

トポロジのモジュール性がもたらすTCPトラヒックのダイナミクス

平山 孝弘[†] 荒川 伸一[†] 新井 賢一^{††} 村田 正幸[†]

[†]大阪大学 大学院情報科学研究科 〒565-0871 大阪府吹田市山田丘 1-5

^{††}NTT コミュニケーション科学基礎研究所 〒619-0237 京都府相楽郡精華町光台 2-4

E-mail: [†]{t-hirayama,arakawa,murata}@ist.osaka-u.ac.jp, ^{††}ken@cslab.kecl.ntt.co.jp

あらまし インターネットトポロジを計測した結果、出線数分布がべき則に従うことが明らかにされている。出線数分布がべき則に従うトポロジでは、出力リンク数が k であるノードの出現確率が $k^{-\gamma}$ (γ は定数) に近似できる。出線数分布がべき則に従うトポロジを生成する手法は多数提案されているが、出線数分布が同じであっても、生成手法に基づき確率的に生成されたトポロジではISPのルータレベルトポロジの構造は再現できず、トポロジが持つ構造的特徴の違いによりネットワークの性能も大きく異なることが指摘されている。本稿では、べき則の性質を有するISPルータレベルトポロジが持つ構造と、エンドホスト間フロー制御の相互作用に起因するトラヒックダイナミクスを評価する。計算機シミュレーションの結果、ISPルータレベルトポロジが有するモジュール構造により、通信要求の増大に対し経路によるトラヒックの時間変動が抑制されることがわかった。

キーワード べき則, ISPルータレベルトポロジ, モジュール性, トラヒックダイナミクス, TCP

Effect of Modularity Structure on Traffic Dynamics

Takahiro HIRAYAMA[†], Shin'ichi ARAKAWA[†], Ken-ichi ARAI^{††}, and Masayuki MURATA[†]

[†] Graduate School of Information Science and Technology, Osaka University, Japan

^{††} NTT Communication Science Laboratories, Japan

E-mail: [†]{t-hirayama,arakawa,murata}@ist.osaka-u.ac.jp, ^{††}ken@cslab.kecl.ntt.co.jp

Abstract Measurement studies of the Internet topology revealed that the degree distributions of the Internet topologies follow a power law. That is, existing probability of nodes having k links is proportional to $k^{-\gamma}$. However, only the power-law degree distribution does not determine network-level performance of networking methods. That is, the structural characteristics of topologies other than the degree distribution are essential to discuss the behavior of networking methods. In this paper, we investigate traffic dynamics on ISP's router-level topology where the degree distribution exhibits power-law attribute and each of the nodes interacts via end-to-end flow control functionality. We show packet delay dynamics on the BA topology generated by the BA model and the ISP's router-level topology. Simulation results show that the number of links that are highly fluctuated is more than twice comparing to the results of the stop-and-wait model. Even in this case, the modularity structure of the ISP topology reduces the number of highly fluctuated link comparing to the results of the BA topology.

Key words Power-law Networks, ISP's Router-level Topology, Modularity, Traffic Dynamics, TCP

1. Introduction

Dynamic interactions among various network-related protocols as a result of functional partitioning make the Internet a complicated system whose details are difficult to confirm because of its large-scale, heterogeneous structure. One of the complex behaviors of the Internet is traffic dynamics. For example, flow control and congestion control performed by end hosts can show short-range and long-range dependence on traffic [1]. The Internet is faced with

ever-changing networking technologies and applications; thus, understanding and controlling the complex behavior of the Internet are important for designing future networks.

Although the statistical properties of network traffic are hard to capture, studies have revealed that the degree distribution of the Internet topology follows a power law. That is, the probability of existence of nodes having k links is proportional to $k^{-\gamma}$ (γ is constant). Barabási et al. proposed the Barabási-Albert (BA) model to generate power-law topologies having power-law degree distri-

butions [2]. Li et al. showed several topologies that have different structures but have the same degree distribution [3]. They pointed out that differences in structures lead to differences in the amount of traffic that the network accommodates. Moreover, the structures of power-law topologies also affect the performance of some networking mechanisms, such as routing mechanisms [4]. These studies indicate that the power-law degree distribution alone does not determine network-level performance. That is, topological structure properties other than the degree distribution are essential to discuss the behavior of some networking studies.

