Characteristic Timescales: Analyzing End-to-end Delay in Networks under High Loads

Anindya Basu
Bell Laboratories, Lucent Technologies
600-700 Mountain Avenue
Murray Hill, NJ 07974, USA
anindya@gmail.com

Debojyoti Dutta
Department of Computer Science
University of Southern California
Los Angeles, CA 90089, USA
ddutta@usc.edu

Sharad Ramanathan
Bell Laboratories, Lucent Technologies
600-700 Mountain Avenue
Murray Hill, NJ 07974, USA
sharadr@lucent.com

Abstract

We develop a theory of network behavior under high load conditions and rapidly changing traffic demand matrices, using techniques from spectral graph theory and the physics of dynamical systems. We focus on the mean end-to-end delay as the performance metric and show that at high loads, there is a "characteristic timescale" of the network that determines the mean end-to-end delay. We show how this timescale can be analytically determined for a given network and validate our theory using simulations. We envisage that such a characterization can be used to compare network performance among various topologies, and can potentially lead to new network optimization techniques.

1 Introduction

Internet Service Providers have traditionally responded to demands for increasing bandwidth using various techniques including traffic engineering and over-provisioning. However, the possibility that the network may enter periods of high congestion due to traffic overload has not entirely been eliminated because of the bursty nature of Internet traffic. Furthermore, for wireless and adhoc networks, over-provisioning may not always be possible due to the limited availability of bandwidth, and overload conditions may be more frequent.

Unfortunately, there has been very little investigation into how networks behave when high congestion conditions occur for extended periods of time. In particular, there is no concise analytical description of how end-to-end delays increase as a network becomes more and more congested due to traffic load. The difficulties encountered when attempting such an analysis are manifold: first, multiple dynamic variables are required to describe the behavior of a network. While techniques have been developed to analyze network behavior using queueing theory and fixed point approximations [6, 26], these are not very accurate at high loads [13]. Second, the nature of network traffic has to be taken into account, and this can be arbitrarily complex. Finally, the routing algorithm used to direct packets and its interaction with network traffic must also be considered.

In this paper, we try to understand network behavior at high loads close to network breakdown. For the purposes of this paper, we define network breakdown to be the condition where the high traffic load causes the goodput to reduce to zero. In other words, network breakdown occurs when packets spend (almost) all their time in large queues and eventually get dropped such that no traffic gets through. To study network behavior close to breakdown, we develop techniques to describe such behavior using fewer (ideally, one) dynamic variables. In addition to making the problem tractable, our approach provides a single metric for characterizing network performance. This implies that the single metric could be used as a natural way of comparing the performance of a variety of network topologies (over a range of traffic distributions).

Such a program of study leads to the following questions that we attempt to answer in this paper. First, what is a good metric to characterize network breakdown? Second, does network breakdown at very high loads happen suddenly or in a continuous manner? In other words, as the network load increases, does the network goodput go to zero (which corresponds to the mean end-to-end delay going to infinity) continuously, or show a discontinuous jump? If the behavior is continuous, can we model it as a second order phase transition? If so, again motivated by other physical models, can we describe the behavior of the network by fewer and fewer degrees of freedom as we approach breakdown loads?

In order to address these questions, we develop a model for data transmission on packet switched networks and ana-

lyze it using techniques inspired by spectral graph theory [8] and physics of dynamical systems [14]. We focus on the mean end-to-end delay[1] and use our model to analytically compute a *characteristic timescale* of the network. This timescale can be used to characterize the mean end-to-end delay. We consider a regime where the demand matrix changes rapidly and the traffic generation processes at the nodes are unknown, and show that the characteristic timescale depends on both the network topology and the traffic. At high loads, we demonstrate that the effect of the network traffic can be incorporated into our model by using the queue lengths at the various nodes. This is an advantage for modern data networks, where the demand matrix is either unknown or changing rapidly. Using a combination of analytic techniques and simulations, we show that the mean end-to-end delay in a highly loaded network is strongly correlated to the characteristic timescale of the network. While it may not always be possible to exactly determine the mean end-to-end delay from the characteristic timescale, we show that the characteristic timescale can be computed easily, and can be used to reliably predict the behavior of the mean end-to-end delay on different networks under various traffic conditions. Furthermore, such a characterization of network performance can ultimately lead to new network optimization techniques by developing algorithms to minimize the characteristic timescale.

The main contributions of this paper are as follows. First, we develop a model for analyzing network behavior at high traffic loads with rapidly varying or unknown demand matrices. Second, we show that various non-linearities introduced by packet interactions at queues and details such as the traffic generation statistics become irrelevant at high loads. Third, from the spectral analysis of this model, we extract the characteristic timescale of the network. We show that this characteristic timescale is highly correlated with the mean end-to-end delay at high loads. Thus, this timescale can be used as a metric to describe network behavior close to breakdown. Fourth, we show that this correlation increases with network load, thereby providing evidence that our analytical results become more and more accurate as the network approaches breakdown. Finally, we show that networks exhibit second order critical phenomena at high loads. In other words, the end-to-end delays go continuously to infinity as the breakdown (or critical) point is reached. This is in contrast to a first order critical phenomenon where the mean end-to-end delay jumps discontinuously to infinity at the breakdown point. While our approach is related to existing work [15, 21, 31] that has modeled network congestion (i.e., highly loaded networks) in terms of *sparsest cuts* or *conductance*, we note that these models are applicable in a regime where the demand matrix is known and fixed. Furthermore, they do not typically incorporate traffic generation statistics or packet interactions at queues. We discuss the differences in more detail in Section 4.

The rest of the paper is organized as follows. In the next section, we present our network model and analytically derive the characteristic timescale. In addition, we prove various properties of the characteristic timescale including its independence of the traffic statistics. This is followed by Section 3 where we provide experimental evidence (using the *ns* simulator) for the theory presented in Section 2. We discuss related work in Section 4 and conclude in Section 5.

[1]We discuss how the mean end-to-end delay relates to the goodput later in this paper.

2 Characterizing End-to-end Delay

In this section, we analytically derive the characteristic timescale for a network. To do so, we assume that the traffic generation processes and the resulting queue length processes are wide sense stationary[2] (WSS) [29]. This implies that the queue lengths have well defined long time means and variances [27]. We now look at the timescales over which the fluctuations about the mean queue lengths become uncorrelated. Intuitively, fluctuations about the mean are caused by packet generation in some part of the network, and these fluctuations remain correlated till those packets have reached their destination. Thus, the time to get uncorrelated represents a measure of end-to-end delays. We refer to this timescale as the *relaxation time* for the network. As we shall see later, the longest lived fluctuation about the mean dominates the relaxation time. The time for this fluctuating mode to get uncorrelated provides a measure of the characteristic timescale of the network. As this timescale diverges, the network *breaks down*.

More formally, we start by describing our network model including our assumptions about the packet dynamics. From the equations describing the packet dynamics, we extract the equations governing the dynamics of queue lengths (which in turn, determine the mean end-to-end delay). We show that at high traffic loads, the equation for queue length dynamics can be simplified substantially — the resulting equation is dominated by a long timescale which we identify as the characteristic timescale. We show that at high loads, the characteristic timescale is independent of the statistics of traffic generation as well as instantaneous queue lengths. In the next section, we use simulations to show that the characteristic timescale is highly (positively) correlated to the mean end-to-end delay on the network, and that the correlation increases with increasing traffic load.

2.1 System Model

Network. We model a network as a connected graph $G = (V, E)$ where the set of vertices V represents the network nodes, and the set of undirected edges E represents the network links. Each node $v \in V$ has a set of $Z(v)$ neighbors denoted by

[2]Note that there is existing work [7] that shows that Internet traffic is non-stationary. However, the non-stationarity that the authors refer to reflects the change in the nature of the traffic distribution as the network load increases. The packet arrival process at a link is Weibull at low loads, and tends locally to Poisson at high loads. Since our regime is the high load regime, our assumptions about stationarity are still valid.

$nbr(v)$. In other words, we have

$$nbr(v) = \{u | u \in V, (v,u) \in E \text{ or } (u,v) \in E\} \quad (1)$$

We assume that each node v has a queue Q_v associated with it,[3] whose maximum size is unbounded. By assuming that queue lengths are unbounded, we ensure that no packet drops occur. Hence, as the traffic load increases, the average queue length increases, and the goodput decreases. We define network breakdown as the point where the average queue length goes to infinity, and the network goodput goes to zero. A similar effect is observed for the network goodput if the maximum queue size is finite — in this case, the network goodput goes to zero because of packet drops. We use the "unbounded queue size" model since it makes the exposition simpler. As stated earlier, we assume that the queue length processes are WSS, and therefore have well-defined long time means and variances. We also assume that the network carries a single traffic class, i.e., there are no priorities assigned to packets. Finally, we use a discrete time model for the system where events happen at discrete time steps. This can be easily generalized to a (more realistic) continuous time model.

Traffic. We make no specific assumptions about the traffic generation statistics at each node except that the traffic generation process is WSS (stated earlier) and has finite correlation times. This means that the traffic generation process has a time independent mean and finite memory. We also assume that the traffic demand matrix can vary rapidly or may be unknown.

Routing. We model routing as a diffusion-like process with some non-linearities. The non-linearities are introduced by the probability that a packet remains enqueued at a node once it has arrived there. As we shall see in Section 2.3, these probabilities depend on the queue lengths at the different nodes, resulting in effective packet interactions. It can be argued that approximating routing by a diffusion-like process may not closely model reality. However, we point out that at high loads, the behavior predicted by our theoretical model is borne out by experimental evidence where the standard shortest path routing algorithm is used (see Section 3). Therefore, we contend that as a first cut approximation, a diffusion model is adequate — further refinements of the routing model are left as future work.

