Network Congestion as an Emergent Phenomena in Internet Traffic

Sean Vig
Physics 569 Term Essay

December 15, 2011

Abstract

This paper investigates the onset of congestion as an emergent phase in Internet networks at the network (IP) layer as a function of packet creation rate, which is the order parameter for these system. This paper addresses the theoretical background of traffic flow, the computational modeling used to investigate the emergence of congestion and the comparison of the network models to actual Internet traffic flow problems. Finally, this paper addresses some of the criticisms of these models, and notes ways which these models could be improved to better model Internet congestion.

1 Introduction

In the last several years, our modern society has seen an enormous growth in our reliance on the Internet, a computer network serving an ever increasing amount of media, from television and movies to news and messages to users around the world. As the Internet grows and encompasses more components of everyday life, it becomes increasingly important to understand the dynamics of network traffic across the Internet, both to optimize the traffic and ensure it does not surpass critical levels. While such a problem appears at first sight to be in the domain of computer and networking engineers, as far back as 1994, physicists began to realize the methods of statistical mechanics and dynamical systems analysis could be applied to the understanding and describing of network traffic [1]. While this has been used to show collective behavior in complex networks, such as the structure of the Internet [2] and the organization of web page links [3], this paper focuses on the treatment of Internet traffic and the understanding of this traffic as having a phase transition to an emergent state of congestion determined solely by packet creation rate. This will address both the formulation of this understanding through simulation and the comparison to samples of actual Internet traffic. This will then conclude with brief criticisms of this method and ways in which these models can be improved.

One of the first studies to show the self-similar behavior of the Internet, prompting the

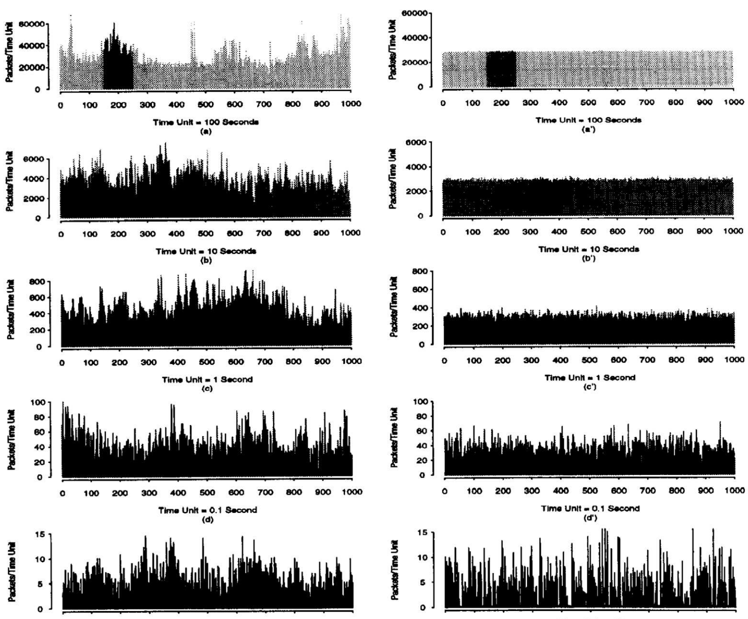

Figure 1: Showing the self-similarity of Internet traffic as a function of packet transmission rate. The figures on the left show actual network traffic and the figures on the right show the same time scales for simulated Poisson distributed traffic [4]

discussion of the Internet as a dynamical system, was a study of the round-trip times (RTT) of pings, which is the time it takes to send a message out from a local client and receive a response back from a remote server [1]. This study found that these response times followed a 1/f-like behavior in the power spectrum, which was a novel result in terms of understanding the dynamics of network traffic as being self-similar. This paper goes on to note, because of this power law spectrum, methods used to describe highway traffic, which was known at the time to show similar power-law spectra may be used to model Internet traffic. Also in this year, a study was performed the Bellcore Morris Research and Engineering Center collecting data on the rate of packet transmission [4]. While models and algorithms at the time assumed Poisson distributions of packet traffic, this study showed a self-similar traffic pattern, see Figure 1. Not only did this overturn the then current notions of Internet traffic as merely functions of individual host traffic, but it opened the door to the understanding of the collective effects shown in Internet traffic.

