社団法人　電子情報通信学会
THE INSTITUTE OF ELECTRONICS,
INFORMATION AND COMMUNICATION ENGINEERS

信学技報
TECHNICAL REPORT OF IEICE.

[奨励講演] べき則の性質を有するトポロジにおけるフィードバック型フロー制御に起因するパケット転送遅延の評価

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により，TCP のフロー制御によってエンドホスト間パケット転送遅延時間の分布が増大し，さらに経由するトラヒックが大きく変動するリンクの数が増大する一方，ISP ルータレベルトポロジが，通信要求の増大に対し経由するトラヒックの時間変動を抑制する性質を有していることを示す．

キーワード　べき則，ISP ルータレベルトポロジ，トラヒックダイナミクス，ストップアンドウェイト，TCP

Dynamics of Feedback-induced Packet Transfer Delay in Power-law Networks

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

[†] Graduate School of Information Science and Technology, Osaka University
1–5 Yamadaoka, Suita, Osaka 565–0871, Japan
[††] NTT Communication Science Laboratories
2-4 Hikaridai Seika-cho Soraku-gun, Kyoto, 619–0237, 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 end-to-end packet delay distributions exhibit a heavy tail by the TCP model. Moreover, 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, Traffic Dynamics, Stop and Wait, TCP

1. Introduction

Dynamic interactions of various network-related protocols induced by the functional partition make the Internet to be complicated systems; details are hardly captured because of large-scale and heterogeneous structure. One of complex behaviors of the Internet is traffic dynamics. For example, flow control and congestion control performed by end host can show short-range and long-range dependence of traffic [1]. The Internet is facing with ever-changing networking technologies and applications, thus understanding and controlling the complex behaviors of the Internet are important for designing future networks.

Although the statistical property of network traffic is hard to capture, 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 $k^{-\gamma}$. Barabási et al. propose the BA model to generate power-law topologies having power-law degree distribution [2]. However, structures of topologies are not determined only by the degree distribution, but also by other factors. Li et al. shows several topologies that have different structures but have the same degree distribution [3]. They pointed out that differences in structures leads to differences in the amount of traffic that the network accommodates. Moreover, structures of power-law topologies also affect the performance of some networking mechanisms such as the routing mechanisms [4]. These works indicate that only the power-law degree distribution does not determine network-level performance. That is, the structures of topologies other than the degree distribution are essential to discuss the behavior of some 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. In previous studies, it has discussed about the relationship between statistical properties of the Internet traffic and end-to-end flow control. In Refs. [5, 6], it is revealed that the Internet traffic exhibits long-range dependence (LRD) where the traffic fluctuation appears independent from time scale of measuring. Various studies have investigated the reason to cause the LRD traffic in the Internet. One of reasons is flow control on the transport layer, such as TCP [7–9]. However, these researches deal with simple and small topologies. We therefore investigate how structure of topologies impacts on traffic dynamics. More specifically, 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 first show packet delay dynamics and compare it with the BA topology generated by the BA model and the ISP's router-level topology. We then discuss how structure of topologies and flow controls affect the end-to-end packet delay distribution and the appearance of long-range dependence in queue length for each link. The results show that increased traffic caused by TCP makes queue length of links in the network fluctuate. However, we show that the queue length does not fluctuate against traffic caused by TCP in the ISP topology.

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

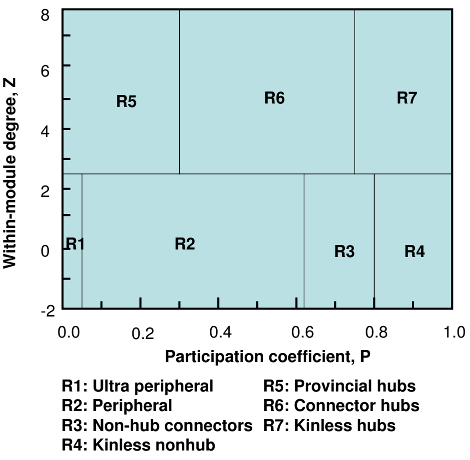


Fig. 1 Classification of node function by participation coefficient and within-module degree