Scheduling. We assume that packets are scheduled at each node in a first come, first served basis. In other words, when a packet p arrives at a node, it is enqueued at the end of a queue. Once all the packets ahead of p have been forwarded, p is scheduled for transmission.

Flow and Congestion Control. We do not make any assumptions about the flow or congestion control schemes used in the network except that they be sender-based, which is the norm in today's networks. More precisely, any action related to flow or congestion control is executed at the sender. Such actions typically entail the retransmission of lost packets and pacing the transmission of new packets. In other words, the effect of flow or congestion control mechanisms is to change the traffic generation statistics at the sender.

2.2 Deriving Characteristic Timescales

Let $p^n_{sd}(v,t)$ be the probability that a packet n going from source s to destination d is at node v at time t. We assume that the probability of the packet n being at node v at time $t+1$ is governed by a Markov process, described by the following equation

$$p^n_{sd}(v,t+1) = \theta^{sd}_{vn} p^n_{sd}(v,t) + \sum_{u \in nbr(v)} \frac{1-\theta^{sd}_{un}}{Z(u)} p^n_{sd}(u,t) \quad (2)$$

and the boundary condition

$$p^n_{sd}(v,t_0) = \delta_{v,s} \quad (3)$$

where $\delta_{v,s}$ is the kronecker-delta function and t_0 is the time at which packet n was injected into the network. The term θ^{sd}_{vn} represents the probability that packet n going from source s to destination d remains enqueued at node v at the next time step. We refer to this as the *staying probability* at node v. The staying probability depends on the queue length at node v and thus incorporates the effects of network traffic and packet interaction at the queues. We compute this quantity in Section 2.3.

Following the diffusion-like routing model, we say that a packet n at node v leaves v with a probability $(1-\theta^{sd}_{vn})$ and proceeds with equal probability to one of the $Z(v)$ neighbors of node v. Thus, the first term on the RHS of equation (2) represents the probability that n was at node v in the previous time step. The second term represents the probability that the packet n arrived at node v from one of its neighbors at the previous time step. The boundary condition in equation (3) states that the packet n has a probability 1 of being found at its source at the time it was created. When a packet n reaches its destination d, it stays there forever — we model this by assigning $\theta^{sd}_{dn} = 1$. For the rest of this paper, we assume that the staying probability of a packet n at node v is independent of n, s, and d and denote it by θ_v.[4] This follows from our assumption that all packets are treated as equal (no priorities). In addition, note that the term θ_v also depends on the time t, since the queue length (which determines θ_v) varies over time. We have not shown this dependence explicitly for reasons that will become clear in Section 2.3.

We now consider all the packets with the same destination d and define the queue length variable

$$Q_d(v,t) = \sum_{n,s} p^n_{sd}(v,t) \quad (4)$$

[3] Strictly speaking, queues are associated with links, and our results can be easily extended to take that into account. We adopt the current model in order to simplify our presentation.

[4] Strictly speaking, θ_v does depend on n and d since $\theta_d = 1$ if the destination of n is d. We gloss over this fact for a more convenient notation, and it does not affect our results.

Then the variable $Q_d(v,t)$ satisfies the equation

$$Q_d(v,t+1) = Q_d(v,t)\theta_v + \sum_{u \in nbr(v)} \frac{1-\theta_u}{Z(u)} Q_d(u,t) + f_v^d(t) \quad (5)$$

where $f_v^d(t)$ is the probability that a packet generated at node v with destination d gets added to the queue at v at time t. Note that equation (5) holds independent of the network load. In fact, we assume that the network load is high only in Sections 2.3, 2.4, and 2.5. Next, we define the total queue variable at a node v to be

$$Q(v,t) = \sum_{u \in V} Q_u(v,t) - Q_v(v,t) \quad (6)$$

The expression above is the sum over all the packets at node v at time t that are destined for nodes other than v. This is because packets destined for node v have already reached their destination and are not part of the queue at v. Summing over all nodes d in equation (5), and using equation (6) and the fact that $\theta_v = 1$ for all packets with destination v at node v, we have

$$Q(v,t+1) = Q(v,t)\theta_v + \sum_{u \in nbr(v)} \frac{1-\theta_u}{Z(u)} Q(u,t) + f_v(t) - (Q_v(v,t+1) - Q_v(v,t)) \quad (7)$$

where $f_v(t) = \sum_d f_v^d(t)$ represents the packet generation process at node v. We thus have the equation of motion for the variable $Q(v,t)$. The expression in the second line of the equation denotes the noise term $\eta_v(t)$ which is defined as

$$\eta_v(t) = f_v(t) + \psi_v(t) \quad (8)$$

where $\psi_v(t) = Q_v(v,t) - Q_v(v,t+1)$. The noise term is the sum of two quantities. The first, $f_v(t)$, is the packet generation process at node v, while the second, $\psi_v(t)$, arises because of packets reaching their destination v and exiting the network. Note that the noise term subsumes the traffic generation term $f_v(t)$. Now equation (7) can be rewritten as

$$|Q(t+1)\rangle = \mathbf{M}\,|Q(t)\rangle + |\eta(t)\rangle \quad (9)$$

where $|Q(t)\rangle$ and $|\eta(t)\rangle$ are $|V|$ dimensional column vectors[5] with entries $Q(v,t)$ and $\eta_v(t)$, respectively, and the matrix $\mathbf{M}$ is defined as

$$M_{vu} = \begin{cases} \theta_v & \text{if } v = u, \\ \frac{1-\theta_u}{Z(u)} & \text{for all } u \in nbr(v), \\ 0 & \text{otherwise.} \end{cases} \quad (10)$$

We note that the columns of the matrix $\mathbf{M}$ add to one. Thus, it has a set of $|V|$ right eigenvectors $\left|ev^{(\alpha)}\right\rangle$ and a set of $|V|$ left eigenvectors $\left\langle ev^{(\alpha)}\right|$. Each of these eigenvectors has a corresponding eigenvalue ($|V|$ in all) $|\lambda^{(1)}|, |\lambda^{(2)}|, \cdots, |\lambda^{(|V|)}|$. Using the property that the network is connected[6], it can be shown [24] that these eigenvalues satisfy the property

$$1 = |\lambda^{(1)}| \geq |\lambda^{(2)}| \geq \cdots \geq |\lambda^{(|V|)}| \quad (11)$$

Observe that if we take averages over long enough periods of time such that transient effects (or queue length fluctuations) die away, we get

$$\overline{|Q\rangle} = \mathbf{M}\,\overline{|Q\rangle} + \overline{|\eta(t)\rangle}, \quad (12)$$

where the bars denote long time averages. We now analyze the fluctuations about this mean queue length vector. We define fluctuations of the total queue length variable around the mean as

$$|\delta Q(t)\rangle = |Q(t)\rangle - \overline{|Q\rangle} \quad (13)$$

or, using equations (9), (12) and (13), we get

$$|\delta Q(t+1)\rangle = \mathbf{M}\;|\delta Q(t)\rangle + |\delta\eta(t)\rangle \quad (14)$$

where $|\delta\eta(t)\rangle = |\eta(t)\rangle - \overline{|\eta(t)\rangle}$.

We then examine the timescales over which the queue length fluctuations are correlated. By doing so, we can identify the longest timescale T_c after which the correlation in the queue length fluctuations (i.e. the variables $|\delta Q(t)\rangle$) die out rapidly. This means that the timescale T_c corresponds to the characteristic timescale of the network.

To compute the relaxation timescales, we decompose the fluctuations $|\delta Q(t)\rangle$ in terms of the eigenvectors of the matrix $\mathbf{M}$ as

$$|\delta Q(t)\rangle = \sum_\alpha \delta A^{(\alpha)}(t) \left|ev^{(\alpha)}\right\rangle = \sum_\alpha |\delta Q_\alpha(t)\rangle \quad (15)$$

where $|\delta Q_\alpha(t)\rangle$ is the contribution due to the αth eigenvector of $\mathbf{M}$. Substituting the expression in equation (15) into equation (14), and left multiplying by $\left\langle ev^{(\alpha)}\right|$ on both sides, we get

$$|\delta Q_\alpha(t)\rangle = \sum_{t'<t} e^{(t-t')\ln|\lambda^{(\alpha)}|} \left\langle ev^{(\alpha)}\right| \eta(t')\right\rangle \left|ev^{(\alpha)}\right\rangle \quad (16)$$

where $\left\langle ev^{(\alpha)}\right| \eta(t')\right\rangle$ represents the dot product of the row vector $\left\langle ev^{(\alpha)}\right|$ with the column vector $|\eta(t')\rangle$. Now consider a traffic fluctuation $|\delta\eta(t)\rangle$ that occurs at time t. This will result in fluctuations in the queue lengths (represented by $|\delta Q(t)\rangle$, and its modes $|\delta Q_\alpha(t)\rangle$). The timescale over which the fluctuation $|\delta Q_\alpha(t)\rangle$ occurs is given by $-\frac{1}{\ln|\lambda^{(\alpha)}|}$. Beyond this timescale, the fluctuation decays to a fraction $\frac{1}{e}$ of its original value and vanishes rapidly. Note that it is possible to choose a longer timescale as the characteristic timescale, during which the fluctuation decays to a smaller fraction $\frac{1}{e^c}$ ($c > 1$ constant) of its original value. This simply makes the resulting timescale a constant multiple of the value we have calculated.

