Analysis, Modeling, and Design of ISP Topologies Considering Traffic Dynamics

January 2013

Takahiro HIRAYAMA

Analysis, Modeling, and Design of ISP Topologies Considering Traffic Dynamics

Submitted to
Graduate School of Information Science and Technology
Osaka University
January 2013

Takahiro HIRAYAMA

List of Publications

Journal Papers

1. T. Hirayama, S. Arakawa, S. Hosoki, and M. Murata, “Models of link capacity distribution in ISP’s router-level topologies,” International Journal of Computer Networks & Communications (IJCNC), vol. 3, pp. 205-216, September 2011.

2. T. Hirayama, S. Arakawa, K. Arai, and M. Murata, “Dynamics of feedback-induced packet delay in ISP router-level topologies,” IEICE Transactions on Communications, vol. E95-B, pp. 2785-2793, September, 2012.

Refereed Conference Papers

1. T. Hirayama, S. Arakawa, K. Arai, and M. Murata, “On the Packet Delay Distribution in Power-law Networks,” in *Proceedings of International Conference on Evolving Internet (INTERNET 2009)*, August 2009.

2. T. Hirayama, S. Arakawa, K. Arai, and M. Murata, “Queue dynamics in power-law networks,” in *Proceedings of International Workshop on Sensor Networks and Ambient Intelligence (SeNAmI 2009)*, December 2009.

3. T. Hirayama, S. Arakawa, K. Arai, and M. Murata, “Dynamics of Feedback-induced Packet Delay in Power-law Networks,” in *Proceedings of International Conference on Systems and Networks Communications (ICSNC2010)*, August 2010.

4. T. Hirayama, S. Arakawa, K. Arai, and M. Murata, “Modularity structure and traffic dynamics of ISP router-level topologies,” in *Proceedings of International Symposium on Nonlinear Theory and its Applications (NOLTA 2012)*, October 2012.

Non-Refereed Technical Papers

1. T. Hirayama, S. Arakawa, K. Arai, and M. Murata, “Congestion Propagation in ISP Topologies Having Power-law Degree Distribution,” *Technical Report of IEICE (NS2008-68)*, vol. 108, pp. 1–6, October 2008. (in Japanese).

2. T. Hirayama, S. Arakawa, K. Arai, and M. Murata, “Modularity structure and traffic dynamics of ISP router-level topologies,” *Technical Report of IEICE (NS2009-82)*, vol. 109, pp. 31–36, October 2009. (in Japanese).

3. T. Hirayama, S. Arakawa, K. Arai, and M. Murata, “Dynamics of Feedback-induced Packet Transfer Delay in Power-law Networks,” *Technical Report of IEICE (NS2010-38)*, vol. 110, pp. 1–6, July 2010.

4. T. Hirayama, S. Arakawa, K. Arai, and M. Murata, “Effect of Modularity Structure on Traffic Dynamics,” *Technical Report of IEICE (NwGN2010-39)*, pp. 15–20, January 2011.

Preface

As the Internet becomes a social infrastructure of information exchange and global communication, many network devices are connected to the Internet and many users communicate via the Internet. Nowadays, the Internet faces several problems. One of problems is complex traffic behavior of the Internet. The development of new networking applications contributes to an increase in complexity of Internet traffic behavior. For instance, conflicts between selfish behavior in overlay networks and traffic engineering in underlay networks lead to performance degradation. Furthermore, with the advancement of cloud computing, data migration between data centers causes unexpected traffic fluctuation and unexpected traffic increase. Such the Internet traffic behavior could lead to excessive expansion or resource shortage. Thus, understanding the Internet traffic behavior is important for designing cost effective networks.

In this thesis, we investigate design of cost effective networks under the traffic dynamics. For this purpose, understanding of traffic dynamics and models of capacity allocation are both essential.

We begin this thesis with discussion about traffic dynamics in ISP router-level topologies (hereinafter we call ISP topologies). We try to uncover the relationships between traffic behavior and topological structures of ISP topologies. Understanding traffic dynamics and its causal structure is expected to be applied to the resource allocation and network design. For example, understanding traffic dynamics in ISP topologies is useful to absorb Internet traffic behavior, such as allocating bandwidth to links having the potential of large fluctuation. For the other example, detecting how topological structures affect traffic dynamics is contributory to develop a design method for cost effective networks that suppress traffic fluctuation. We first analyze traffic dynamics in large-scale topologies, such as ISP topologies, where TCP sessions interact with each other over all links. Then, we investigate how the topological structure affects the network performance in detail. Our results show that flow control functionality of TCP causes high

traffic fluctuation. However, the high-modularity structure of ISP topologies reduces the number of links where traffic fluctuation occurs.

Our next concern is how the link bandwidth of ISP topologies was designed under the traffic dynamics induced by flow control functionality. Because link bandwidth is a characteristic particular to communication networks and is essential to the network performance, modeling link bandwidth distribution of ISP networks is crucial to establish new network control methods. First of all, we confirm the link bandwidth distribution of an actual ISP network in Japan follows a power law, and topologies with power-law link bandwidth distribution can accommodate traffic efficiently. To clarify the reason why the link bandwidth distribution shows a power-law characteristic, we focus on design principles in ISPs. That is, ISPs enhance the bandwidth of a bottleneck link, either periodically or occasionally, through the traffic monitoring. However, the bottleneck link is dynamically changing due to the dynamic interaction between TCP and structure of ISP topologies. We therefore investigate the evolution process of link bandwidth and its distribution to accommodate TCP traffic in ISP topologies. We conduct a simulation for periodic bandwidth enhancement with consideration of traffic dynamics induced by TCP and show that the periodic bandwidth enhancement reproduces the similar characteristics of bandwidth distribution of the actual ISP network in Japan. These findings are in turn applied to develop our modeling method of link bandwidth of ISP networks. We propose a bandwidth allocation model by considering the both the topological structures of ISP topologies and TCP traffic dynamics, and show that our model generates a bandwidth distribution with the similar properties to that obtained by the periodic bandwidth enhancement. We believe that our model provides simulation environments for new control methods of the Internet.

Finally, we discuss about how interconnection topologies between constitutive subnetworks affect network performance. Interconnection topology between ISPs is important to design cost effective networks since ISPs should consider how to accommodate external traffic from neighboring ISPs. We therefore investigate the relationship between network performance and structure of interconnection topology. Our results show that the multiscale structure, where subnetworks are connected by links of various distance and link density changes depending on the hierarchy level, suppresses the total amount of bandwidth to accommodate traffic demand in interconnected networks. These results show possibility of a solution of the Internet design. Our findings feed new insight design of the Internet topology composed by interconnecting ISP topologies.

Acknowledgments

First and foremost, I would like to express my sincere appreciation to Professor Masayuki Murata of Osaka University for introducing me to the area of complex networks. His creative suggestions, insightful comments, and patient encouragement have been essential for my research activity. I also thank him for providing me with the opportunity to research with a talented team of researchers.

I am heartily grateful to the members of my thesis committee, Professor Koso Murakami, and Professor Teruo Higashino of Graduate School of Information Science and Technology, Osaka University, and Professor Hirotaka Nakano of Cyber Media Center, Osaka University, for their critical reviews and comments from various angles.

I am also deeply grateful to Associate Professor Shin'ichi Arakawa of Osaka University for his much appreciated comments and support. His critical comments and unerring guidance are always informative and helpful. His kindness on my behalf were invaluable, and I am forever in debt.

I would like to thank Professor Naoki Wakamiya and Associate Professor Go Hasegawa, Assistant Professor Yuichi Ohsita, and Assistant Professor Yuki Koizumi of Osaka University for enlightening discussions. I am thankful for the help of Assistant Professor Shinsuke Kajioka of Information Technology Center, Nagoya Institute of Technology. In addition, I would also like to thank Dr. Ken-ichi Arai of of NTT Communication Science Laboratories for fruitful discussions.

