2 simulation results for these numerical examples are given in [6].

This has rather serious practical implications. First, if the instability propagates from one bottleneck to another, it is difficult to track down its source since every bottleneck that exhibits an instability needs to be checked. Moreover, even if every bottleneck is checked, it may still not be possible to locate the source. Second, even when a network manager can somehow estimate all system parameters and satisfy the locally sufficient conditions for stability based on a single node analysis [7], [10], these locally sufficient conditions at individual nodes do not guarantee the global stability of the network. These issues on parameter selection pose a serious problem since different sets of bottlenecks connections traverse (determined by the routing algorithms) may belong to different domains, and ensuring the global network stability may require communication and cooperation between these domains.

References

[1] V. Firoiu and M. Borden, "A study of active queue management for congestion control", *Proc. of INFOCOM 2000*, Tel Aviv, Israel, March 2000.

[2] S. Floyd and V. Jacobson, " Random Early Detection Gateways for Congestion Avoidance," *IEEE Trans. on Networking*, Vol.1, no. 7, pp. 397-413, 1993.

[3] S. Floyd, "TCP and explicit congestion notification", *ACM Compute Communication Review*, vol. 24, pp. 10-23, October 1994.

[4] J. Guckenheimer, and P. Holmes, "Nonlinear oscillations, dynamical systems, and bifurcations of vector fields", Springer-Verlag, NY, 1st Ed., 1983.

[5] R. J. La, P. Ranjan, and E. H. Abed, "Analysis of Adaptive Random Early Detection", *Submitted to ITC 2003*, available at "http://www.ece.umd.edu/˜ hyongla".

[6] R. J. La, "Instability of a tandem network and its propagation under RED", *To be submitted to IEEE Trans. on Automatic Control*, available at "http://www.ece.umd.edu/˜ hyongla".

[7] S. H. Low, F. Paganini, J. Wang, S. Adlakha, and J. C. Doyle, "Dynamics of TCP/RED and a scalable control", *Proc. of IEEE INFOCOM 2002*.

[8] 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.

[9] 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.

[10] P. Ranjan, E. H. Abed, and R. J. La, "Nonlinear instabilities in TCP-RED", *Proc. of IEEE INFOCOM 2002*.

[11] 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.