[5] We use the "bra" and "ket" notation following [11] instead of the standard mathematical practice of using bold letters to denote column vectors. This was done for notational convenience, especially when denoting left and right eigenvectors.

[6] This implies that the Markov process described by $\mathbf{M}$ is aperiodic and irreducible.

The above assertion holds for all modes except the first mode that corresponds to the eigenvalue $|\lambda^{(1)}| = 1$ (since the quantity $-\frac{1}{\ln|\lambda^{(1)}|} \to \infty$). In this case, from equation (16), and the fact that $\left\langle ev^{(1)}\right| = [11\cdots1]$, we have

$$\begin{aligned} |\delta Q_1(t)\rangle &= \sum_{t'<t} e^{(t-t')\ln 1} \sum_{v\in V} \delta\eta_v(t') \left|ev^{(1)}\right\rangle \\ &= \left(\sum_{t'<t}\sum_{v\in V} \delta\eta_v(t')\right) \left|ev^{(1)}\right\rangle \end{aligned} \quad (17)$$

Since we are analyzing the network below the breakdown point, the queue lengths have not diverged. Therefore, over long time periods, the total number of packets being generated is the same as the total number of packets reaching their destination. Using this, and the definition of $\eta_v(t)$ in equation (8), we can say that the quantity in parenthesis in equation (17) is zero. This implies that $|\delta Q_1(t)\rangle = 0$. In other words, a fluctuation in the traffic does not produce any fluctuation in the $|\delta Q_1(t)\rangle$ corresponding to the first (largest) eigenvalue.

Hence, from the above analysis, we deduce that the longest relaxation timescale corresponds to the $|\delta Q_2(t)\rangle$ mode that is associated with the second largest eigenvalue, $|\lambda^{(2)}|$. The related relaxation timescale is given by

$$T_c = -\frac{1}{\ln|\lambda^{(2)}|} \quad (18)$$

which is the characteristic timescale of the network.

2.3 Computing Staying Probabilities

Since the characteristic timescale is given by the second eigenvalue of the matrix $\mathbf{M}$ in equation (9), we need to determine the matrix entries in order to compute the eigenvalues. The matrix entries depend on the staying probabilities θ_v at each node v. In this section, we establish a relationship between the staying probabilities θ_v and the queue length at node v, thereby incorporating the effects of network traffic load. Without loss of generality, we assume that the bandwidths of all the links are 1, and the queues are defined at nodes instead of links.

Let $Q(v,t')$ be the size of the queue at node v at time t'. Thus, a packet joining the queue at time t' must wait a time $Q(v,t')$ before it leaves, say, at time t'' (assuming that packets have no priorities). Thus, the waiting time for a packet at node v depends on the queue size not at the time when it leaves v, but when it arrives at v. This leads to memory effects since the waiting time depends on past history.

We now model this memory effect by postulating that after waiting for $Q(v,t')$ steps, the staying probability of a packet becomes small (since the packet is dequeued and sent on its way). Let the staying probability of a packet at time t while it is enqueued at the node v be $\theta_v(t)$. Then we have

$$\prod_{t=t'}^{t=t'+Q(v,t')} \theta_v(t) = \frac{1}{e} \quad (19)$$

As in the derivation of the characteristic timescale, we once again use $\frac{1}{e}$ (instead of some higher power thereof) to denote a small quantity. We now approximate equation (19) by

$$\theta_v(t)^{Q(v,t')} = \frac{1}{e} \quad (20)$$

or,

$$\theta_v(t) = \exp\left(-\frac{1}{Q(v,t')}\right) \quad (21)$$

Let the mean (long term) queue length at node v be $\overline{Q_v}$. We can write the queue length at node v as

$$Q(v,t') = \overline{Q_v} + \delta Q(v,t') \quad (22)$$

which implies that

$$\theta_v(t) = \exp\left(-\frac{1}{\overline{Q_v} + \delta Q(v,t')}\right) \quad (23)$$

Now consider the network when the traffic load is high. This means that $\overline{Q_v}$ is large, and we can use Taylor's series expansion to rewrite equation (23) as

$$\theta_v(t) = \exp\left(-\frac{1}{\overline{Q_v}}\right)\left(1 + \frac{\delta Q(v,t')}{\overline{Q_v}^2} + \text{higher order terms}\right) \quad (24)$$

In other words, we have

$$\theta_v(t) = \overline{\theta_v} + \delta\theta_v(t) \quad (25)$$

where $\overline{\theta_v} = \exp(-\frac{1}{\overline{Q_v}})$, and $\delta\theta_v(t)$ represents the fluctuations of the variable due to the fluctuations in the queue length about the mean. We can now rewrite equation (14) as

$$|\delta Q(t+1)\rangle = \mathbf{M}(\overline{\theta})\ |\delta Q(t)\rangle + O\left(\frac{\delta Q(t')}{\overline{Q}}\right)^2 + |\delta\eta(t)\rangle \quad (26)$$

where $O\left(\frac{\delta Q(t')}{\overline{Q}}\right)^2$ represents terms of quadratic order or higher in the fluctuations and $\mathbf{M}(\overline{\theta})$ is defined as

$$M_{vu}(\overline{\theta}) = \begin{cases} \overline{\theta_v} & \text{if } v = u, \\ \frac{1-\overline{\theta_u}}{Z(u)} & \text{for all } u \in nbr(v), \\ 0 & \text{otherwise.} \end{cases} \quad (27)$$

The quadratic and higher order terms in equation (26) have denominations of order $\frac{1}{\overline{Q}^2}$. Now note that at high loads, the mean queue lengths are large. We make the assumption that the fluctuations about the mean queue lengths are small compared to the mean in this regime. This makes the higher order terms in equation (26) vanish.[7] Thus, at high loads, we can rewrite the equation for queue length dynamics as

$$|\delta Q(t+1)\rangle = \mathbf{M}(\overline{\theta})\ |\delta Q(t)\rangle + |\delta\eta(t)\rangle \quad (28)$$

[7] It is possible to do a more careful analysis without approximating the higher order terms, using renormalization group theory [23]. However, empirical evidence (see Section 3) shows a very high correlation between mean end-to-end delays and the second eigenvalues computed using our approximation. Thus, using our approximation does not affect the validity of our results, and at the same time simplifies the presentation.

Thus, we can use the long time averages of queue lengths to compute the staying probabilities at each node. We also point out that this is the reason why we have not explicitly shown the dependence of the staying probabilities on time in Section 2.2. Using the staying probabilities, it is now easy to calculate the value of $|\lambda^{(2)}|$ using standard techniques (see, for example, [28]) and obtain the network's characteristic timescale.

2.4 Divergence of the Characteristic Timescale

We now describe how the characteristic timescale behaves as the traffic load on the network increases. Intuitively, as the network approaches breakdown point (with increasing traffic load), we expect the characteristic timescale to diverge. We now show that the characteristic timescale diverges continuously with the queue lengths.

Lemma 2.1 *The second largest eigenvalue (corresponding to the slowest mode)* $|\lambda^{(2)}| \rightarrow 1$ *continuously with the queue lengths.*

Proof: Consider a regime where the queue length at any node v is very large. Therefore, the staying probability $\overline{\theta_v} \rightarrow 1$. In other words, we have

$$\overline{\theta_v} = e^{-\frac{1}{\frac{Q_v}{}}} = 1 - \gamma_v \epsilon \tag{29}$$

where $\gamma_v = O(1)$ is a strictly positive constant and $\epsilon > 0$ is small and vanishes as the queue lengths go to infinity. In this case, we can write the matrix $\mathbf{M}$ from equation (9) as

$$M_{vu} = \begin{cases} 1 - \gamma_v \epsilon & \text{if } v = u, \\ \frac{\gamma_u \epsilon}{Z(u)} & \text{for all } u \in nbr(v), \\ 0 & \text{otherwise.} \end{cases} \tag{30}$$

As before, it is possible to show that the moduli of all the eigenvalues of this matrix is bounded by one. Consider the second largest eigenvalue of this matrix $|\lambda^{(2)}| = 1 - \delta$ (say). This eigenvalue satisfies the equation

$$det(\mathbf{C}) = 0, \tag{31}$$

with the matrix $\mathbf{C}$ given by

$$\mathbf{C}_{vu} = \begin{cases} \delta - \gamma_v \epsilon & \text{if } v = u, \\ -\frac{\gamma_u \epsilon}{Z(u)} & \text{for all } u \in nbr(v), \\ 0 & \text{otherwise.} \end{cases} \tag{32}$$

We now assume that $\delta = \epsilon S(\epsilon)$ where $S(\epsilon)$ is a term that depends on ϵ. Substituting the value of δ, we can rewrite the matrix $\mathbf{C}$ as $\mathbf{C} = \epsilon \mathbf{C}^{(1)}$ where

$$\mathbf{C}^{(1)}_{vu} = \begin{cases} S(\epsilon) - \gamma_v & \text{if } v = u, \\ -\frac{\gamma_u}{Z(u)} & \text{for all } u \in nbr(v), \\ 0 & \text{otherwise.} \end{cases} \tag{33}$$

Note that $S(\epsilon)$ is the smallest non-zero root of the equation $det(\mathbf{C}^{(1)}) = 0$, or equivalently, $-S(\epsilon)$ is the smallest eigenvalue of the matrix $\mathbf{C}^{(2)}$ with

$$\mathbf{C}^{(2)}_{vu} = \begin{cases} -\gamma_v & \text{if } v = u, \\ -\frac{\gamma_u}{Z(u)} & \text{for all } u \in nbr(v), \\ 0 & \text{otherwise.} \end{cases} \tag{34}$$