I thank all the members of the Advanced Network Architecture Laboratory at the Graduate School of Information Science and Technology, Osaka University, for their support and encouragement and for providing a pleasant and collegial atmosphere.

Last, but not least, I thank my parents for their invaluable support and constant encouragement during my undergraduate and doctoral studies.

Contents

List of Publications **i**

Preface **iii**

Acknowledgments **v**

1 Introduction **1**

1.1 Background and Overview 1

1.2 Outline of Thesis 5

2 Analyzing Dynamics of Feedback-induced Packet Delay in ISP Router-level Topologies **9**

2.1 Introduction 9

2.2 Related Work 11

2.2.1 Structural Properties of Power-law Networks 11

2.2.2 Traffic Dynamics in Power-law Networks 12

2.3 Simulation Model 14

2.3.1 Network Topologies 14

2.3.2 Packet Processing Model of Nodes 14

2.3.3 Flow Control between End Hosts 15

2.4 Dynamics of TCP in Power-law Networks 16

2.4.1 End-to-End Delay Distribution 16

2.4.2 Queue Length Fluctuation 17

2.4.3 Fluctuation Reduction Effect of High-Modularity Structure 18
2.4.4 Effects of TCP 22
2.5 Conclusion 23

3 Performance Evaluation of Link Bandwidth Distribution in ISP Router-level Topologies 31
3.1 Introduction 31
3.2 Related Work 32
3.3 Link Bandwidth Distribution in ISP Router-level Topology 34
3.4 Topologies for Evaluation 34
3.4.1 ISP Topologies 34
3.4.2 The Abilene-inspired topology 35
3.4.3 The BA Topology 35
3.4.4 The FKP Topology 36
3.5 Performance Evaluation on Link Bandwidth Distribution 37
3.5.1 Link Bandwidth Distributions 37
3.5.2 Assigning Link Bandwidth 38
3.5.3 Evaluation 39
3.5.4 Effects of Node Processing Constraint 40
3.6 Conclusion 42

4 Modeling Link Bandwidth Distribution in ISP Router-level Topologies 47
4.1 Introduction 47
4.2 Related Work 49
4.3 Modeling and Validation 50
4.3.1 Simulation Model 50
4.3.2 Effect of TCP Traffic on the Bandwidth Distribution 51
4.3.3 Validation of the Bandwidth Allocation Algorithm with an ISP Topology 53
4.4 Effects of Bandwidth Allocation Considering TCP Traffic Dynamics in Large-scale Topologies 54

4.4.1 Evaluation of Bandwidth Distribution and Network Performance 54
4.4.2 Effect of Topological Structure on Performance 57
4.4.3 Relations between Bandwidth Distribution and ISP Topological Structure 59
4.5 Modeling Bandwidth Distribution with Consideration of TCP Traffic Dynamics 59
4.5.1 Modeling Bandwidth Allocation and Evaluation 60
4.5.2 Modeling Bandwidth Distribution with a Simple Allocation Approach 62
4.6 Conclusion 62

5 Design of Interconnected Networks between ISP Topologies 73
5.1 Introduction 73
5.2 Related Work 75
5.2.1 Robustness against Cascade Failure of Interconnected Networks 75
5.2.2 Robustness of the Multiscale Structure 76
5.3 Network Connecting Model 77
5.4 Performance Evaluation of Interconnection Topology 79
5.4.1 Evaluation of Interconnected Networks Formed by Subnetworks Having Five Classes of Structure 79
5.4.2 Evaluation of Interconnected Networks Formed by ISP Topologies 82
5.5 Conclusion 88

6 Conclusion and Future Work 91

Bibliography 95

List of Figures

1.1 Evolution of capacity of routers, traffic load, and CMOS energy efficiency 2

2.1 Classification of node function by participation coefficient and within-module degree . . . 13
2.2 Classification of node function in each topology 24
2.3 End-to-end delay distribution 25
2.4 The number of congested links 26
2.5 Correlations between Hurst parameter and rank of Hurst parameter 27
2.6 Correlations between Hurst parameter and Hurst parameter rank of generated topologies . 28
2.7 Relationships between link load and queue fluctuation 28
2.8 Correlations between deviations of RTT for each session and deviation rank in the AT&T topology 29

3.1 Node processing capacity distribution in the IIJ network 35
3.2 Link bandwidth distribution in the IIJ network 36
3.3 Normalized link bandwidth distributions 38
3.4 The amount of traffic that the AT&T topology can accommodate 39
3.5 The amount of traffic that the Sprint topology can accommodate 40
3.6 The amount of traffic that the BA topology can accommodate 41
3.7 The amount of traffic that FKP topology can accommodate 42
3.8 The amount of traffic that the Abilene-inspired topology can accommodate 43
3.9 Relationship between node degree and node processing capacity 44

3.10 The network performance of the Sprint topology under constraints 45
3.11 The network performance of Abilene-inspired topology under constraints 46

4.1 The Parking lot topology 53
4.2 Observed bandwidth distribution of the IIJ topology 55
4.3 Bandwidth distributions of the IIJ topology 56
4.4 Bandwidth distribution of each topology 64
4.5 Correlation between bandwidth and number of passing sessions 65
4.6 Transition of maximum drop rate 66
4.7 Transition of network throughput 67
4.8 Schematic of the structure of each topology 68
4.9 Bandwidth distributions of ISP topologies 69
4.10 Schematic of proposed model 70
4.11 Bandwidth distribution of the AT&T topology generated by the proposed model 70
4.12 Throughput comparison for the AT&T topology 71
4.13 Bandwidth distribution obtained by simple bandwidth allocation 72

5.1 Illustrative example of the network construction algorithm 77
5.2 The Abilene-inspired topology 79
5.3 Construction of interconnected network: fundamental form 80
5.4 Construction of interconnected network when two nodes are located on the same point . . 80
5.5 Evaluation of interconnected networks formed by subnetworks having the random structure 83
5.6 Evaluation of interconnected networks formed by subnetworks having the local team structure 84
5.7 Evaluation of interconnected networks formed by subnetworks having the random inter-divisional structure 85
5.8 Evaluation of interconnected networks formed by subnetworks having the core-periphery structure 86
5.9 Evaluation of interconnected networks formed by subnetworks having the multiscale structure 87

5.10 Bandwidth distributions of the core-periphery structure and the multiscale structure 88
5.11 Evaluation of interconnected networks formed by the two Abilene-inspired topologies . . . 89

List of Tables

2.1 The simulation parameters used in the TCP model 16
2.2 The ratio of highly fluctuating links in each topology 19
2.3 Simulation results of ISP topologies 20
2.4 Simulation results of generated topologies 21

4.1 Simulation parameter values 52
4.2 Bottleneck links during simulation of parking lot topology 54

Chapter 1

Introduction

1.1 Background and Overview

As the Internet becomes a social infrastructure of information exchange and global communication, many network devices are connected to the Internet and many users communicate via the Internet. Nowadays, the Internet faces several problems. One of problems is the growth of traffic volume. According to Cisco VNI forecast [1], the total volume of global IP traffic is expected to grow from a volume of 31 Exabytes per month in 2011 to a volume of 110 Exabytes per month in 2016. The traffic volume of the Internet is always increasing. The authors of Ref. [2] compared evolution of high-end IP router's capacity, traffic load, and energy efficiency in silicon (e.g. CMOS) technologies (shown in Fig. 1.1). The increasing trend in the volume of traffic follows Moore's law by doubling every 18 months [3]. Similarly, capacity of high-end IP routers keeps increasing with an increase factor of 2.5 every 18 months [4]. The figure shows that increasing ratio of router capacity exceeds the growth rate of traffic volume. On one hand, there is a possibility that the router capacity will still continue to increase. On the other hand, upgrades or replacements of routers are driven by the increase of traffic volume. Therefore, reducing traffic congestion is important to suppress the use of high performance routers. Similarly, energy consumption of high performance routers becomes more serious high performance routers usually consume lots of energy.. Designing networks in order to minimize the potential for excessive congestion on routers or links reduces the use of high performance routers.