2. Related Work

2.1 Structural Properties of Power-law Networks

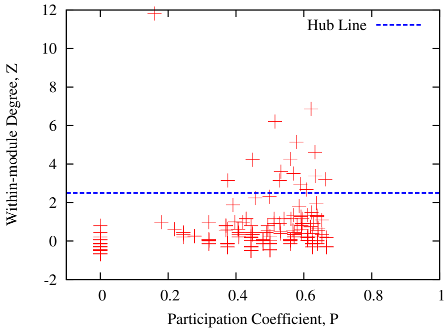
Recently, considerable numbers of studies have investigated power-law networks whose degree distribution follows a power law. Barabàsi et al. introduced the BA model as a power-law topology generating method in Ref. [2]. The BA model generates power-law topology based on two rules, one is incremental growth and the other is preferential attachment. The resulting 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 researches investigate topological properties appeared by the BA model or its variant models. 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 the ISP's router-level topologies. We have compared the structural differences of the AT&T topology measured by Rocketfuel tool [10] and the topology generated by BA model in Ref. [11]. The results indicate that design principles of networks greatly affect the structure of the ISP topologies. Design principles determine a node functionality, which in turn determines the connectivity of nodes.

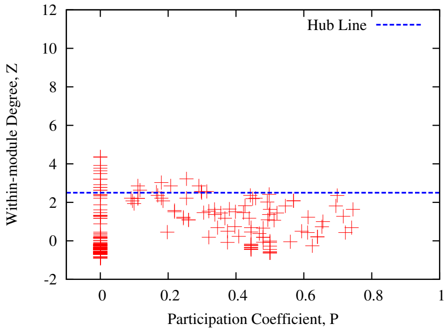
In [12], Guimeara et al. have proposed the classification method of node functions. The method divides a network to multiple modules and defines the within-module degree Z_i, and the participation-coefficient, P_i, for each node i. Assuming that the node i belongs to a module s_i, the within-module degree Z_i of node i is defined as,

$$Z_i = \frac{k_i - < k_{s_i} >}{\sigma_{s_i}}, \tag{1}$$

where k_i is the degree of nodes, $< k_{s_i} >$ represents the average degree in module s_i, and σ_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,



(a) The BA topology



(b) The AT&T topology

Fig. 2 Classification of node function in each topology

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

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

Figure 2(a) and Figure 2(b) show the result of the Guimeara's method to the BA topology and the AT&T topology. The module is calculated from the method in [13]. In Figure 2, the horizontal axis indicates within-module degree Z and the vertical axis the participation coefficient P. Depending on the values of P and Z, the role of 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 into the "Connector hub(s)". "Provincial hub(s)" also takes the larger Z_i but smaller P_i; the node i has many links connecting with nodes in the same module.

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

2. 2 Traffic Dynamics in Power-law Networks

Some papers investigate traffic-level behaviors in topologies having power-law degree distributions [14, 15]. Reference [14] demonstrates that the congestion spreads easily over the topologies in the BA topologies because of the low-diameter of topologies. The low-diameter effects also appear in the queuing delay distribution of topologies. They show that the queuing delay distribution of the BA topology has a long-tail. The effect of end-to-end flow control is also investigated on the topology obtained by the BA model [15]. The authors examined TCP control with long range dependence (LRD) input traffic and Poisson input traffic, and revealed that average of end-to-end packet delay sharply increases for both input traffic since packets more concentrate on the hub nodes in the BA topology.

The previous works have used topologies generated by the BA model or its variant models. However, even if the degree distributions of some topologies are the same, more detailed characteristics are often quite different. As discussed in Section 2. 1, the BA model does not adequately model the structure of the ISP's router-level topologies. This fact clearly indicates that the power-law degree distribution alone does not determine traffic-level behaviors in router-level topologies.