Since the columns of the matrix $\mathbf{C}^{(2)} + \mathbf{I}$ (where $\mathbf{I}$ is the identity matrix) add up to one, it can be shown that for any eigenvalue λ_C of $\mathbf{C}^{(2)}$, we have

$$|\lambda_C| < 2 \tag{35}$$

From this, we conclude that $|S(\epsilon)| < 2$. Therefore, as queue lengths diverge at high loads, $1 - \lambda_2 \leq 2\epsilon$. This implies that $|\lambda^{(2)}| \rightarrow 1$ continuously. In other words, the characteristic timescale of the network (given by $-\frac{1}{\ln |\lambda^{(2)}|}$) diverges continuously with the queue lengths. ∎

2.5 Dependence on Traffic Statistics

Thus far, we have derived the characteristic timescale for a network without taking into account the statistics of the network traffic. In particular, it is possible that the correlation timescales for the traffic generation statistics are longer than our derived value of $-\frac{1}{\ln |\lambda^{(2)}|}$. In such a case, the longest relaxation timescale will not be determined by $-\frac{1}{\ln |\lambda^{(2)}|}$, but by the correlation timescales of the traffic generation statistics. We now show that such is not the case by proving that the traffic statistics are irrelevant at high loads. Recall that we have assumed (see Section 2.1) that any flow or congestion control scheme is sender-based, where the net effect at the sender is to change the traffic generation statistics. Thus at high loads, the characteristic timescale is also not affected by the flow or congestion control scheme used in the network. For the rest of this paper, we assume that a *high load* regime is one where the queue lengths are large, and the magnitude of the traffic fluctuations about the mean is small compared to the mean itself.

Our proof proceeds in several steps. We first compute the time correlation of the noise term at node d, $\eta_d(t)$. This is the d^{th} component of the noise term $|\eta(t)\rangle$ in equation (9). Recall (from equation (8)) that this noise term subsumes the traffic generation term $f_d(t)$. We show that at high loads, the correlation timescales of the traffic generation statistics (from the $f_d(t)$ term) are dominated by the correlation timescales (say T_ψ) from the $\psi_d(t)$ term. This implies that the correlation timescales of the noise term $\eta_d(t)$ are independent of the traffic generation statistics.

We now relate the timescale T_ψ obtained in the previous step to the characteristic timescale $-\frac{1}{\ln |\lambda^{(2)}|}$ that we obtained in Section 2.2. We show that at high traffic loads, the terms T_ψ and $-\frac{1}{\ln |\lambda^{(2)}|}$ are roughly equal, and the difference between them diminishes as the traffic load goes higher. We can then infer the following theorem.

Theorem 2.1 *At high traffic loads, the characteristic timescale of a network is given by* $-\frac{1}{\ln|\lambda^{(2)}|}$*, independent of the details of traffic generation statistics at any node.*

We begin with computing how long it takes for the fluctuations in the noise term $\eta_d(t)$ (about the mean) to get uncorrelated (i.e., the correlation timescale). We estimate this quantity by calculating the autocovariance function for the noise term. We know from Section 2.2 that each component $\eta_d(t)$ of the noise term is given by

$$\eta_d(t) = f_d(t) + \psi_d(t) \tag{36}$$

Now the autocovariance term for $\eta_d(t)$ is given by

$$\mu(t_1, t_2) = E\{[\eta_d(t_1) - \overline{\eta_d(t_1)}][\eta_d(t_2) - \overline{\eta_d(t_2)}]\} \tag{37}$$

where the terms with the bars above them denote means. We denote the autocovariance function for any process $g(t)$ by $\langle g(t_1)g(t_2)\rangle, t_2 \geq t_1$. Now, using equations (36) and (37), we get

$$\begin{aligned}\langle \eta_d(t_1)\eta_d(t_2)\rangle &= \langle f_d(t_1)f_d(t_2)\rangle + \langle \psi_d(t_1)\psi_d(t_2)\rangle \\ &\quad + \text{other cross product terms}\end{aligned} \tag{38}$$

Observe that the cross product terms are not important in equation (38). This is because the corresponding correlation timescales will be dominated by the correlation timescales of one of the first two terms on the RHS of equation (38). Therefore, we can say that

$$\langle \eta_d(t_1)\eta_d(t_2)\rangle \sim \langle f_d(t_1)f_d(t_2)\rangle + \langle \psi_d(t_1)\psi_d(t_2)\rangle \tag{39}$$

where $\sim$ denotes "behaves as". We now focus on the term $\langle f_d(t_1)f_d(t_2)\rangle$. Since the traffic generation process is WSS, we know that the autocovariance term is a function of $t_2 - t_1$, and the mean is time independent. Therefore, we get

$$\langle f_d(t_1)f_d(t_2)\rangle = \mu(t_2 - t_1) = E\{[f_d(t_1)][f_d(t_2)]\} - \overline{f_d}^2 \tag{40}$$

where $\overline{f_d}$ is the time independent mean of the traffic generation process. We also know that the correlation timescale T_{gen} for $f_d(t)$ is finite. Therefore, we can write

$$\langle f_d(t_1)f_d(t_2)\rangle \sim e^{-\frac{t_2-t_1}{T_{gen}}} \tag{41}$$

Note that in the time T_{gen}, the correlation of the traffic fluctuations about the mean decays to a small value (namely, $\frac{1}{e}$). Our definition thus follows the spirit of Section 2.2, where we defined the characteristic timescale in a similar manner. Using equations (39) and (41), we get

$$\langle \eta_d(t_1)\eta_d(t_2)\rangle \sim e^{-\frac{t_2-t_1}{T_{gen}}} + \langle \psi_d(t_1)\psi_d(t_2)\rangle \tag{42}$$

We now need to compute the term $\langle \psi_d(t_1)\psi_d(t_2)\rangle$. By definition, we have

$$\psi_d(t) = Q_d(d, t) - Q_d(d, t+1) \tag{43}$$

or, using equation (5), we have,

$$\psi_d(t) = -\left(\sum_{u \in nbr(d)} \frac{1-\theta_u}{Z(u)} Q_d(u,t) + f_d^d(t)\right) \tag{44}$$

Since $f_d^d(t) = 0$ (the node d does not send packets to itself), we can now write the autocovariance term for $\psi_d(t)$ as

$$\begin{aligned}&\langle \psi_d(t_1)\psi_d(t_2)\rangle \\ &\quad \sim \sum_{\substack{u \in nbr(d) \\ u' \in nbr(d)}} \frac{1-\theta_u}{Z(u)} \frac{1-\theta_{u'}}{Z(u')} \langle Q_d(u,t_1)Q_d(u',t_2)\rangle\end{aligned} \tag{45}$$

Once again, we have ignored the cross product terms as in equation (39). We now compute the autocovariance for $Q_d(u,t)$, and then use equation (45) to compute the autocovariance for $\psi_d(t)$. The autocovariance for $Q_d(u,t)$ can be computed by rewriting equation (5) in the matrix form as

$$|Q_d(t+1)\rangle = \mathbf{M}^d |Q_d(t)\rangle + \left|f^d(t)\right\rangle \tag{46}$$

where $|Q_d(t)\rangle$ and $\left|f^d(t)\right\rangle$ are $|V|$ dimensional column vectors with entries $Q_d(v,t)$ and $f_v^d(t)$, respectively. The matrix $\mathbf{M}^d$ is defined in a manner similar to the matrix $\mathbf{M}$, except for entries in the column corresponding to node d. More precisely, we have

$$\mathbf{M}^d_{vu} = \begin{cases} 1 & \text{if } v = u = d, \\ \theta_v & \text{if } v = u \neq d, \\ \frac{1-\theta_u}{Z(u)} & \text{for all } u \in nbr(v), u \neq d, \\ 0 & \text{otherwise.} \end{cases} \tag{47}$$

Note that all the entries in the column corresponding to the node d (i.e., the destination node) are zero except for $\mathbf{M}^d_{dd}$ which is 1. In other words, once a packet reaches its destination d, it stays there (with probability 1). We now prove the following lemma.