2 Technical Background

Before describing Internet models, it is important to understand the basic technical specifications of TCP/IP traffic and its relationship to the Open Systems Interconnection (OSI) network model, as the network simulations and modeling of phase dynamics must model traffic at some level of the OSI model. The OSI model describes the various levels of the network framework and how network traffic can be framed at various levels of abstraction. Each of the seven layers of the model builds upon the layer below it, often adding some additional semantics to define the protocol and allow for data transfer. The model starts at the lowest level with the physical network connection protocol allowing for the transmission of individual bits, such as Ethernet or Bluetooth protocols, up to the highest level describing application protocols, such as FTP and VT. This background will only cover network specifications and protocols as is necessary to the study of phase transitions, for a general review of network dynamics as it pertains to network engineering and how this is related to the study of phase transitions, see [5].

For our purposes, we focus on the network and transport layers, the third and fourth layers of the OSI model, respectively. The network level describes the grouping of network data into packets with various identifying information, such as the address of the source and destination hosts and hash checking features. The IP protocol and its replacement, the IPv6 protocol, dominate internet traffic at this level of the OSI model. More importantly, in addition to the additional header information, the creation of packets allows for the transmission of data (apart from the overhead of the various headers and syntax elements) the size of which is specified by the packet header.

The other layer that is of interest in network modeling is the transport layer, which provides a means of transparent data transfer for end clients. This allows higher layers to pass data to the transport layer, and the transport protocol will handle the creation and exchange of packets. Some transport layer protocols, most notably TCP, are able to provide reliable connections by ensuring data arrival through an exchange of synchronization (SYN)

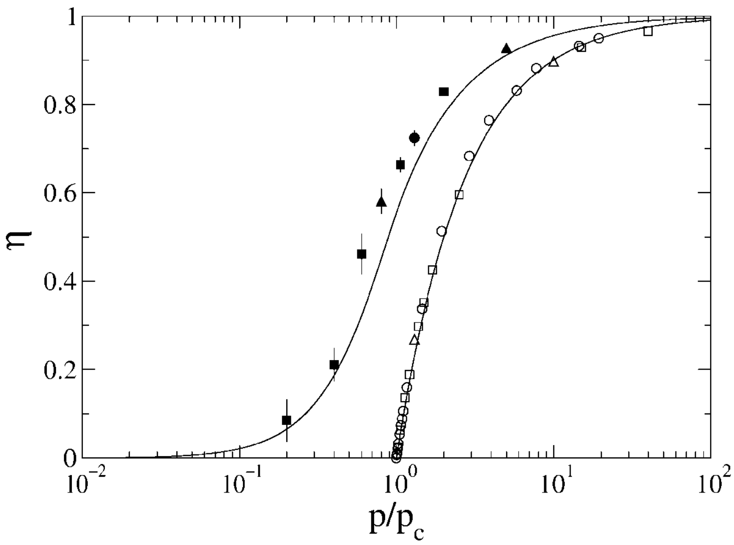

Figure 2: This shows the variation of the order parameter, ν, as a function of the packet creation rate, here denoted p, and the critical packet creation rate, p_c. This simulation used a Caley tree model. The shaded shapes show the simulations with noise and the open shapes without noise. The lines show analytical results for the 2 node case [6].

and acknowledgement (ACK) packets, and is the primary constituent in internet traffic, accounting for traffic sources such as HTTP and FTP. Other protocols, such as UDP, may forgo ensuring transmission reliability and establishing connections in favor of low overhead when such checks are not critical, as is the case with DNS traffic. Of particular interest, TCP, which is the largest contributor to Internet traffic, implements a congestion control mechanism to help ensure reliable data transmission. This congestion control is implemented by varying the packet size and the number of packets transmitted based on feedback from the SYN/ACK process, increasing data throughput with continued successful transmission and decreasing throughput on failed transmission.