The traffic dynamics has received great interest from the networking research community. In the Internet, it has revealed that the Internet wide-area and local-area traffic can have the long-range dependence or self-similarity [5]. That is, network traffic exhibits a large variability even at a wide range of time scales. Recently, these statistical properties has again observed in Peer-to-Peer traffic [6]. In Ref. [7, 8], TCP flow control is considered to be cause of long-range dependence in the Internet traffic. Simulation results of Ref. [9] show that long-range dependence is caused by feedback flow control functionality only, because when stop-and-wait protocol is used instead of TCP, long-range dependence is still observed in traffic. However, these studies deal with simple small topologies, we therefore investigate how structure of topologies impacts on traffic dynamics.

3. Simulation Model

3. 1 Network Topologies

We use the AT&T topology measured by Rocketfuel tool [10] as ISP's router-level topology. For comparison purpose, we use the BA topology generated by the BA model. The BA topology is generated such that the numbers of nodes and links of it are the same as that of the AT&T topology.

3. 2 Packet Processing Model at Node

Each node has limited buffers at each out-going links. 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 out-going link connecting to the next node. Each out-going link sends packets to the next node based on FIFO and a drop-tail queuing discipline, and delivers n packets per unit of time. Here, we do not use the 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.

3.3 Flow Control between End Hosts

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 longer than retransmission time out (RTO), the source node thinks the packet loss has occurred and resends a packet. Time-out period is defined based on round trip time (RTT) and doubled 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 slow start threshold, 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 window size exceeds the slow start threshold ($ssthresh$), 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 the adequate number of DATA packets to the destination node. In our model, congestion window size does not exceed the pre-decided maximum window size. If the source node does not receive any ACK packet longer than RTO, the source node detects serious congestion and recognizes serious congestion has occurred. The source node resends the lost DATA packet, and reduces the congestion window to one packet size. Time-out period is defined in the same way as the stop-and-wait model.

In addition, we use the fast retransmit and fast recovery algorithms defined by RFC 2581 [16]. The source node uses the fast retransmit algorithm when it detects packet loss and light congestion by 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 extends the congestion window by one packet size as long as it receives the same duplicate ACKs.

4. Dynamics of TCP in Power-law Networks

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

4.1 End-to-End Packet Delay Distribution

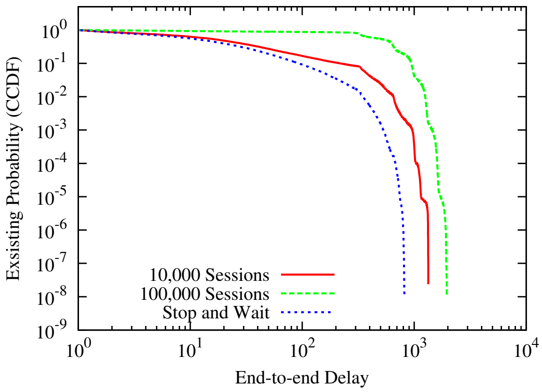
Figure 3 shows the end-to-end packet delay distribution of the BA topology and the AT&T topology. In this figure, the end-to-end packet delay distribution by the stop-and-wait model (the number of sessions is 100,000) is also plotted. In the AT&T topology, when the number of sessions is 10,000, packets spend longer time to reach their destination than the time in the BA topology. In contrast, when the number of sessions is 100,000, end-to-end packet delay distribution of the BA topology changes drastically. It has a long-tail distribution, i.e., many packets take a long time to go to their destination. End-to-end delay under 500 time units is hardly observed in the BA topology. However, end-to-end delay distribution of the AT&T topology does not change widely when the number of sessions increases.