Lemma 2.2 *The correlation timescale for* $Q_d(u,t)$ *is given by* $-\frac{1}{\ln|\lambda_d^{(2)}|}$*, where* $|\lambda_d^{(\alpha)}|$ *is the modulus of the* α*th eigenvalue of the matrix* $\mathbf{M}^d$*.*

Proof: Using an analysis similar to that in Section 2.2, it is possible to show that the $|V|$ eigenvalues of the matrix $\mathbf{M}^d$ can be written as

$$1 = |\lambda_d^{(1)}| \geq |\lambda_d^{(2)}| \geq \cdots \geq |\lambda_d^{(|V|)}| \tag{48}$$

We denote the right eigenvectors of $\mathbf{M}^d$ as $\left|ev_d^{(\alpha)}\right\rangle$ and the left eigenvectors as $\left\langle ev_d^{(\alpha)}\right|$. We can now decompose the column vector $|Q_d(t)\rangle$ into a linear sum of its right eigenvectors

$$|Q_d(t)\rangle = \sum_\alpha B_d^{(\alpha)}(t) \left|ev_d^{(\alpha)}\right\rangle \tag{49}$$

where the $B_d^{(\alpha)}(t)$s are scalars. Focusing on the u^{th} component ($1 \leq u \leq |V|$) of $|Q_d(t)\rangle$, we have

$$Q_d(u,t) = \sum_{\alpha} B_d^{(\alpha)}(t) ev_d^{(\alpha)}(u) \tag{50}$$

where $ev_d^{(\alpha)}(u)$ is the u^{th} component of $\left|ev_d^{(\alpha)}\right\rangle$. Therefore, we have

$$\begin{aligned}
&\langle Q_d(u,t_1) Q_d(u',t_2)\rangle \\
&= \left\langle \sum_{\alpha>1} B_d^{(\alpha)}(t_1) ev_d^{(\alpha)}(u) \sum_{\beta>1} B_d^{(\beta)}(t_2) ev_d^{(\beta)}(u') \right\rangle \\
&= \sum_{\alpha,\beta>1} \left\langle B_d^{(\alpha)}(t_1) B_d^{(\beta)}(t_2) \right\rangle ev_d^{(\alpha)}(u) ev_d^{(\beta)}(u') \qquad (51)
\end{aligned}$$

Note that the terms for $\alpha = 1$ and $\beta = 1$ are absent since they represent the long term mean, and we are looking at the autocovariance, or the fluctuations *about* the mean. The scalar coefficients $B_d^{(\alpha)}(t)$ can be computed by substituting equation (49) into equation (5) and left multiplying both sides by $\left\langle ev_d^{(\alpha)}\right|$. This yields

$$B_d^{(\alpha)}(t) = \sum_{t'<t} e^{(t-t')\ln|\lambda_d^{(\alpha)}|} \left\langle ev_d^{(\alpha)} \middle| f^d(t') \right\rangle \tag{52}$$

We now use Fourier transforms to compute the spectral densities of the time series described by the coefficients $B_d^{(\alpha)}(t)$. The Fourier transform of any time dependent function $g(t)$ is given by

$$g(\omega) = \int_{-\infty}^{\infty} g(t) e^{-i\omega t} dt \tag{53}$$

where $i^2 = -1$. Applying the Fourier transform to equation (52), we get

$$\begin{aligned}
&\left\langle B_d^{(\alpha)}(\omega_1) B_d^{(\beta)}(\omega_2) \right\rangle \\
&\sim \sum_{u\in V} \frac{1}{-i\omega_1 + \ln|\lambda_d^{(\alpha)}|} \frac{1}{i\omega_1 + \ln|\lambda_d^{(\beta)}|} \frac{\delta(\omega_1+\omega_2)}{-i\omega_1 + 1/\tau_u} \\
&+ \sum_{u\in V} \frac{1}{i\omega_1 + \ln|\lambda_d^{(\alpha)}|} \frac{1}{-i\omega_1 + \ln|\lambda_d^{(\beta)}|} \frac{\delta(\omega_1+\omega_2)}{-i\omega_1 + 1/\tau_u} \qquad (54)
\end{aligned}$$

where δ is the dirac-delta function, and τ_u is given by

$$\left\langle f_u^d(t_1) f_u^d(t_2) \right\rangle \sim e^{-\frac{t_2-t_1}{\tau_u}} \tag{55}$$

In other words, τ_u is the correlation timescale for the packets being generated at node u with destination d. Applying the Wiener-Khintchine theorem [29], we get the autocovariance from the inverse Fourier transform of this expression. Thus, we have

$$\begin{aligned}
&\left\langle B_d^{(\alpha)}(t_1) B_d^{(\beta)}(t_2) \right\rangle \\
&\sim e^{(t_2-t_1)\ln|\lambda_d^{(\alpha)}|} + e^{(t_2-t_1)\ln|\lambda_d^{(\beta)}|} \\
&\quad + \sum_{u\in V} e^{-\frac{t_2-t_1}{\tau_u}} \qquad (56)
\end{aligned}$$

Going back to equation (51), we can conclude that

$$\begin{aligned}
&\langle Q_d(u,t_1) Q_d(u',t_2)\rangle \\
&\sim \sum_{2\leq\alpha,\beta\leq|V|} e^{(t_2-t_1)\ln|\lambda_d^{(\alpha)}|} + e^{(t_2-t_1)\ln|\lambda_d^{(\beta)}|} \\
&\quad + \sum_{u\in V} e^{-\frac{t_2-t_1}{\tau_u}} \qquad (57)
\end{aligned}$$

Now consider the sum

$$\sum_{2\leq\alpha,\beta\leq|V|} e^{(t_2-t_1)\ln|\lambda_d^{(\alpha)}|} + e^{(t_2-t_1)\ln|\lambda_d^{(\beta)}|} \tag{58}$$

The term with the longest correlation timescale in this sum occurs when $\alpha = \beta = 2$ (from equation (48)). Thus, we have

$$\langle Q_d(u,t_1) Q_d(u',t_2)\rangle \sim e^{(t_2-t_1)\ln|\lambda_d^{(2)}|} + e^{-\frac{t_2-t_1}{\tau}} \tag{59}$$

where $\tau = \max \tau_u$. Assuming $|\lambda_d^{(2)}| \to 1$ at high loads (we prove this next), it is easy to see that the correlation timescale for the dominant term is given by $-\frac{1}{\ln|\lambda_d^{(2)}|}$. ∎

Using Lemma 2.2 and equation (45), we have

$$\langle \psi_d(t_1) \psi_d(t_2)\rangle \sim e^{(t_2-t_1)\ln|\lambda_d^{(2)}|} \tag{60}$$

This, together with the equation (42) implies that

$$\langle \eta_d(t_1) \eta_d(t_2)\rangle \sim e^{-\frac{t_2-t_1}{T_{gen}}} + e^{(t_2-t_1)\ln|\lambda_d^{(2)}|} \tag{61}$$

Once again, assuming $|\lambda_d^{(2)}| \to 1$ at high loads, it is easy to see that the correlation timescale for the noise term $\eta_d(t)$ is given by $-\frac{1}{\ln|\lambda_d^{(2)}|}$. We therefore conclude that the correlation timescale for the noise term $|\eta(t)\rangle$ in equation (9) is given by

$$T_{corr} = \max_{1\leq d\leq|V|} -\frac{1}{\ln|\lambda_d^{(2)}|} \tag{62}$$

We now have to show that for any node d, $|\lambda_d^{(2)}| \to 1$ at high loads and relate the $-\frac{1}{\ln|\lambda_d^{(2)}|}$ timescale to the characteristic timescale given by $-\frac{1}{\ln|\lambda^{(2)}|}$. This will complete the proof that the correlation timescale of the noise term in equation (9) is dominated by the characteristic timescale of the network. We accomplish this using the following lemma.

Lemma 2.3 *For any node d, at high loads, the eigenvalues $|\lambda_d^{(2)}|$ and $|\lambda^{(2)}|$ approach each other, and $|\lambda_d^{(2)}| \to 1$.*

Proof: Let d be a typical node in the network. Consider the matrix $\mathbf{m}^d$ defined as

$$\mathbf{m}^d = \mathbf{M}^d - \mathbf{M} \tag{63}$$

Then, the elements of the matrix $\mathbf{m}^d$ are given by

$$m_{vu}^{(d)} = \begin{cases} 1-\theta_d & \text{if } v = u = d, \\ \frac{\theta_u - 1}{Z(u)} & \text{for all } u \in nbr(v), v = d, \\ 0 & \text{otherwise.} \end{cases} \tag{64}$$

where the matrices $\mathbf{M}$ and $\mathbf{M}^d$ are defined by equations (10) and (47), respectively. Left and right multiplying equation (63) by the second left and right eigenvectors of $\mathbf{M}$, we get

$$\left\langle ev^{(2)} \right| \mathbf{M}^d \left| ev^{(2)} \right\rangle = |\lambda^{(2)}| + \left\langle ev^{(2)} \right| \mathbf{m}^d \left| ev^{(2)} \right\rangle \quad (65)$$

At high loads, the staying probability $\theta_v \to 1$ for any node v. Under such conditions, from the definition of matrix $\mathbf{m}^d$, we can deduce that $\mathbf{m}^d \to \mathbf{0}$, where $\mathbf{0}$ is the matrix with all zero entries. Thus, we can say that as $\theta_v \to 1$, we have

$$\left\langle ev^{(2)} \right| \mathbf{M}^d \left| ev^{(2)} \right\rangle \to |\lambda^{(2)}| \quad (66)$$

Therefore, as the network load rises, the second eigenvalue (and the second eigenvector) of the matrices $\mathbf{M}$ and $\mathbf{M}^d$ approach each other. Using Lemma 2.1, we can conclude that $|\lambda_d^{(2)}| \to 1$ at high loads. ∎

2.6 Summary

In this section, we have analytically shown three main results. First, we have proved that the mean end-to-end delay is dominated by the longest relaxation mode of the network, which we call the characteristic timescale of the network. Therefore, we expect the mean end-to-end delay to correlate well with the characteristic timescale. We have also shown how to compute the characteristic timescale from the second largest eigenvalue of the matrix $\mathbf{M}$. Second, the equation describing the dynamics for the queue lengths is given by equation (28). We have shown that this equation becomes a better and better approximation of the packet dynamics at higher and higher loads. This is because the quadratic and higher order terms in equation (26) vanish at high loads when the fluctuations in queue lengths are small compared to the mean queue lengths. Finally, we have shown that the statistics of the traffic generation do not affect the characteristic timescale of the network at high loads.