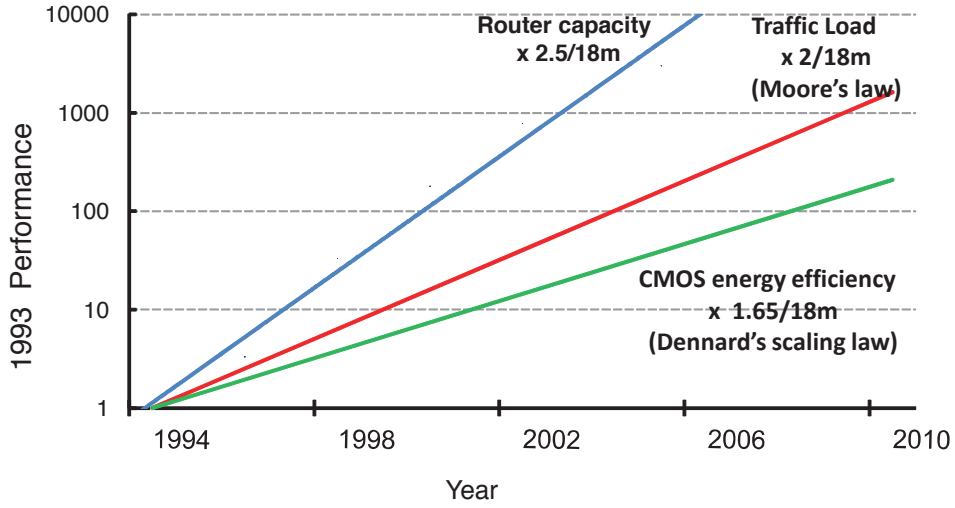


Figure 1.1: Evolution of capacity of routers, traffic load, and CMOS energy efficiency: This figure is referred from Ref. [2]

For designing cost effective networks, understanding Internet traffic dynamics is essential. Understanding traffic dynamics and its causal structure is expected to be applied to the resource allocation and network design. For example, understanding traffic dynamics is useful to absorb Internet traffic behavior, such as allocating bandwidth to links having the potential of large fluctuation. For the other example, detecting how topological structure affects traffic dynamics is contributory to develop a design method for cost effective networks that suppress traffic fluctuation. However, the Internet traffic behavior is hard to capture precisely because of its complexity. The development of new networking applications contributes to an increase in complexity of Internet traffic behavior. For instance, conflicts between selfish behavior in overlay networks and traffic engineering in underlay networks lead to performance degradation [5]. Furthermore, with the advancement of cloud computing [6], data migration between data centers causes unexpected traffic fluctuation and unexpected traffic increase [7]. In addition, the increasing number of Internet connected devices induced by the fast popularization of smartphones and laptops [8] and the development of Internet of things [9] causes increase in complexity.

We therefore focus on the relations between complex behavior and topological structure, rather than precisely capturing the Internet traffic behavior. For this purpose, we use the knowledge of complex

networks science. Complex networks are ubiquitous real world including the Internet and, for example, transportation networks, social networks, food web. Many of the findings in complex networks have been accomplished over broad areas such as physics [10–13], biology [14], social science [15, 16]. One of the features of complex networks is a power-law degree distribution. In the power-law topologies, existence probability $P(k)$ that nodes having k links is proportional to $k^{-\gamma}$ (γ is constant). Actually, measurement studies revealed that the degree distribution of the Internet topology follows a power law [17,18]. Recently, many researchers discussed about generation methods of power-law topologies [19–23] and topological properties of complex systems [14, 24–28]. For example, traffic dynamics in complex networks has been discussed in Refs. [24–26, 29–31]. These studies will be helpful to understanding the traffic dynamics in ISP topologies. Findings from these studies are helpful to designing cost effective networks. However, the network performance of power-law topologies is affected not only by their degree distribution but also by their topological structure. Li et. al. [21] introduced topologies having different structure regardless of the same degree distribution. They pointed out that the difference in structure leads to the difference in network performance, such as total traffic which network can accommodate. Previously, we have compared the structural differences of the AT&T topology measured using the Rocketfuel tool [32] and the topology generated by the Barabási- Albert (BA) model (introduced in Ref. [19]) [33, 34]. We separate the two topologies into modules, the set of densely-connected nodes, by using the method in Ref. [15]. We then examine class of node functions determined by Ref. [14]. The comparison between the two topologies revealed the AT&T topology has highly-modularity structure, that is, some modules are connecting with each other by a few inter-module links. We confirm the high-modularity structure is a common characteristic of ISP topologies. We therefore focus on the high-modularity structure of ISP topologies.

Then, we begin this thesis with discussion about traffic dynamics in ISP topologies. We try to uncover the relationships between traffic behavior and the topological structure of ISP topologies. Toward understanding Internet traffic behavior, we first analyze traffic dynamics in large-scale topologies, such as ISP topologies. In previous studies, the relationships between the statistical properties of Internet traffic and end-to-end flow control have been discussed. In Refs. [35, 36], it is revealed that Internet traffic exhibits long-range dependence (LRD), where traffic fluctuation appears to be independent of measurement time scale. Various studies have investigated the reasons for LRD. One of the reasons is flow control in

the transport layer, such as TCP [37–39]. However, these studies deal with small, simple topologies. In large-scale topologies, TCP sessions interact with each other over all links. Then, we investigate how the high-modularity structure affects the network performance in detail. Our results show that flow control functionality of TCP causes high traffic fluctuation. However, the high-modularity structure of ISP topologies reduces the number of links where traffic fluctuation occurs.

Our next concern is how the link bandwidth of ISP topologies was designed under the traffic dynamics induced by flow control functionality. Because link bandwidth is a characteristic particular to communication networks and is essential to the network performance, modeling link bandwidth distribution of ISP networks is crucial to establish new network control methods. First of all, we confirm the link bandwidth distribution of an actual ISP network in Japan follows a power law and topologies with power-law link bandwidth distribution can accommodate traffic efficiently. To clarify the reason why the link bandwidth distribution shows a power-law characteristic, we focus on design principles in ISPs. ISP networks are designed to accommodate dynamic traffic. ISPs monitor the traffic, and forecast the peak demand in the near future. On the basis of this prediction, they then enhance the bandwidth of a bottleneck link, either periodically or occasionally [40]. The current bandwidth distributions of ISP topologies thus result from the accumulation of link bandwidth enhancements. We consider that there would be a close relationship between link bandwidth and structure of ISP topologies. We therefore investigate the evolution process of bandwidth distributions to accommodate TCP traffic in ISP topologies. We conduct a simulation for periodic bandwidth enhancement with consideration of traffic dynamics induced by TCP and show that the periodic bandwidth enhancement reproduces the similar characteristics of bandwidth distribution of the actual ISP network in Japan. Comparing the results of ISP topologies and model-based topologies, we conclude interaction between the topological structure of ISP topologies and TCP traffic contributes a power-law bandwidth distribution. These findings are in turn applied to develop our modeling method of link bandwidth of ISP networks. We propose a bandwidth allocation model by considering the both the topological structure of ISP topologies and TCP traffic dynamics, and show that our model generates a bandwidth distribution with the similar properties to that obtained by periodic enhancement. We believe that our model provides simulation environments for new control methods of the Internet.