Table 1 The simulation parameters used in the TCP model

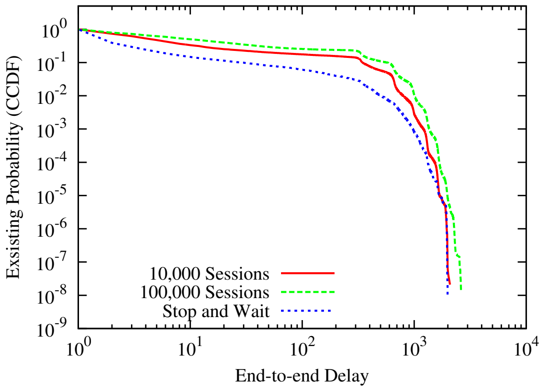
Buffer size	1,000 packets
Session arrival rate	1 session / unit of time
Maximum window size	10 packets
Link capacity	3 packets / unit of time
Simulation time	300,000 units of time

The reason comes from that congestion tends to occur in the BA topology when TCP is applied. The term "congestion" means that a buffer is occupied by packets and cannot receive more packets. We confirm the frequency of congestion in the BA topology is about 4 times larger than the frequency in the AT&T topology (but not presented due to space restriction). Since packets concentrate in a variety of places and queuing delay gets longer, the end-to-end packet delay distribution of the BA topology exhibits a long tail.

To clarify the reason why congestion tends to occur in the BA topology, we focus on the difference in the structures between the BA topology and the AT&T topology. As discussed in Section 2.1, the BA topology has many "Connector hubs" and the AT&T topology has many "Provincial hub" nodes. In the BA topology, "Connector hub" nodes transfer large amount of packets between modules. However, in the AT&T topology, packets traveling between two modules, i.e., between two regions are first aggregated at "Provincial hub" nodes and then forwarded via inter-module links. To confirm this assumption, we investigate composition percentages of congested links in each topology. We classify the link function based on function of nodes attached to the link. Table 2 shows the function of congested links of the BA topology and the AT&T topology. Each value represents the percentages of congested links observed in the simulation period between 290,000 and 300,000 units of time. We classify the link function based on function of two nodes attached to the link. In the table, "Connector" means the ratio of congested links which connect to "Connector hub" nodes, "Provincial" means the ratio of congested links which connect to "Provincial hub" nodes, and "Other" means the ratio of the rest. We observe that the links connecting to "Connector hubs" tend to be congested in the BA topology. Packets concentrate at "Connector hub" nodes. Thus, the links forwarding packets to "Connector hub" nodes tend to be congested. Since congestion tends to occur near

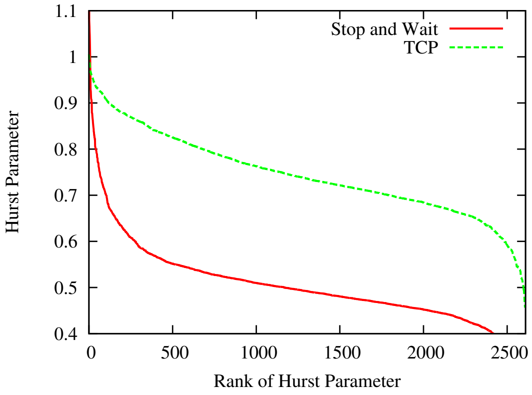


(a) The BA topology

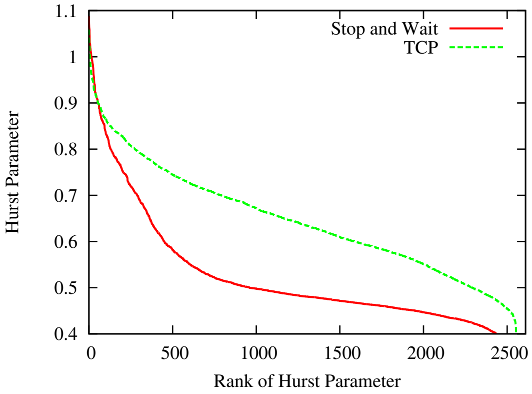


(b) The AT&T topology

Fig. 3 End-to-end packet delay distribution



(a) The BA topology



(b) The AT&T topology

Fig. 4 Correlations between Hurst parameter and Hurst parameter rank

Table 2 Component of congested links

Topology	Connector	Provincial	Other
The BA topology	0.4834	0.1030	0.4199
The AT&T topology	0.0123	0.3158	0.6717

“Connector hub” nodes, the end-to-end delay distribution of the BA topology has a long tail.