3 Simulations and Results

Now, we start to look at the experimental results and the theoretical analysis of the simulated models. In all the models, the system is analyzed at the network layer, that is there is a set rate of packet creation, denoted as λ, and they find a critical rate of packet creation, λ_c, which gives a phase transition between free flowing and congested states. The order parameter which arises in the congested state is ν, a measure of the rate of increase of the number of packets in the system, written $\nu = \frac{1}{S\lambda}\frac{dN}{dt}$, where S is a measure of the size of the system and N is the total number of packets in the system at a given time. An example showing the onset of non-zero values for this order parameter at λ_c is shown in Figure 2

While the general methodology is the same, each model has a varying degree of complex-

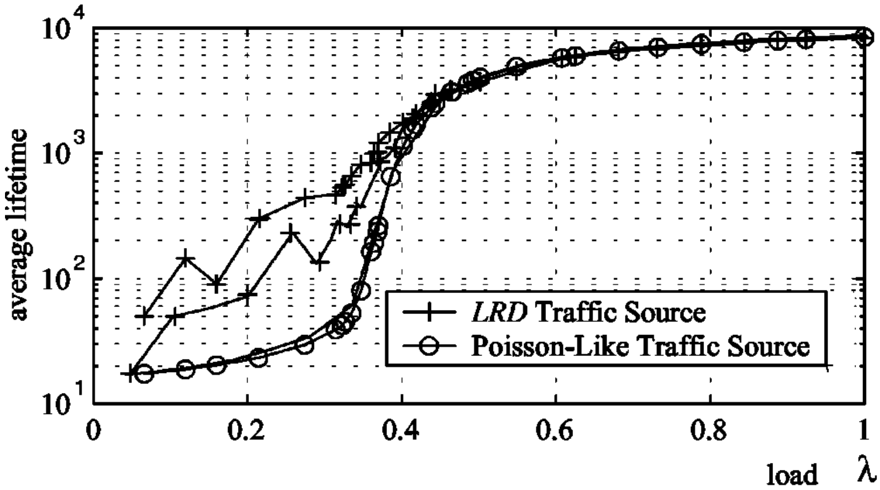

Figure 3: The lifetime of packets in a square lattice model as a function of the packet creation rate λ for traffic sources with either long range dependence (LRD) or Poisson statistics. Note the onset of congestion at the critical value $\lambda = 0.30$ [8].

ity, with the obvious trade-off between giving solvable theoretical predictions and presenting a more realistic model of the Internet. The dynamics of the network structure and setup of these models vary in a few key ways, each of which will be explored. The first is the structure, topology and size of the network, which specifies both the number of nodes used in the model and how the nodes are arranged. Also with this consideration is the routing strategy used, including how the routers in the model handle packets and any real-world limitations, such as finite buffer size, imposed on the routers. Finally, different models can also simulate various congestion handling schemes implemented in real-world Internet hardware and protocol, such as TCP congestion avoidance and router queuing and scheduling algorithms. We will look at these factors and see their impact on the resulting network traffic and the theoretical understanding of the network dynamics. For a more detailed review of critical behavior in various models, including network traffic, see [7]; here we cover the primary models used in the modeling of Internet traffic.

3.1 Network topology

The simplest model that can be used to describe a network is a one-dimensional chain, with packets passed back and forth along the chain. Some studies have implemented one-dimensional chains of two and three nodes to look at the effect of the TCP back-off mechanism [9, 10]. By simplifying the model to the bare minimum number of hosts, the results show primarily the effect of the TCP congestion avoidance mechanism on the phase dynamics without having to account for the overall network structure, which could give rise to more complex traffic dynamics. In these studies, the model showed a phase transition when an exponential back-off algorithm was used, with the data showing a 1/f-type fluctuation about the critical rate of input packets, as would be expected for a collective phenomenon, despite the simplicity of the model. These simple models show the importance of the TCP congestion avoidance mechanism in modeling network dynamics.