Finally, we discuss about how interconnection topologies between constitutive subnetworks affect network performance. In this thesis, we call a network constructed by subnetworks “interconnected network”

and call a topology composed by links between subnetworks "interconnection topology." Until the previous paragraph, we have focused on the relations between traffic dynamics and bandwidth distribution in a single ISP topology. However, the Internet is composed by a large number of ISP topologies. Recently, huge content provider called "Hyper Giants" such as Google contribute a large amount of Internet inter-domain traffic. According to one estimate, the total Internet inter-domain traffic reaches around 120Tbps in the mid of 2012 [41]. To achieve cost effective networks, ISPs should consider how to accommodate external traffic from neighboring ISPs. In the past, ISPs have constructed peering and transit connections with other ISPs at their own policies in order to increase their profit. However, communication efficiency obtained by the discretion of each ISP does not lead to globally optimized efficiency. Therefore, interconnection topology between ISPs is important to design cost effective networks. We investigate the performance of interconnection topology with graph-theoretical approach. It is expected that topological structure of interconnection topology between subnetworks affect the performance of interconnected networks. We construct interconnection topologies having five classes of structure determined by Ref. [16]. We then evaluate network performance of interconnected topologies composed by subnetworks having each class of structure. Our results show that the multiscale structure, where subnetworks are connected by links of various distances in terms of relative position in subnetworks, and link density changes depending on the hierarchy level, suppresses the total amount of bandwidth to accommodate traffic demand regardless of the difference in structure of subnetworks and the difference in degree distributions of subnetworks. These results show the possibility of a solution of the Internet design. Our findings feed new insight design of the Internet topology composed by interconnecting ISP topologies.

1.2 Outline of Thesis

Analyzing Dynamics of Feedback-induced Packet Delay in ISP Router-level Topologies [42–50]

In Chapter 2, we investigate packet delay dynamics and traffic fluctuation in ISP topologies where the degree distribution exhibits a power-law nature, and the nodes interact via end-to-end feedback control functionality. Measurement studies have revealed that the degree distribution of the Internet topology

follows a power law [17, 18]. That is, the existence probability $P(k)$ of node with k out-going links approximates to $k^{-\gamma}$ (γ is constant). The structure of power-law topologies also affect the performance of networking mechanisms, such as routing mechanisms [33]. These studies indicate that the power-law degree distribution alone does not determine network-level performance. Thus, the topological structure of ISP topologies has the large potential to affect the network performance. We first show the packet delay dynamics of the BA topologies generated by the Barabási- Albert (BA) model and the ISP topologies. Simulation results show that the end-to-end delay distributions exhibit a heavy tail in the TCP model. Moreover, the number of links with highly fluctuating queue length increases dramatically compared to that in the stop-and-wait model. However, the high-modularity structure of the ISP topologies reduces the number of highly fluctuating links compared with the BA topologies.

Performance Evaluation of Link Bandwidth Distributions in ISP Router-level Topologies [51]

One of our main motivations in Chapter 3 and 4 is to investigate the modeling methodology for realistic ISP networks. In Chapter 3, we investigate impacts of link bandwidth distributions in the ISP topologies. The performance of network control methods strongly depends on Internet topology and bandwidth of links, a proper network model is necessary to show effectiveness of those methods. However, many researches about modeling methods of topology have been discussed only focused on the power-law characteristic of the degree distribution. In these researches, there was an assumption that link bandwidth can be ignored, or they are identical. Therefore, it is important to reveal characteristics of a communication network and origin of them. We first reveal that the link bandwidth distribution of an actual ISP network in Japan obeys a power-law with exponent -1.1. To clarify the reason for the link bandwidth distribution, we evaluate throughput of networks having various kinds of link bandwidth distributions. We generate three distributions of link bandwidth; power-law, exponential and identical. Then we assign the link bandwidth to ISP topologies, and compare the network performance. Results show that topologies with power-law link bandwidth distribution can accommodate more traffic than that with other two distributions. We also reveal that it is insufficient to consider only the constraint of node processing capacities, even though it is discussed in Ref. [21], in evaluating the network performance.

Modeling Link Bandwidth Distribution in ISP Router-level Topologies [52]

In the previous Chapter 3, we show that topologies with power-law link bandwidth distribution can accommodate more traffic than that with other two distributions. We consider that there would be a close relationship between link bandwidth and structure of topology and reveal the relationship between link bandwidth and structure in ISP router-level topology.

In Chapter 4, we investigate the evolution process of bandwidth distributions to accommodate TCP traffic in ISP topologies. We focus on design principles of ISPs. ISPs monitor the traffic, and forecast the peak demand in the near future. ISPs predict future demand from monitored traffic enhance the bandwidth of a bottleneck link, either periodically or occasionally [40]. The current bandwidth distributions of ISP topologies have the possibility to be affected by the traffic dynamics induced by a dynamic flow control. We confirm the resulting bandwidth distributions obtained by periodic enhancement has the similar properties to the bandwidth distribution of an actual ISP topology in Japan with a power-law exponent of around -1. The results of other ISP topologies indicate that bandwidth distributions of ISP and BA topologies follow a power-law. We finally propose a bandwidth allocation model that reproduces the similar properties to the bandwidth distribution obtained through TCP traffic simulations.

Design of Interconnected Networks between ISP Topologies [53]

In Chapter 5, we focus on how interconnection topologies between interacting subnetworks affect network performance. The Internet is composed by connected a large number of autonomous systems (ASs), such as ISPs. ISPs have constructed peering and transit connections with other ISPs at their own policies in order to increase their profit. However, communication efficiency obtained by the discretion of each ISP does not lead to globally optimized efficiency. Therefore, designing interconnection topology between ISPs will be essential to accommodate increased traffic in the near future. We categorize interconnection topologies into five classes determined by Ref. [16]. We investigate performance of interconnected networks constructed by subnetworks having each structure. We show that performance of interconnected networks is having the multiscale structure, suppresses construction cost regardless of the difference in structure of subnetworks and the difference in degree distributions of subnetworks.

Chapter 2

Analyzing Dynamics of Feedback-induced Packet Delay in ISP Router-level Topologies

2.1 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 of TCP can cause short-range and long-range dependence of traffic [54]. Ever-changing networking technologies and applications make the behavior of the Internet more complex [55]; 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 that nodes having k links exist 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 distributions [19]. Li et al. showed several topologies that have different structures but have the same degree distribution [21, 56]. 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 networking mechanisms, such as routing mechanisms [33]. 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 performance of networks [21, 57].

In previous studies, the relationship between the statistical properties of Internet traffic and end-to-end flow control has been discussed. In Refs. [35, 36], it is revealed that Internet traffic exhibits long-range dependence (LRD), where traffic fluctuation appears to be independent of measurement time scale. Various studies have investigated the reasons for LRD. One of the reasons is flow control in the transport layer, such as TCP [37–39]. However, these studies deal with small, simple topologies.

It has been revealed that end-to-end flow control functionality of TCP sessions affects traffic dynamics through an interaction at a single link. In large-scale topologies, TCP sessions interact with each other over all links. Nevertheless, previous studies on design and performance evaluation of a large-scale topology [24, 26] have not considered the end-to-end feedback control functionality provided by transport layer. We therefore investigate the traffic dynamics in large-scale topologies where the topological structure greatly affects the network performance. More specifically, we investigate traffic dynamics on ISP router-level topologies (ISP topologies) where the degree distributions exhibit a power-law nature, and the nodes interact via end-to-end feedback flow control functionality. First, we show the packet delay dynamics and compare the results against the BA topologies generated by the BA model and the ISP topologies. We then discuss how the interaction between the structures of topologies and the flow controls affects the end-to-end delay distribution and the appearance of LRD in the queue length for each link. The results show that the TCP flow control makes the queue length of links in the network fluctuate because TCP tries to use available bandwidth. However, we show that the queue length of many links does not fluctuate regardless of TCP in the ISP topologies. This phenomenon is due to the high-modularity structure of the ISP topologies. We investigate the relationship between the modularity of topologies and queue fluctuation. We find that topologies with high-modularity structures have a lower number of highly fluctuating links.

This chapter is organized as follows. We introduce related work in Section 2.2. In Section 2.3, we show the network model that we used for the simulations. In Section 2.4, we evaluate the influence of the

power-law topologies and TCP flow control. Finally, in Section 2.5, we conclude this chapter and mention future work.