We have also shown in Section 2.4 that the characteristic timescale of the network diverges continuously with the queue lengths. This behavior has the hallmarks of a *second order phase transition*, where our description of the system using a single characteristic timescale increases in accuracy with higher loads. In contrast, a *first order phase transition* produces a discontinuous jump at the breakdown point. For such a case, one might never be able to get to high enough loads at which the approximations we have used are applicable. In the next section, we provide experimental evidence for second order phase transition behavior close to network breakdown.

3 Experimental Validation

In this section, we provide experimental validation of the theory presented in Section 2 using packet-level simulations. The simulations described here were all carried out using the network simulator *ns* [25]. We first describe the design of the experiments themselves, followed by an analysis of the results.

topology	# nodes	# links
WL	10	variable (avg: 33)
WAX	35	70
BA	35	99
ISP1	108	154
ISP2	83	157
ISP3	161	328
ISP4	79	147

Table 1: Topology Details

3.1 Design of Experiments

We designed the simulation experiments in order to gather evidence that would reinforce the theoretical analysis in the previous section. More specifically, we wanted to examine, using simulations, if network behavior (in particular, the mean end-to-end delay) at high loads is indeed determined by a single characteristic timescale predicted by our model. We first focus on the network topologies that we used, followed by the experiments that we ran.

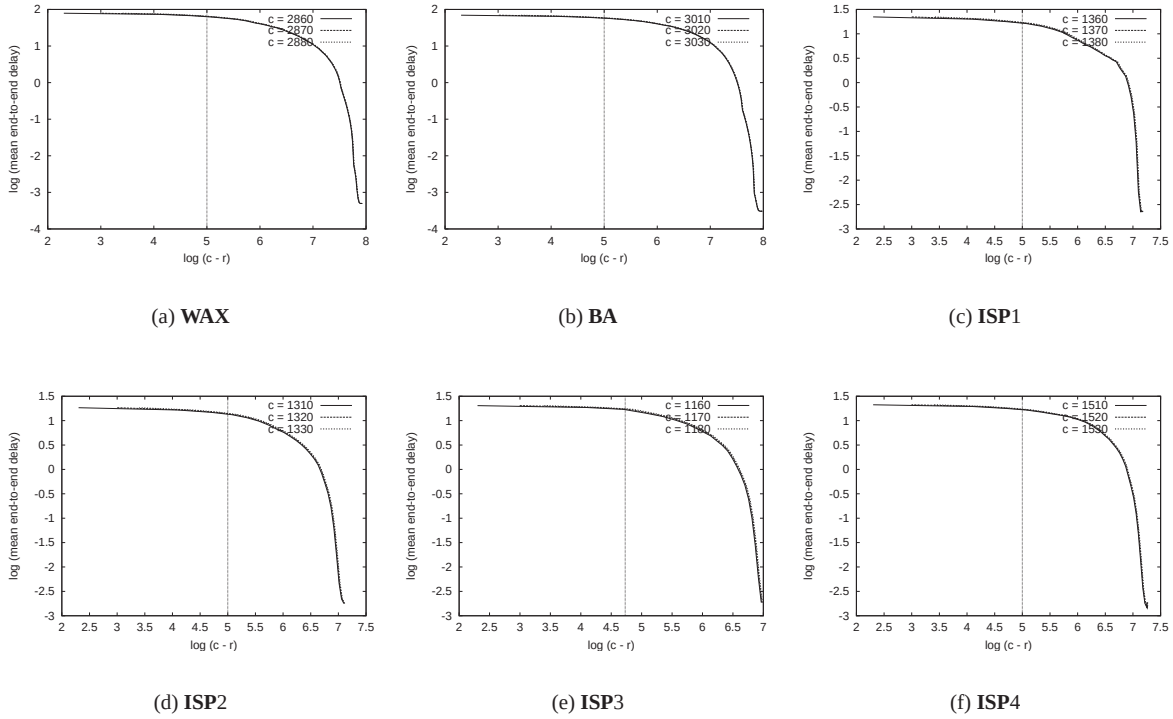
Network Topologies For this paper, we used three different sets of network topologies. The first set (labeled **WL**) consisted of 20 topologies, each of which was constructed to resemble an adhoc wireless network. Each topology in this set consisted of 10 nodes — the number of links varied from 22 to 72 (for reasons that will become clear later). We refer to this class of topologies as "wireless-like" topologies.

The topologies in the second set (labeled **ISP**) were based on real ISP topologies. Each of these 4 ISP topologies were obtained from the Rocketfuel [30] web site. Owing to scaling limitations of *ns*, and the computational power available to us, we have used a representative subset of the entire topology in each case. This subset usually consisted of the nodes in the core of the network — nodes and links at the edge of the network were removed. The details of each topology are shown in Table 1.

The topologies in the third set were synthetically generated using two different models. The first topology (labeled **WAX**) is based on the Waxman model [32], with randomly placed nodes on a 2-dimensional plane. The values of the Waxman-specific parameters α and β were set to 0.15 and 0.2, respectively. The second topology (labeled **BA**) used the Barabasi-Albert model [4] which uses a scale-free power-law to generate vertex connectivities. Once again, the nodes in the **BA** topology were randomly placed on a 2-dimensional plane, with each node connecting to 3 other nodes.

In each of these cases, for uniformity of comparisons, we set the link cost metrics to 1, the link bandwidths to 1Mbps and the end-to-end link propagation delays to 10ms.

Experimental Methodology The goal in our experiments was to study the effect of high loads on the mean end-to-end delay. To do so, we wanted to progressively increase the traffic



(a) **WAX** (b) **BA** (c) **ISP**1

(d) **ISP**2 (e) **ISP**3 (f) **ISP**4

Figure 1: Power Law Behavior Near Critical Points

on the network in such a way that each link was as uniformly loaded as possible. This would eliminate any effects due to pathological traffic mixes that overutilize bottleneck links while other parts of the network remain unutilized. In this manner, we would be able to obtain a lower bound on (or the best case estimate of) the network breakdown point and study the network behavior in its neighborhood. This gave rise to two main design issues.

First we needed a way of generating the traffic load. For this purpose, we used Constant Bit Rate (CBR) sources to generate traffic for loading the network. While it may be argued that real Internet traffic is not CBR, but more bursty in nature, we felt that CBR sources were adequate for our purposes. The reason (see Theorem 2.1) is that the queue length behavior (and hence, the mean end-to-end delay) is independent of the statistics of packet generation at high loads[8], which is the regime we want to study. Hence, for our purposes, CBR is as good a choice as any other traffic generator.

Second, we had to choose a routing algorithm. The obvious choice was the standard shortest path routing algorithm [10] (or **SPP**) that is used in today's Internet. However, one of the drawbacks of the **SPP** algorithm is that it does not adapt to traffic conditions — packets continue to get routed through areas of high congestion even though alternate, non-optimal routes may exist. Consequently, the traffic mix may have to be carefully selected in order to avoid the pathological scenarios described earlier in this section.

We have therefore chosen to use the **PBTA** algorithm (described in [5]) as our routing algorithm. The **PBTA** algorithm attempts to avoid highly congested areas by utilizing alternate paths, thereby distributing the traffic load more evenly. Thus, even if the traffic mix is not uniform, the network will still exhibit breakdown behavior as if all the network links were loaded more or less uniformly. We have also verified experimentally that there is no major qualitative change in the network behavior when the **PBTA** algorithm is used instead of the **SPP** algorithm. In the rest of this section, all the experiments that we describe were run using the **PBTA** routing algorithm, unless otherwise indicated.

Description of Experiments We ran each experiment for a simulation time of up to 6000 seconds. During this time half the nodes in the network sent traffic to the other half, where the source-destination pairs were chosen at random. The source-destination pairs were changed periodically, with the time period varying between 5 and 20 seconds. This ensured that the traffic demand matrix changes rapidly.

In order to estimate the mean end-to-end delay, we decided to let the queues grow without bounds by setting the maximum queue lengths to infinity. Otherwise, if queue sizes were finite, the end-to-end delays for traffic rates above some threshold would be bounded by the maximum queue size, while more and more packets would be dropped. This would make it difficult to observe how end-to-end delays increase with increasing

[8]Recall that by high loads, we mean loads where the magnitude of the traffic fluctuations about the mean are small compared to the mean itself.

traffic load. Once the experiments had completed, we computed the mean end-to-end delay for each sending rate, and network topology. This data was further processed to compute other metrics as described in the next few subsections.

3.2 Phase Transitions

We first attempted to understand the network behavior at high loads in terms of dynamic phase transitions near critical points. This is important because the validity of the approximations used in our theoretical analysis depends on the nature of the phase transitions (see Section 2.6 for details). The literature (see, for example, [20]) identifies two different kinds of phase transitions. A *first order* phase transition is one where a measured physical quantity shows a discontinuous change at the critical point. In our case, the measured physical quantity is the mean end-to-end delay, and the critical point is the point where network breakdown occurs.

In contrast, a *second order* phase transition occurs when the measured physical quantity changes continuously as the critical point is approached. In particular, it has been observed that such continuous phase transitions can be characterized by parameters that are known as *critical exponents*. These exponents express the behavior of the measured physical quantity in terms of a power law near the critical point. Another feature of dynamic critical phenomena is critical slowing down, where the dynamics of the system slows down tremendously as the critical point is approached. This is often characterized by the timescale of the system, T behaving as

$$T \sim (c - r)^{-z} \tag{67}$$

where c is called the *critical rate*, and $c - r$ measures the distance from the critical point and z is called the *dynamic critical exponent* [17]. Observe that as the sending rate $r \to c$, the system timescale $T \to \infty$. In other words, the mean end-to-end delay diverges continuously (and the goodput goes to zero) as we approach the critical rate from below. Such a divergence in the mean end-to-end delay signals network breakdown.