Expanding the network to more complex and realistic models gives data more easily generalized to real-world Internet traffic. One prevalent model is a square lattice of nodes and routers. While this model does not have the same self-symmetric branching that is seen in actual Internet topology, by setting a periodic boundary condition on the lattice, it can map to a system with toroidal topology. One such study, which expanded upon similar previous models, studied such a lattice network [8]. With additional complex network dynamics, such as implementing packet sources that have long-range dependence, while still being able to perform a mean field theoretical analysis of the model, which matches well with the Poisson data. The resulting dynamics is summarized in Figure 3. This model allows theoretical understanding through mean field analysis and implements the behavior noted in the previous paragraph, showing dynamics both above and below the critical point better than the one-dimensional chain model despite lacking the self-similar structure that we see in the Internet.

While these simple models allow for some tractable analysis, it has been shown that the Internet is most accurately depicted as a scale-free network[3], and so to understand the full dynamics of the Internet, many simulations are done with scale-free network topology. Because of the difficulty in describing these systems theoretically, due to both size and topology these simulations usually focus on investigating some other parameter through simulation, such as routing parameters or TCP behavior. One paper that uses simulations to analyze the effects of buffer size in a scale-free network topology is [11]. In this study, in studying the variation of routing parameters having to do with the queue and the queue handling, the data showed a critical packet creation rate, λ_c, which would vary with these queue parameters, see Figure 4.

It is important to note that while each choice of network topology shows the same general dynamics, the details of network dynamics is not the same for the various topologies. This was shown in a study of the effects of different network topologies on the order parameter dynamics as a function of packet creation [12]. The figure showing the order parameter for various structures is shown in Figure 5. A review of various network structures used in modeling networks and their resulting dynamics in the Internet as well as other complex systems can be seen in [13].

4 Analysis of Internet Traffic

With being shown in network traffic simulations, it is important to make comparisons with actual Internet traffic data. The foundational work in this area of study demonstrated that Internet traffic shows self-similar and long-range behavior [1] and several of these papers have comparisons to actual data, again see Figure 1. For example, data traces were performed along a T1 link between Japan and the United States [14], which showed the power-law spectrum of the congestion duration and the auto-correlation analysis confirmed the exis-

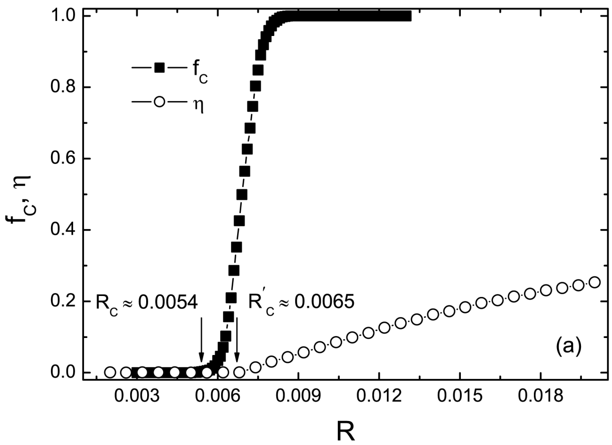

Figure 4: Probability of congestion, f_c, and the order parameter, ν, shown as a function of the packet creation rate, R. This shows a first-order phase transition at the critical value $R_c = 0.005$ [11].

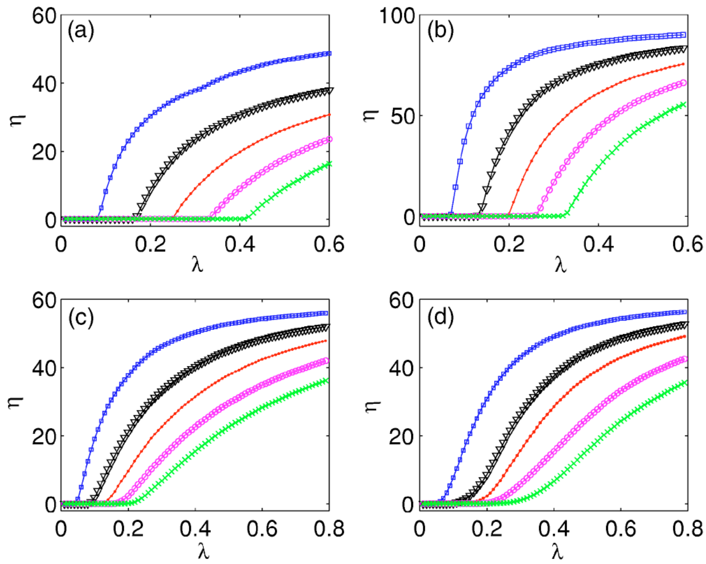

Figure 5: The variation of the order parameter, ν, as a function of the packet creation rate for (a) Cayley tree (b) homogeneous network (c) scale-free network and (d) random network models. The β parameter varies the routing capacity of the routing nodes [12].