In previous studies, the relationship between the statistical properties of Internet traffic and end-to-end flow control has been discussed. In Refs. [5, 6], it is revealed that Internet traffic exhibits long-range dependence (LRD), where traffic fluctuation appears independently of the measurement time scale. Various studies have investigated the reasons for such LRD on the Internet. One of the reasons is flow control in the transport layer, such as TCP [7–9]. However, these studies deal with small, simple topologies. We therefore investigate how the structures of topologies impact the traffic dynamics. More specifically, we investigate traffic dynamics on an ISP router-level topology where the degree distribution exhibits a power-law nature, and the nodes interact via end-to-end flow control functionality. Previously, we showed the packet delay dynamics on the BA topology generated by the BA model and on the ISP router-level topology [10]. We then discussed how the interaction between the structures of topologies and the flow controls affect the end-to-end packet delay distribution and the appearance of LRD in the queue length for each link. The results show that the increased traffic due to TCP makes the queue length of links in the network fluctuate. However, the queue length does not fluctuate for traffic due to TCP in the ISP topology. In this paper, we show that the above results are due to the high-modularity structure of the ISP topology. We investigate the relationship between the modularity of topologies and queue fluctuation. We found that topologies with high-modularity structures have a lower number of highly fluctuating links.

This paper is organized as follows. We introduce related work in Section 2. In Section 3., we show the network model that we used for the simulations. In Section 4., we evaluate the influence of the power-law topologies and TCP flow control. Finally, in Section 5., we conclude this paper and mention future work.

2. Related Work

2.1 Structural Properties of Power-law Networks

Recently, there has been a considerable number of studies investigating power-law networks whose degree distribution follows a power law. Barabási et al. introduced the BA model as a method for generating a power-law topology in Ref. [2]. The BA model generates a power-law topology based on two rules: one is incremental growth, and the other is preferential attachment. The re-

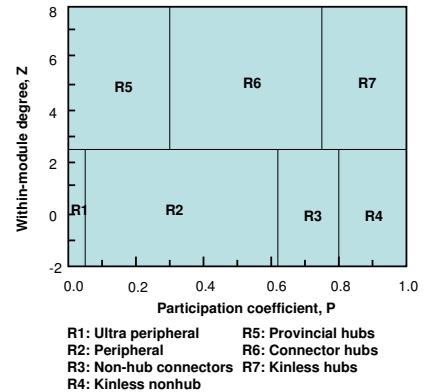


Fig. 1 Classification of node function by participation coefficient and within module degree

sulting power-law networks have two main characteristics. First, many nodes have a few links, and a few nodes, so-called hub nodes, have many links. Second, the average length of the shortest paths between nodes is small. Many studies have investigated topological properties appearing in the BA model or its variants. However, when router-level topologies are concerned, the BA model, where links are attached based on a preferential probability, does not emulate the structure of ISPs’ router-level topologies. We have compared the structural differences of the AT&T topology measured using the Rocketfuel tool [11] and the topology generated by the BA model [12]. The results indicate that the design principles of networks greatly affect the structure of the ISP topology: Design principles determine the node functionality, which in turn determines the connectivity of nodes.

In [13], Guimerà et al. proposed a classification method of node functions. In this method, a network is divided into multiple modules, and the within-module degree, Z_i , and the participation coefficient, P_i , are defined for each node. Assuming that the node i belongs to a module s_i , the within-module degree Z_i of node i is defined as

$$Z_i = \frac{k_i - \langle k_{s_i} \rangle}{\sigma_{s_i}}, \quad (1)$$

where k_i is the degree of the node, $\langle k_{s_i} \rangle$ represents the average degree in module s_i , and σ_{s_i} is the variance of the degree distribution of nodes in module s_i . The participation coefficient P_i of node i is also defined as

$$P_i = 1 - \sum_{s_i=1}^{N_m} \left(\frac{k_{is}}{k_i} \right)^2, \quad (2)$$

where k_{is} represents the fraction of links connecting with nodes in module s_i and N_m represents the number of modules. That is, when all the links of node i connect with nodes belonging to the same module s_i , P_i becomes 0. Figure 1 shows that the roles of nodes are categorized by the values of Z_i and P_i .