2.2 Related Work

2.2.1 Structural Properties of Power-law Networks

Recently, there have 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. [19]. The BA model generates a 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 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 ISP topologies. We have compared the structural differences of the AT&T topology measured using the Rocketfuel tool [32] and the topology generated by the BA model [34]. The results indicate that the design principles of networks greatly affect the structure of the ISP topologies: Design principles determine the node functionality, which in turn determines the connectivity of nodes.

In [14], 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

$$Z_i = \frac{k_i - <k_{s_i}>}{\sigma_{s_i}}, \tag{2.1}$$

where k_i is the degree of the node, $<k_{s_i}>$ 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.

We consider hub nodes as the nodes with Z_i greater than or equal to 2.58. That is, when we assume the standard normal distribution for Z_i, top 0.5% nodes are hub nodes and the remaining nodes are non-hub

nodes. Note that the distribution of Z_i is not always the normal distribution. However, we use 2.58 to distinguish hub nodes having a much larger number of out-going links in the module and non-hub nodes.

The participation coefficient P_i of node i is defined as

$$P_i = 1 - \sum_{s_i=1}^{N_m} \left(\frac{k_{is}}{k_i} \right)^2, \tag{2.2}$$

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 2.1 shows that the roles of nodes are categorized by the values of Z_i and P_i. 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 large Z_i and small P_i; that is, it has many links connecting with nodes in the same module.

Figure 2.2(a) and Fig. 2.2(b) show the results of applying Guimerà's method to the BA topology and the AT&T topology. The module is calculated using the method in [15]. In Fig. 2.2, the vertical axis indicates within-module degree Z, and the horizontal axis the participation coefficient P. Looking at Fig. 2.2(a), we observe that the BA topology has many connector hub nodes that connect between modules. However, Fig. 2.2(b) shows that there are few connector hubs in the AT&T topology.

2.2.2 Traffic Dynamics in Power-law Networks

Some studies have investigated traffic-level behavior in topologies having power-law degree distributions [24, 29]. Reference [24] 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 [29]. The authors examined TCP control with LRD input traffic and Poisson input traffic and revealed that the average end-to-end delay sharply increases for both types of input traffic since packets concentrate more at hub nodes in the

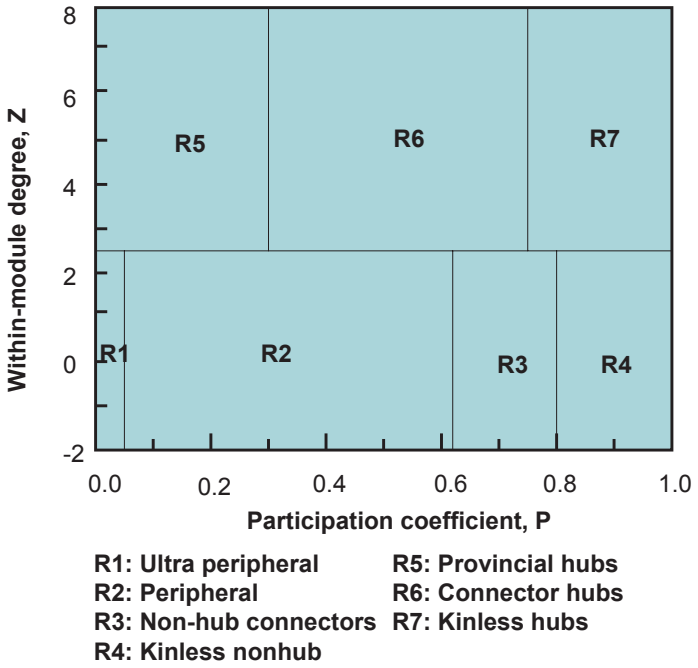


Figure 2.1: Classification of node function by participation coefficient and within-module degree

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 2.2.1, the BA model does not adequately describe the structure of ISP topologies. This clearly indicates that the power-law degree distribution alone does not determine traffic-level behavior in ISP 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 [35]. 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 [36]. In Refs. [37, 38], TCP flow control is considered to be the cause of LRD in Internet traffic. Simulation results of Ref. [39] show that LRD is still observed when the stop-and-wait protocol is used instead of TCP. That is, the flow control functionality is

an essential factor for the cause of LRD. However, these studies deal with small, simple topologies. Thus, we investigated how the structures of topologies impact the traffic dynamics.

2.3 Simulation Model

2.3.1 Network Topologies

For the simulation, we use several ISP topologies using the Rocketfuel tool [32] and BA topologies generated by the BA model. BA topologies are generated such that the numbers of nodes and links are the same as the corresponding ISP topologies. In Ref. [21, 57], it is shown that the BA model is insufficient to model ISP topologies. However, we use the BA topology in this chapter, because one of our purposes in the current paper is to clarify the essential structure of ISP topologies that characterizes their traffic dynamics by comparing with the BA topology. The Rocketfuel is a topology measurement tool based on `traceroute`. Note that the authors of Ref. [32] pointed out that the Rocketfuel might not cover some parts of ISP topologies. Even when the Rocketfuel misses some of routers, it is sufficient to investigate traffic dynamics because the Rocketfuel provides active paths, including MPLS-based paths, between routers. The actual ISP networks may have some routers invisible from `traceroute`. We believe that the routers are used for a backup purpose because they are placed on the inactive path for Internet users.

In the following simulation, we first evaluate and compare the AT&T topology with the corresponding BA topology to identify the essential structure of ISP topologies that characterizes their traffic dynamics.

2.3.2 Packet Processing Model of Nodes

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 actual ISP topologies. However, we use the same link bandwidth to evaluate traffic dynamics induced only by topological structure. 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 by a packet's source node. According to Ref. [58], 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.

2.3.3 Flow Control between End Hosts

We examined stop-and-wait model and TCP model for flow control between end hosts. The stop-and-wait model is introduced to clarify the impact of flow control functionality of TCP. In the simulation, pre-specified numbers of sessions are created between nodes. Source and destination nodes are randomly selected and each session arrives at a node pair according to the Poisson process with mean rate λ. Each session always has data to send during the simulations, i.e., once a session is generated, it continues to send data to destination node until the simulation ends.

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.

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

Table 2.1: The simulation parameters used in the TCP model

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

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 [59]. 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.

2.4 Dynamics of TCP in Power-law Networks

In this section, we show the results of simulation for TCP and discuss the end-to-end delay and queue-length fluctuation in detail. In the simulation, each link can transfer three packets per unit time. The other parameters are summarized in Table 2.1.

2.4.1 End-to-End Delay Distribution

Figure 2.3 shows the end-to-end delay distribution of the BA (AT&T) topology and the AT&T topology. The end-to-end delay represents one-way delay for DATA or ACK packets. In this figure, the end-to-end delay distribution in the stop-and-wait model (the number of sessions is 100,000) is also plotted. From

comparison with the stop-and-wait model, end-to-end delay gets larger in the TCP model. In the AT&T topology, when the number of sessions is 10,000, packets take longer to reach their destinations compared with the BA topology. In contrast, the end-to-end delay distribution of the BA topology changes drastically when the number of sessions is 100,000. It has a long-tail distribution; i.e., many packets take a long time to reach their destinations. End-to-end delay of less than 500 time units is hardly observed in the BA topology. However, the end-to-end delay distribution of the AT&T topology does not vary widely when the number of sessions increases.

The reason for this is that the 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 any more packets. To explain the reason more clearly, we show the number of congested links dependent on each time step in Fig. 2.4. The figure shows that the frequency of congestion in the BA topology is about 4 times larger than that in the AT&T topology. As discussed in Section 2.2.1, the BA topology has many connector hub nodes, and the connector hub nodes transfer a large amount of packets between modules. Packets concentrate at connector hub nodes in the BA topology. Thus, the links connected to connector hub nodes tend to be congested. Since most of sessions concentrate on the connector hub nodes and the congested links, the end-to-end delay distribution of the BA topology has a long tail.