tence of the phase transition between free flow and congested phases predicted by network simulations.

Despite increased attempts to verify the results of the simulations by comparisons to Internet traffic data, there are technical boundaries limiting the ability to collect the data that would be necessary for such an exercise. The main difficulty with this is the complication in obtaining traffic data over large sections of the network at a single time. Much of the data collected is data traces for traffic through nodes that receive large traffic volume, such along the Internet backbone or other major traffic points, but this can only give data for a single node. Another difficulty in collecting data is the periodic variation of the volume and type of Internet traffic, which may occur over days or weeks, as a function of when and how people use the Internet, rather than resulting from any dynamics of the network itself. Accounting for this poses a challenge in terms of collecting data and comparing simulations to actual Internet traffic.

5 Criticisms of Methods

Despite successes of the simulation and modeling of network traffic revealing some key elements of Internet traffic flow, there are several criticism as to the methods used in these simulations and the applicability of these models to actual networking problems. One of the key difficulties in modeling the Internet is the size and complexity of the Internet compared to the given models. This complexity is seen not just in the size of the network, but the heterogeneity of the network at every level, in addition to the constant fluctuations and growth of the network itself through various nodes connecting, disconnecting and moving to new locations in the network [15].

Another major criticism of the simulations is, for the most part, the models ignore the underlying protocols that govern the transfer of data on networks, that is, things like intelligent queuing and TCP congestion control are simulated poorly, if at all. Because the level of description when discussing the emergent congested phase is on the network level, all the effects from higher levels, up to the application level, but most importantly the transport level, where we see things such as TCP congestion avoidance protocols. These models ignore factors which could govern other emergent phenomena including emergent behavior. While the use of emergence in describing traffic flow is designed to ignore the underlying microscopic interactions, these mechanisms may give rise to an additional set of rich dynamics. Several studies have been done focusing on the TCP protocol which suggest there may be additional dynamics, specifically on very short time scales less (than the RTT) [16]. In addition, the Internet shows collective behavior in stateless protocols, such as UDP and as opposed to TCP, which the current methods of TCP networks cannot account for [17].

The last key criticism of these methods is the lack of a full description of Internet traffic dynamics, especially at low traffic levels far from criticality. While the models are in accordance with data traces at near critical traffic levels and can predict the long-range correlation, at traffic levels far below critical, the models break down. In the models, the long-range correlation of traffic only occurs near the critical value, however, in Internet traffic, even at low traffic levels, long-range correlations are observed in the traffic flow. To see this see the data traces presented in [14]

6 Conclusion

The key drive of this work is to leverage a better understanding of how large, complex networks interact in developing Internet systems through increased knowledge of their traffic dynamics. While the models and simulations studying the phase transition as developed in this paper has yet to make a significant impact on our development of the Internet, there are some studies that explore some aspect of routing or a real-life scenarios. For example, one paper examined the behavior of a network under a simulated distributed denial-of-service (DDoS) attack [18]. A DDoS is an attack carried out by leveraging many peers in an attempt to make a single host unavailable by flooding it with traffic.

In conclusion, this paper has examined some of the various aspects of Internet dynamics as it relates to the emergence of congestion as a collective phase, generated beyond a critical packet creation rate where the network becomes unable to deliver packages at the same rate they enter the system, leading to long-range dependence on models traffic flow. The description of these dynamics, while observable in actual Internet traffic, is best understood through the use of network simulations, which is done on networks of varying size and complexity to better grasp the underlying network dynamics. Despite it successes, however, there are serious concerns and criticisms from networking engineers which must be addressed to ensure the validity and applicability of these simulations to dealing with real-world Internet scenarios. Over the last 15 years, as this phenomena of Internet self-similarity has been studied, we have come to understand a great deal about networks and the dynamics of traffic on the Internet, and while there is still a lot to learn, these studies show great promise in helping us deal with an ever expanding digital network.