Figure 2(a) and Fig. 2(b) show the results of applying Guimerà’s method to the BA topology and the AT&T topology. The module is

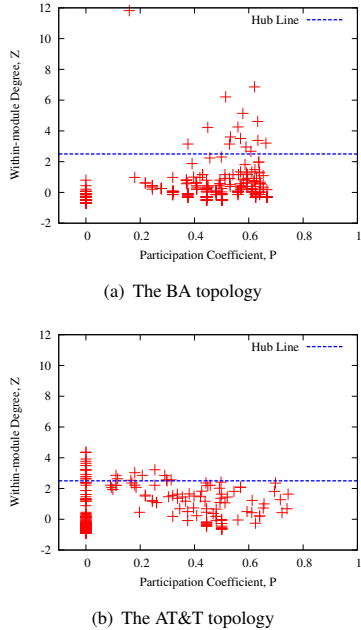


Fig. 2 Classification of node function in each topology

calculated using the method in [14]. In Fig. 2, the horizontal axis indicates within-module degree Z , and the vertical axis indicates the participation coefficient P . Depending on the values of P and Z , the role of the node is categorized into several classes. For example, when Z_i is large and P_i is relatively large, the node i has many links connecting to other modules. Thus, the node i is categorized as a *connector hub*. On the other hand, a *provincial hub* node has larger Z_i but smaller P_i ; that is, it has many links connecting with nodes in the same module.

Looking at Fig. 2(a), we observe that the BA topology has many connector hub nodes that connect between modules. However, Fig. 2(b) shows that there are no connector hubs in the AT&T topology. This means that the AT&T topology has few inter-module links that connect different modules.

2.2 Traffic Dynamics in Power-law Networks

Some studies have investigated traffic-level behavior in topologies having power-law degree distributions [15, 16]. Reference [15] demonstrates that congestion spreads easily over BA topologies because of their low diameter. Low-diameter effects also appear in the queuing delay distribution of these topologies. They show that the queuing delay distribution of BA topologies has a long tail. The effect of end-to-end flow control in the topology obtained by the BA model has also been investigated [16]. The authors examined TCP control with LRD input traffic and Poisson input traffic and revealed that the average end-to-end packet delay sharply increases for both types of input traffic since packets concentrate more at hub nodes in the BA topology.

Previous studies have used topologies generated by the BA model or its variants. However, even when the degree distributions of some topologies are the same, their more detailed characteristics are often quite different. As discussed in Section 2.1, the BA model does

not adequately describe the structure of ISP router-level topologies. This clearly indicates that the power-law degree distribution alone does not determine traffic-level behavior in router-level topologies.

Traffic dynamics has received great interest from the networking research community. It has been revealed that Internet wide-area and local-area traffic can show LRD or self-similarity [5]. That is, network traffic exhibits a large variability even at a wide range of time scales. Recently, these statistical properties have also been observed in peer-to-peer traffic [6]. In Refs. [7, 8], TCP flow control is considered to be the cause of LRD in Internet traffic. Simulation results given in Ref. [9] show that LRD is caused by feedback flow control functionality only, because when the stop-and-wait protocol is used instead of TCP, LRD is still observed in traffic. However, these studies deal with small, simple topologies. Thus, we investigated how the structures of topologies impact the traffic dynamics.

3. Simulation Model

3.1 Network Topologies

As the ISP router-level topology, we use the AT&T topology measured using the Rocketfuel tool [11]. The Rocketfuel is a topology measurement tool based on `traceroute`. The authors pointed out that the Rocketfuel might not cover some parts of ISP topologies. For example, ISPs make some of routers invisible and use them as a backup purpose. Even when the Rocketfuel misses some of routers, it is sufficient to investigate traffic dynamics because the Rocketfuel provides active paths between routers. For comparison, we use the BA topology generated by the BA model. The BA topology is generated such that the numbers of nodes and links are the same as those of the AT&T topology.