2.4.2 Queue Length Fluctuation

Next, we evaluate the fluctuation of queue length. If queue length of a link fluctuates drastically, a session encounters a temporal congestion on the link, which leads to a packet drop and a longer delay. We evaluate the fluctuation using the Hurst parameter ($H, 0.5 < H < 1$) by applying the R/S plot method [60]. The Hurst parameter represents the degree of LRD. High Hurst parameter means that queue length of the link highly fluctuates.

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

$$\begin{aligned} R/S(n) &= 1/S_n[\max(0, W_1, W_2, \cdots, W_n) - \min(0, W_1, W_2, \cdots, W_n)], \\ W_k &= (X_1 + X_2 + \cdots + X_k) - kX(n), \end{aligned} \tag{2.3}$$

where $(X_k : k = 1, 2, \cdots, n)$ 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. We set the observation scale of n from $10^5/2^{13}$ to 10^5 time units and use the time series of queue length extracted from the last 100,000 time units of simulation. The slope of the fitted curve of the correlation function, αn^H, represents the Hurst parameter of a set X.

Figure 2.5 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 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. Besides, we observe that the AT&T topology reduces the number of fluctuating links compared with the results of the BA topology.

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 2.2. 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.

2.4.3 Fluctuation Reduction Effect of High-Modularity Structure

In the previous section, we showed that the 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. [61] defined a modularity value (Q) as

$$Q = \sum_i (e_{ii} - a_i^2), \tag{2.4}$$

Table 2.2: The ratio of highly fluctuating links in each topology

Topology	10,000 sessions			100,000 sessions		
	Total	Inter	Intra	Total	Inter	Intra
BA topology	0.54	0.19	0.35	0.88	0.30	0.58
AT&T topology	0.39	0.080	0.31	0.50	0.066	0.44

where e_{ii} is defined as the number of links connecting nodes of module i divided by the number of all links. While, a_i is defined as the number of links that have one or both vertices inside of module i divided by the number of all links. The module is calculated using the method in [15]. According to this definition of Q, high Q value means that the topology has high modularity structure, that is, some modules are connecting with each other by a few inter-module links. The modularity value of the BA topology is nearly 0.32, and that of the AT&T topology is about 0.68. This result indicates that a high-modularity structure reduces the number of highly fluctuating links.

To confirm that ISP topologies reduce the number of fluctuating links, we conduct simulations with other ISP topologies and compare them with BA topologies. Three ISP topologies, Sprint, Verio, and Telstra, are used for the simulations because the topologies have the similar numbers of nodes and links to the AT&T topology. The topological properties of ISP topologies and its corresponding BA topologies are summarized in Table 2.3. The simulation parameters are the same as the evaluation of the AT&T topology and are summarized in Table 2.1. Note that the number of sessions is 87,320 in the Telstra topology and the BA Telstra topology due to limit of the number of node pairs. In Table 2.3, the ratio of highly fluctuating links for each topology is also presented. We observe that each ISP topology reduces the number of fluctuating links compared with the results of the corresponding BA topologies. We also observe that each ISP topology has higher modularity value than the corresponding BA topology. These results indicate that topologies having high-modularity structure reduce the number of highly fluctuating links, as we observed in the AT&T topology.

To see the impact of modularity structure more clearly, we investigate queue fluctuation on three topologies that have the same number of nodes and links to the AT&T topology, but have different modularity values. We generated three topologies having N (=523) nodes and E (=1,304) links, m inter-modules

Table 2.3: Simulation results of ISP topologies

Topology	Nodes	Links	Modularity	Ratio of highly fluctuating links
AT&T	523	1,304	0.68	0.50
BA (AT&T)			0.32	0.88
Sprint	467	1,280	0.66	0.53
BA Sprint			0.30	0.81
Verio	817	1,874	0.70	0.46
BA Verio			0.25	0.77
Telstra	296	594	0.74	0.46
BA Telstra			0.29	0.78

links, and then control the modularity value by changing m. More precisely, we generated the three topologies through following steps.

1. Generates 10 sub-networks each of which consists of $\frac{N}{10}$ nodes and $\frac{E-m}{10}$ links and is connected based on the BA model.

2. Considers each sub-network as a module.

3. Adds m inter-module links by randomly selecting two nodes belonging to different modules and connecting the nodes.

As the number of inter-module links increases, the modularity value decreases. The modularity values of the generated topologies were 0.86, 0.81, and 0.76, and we confirmed that these topologies had power-law degree distributions.

Figure 2.6 shows the Hurst parameters of each link from the results of the simulations with the TCP model. The x-axis and y-axis represent the same parameters as those in Fig. 2.5. In the simulations, we again used the parameters in Table 2.1 and again randomly select the source and destination nodes for each session. In Fig. 2.6, 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 2.4. Thus, topological structures that have high modularity values prevent the appearance of highly fluctuating links.

Table 2.4: Simulation results of generated topologies

Modularity	Ratio of highly fluctuating links	Number of arriving packets		
		Total	Intra	Inter
0.86	0.62	4.7×10^8	4.5×10^8	1.7×10^7
0.81	0.78	3.6×10^8	2.9×10^8	7.4×10^7
0.76	0.84	3.2×10^8	2.2×10^8	1.0×10^8

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. 2.7. 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. 2.7(a)). 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. 2.7(c)). Although the queue length keeps nearly-constant, each of TCP sessions dynamically changes its sending rate. Thus, the queue length of the tributary links to the inter-module links is governed by the dynamics of each TCP session and thus fluctuates like Fig. 2.7(b). The situation in Fig. 2.7(b) occurs when the links are close to the inter-module links because the large number of TCP sessions is aggregated. More precisely, connector hub nodes or non-hub connector nodes in Fig. 2.1 perform the aggregation of TCP sessions. However, connector hub nodes rarely exist in topologies with the high-modularity structure: most of hub nodes are provincial hub nodes. Thus, the situation in Fig. 2.7(b) occurs around the connector non-hub nodes, and therefore the number of such links is small. Note that these results are induced by the shortage of buffer of inter-module links. Of course, if buffer size is unlimited, queue length of bottleneck links may fluctuate more drastically. However, such case hardly occurs in real situations because end hosts send traffic to use available bandwidth of bottleneck links as efficiently as possible.

The number of arriving packets forwarded between different modules also decreases in a topology that has a high modularity value. Table 2.4 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. That is, the throughput of intra-module sessions gets

high in the high modularity topology. 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.

2.4.4 Effects of TCP

In Section 2.4.1, 2.4.2, and 2.4.3, we discussed packet delay dynamics in the TCP model. In this section, we explain the merits and demerits of TCP. One of the merits of TCP is improved throughput. Comparing the results of the stop-and-wait model with those of the TCP model, we found that network throughput is improved in the AT&T topology due to the function of TCP that adjusts the packet sending rate according to the network conditions. In the AT&T topology, the ratio of links having utilization larger than 80% is 74% by the TCP model, while the ratio is 58% by the stop-and-wait model. Note that, the number of sessions is 100,000 and other parameters are the same as used in previous subsections (see Table 2.1 for details).