The plots in Figure 1 show log-log plots of the mean end-to-end delay as a function of the difference $c - r$, where r is the sending rate. Since we cannot analytically determine c for a given network, we have used a few different values for c to plot the data for each network. For example, in Figure 1(a), the values for c that we have used are 2860Kbps, 2870Kbps, and 2880Kbps. In each case, we see that the log-log plots are linear at higher loads — this regime is shown in each graph by all the points to the left of the vertical line.

The above observation implies two things: first, the nature of the phase transition is continuous or second order, and second, there is a power law that describes this behavior. The dynamic critical exponent (of the power law) can be determined from the slope of the linear portion of the graph. While this exponent depends on the value of the critical rate c that we have used for the fit, we see that the plots for each individual value of c are more or less identical. We conclude that the qualitative behavior of the mean end-to-end delay (at high loads) obeys a power law robustly and is not very sensitive to the value of the critical rate c.

3.3 Mean End-to-end Delay and Eigenvalue Correlations

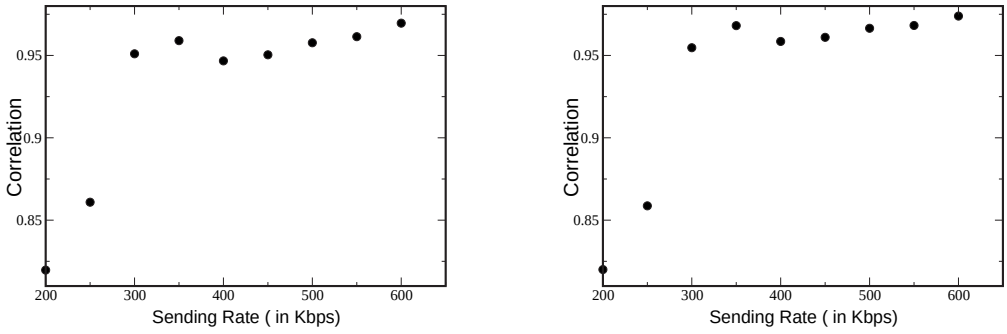
Having satisfied ourselves that the mean end-to-end delay diverges continuously as the critical load is approached, we proceeded to examine if there was indeed a characteristic network timescale that correlates to the mean end-to-end delay. We used the **WL** network topologies for this purpose. These topologies were constructed by starting with one randomly constructed topology and then changing it by adding/subtracting links to/from it. The **WL** topologies thus gave us a good data set for examining how the mean end-to-end delay changes as the network topology (and therefore, the characteristic timescale) changes, while the number of network nodes remain invariant.

For each **WL** topology and each sending rate from 200Kbps to 600Kbps, we computed the mean end-to-end delay over all source-destination pairs. We also computed the second largest eigenvalue (λ_2) for each topology and sending rate combination using Matlab. This was done by feeding in the average queue lengths from the *ns* simulations into the computation procedure described earlier. The queue length averages were taken over a time period larger than the mean end-to-end delay. This ensured that these averages were a good approximation of the long time queue length averages described in Section 2.3. Given the topology of the network and the mean queue lengths, we were then able to compute the matrix $\mathbf{M}$ in equation (9) and its second largest eigenvalue, $|\lambda^{(2)}|$. Finally, for each sending rate, we computed the correlation between $-\frac{1}{\ln|\lambda^{(2)}|}$ and the mean end-to-end delay over all the networks.

The results are shown in Figure 2. We see that there is indeed a high correlation between $-\frac{1}{\ln|\lambda^{(2)}|}$ and the mean end-to-end delay over a wide range of sending rates. As the rate increases, the correlation increases, which implies that the characteristic timescale becomes a better predictor of the mean end-to-end delay with increasing load. Note that this holds irrespective of the routing algorithm since we have almost identical curves for both the **SPP** and the **PBTA** algorithms. Furthermore, our results indicate that both the diffusion model approximation (Section 2.1) and neglecting the non-linearities in the queue length fluctuations (Section 2.3) are justified at high loads.

3.4 Spectral Analysis

In order to further reinforce our claim about a single characteristic timescale, we examined how the total number of packets in transit across the network varies over time. In particular, we were interested in looking at the time correlations of the total number of packets in the network. If there is indeed a single characteristic timescale that dominates at high loads, the time correlations should be the highest for a time lag equal to the



(a) **SPP** (b) **PBTA**

Figure 2: Correlation Between $-\frac{1}{\ln|\lambda^{(2)}|}$ and Mean End-to-end Delay at Different Rates

characteristic timescale, and die out for time lags larger than this timescale.

Another way of looking at the problem is to analyze the packet time series data in the frequency domain. As explained in Section 2.5, we know that the time correlations and the power spectra of the time series are Fourier transforms of each other. We therefore looked at the power spectra for the time series data corresponding to the number of packets in the network for a variety of networks at different traffic loads.

The time series data was obtained using instantaneous snapshots of the network at 20 second intervals. The total number of packets in the system was computed as the sum of all the instantaneous queue lengths and the packets "in transit" on a link. The results of the spectral analysis of the time series data are shown in Figure 3. Each graph represents a different network topology, and consists of three subgraphs, each subgraph corresponding to a different sending rate. In each subgraph, the X-axis shows the frequency, and the Y-axis shows the percentage of total energy that each frequency component contributes. The sending rate is shown at the top right hand corner of each subgraph. From the graphs, it is clear that as the traffic load (sending rate) increases, more and more energy shifts to the low frequency components (i.e., to the slower modes). In addition, a single frequency separates out as the dominant energy component while the others have very low energy contributions. This indicates that there is indeed a single characteristic timescale above which the time correlations for the packet statistics vanish, which is precisely what the theory predicts.

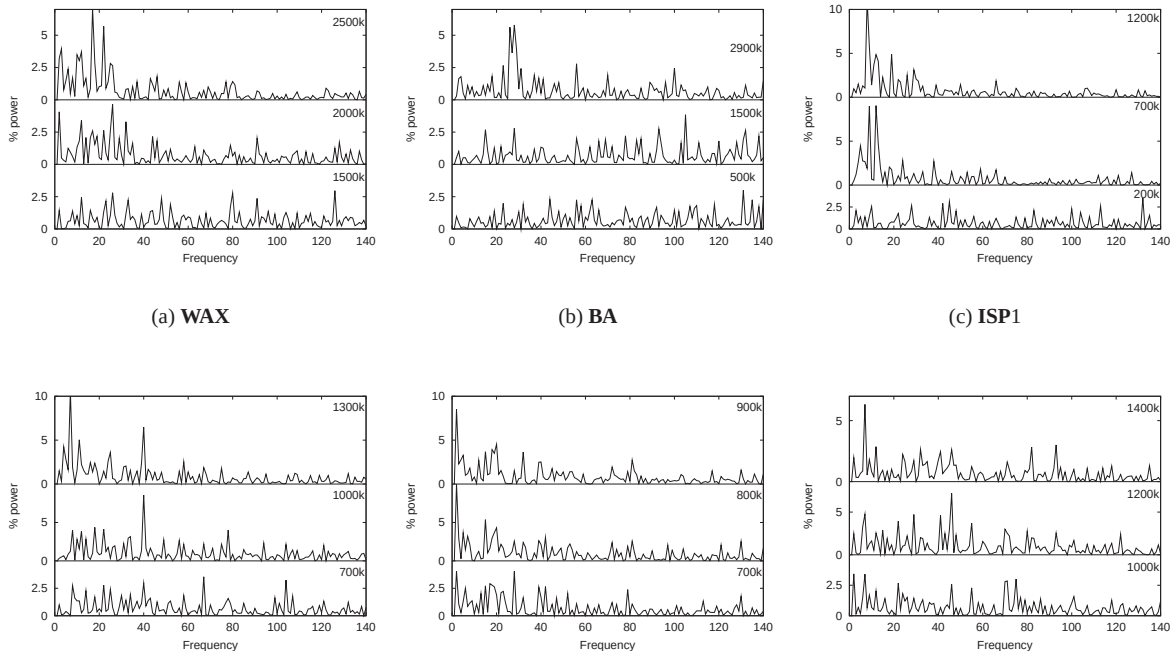
3.5 Summary

In this section, we have used simulations to present empirical evidence for the theory outlined in Section 2. First, we have provided experimental evidence that the network behavior close to breakdown exhibits a second order like phase transition. In other words, the mean end-to-end delay diverges continuously as the load approaches the critical rate from below. Next, we have shown that there is a single characteristic timescale that is related to the mean end-to-end delay at high loads using two different methods. First, we have demonstrated that the correlation between the analytically computed parameter $-\frac{1}{\ln|\lambda^{(2)}|}$ and the mean transit time is high. This correlation increases as the traffic load on the network goes up and is independent of the routing algorithm. In addition, we have been able to compute the characteristic timescale without any knowledge about the nature of the rapidly changing traffic demand matrix, or the statistics of traffic generation. Second, we have looked at the power spectrum for the time series corresponding to the total number of packets in the network. At high loads, we have found that there is a single low frequency that dominates the power spectrum. This is consistent with the analytics in Section 2.