References

[1] Istvàn Csabi. 1/f noise in computer network traffic. *Journal of Physics A: Mathematical and General*, 27(12):L417, 1994.

[2] Michalis Faloutsos, Petros Faloutsos, and Christos Faloutsos. On power-law relationships of the internet topology. In *Proceedings of the conference on Applications, technologies, architectures, and protocols for computer communication*, SIGCOMM '99, pages 251–262, New York, NY, USA, 1999. ACM.

[3] Albert-Lászó Barabási and Réka Albert. Emergence of scaling in random networks. *Science*, 286:509–512, Oct 1999.

[4] W.E. Leland, M.S. Taqqu, W. Willinger, and D.V. Wilson. On the self-similar nature of ethernet traffic (extended version). *Networking, IEEE/ACM Transactions on*, 2(1):1–15, Feb 1994.

[5] Reginald D. Smith. The dynamics of internet traffic: Self-similarity, self-organization, and complex phenomena. 2008. arXiv:nlin/0807.3374v4.

[6] A. Arenas, A. Díaz-Guilera, and R. Guimerà. Communication in networks with hierarchical branching. *Phys. Rev. Lett.*, 86:3196–3199, Apr 2001.

[7] S. N. Dorogovtsev, A. V. Goltsev, and J. F. F. Mendes. Critical phenomena in complex networks. *Rev. Mod. Phys.*, 80:1275–1335, Oct 2008.

[8] M. Woolf, D. K. Arrowsmith, R. J. Mondragón-C, and J. M. Pitts. Optimization and phase transitions in a chaotic model of data traffic. *Phys. Rev. E*, 66:046106, Oct 2002.

[9] Simon Gàbor and Istvàn Csabai. The analogies of highway and computer network traffic. *Physica A: Statistical Mechanics and its Applications*, 307(3-4):516–526, 2002.

[10] Kensuke Fukuda, Hideki Takayasu, and Misako Takayasu. Origin of critical behavior in ethernet traffic. *Physica A: Statistical Mechanics and its Applications*, 287(1-2):289–301, 2000.

[11] Zhi-Xi Wu, Wen-Xu Wang, and Kai-Hau Yeung. Traffic dynamics in scale-free networks with limited buffers and decongestion strategy. *New Journal of Physics*, 10(2):023025, 2008.

[12] Liang Zhao, Ying-Cheng Lai, Kwangho Park, and Nong Ye. Onset of traffic congestion in complex networks. *Phys. Rev. E*, 71:026125, Feb 2005.

[13] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D.-U. Hwang. Complex networks: Structure and dynamics. *Physics Reports*, 424(4-5):175–308, 2006.

[14] Kensuke Fukuda, Misako Takayasu, and Hideki Takayasu. A cause of self-similarity in tcp traffic. *International Journal of Communication Systems*, 18(6):603–617, 2005.

[15] S. Floyd and V. Paxson. Difficulties in simulating the internet. *Networking, IEEE/ACM Transactions on*, 9(4):392 –403, aug 2001.

[16] L Guo and M Crovella. How does tcp generate pseudo-self-similarity? In *IEEE International Workshop on Modeling, Analysis, and Simulation of Computer and Telecommunication Systems*, pages 215–223, 2001.

[17] C.-Y. Lee. Higher-order correlations in data network traffic. *Journal of the Korean Physical Society*, 45(6):1664–1670, 2004.

[18] Jian Yuan and Kevin Mills. Macroscopic dynamics in large-scale data networks. In Ljupco Kocarev and Gábor Vattay, editors, *Complex Dynamics in Communication Networks*, volume 5 of *Understanding Complex Systems*, pages 191–211. Springer Berlin / Heidelberg, 2005. 10.1007/10973509_8.