In contrast, when the number of sessions gets larger, TCP causes fluctuation of the queue length, as depicted in Fig. 2.5, which leads to the high fluctuation in the end-to-end delay of packets. Figure 2.8 shows deviations of RTT, RTT_{dev}, for each session in the AT&T topology. The y-axis represents the mean deviation of RTT for each session, defined as

$$\mathrm{RTT}_{dev} \quad = \quad (1-\alpha) \times \mathrm{RTT}_{dev} + \alpha \times |e\mathrm{RTT} - \mathrm{RTT}|, \tag{2.5}$$

where $\alpha = 0.125$, and the x-axis represents its rank. Every time the source node receives the ACK, the source node measures an instantaneous round-trip time (RTT in Eq. 2.5) and estimates the average round-trip time (eRTT) by the following equation:

$$e\mathrm{RTT} \quad = \quad (1-\alpha) \times e\mathrm{RTT} + \alpha \times \mathrm{RTT},$$

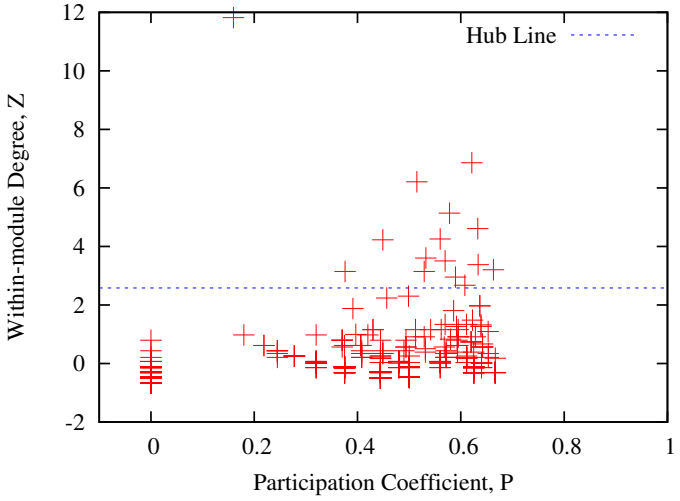
In the figure, the results of the stop-and-wait model are also plotted for comparison. As shown in Fig. 2.8, in the TCP model, the number of sessions that have a large deviation increases as compared with that in the stop-and-wait model.

The complex functionality of TCP, such as flow control, congestion control, and fast retransmit functionalities, causes fluctuation of the queue length. This is one of the disadvantages of TCP. We confirmed that the fraction of links having a high H value in the stop-and-wait model is smaller than the number of links in the TCP model. As shown in Fig. 2.5(b), when the number of sessions is 100,000, 50% of all links are highly fluctuating in the TCP model, whereas 9% of all links are highly fluctuating in the stop-and-wait model.

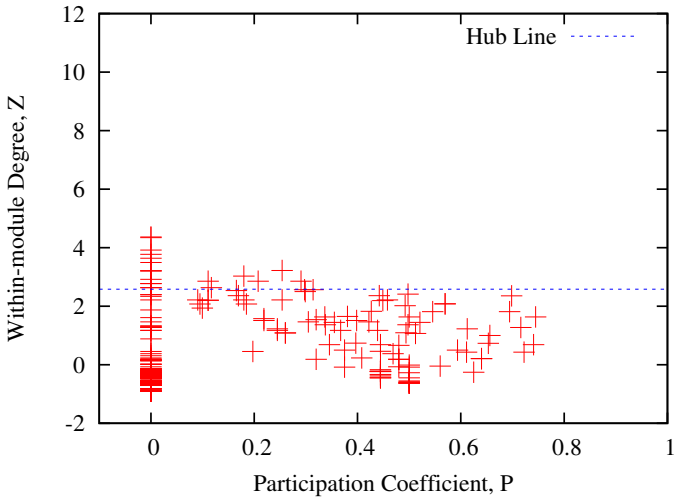
2.5 Conclusion

In this chapter, we investigated the interaction between the structures 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 improves network throughput in the AT&T topology. On the other hand, the functionality of TCP makes the queue length fluctuate. Even in this case, the highly modular structure of the AT&T topology reduces the number of highly fluctuating links compared with the BA topology. We also evaluated the queue length on other topologies and confirmed that the modularity structure can reduce the number of highly fluctuating links.

Our results suggest that reproducing the modularity structure is important when evaluating the performance of transport protocols. Our future work is to develop a topology generation method that reproduces the modularity structure and apply it to performance evaluations.