3.2 Packet Processing Model at Node

Each node has limited buffers at each outgoing link. When a packet arrives at a given node and when the node is the packet's destination, the node removes the packet from the network. Otherwise, the node selects the next node based on a minimum hop routing algorithm and forwards the packet to a buffer of an outgoing link connecting to the next node. Each outgoing link sends packets to the next node based on FIFO and a drop-tail queuing discipline, delivering C packets per unit time. Note that capacity of links is heterogeneous in real ISP topologies. However, we use the same link capacity to evaluating traffic dynamics induced by topological structure more clearly. Here, we do not use dynamic routing; i.e., each packet traverses the shortest path calculated beforehand. When multiple shortest paths to reach the destination are found, the next node is determined based on a packet's source node. According to Ref. [17], traffic fluctuation caused by TCP-like flow control appears in short time scales such as the round trip time (RTT). So, we use a static routing strategy.

3.3 Flow Control Between End Hosts

We examined stop-and-wait model and TCP model for flow control between end hosts. In the simulation, pre-specified numbers of

Table 1 The simulation parameters used in the TCP model

Buffer size	1,000 packets
Session arrival rate (λ)	1 session / unit of time
Maximum <i>cwnd</i>	10 packets
Link capacity (C)	3 packets / unit of time
Simulation time	300,000 units of time

sessions are created between nodes. Source and destination nodes are randomly selected and each session arrives according to the Poisson process with mean rate λ . Each session always has data to send during the simulations.

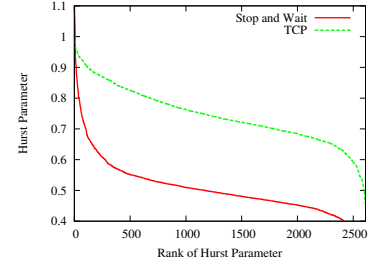
3.3.1 Stop-and-Wait Model

In this model, when a source node sends a DATA packet to its destination node, the source node stops sending a new packet until the source node receives the acknowledgement (ACK) packet from the destination node. If a source node does not receive the ACK packet within the retransmission time out (RTO) period, the source node thinks that packet loss has occurred and resends the packet. The time-out period is defined based on the RTT and is doubled for every time out.

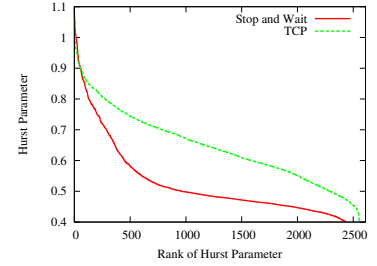
3.3.2 TCP Model

In this model, source nodes control the amount of DATA packets based on the slow start and congestion avoidance algorithms. The slow start and congestion avoidance algorithms are basic flow-control functions of TCP. If the window size is lower than the slow start threshold (*ssthresh*), the source node uses the slow start algorithm. When the source node receives an ACK packet, it extends the congestion window (*cwnd*) by one packet size (= segment size, *smss*) and sends two new DATA packets to the destination node. If the window size exceeds *ssthresh*, the source node uses the congestion avoidance algorithm. When the source node receives an ACK packet, it extends the congestion window by $1/cwnd$, and it sends an adequate number of DATA packets to the destination node. In our model, the congestion window size does not exceed a pre-decided maximum window size. If the source node does not receive any ACK packet within the RTO period, the source node recognizes that serious congestion has occurred. The source node resends the lost DATA packet and reduces the congestion window by one packet size. The time-out period is defined in the same way as in the stop-and-wait model.

In addition, we use the fast retransmit and fast recovery algorithms defined by RFC 2581 [18]. The source node uses the fast retransmit algorithm when it detects packet loss and light congestion by the arrival of three duplicate ACKs. When the source node receives the third duplicate ACK, it reduces the congestion window to half and resends the lost DATA packet. After the retransmission, the source node extends the congestion window based on the fast recovery algorithm. The source node keeps extending the congestion window by one packet size as long as it receives the same duplicate ACKs.



(a) The BA topology



(b) The AT&T topology

Fig. 3 Correlations between Hurst parameter and Hurst parameter rank

4. Dynamics of TCP in Power-law Networks

In this section, we show the results of simulation for TCP and discuss queue-length fluctuation in detail. In the simulation, each link can transfer three packets per unit time. The other parameters are summarized in Table 1.