4 Related Work

Our work is closest in spirit to a body of work that has used multicommodity flows to study network congestion. Graph properties such as the sparsest cut (or the minimum cut ratio) [3, 21] and conductance [15, 31] have been used to characterize congestion in networks with traffic demand matrices that are known apriori and are time invariant. The worst congestion in a graph is inversely related to the maximum concurrent flow in a graph with unit edge capacities. In [3], it has been shown that the maximum concurrent flow that can be achieved in an arbitrary graph is within a factor $O(\log k)$ of the minimum cut ratio, where k is the number of commodities. Conductance has been used in [15] to show that the maximum congestion in a power law random graph with n nodes is at best $O(n \log^2 n)$. When the shortest path routing algorithm is used, the maximum congestion is $\Omega(n^{1+O(1)})$ [2].

Our approach complements the work described above in multiple ways. First, we analyze the network in a different regime, where the traffic demand matrices change frequently or are unknown. At high loads with changing demand matrices, the current traffic patterns depend on both the current demand matrix and its past history. Under such conditions, it is not possible to use quasi-static approximations by assuming that de-



(a) **WAX** (b) **BA** (c) **ISP**1

(d) **ISP**2 (e) **ISP**3 (f) **ISP**4

Figure 3: Spectral Analysis Near Critical Points

mand matrices change slowly enough such that a steady state traffic pattern has been established. This means that the techniques related to multicommodity flows cannot be applied to the regime we study in a straightforward manner. Second, our goal is to study the effect of high loads on the mean end-to-end delay. In contrast, the work on multicommodity flows focuses on computing congestion levels in a network for a given demand matrix. Therefore, it is not necessary to explicitly model the behavior of network queues that affects end-to-end delays. On the other hand, our analysis must take into account the non-linearities introduced by packet interactions at the queues. At high loads, we have shown how network behavior can be predicted accurately by modeling these packet interactions using mean queue lengths.

Techniques from spectral graph theory [8] have also been used to study structural properties of network topologies, such as clustering [16, 18], degree distributions [1] and for characterizing distribution of information flow times in adhoc networks [12]. The relationships between graph topologies and their eigenvalues has also been studied (see, for example, [9]). Once again, our work complements these by studying the mean end-to-end delay as a function of both the static network topology as well as the traffic dynamics.

Finally, the treatment of physical systems by understanding their long wavelength physics has a long history (see, for example, [14]). So does the field of dynamic critical phenomena and phase transitions [17]. It is also possible to model network traffic as a fluid flow problem and study its dynamics [19, 22]. However, in our model, the effective interactions between packets through waiting times at a node (which in turn depends on all other packets at that node) produces non-linearities. This kind of non-linearity may be accounted for by postulating that the viscosity of the fluid is some complicated function of the fluid density. We are unaware if postulating a divergent viscosity near the network breakdown point might map our model to a complex fluid flow model.

5 Conclusions

In this paper, we have shown that the mean end-to-end delay in networks under high loads can be characterized by a single parameter, the characteristic timescale. This parameter is associated with the time it takes for the traffic fluctuations on the network to become uncorrelated. Such a characterization becomes arbitrarily accurate as the load increases, and we have proved that it is independent of the details of traffic generation and of the evolution of demand matrices. Our simulations using real network topologies have corroborated the analytical results. Furthermore, our simulations have shown that from time averaged queue length information, and an eigenvalue computation, we can reliably predict the changes in the mean end-to-end delay as a function of the network topology and the traffic load.

In the future, we would like to extend this work in two ways. First, we will further explore methods to minimize packet delay in networks by deforming them (subject to constraints) in

a manner that reduces the characteristic timescale. Second, we would like to refine our model to take into account the effects of feedback on the traffic generation process as evidenced by protocols such as TCP.

References

[1] W. Aiello, F. R. K. Chung, and L. Lu. Random Evolution in Massive Graphs. In *Proceedings of the 42nd IEEE Symposium on Foundations of Computer Science*, pages 510–519, Las Vegas, NV, October 2001.

[2] A. Akella, S. Chawla, A. Kannan, and S. Seshan. Scaling Properties of Internet Graphs. In *Proceedings of the 22nd ACM Symposium on Principles of Distributed Computing*, pages 337–346, Boston, MA, July 2003.

[3] Y. Aumann and Y. Rabani. An $O(\log k)$ Approximate Min-Cut Max-Flow Theorem and Approximation Algorithm. *SIAM Journal of Computing*, 27(1):291–301, 1998.

[4] A. L. Barabási and R. Albert. Emergence of Scaling in Random Networks. *Science*, 286:509–512, 1999.

[5] A. Basu, A. Lin, and S. Ramanathan. Routing Using Potentials: A Dynamic Traffic-Aware Routing Algorithm. In *Proceedings of SIGCOMM '03*, pages 37–48, Karlsruhe, Germany, August 2003.

[6] T. Bu and D. Towsley. Fixed Point Approximations for TCP Behavior in an AQM Network. In *Proceedings of SIGMETRICS '01*, pages 216–225, Cambridge, MA, June 2001.

[7] J. Cao, W. S. Cleveland, D. Lin, and D. X. Sun. On The Nonstationarity of Internet Traffic. In *Proceedings of SIGMETRICS '01*, pages 102–112, Cambridge, MA, June 2001.

[8] F. R. K. Chung. *Spectral Graph Theory (CBMS Regional Conference Series in Mathematics, No. 92)*. American Mathematical Society, Providence, RI, May 1997.

[9] F. R. K. Chung, L. Lu, and V. Vu. Eigenvalues of Random Power Law Graphs. *Annals of Combinatorics*, 7(1):21–33, 2003.

[10] E. W. Dijkstra. A Note on Two Problems in Connexion with Graphs. *Numerische Mathematik*, 1:269–271, 1959.

[11] P. A. M. Dirac. *Principles of Quantum Mechanics*. Clarendon Press, Oxford, UK, 4th edition, March 1982.

[12] R. D'Souza, S. Ramanathan, and D. Temple-Lang. Measuring Performance of Adhoc Networks Using Timescales for Information Flow. In *Proceedings of Infocom '03*, pages 1564–1574, San Francisco, CA, March – April 2003.

[13] D. Dutta, A. Goel, and J. Heidemann. Faster Network Simulation with Scenario Pre-Filtering. In *Proceedings of the 10th IEEE/ACM International Symposium on Modeling, Analysis and Simulation of Computer and Telecommunication Systems*, pages 303–314, Fort Worth, TX, October 2002.

[14] D. Forster. *Hydrodynamic Fluctuations, Broken Symmetry, and Correlation Functions*. Perseus Publishing, New York, NY, January 1994.

[15] C. Gkantsidis, M. Mihail, and A. Saberi. Conductance and Congestion in Power Law Graphs. In *Proceedings of SIGMETRICS '03*, pages 148–159, San Diego, CA, June 2003.

[16] C. Gkantsidis, M. Mihail, and E. Zegura. Spectral Analysis of Internet Topologies. In *Proceedings of Infocom '03*, pages 364–374, San Francisco, CA, March – April 2003.

[17] P. C. Hohenberg and B. I. Halperin. Theory of Dynamic Critical Phenomena. *Reviews of Modern Physics*, 49(3):435–479, 1977.

[18] J. M. Kleinberg. Authoritative Sources in A Hyperlinked Environment. *Journal of the ACM*, 46(5):604–632, 1999.

[19] K. Kumaran and D. Mitra. Performance and Fluid Simulations of A Novel Shared Buffer Management System. In *Proceedings of Infocom '98*, pages 1449–1461, San Francisco, CA, March – April 1998.

[20] L. D. Landau, J. B. Sykes (Translator), M. J. Kearsley (Translator), E. M. Lifshitz, and L. P. Pitaevskii (Editor). *Statistical Physics (Course of Theoretical Physics)*, volume 5. Butterworth-Heinemann, 3rd edition, December 1999.

[21] F. T. Leighton and S. Rao. Multicommodity Max-Flow Min-Cut Theorems and Their Use in Designing Approximation Algorithms. *Journal of the ACM*, 46(6):787–832, 1999.

[22] Y. Liu, F. L. Presti, V. Misra, D. Towsley, and Y. Gu. Fluid Models and Solutions for Large-Scale IP Networks. In *Proceedings of SIGMETRICS '03*, pages 91–101, San Diego, CA, June 2003.

[23] S. K. Ma. *Modern Theory of Critical Phenomena*. W. A. Benjamin Inc., Reading, MA, 1976.

[24] J. R. Norris. *Markov Chains*. Cambridge University Press, Cambridge, UK, August 1998.

[25] The Network Simulator — ns-2. http://www.isi.edu/nsnam/ns/.

[26] J. Padhye, V. Firoiu, D. Towsley, and J. Kurose. Modeling TCP Throughput: A Simple Model and its Empirical Validation. In *Proceedings of SIGCOMM '98*, pages 303–314, Vancouver, Canada, September 1998.

[27] R. K. Pathria. *Statistical Mechanics*. Pergamon Press, New York, NY, 1972.

[28] W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling. *Numerical Recipes in C*. Cambridge University Press, Cambridge, UK, 2nd edition, January 1993.

[29] M. B. Priestley. *Spectral Analysis of Time Series*. Academic Press, London, UK, June 1981.

[30] Rocketfuel: An ISP Topology Mapping Engine. http://www.cs.washington.edu/research/networking/rocketfuel/.

[31] A. Sinclair. *Algorithms for Random Generation and Counting: A Markov Chain Approach*. Birkhäuser Boston, Boston, MA, February 1993.

[32] B. Waxman. Routing of Multipoint Connections. *IEEE Journal on Selected Areas in Communications*, 6(9):1617–1622, December 1998.