4. 2 Queue Length Fluctuation

We next evaluate the fluctuation of queue length. If queue length of a link fluctuates drastically, packets that traverse the link experience non-constant queuing delay, which leads to the performance degradation. We evaluate the fluctuation using the measurement Hurst parameter (H, 0.5< H <1) by applying the R/S plot method [17]. The measurement Hurst parameter represents the degree of strength of long-range dependence. The detail of R/S plot method is as follows. First, we define R_n/S_n value as,

$$R_n/S_n = 1/S_n[\max(0, W_1, W_2, \cdots, W_n) - \min(0, W_1, W_2, \cdots, W_n)], \quad (3)$$

where

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

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

Figure 4 shows Hurst parameters for each link. Y-axis represents Hurst parameters and X-axis represents its rank. In this figure, the result of the stop-and-wait model is also added. Looking at this figure, we observe that the number of links that take high Hurst parameter increases by 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 value ($H \geqq 0.8$) in Table 3. When the number of sessions is small, queue length of the links that connect two regions fluctuates drastically. That is, inter-module links tend to be highly fluctuated. This is because many packets concentrate on inter-module links. As the number of sessions gets higher, queue length of the links that connect inside a region fluctuates while the Hurst parameter of inter-module links decreased. That is, fluctuation spreads to tributary links of bottleneck.

4. 3 Positive and Negative Effects of TCP

In Section 4. 1 and 4. 2, we discussed about packet delay dynamics in the TCP model. In this section, we explain about the positive and negative effects of TCP. One of positive effects of TCP is

Table 3 The ratio of high fluctuated links in each topology

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

improvement of throughput. We compare the results of the stop-and-wait model with that of the TCP model, network throughput is improved in the AT&T topology due to the functionality of TCP that adjusts packet sending rate to network condition (but not presented due to space restriction).

In contrast, when the number of sessions gets larger, TCP become a factor of fluctuation of queue length as depicted in Figure 4, which leads to the high fluctuation in end-to-end delay of packets. If queue length of a given link fluctuate drastically, end-to-end delay of packets that pass through the link vary significantly. Since end hosts estimate network condition with Round Trip Time (RTT), when end-to-end packet delay fluctuate, end hosts cannot know network condition accurately and thus flow control becomes difficult.

Queue length of the inter-module links fluctuates drastically in the stop-and-wait model, while queue length of the intra-module links does not fluctuate in the AT&T topology. We confirm, but not shown due to the space limitation, that the number of links having high H value in the stop-and-wait model is smaller than the number of links in the TCP model. In the TCP model, 11% of total links are highly fluctuated when the number of sessions is 100,000, while 5% of total links are highly fluctuated in the stop-and-wait model. These results indicate complex functionality of TCP, such as flow control, congestion control, and fast retransmit functionalities, causes fluctuation of queue length. This is one of negative effects of TCP.

5. Conclusion

In this paper, we investigated on the interaction between structures of topologies and flow control between end hosts. Comparing with the simulation results of the stop-and-wait model and the TCP model, functionality of TCP improves network throughput in the AT&T topology. On the other hand, functionality of TCP makes the queue length fluctuated. Even in this case, the modularity structure of the AT&T topology reduces the number of highly fluctuated link comparing to the results of the BA topology.

In the future work, we will conduct evaluations of packet delay dynamics on topologies that have heterogeneous link capacity, and evaluation of combination 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 Early-concept Grants for Exploratory Research on New-generetion Network of 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] 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.

[11] 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.

[12] R. Guimerà and L. A. N. Amaral, "Functional cartography of complex metabolic networks," *Nature*, vol. 433, pp. 895–900, Feb. 2005.

[13] 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.

[14] 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.

[15] 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.

[16] M. Allman, V. Paxson, and W. Stevens, "TCP congestion control," *RFC 2581*, Apr. 1999.

[17] 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.