4.1 Queue Length Fluctuation

We evaluated the fluctuation of queue length. If the queue length of a link fluctuates drastically, packets that traverse the link experience non-constant queuing delay, which leads to performance degradation. We evaluated the fluctuation using the Hurst parameter (H , $0.5 < H < 1$) by applying the R/S plot method [19]. The Hurst parameter represents the degree of LRD.

Details of the R/S plot method are as follows. First, we define $R/S(n)$ as

$$R/S(n) = 1/S_n [\max(0, W_1, W_2, \dots, W_n) - \min(0, W_1, W_2, \dots, W_n)], \quad (3)$$

where

$$W_k = (X_1 + X_2 + \dots + X_k) - kX(n),$$

where ($X_k : k = 1, 2, \dots, n$) is an observation data set. $X(n)$ represents the mean value, and S_n represents the standard deviation of the data set X . By estimating $R/S(n)$ for an observation scale n and plotting the correlation between n and $R/S(n)$, we can observe the statistical dependence over time. The slope of the fitted curve of the correlation function, αn^H , represents the Hurst parameter of a set X .

Figure 3 shows Hurst parameters for each link. The y-axis represents the Hurst parameter and the x-axis represents its rank in a descending order. In this figure, the results for the stop-and-wait

Table 2 The ratio of highly fluctuating links in each topology

Topology	10,000 sessions		100,000 sessions	
	Total	Inter-module	Total	Inter-module
BA topology	0.257	0.21	0.26	0.21
AT&T topology	0.08	0.025	0.11	0.017

model are also added. Looking at this figure, we observe that the number of links that take high Hurst parameters increases in the TCP model.

To see the relation between the Hurst parameter and topological structure in the AT&T topology clearly, we show the ratio of links that take high H values ($H \geq 0.8$) in Table 2. When the number of sessions is small, the queue length of the links that connect two regions fluctuates drastically. That is, inter-module links tend to have highly fluctuating queue lengths. This is because many packets concentrate at inter-module links. As the number of sessions gets higher, the queue length of the links that connect inside a region fluctuates, whereas the Hurst parameter of inter-module links decreases. That is, the fluctuation spreads to tributary links of the bottleneck.

4.2 Fluctuation Reduction Effect of High-Modularity Structure

In the previous section, we showed that the modularity structure of the AT&T topology reduces the number of fluctuating links. In this section, we examine the relationship between the modularity of topologies and the fluctuation reduction effects of those topologies.

Newman et al. [20] defined a modularity value ($Q, 0 < Q < 1$) as

$$Q = \sum_i (e_{ii} - a_i^2), \quad (4)$$

where e_{ii} is the fraction of links in the network that connect vertices in the same module, and a_i represents the fraction of links that connect between different modules. According to this definition, the modularity value of the BA topology is nearly 0.63, and that of the AT&T topology is about 0.89. This result indicates that a high-modularity structure reduces the number of highly fluctuating links.

To confirm this assumption, we generated three topologies having different modularity values. These topologies have the same number of nodes and links. We generated these topologies via the following steps.

- (1) Uniformly divide the nodes into 10 modules and connect the nodes in the same module based on the BA model.
- (2) Add inter-module links between randomly selected modules.

We can set the modularity value of the generated topology based on the number of inter-module links in Step 2. The modularity values of the three topologies were 0.96, 0.89, and 0.84. We confirmed that these topologies had power-law degree distributions. Figure 4 shows the Hurst parameters of each link from the results of the sim-

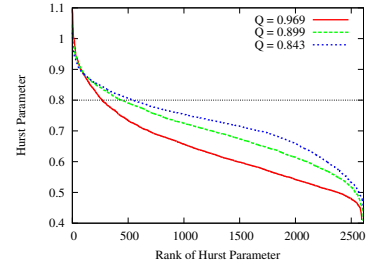


Fig. 4 Correlations between Hurst parameter and Hurst parameter rank of generated topologies

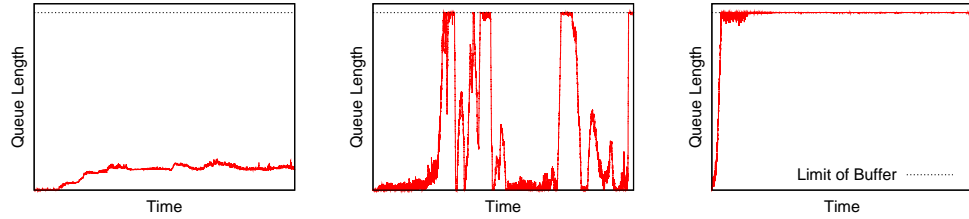
ulations with the TCP model. The x-axis and y-axis represent the same parameters as those in Fig. 3. In the simulations, we used the parameters summarized in Table 1. In Fig. 4, as the modularity value increases, the number of links that take high Hurst parameters decreases. This result is confirmed by the ratio of highly fluctuating links that have a Hurst parameter larger than 0.8 shown in Table 3. Thus, topological structures that have high modularity values prevent the appearance of highly fluctuating links.

A question is why the high-modularity structure reduces the number of highly fluctuating links. To answer this, we show queue fluctuation on links having different link load in Fig. 5. Here, the link load is the number of end-to-end sessions that pass through the link. When a link load is low, the queue length does not fluctuate (Fig. 5(a)). As a link load increases, the queue fluctuation becomes large (Fig. 5(b)). However, if the link load exceeds a certain level, the queue length is mostly constant due to a limit of buffer size in queue (Fig. 5(c)). In topologies that have high-modularity structure, packets tend to concentrate at inter-module links. When the number of sessions is large, queue length of some of intra-module links fluctuates like Fig. 5(b) and queue length of inter-module links keeps nearly-constant value like Fig. 5(c). In summary, the number of links having highly fluctuating queue does not increase even when the large number of sessions concentrates on these links. So, highly-modularity structure of topologies suppresses queue fluctuation.

At the same time, the number of arriving packets forwarded between different modules decreases in a topology that has a high modularity value. Table 3 shows the number of arriving packets. Here, *Intra-module* means the number of arriving packets between two nodes in the same module, and *Inter-module* means the number of arriving packets between two nodes belonging to different modules. According to this table, as the modularity value gets higher, the number of arriving inter-module packets decreases, whereas the number of arriving intra-module packets increases. In topologies having a high modularity value, a few inter-module links tend to become congested, and it becomes more difficult for inter-module packets to arrive at their destination.

5. Conclusion

In this paper, we investigated the interaction between the struc-



(a) When a link load is low, queue length does not fluctuate. (b) As a link load increases, the queue fluctuation becomes large. (c) The queue length is mostly constant due to a buffer limit

Fig. 5 Relationships between link load and queue fluctuation

Table 3 Simulation results of generated topologies

Modularity (Q)	Ratio of highly fluctuating links	Number of arrived packets		
		Total	Intra-module	Inter-module
0.96	0.10	4.7×10^8	4.5×10^8	1.7×10^7
0.89	0.17	3.6×10^8	2.9×10^8	7.4×10^7
0.84	0.21	3.2×10^8	2.2×10^8	1.0×10^8

tures of topologies and flow control between end hosts. Comparing the simulation results of the stop-and-wait model and the TCP model, the functionality of TCP makes the queue length fluctuate. Even in this case, the high-modularity structure of the AT&T topology reduces the number of highly fluctuating links compared with the BA topology.

In future work, we will evaluate packet delay dynamics on topologies that have heterogeneous link capacity, and we will explore combined use of end-host flow control and traffic engineering in more detail.

Acknowledgment

This work was partly supported by Grant-in-Aid for Scientific Research (A) 21240004 of the Japan Society for the Promotion of Science (JSPS) in Japan and the National Institute of Information and Communications Technology (NICT) of Japan.

References

- [1] A. Veres and M. Boda, "The chaotic nature of TCP congestion control," in *Proceedings of IEEE Conference on Computer Communications*, vol. 3, pp. 1715–1723, Mar. 2000.
- [2] A. -L. Barabási and R. Albert, "Emergence of scaling in random networks," *Science*, vol. 286, pp. 509–512, Oct. 1999.
- [3] L. Li, D. Alderson, W. Willinger, and J. Doyle, "A first-principles approach to understanding the Internet's router-level topology," *ACM SIGCOMM Computer Communication Review*, vol. 34, pp. 3–14, Oct. 2004.
- [4] R. Fukumoto, S. Arakawa, and M. Murata, "On routing controls in ISP topologies: A structural perspective," in *Proceedings of ChinaCom*, Oct. 2006.
- [5] V. Paxson and S. Floyd, "Wide-area traffic: The failure of poisson modeling," *IEEE/ACM Transactions on Networking*, vol. 3, pp. 226–244, July 1995.
- [6] N. Basher, A. Mahanti, A. Mahanti, C. Williamson, and M. Arlitt, "A comparative analysis of web and peer-to-peer traffic," in *Proceedings of the international conference on World Wide Web*, pp. 287–296, Apr. 2008.
- [7] K. Park, G. Kim, and M. Crovella, "On the relationship between file sizes, transport protocols, and self-similar network traffic," in *Proceedings of the International Conference on Network Protocols (ICNP)*, pp. 171–180, Oct. 1996.
- [8] A. Feldmann, A. C. Gilbert, W. Willinger, and T. G. Kurtz, "The changing nature of network traffic: Scaling phenomena," *ACM SIGCOMM Computer Communication Review*, vol. 28, pp. 5–29, Apr. 1998.
- [9] K. Fukuda, M. Takayasu, and H. Takayasu, "A cause of self-similarity in TCP traffic," *International Journal of Communication Systems*, vol. 18, pp. 603–617, Aug. 2005.
- [10] T. Hirayama, S. Arakawa, K. Arai, and M. Murata, "Dynamics of Feedback-induced Packet Delay in Power-law Networks," in *The Fifth International Conference on Systems and Networks Communications (ICSNC2010)*, Aug. 2010.
- [11] N. Spring, R. Mahajan, D. Wetherall, and T. Anderson, "Measuring ISP topologies with rocketfuel," *IEEE/ACM Transactions on Networking*, vol. 12, pp. 2–16, 2004.
- [12] R. Fukumoto, S. Arakawa, T. Takine, and M. Murata, "Analyzing and modeling router-level Internet topology," in *Proceedings of the International Conference on Information Networking*, pp. 171–182, Jan. 2007.
- [13] R. Guimerà and L. A. N. Amaral, "Functional cartography of complex metabolic networks," *Nature*, vol. 433, pp. 895–900, Feb. 2005.
- [14] M. Newman, "Modularity and community structure in networks," in *Proceedings of the National Academy of Sciences of the United States of America*, vol. 103, pp. 8577–8582, Apr. 2006.
- [15] B. Tadić, S. Thurner, and G. Rodgers, "Traffic on complex networks: Towards understanding global statistical properties from microscopic density fluctuations," *Physical Review E*, vol. 69, p. 36102, Mar. 2004.
- [16] M. Woolf, D. Arrowsmith, R. Mondragon, J. Pitts, and S. Zhou, "Dynamical modelling of TCP packet traffic on scale-free networks," *Institut Mittag-Leffler*, vol. 6, p. 7, Oct. 2004.
- [17] A. Feldmann, A. C. Gilbert, P. Huang, and W. Willinger, "Dynamics of IP traffic: a study of the role of variability and the impact of control," in *SIGCOMM '99: Proceedings of the conference on Applications, technologies, architectures, and protocols for computer communication*, (New York, NY, USA), pp. 301–313, ACM, 1999.
- [18] M. Allman, V. Paxson, and W. Stevens, "TCP congestion control," *RFC 2581*, Apr. 1999.
- [19] W. E. Leland, M. S. Taqqu, W. Willinger, and D. V. Wilson, "On the self-similar nature of ethernet traffic (extended version)," *IEEE/ACM Transactions on networking*, vol. 2, pp. 1–15, Feb. 1994.
- [20] M. E. J. Newman and M. Girvan, "Finding and evaluating community structure in networks," *Physical Review E*, vol. 69, Feb. 2004.