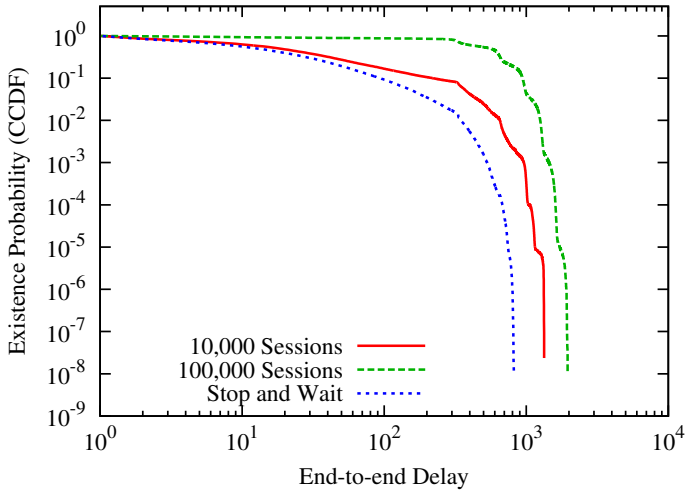


(a) The BA topology

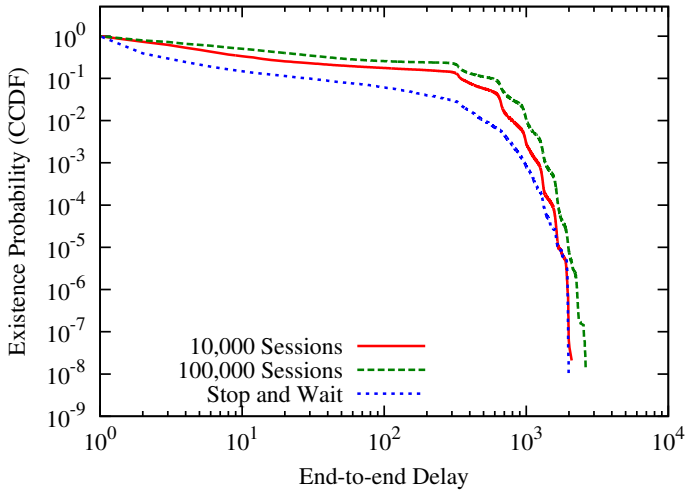


(b) The AT&T topology

Figure 2.2: Classification of node function in each topology

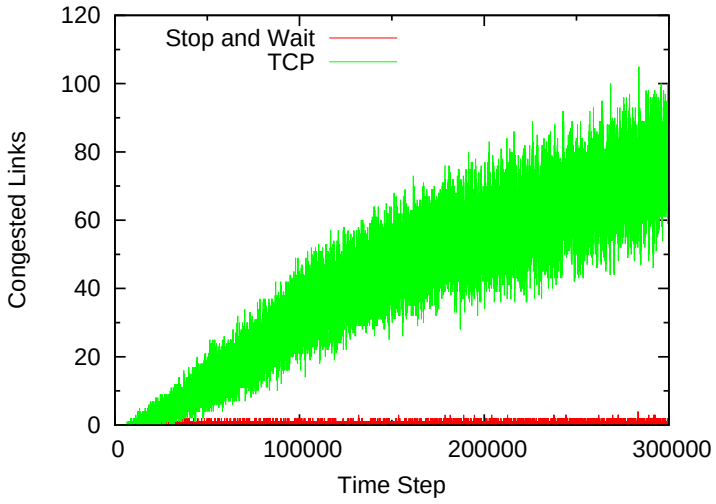


(a) The BA topology

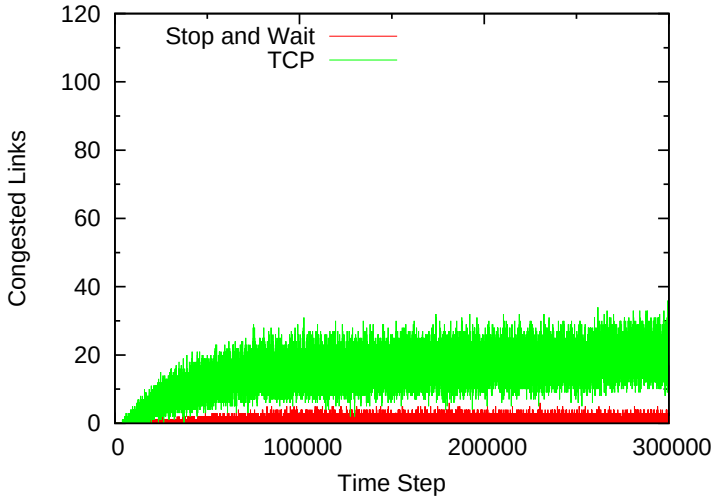


(b) The AT&T topology

Figure 2.3: End-to-end delay distribution

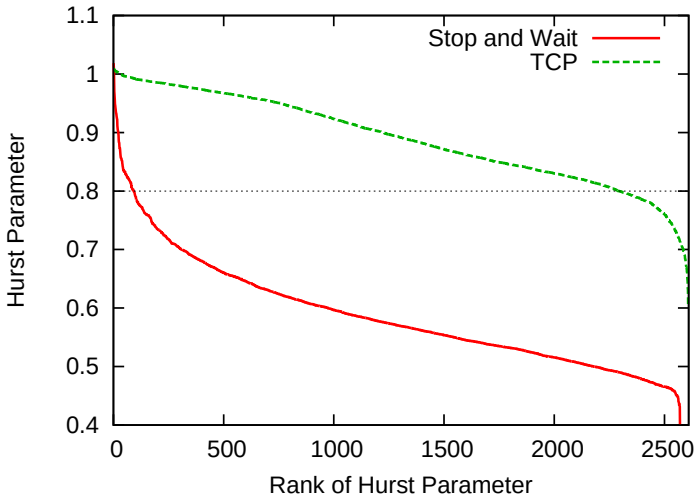


(a) The BA topology

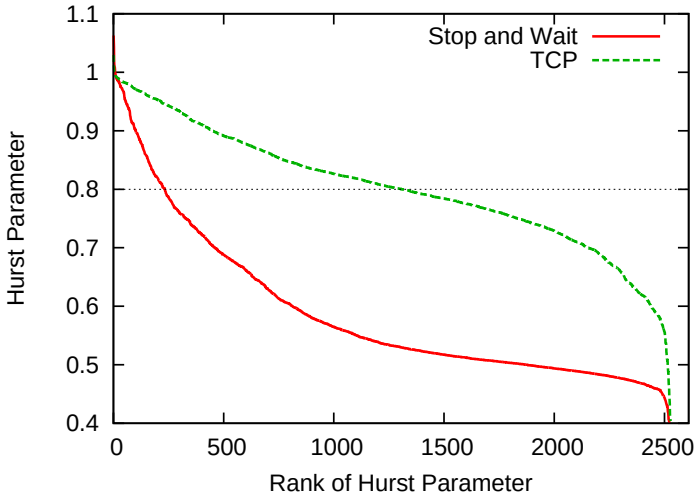


(b) The AT&T topology

Figure 2.4: The number of congested links



(a) The BA topology



(b) The AT&T topology

Figure 2.5: Correlations between Hurst parameter and rank of Hurst parameter

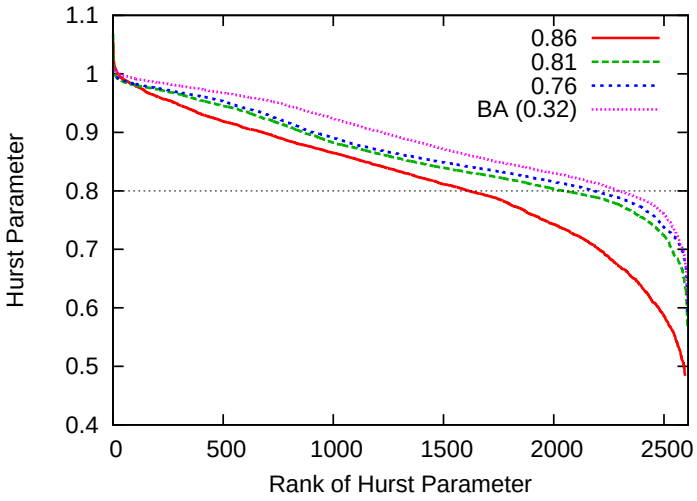
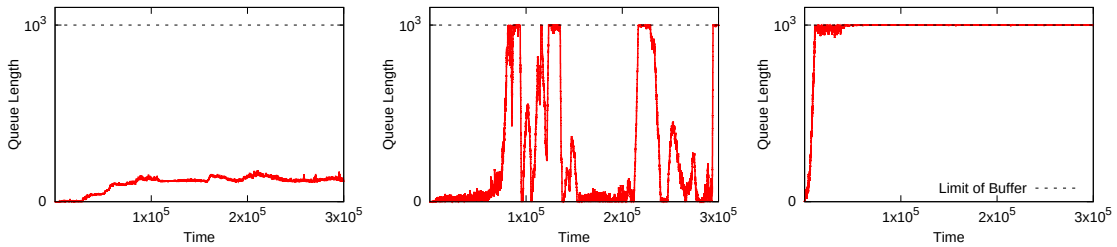


Figure 2.6: Correlations between Hurst parameter and Hurst parameter rank of generated topologies



(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

Figure 2.7: Relationships between link load and queue fluctuation

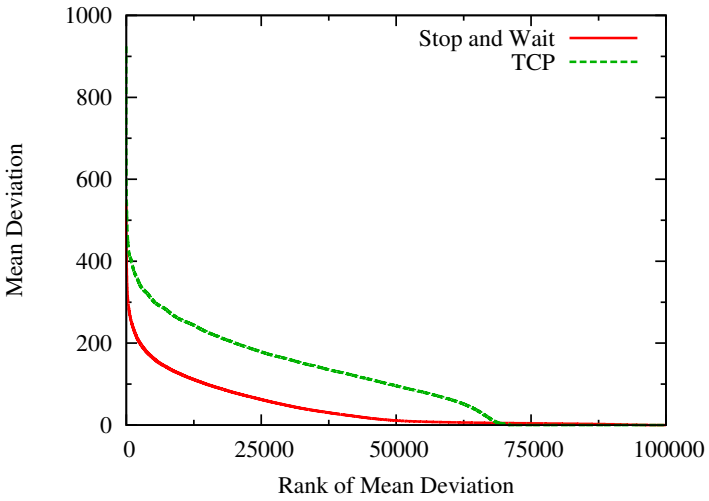


Figure 2.8: Correlations between deviations of RTT for each session and deviation rank in the AT&T topology: The number of sessions is 100,000.

Chapter 3

Performance Evaluation of Link Bandwidth Distribution in ISP Router-level Topologies

3.1 Introduction

Modeling the Internet is one of important issues to evaluate network control methods [62]. Since the performance of network control methods strongly depends on Internet topology and bandwidth of links, a proper network model is necessary to show effectiveness of those methods. Therefore, it is important to reveal characteristics of a communication network and origin of them.

Measurement studies on the Internet topology show that the degree distribution obeys the power-law [17], [18]. That is, the existence probability $P(k)$ of node with k out-going links approximates to $k^{-\gamma}$ (γ is constant). There are many modeling methods for the power-law topology [19–21]. Reference [20] presents FKP model for a topology generation model that consider technical constraints of router-level topologies. However, the topology is much different from the actual ISP topologies [27]. In Ref. [21], the authors enumerate several topologies having the same degree distribution. Then, they evaluate the amount of traffic that the network can accommodate in each topology under the constraint of node processing capacities. The results indicate that the network throughput highly depends on the structure of topologies

even though they have the same degree distribution.

Many researches about the modeling method of router-level topology have been discussed based on the characteristic that the degree distribution obeys the power-law. In these researches, there was an assumption that link bandwidth can be ignored, or they are identical. However, the performance of network strongly depends on link bandwidth because link bandwidth is characteristics particular to communication networks.

In this chapter, we investigate the characteristic of link bandwidth in router-level topology. One of our main motivations in this and next chapters is to investigate the modeling methodology for realistic ISP networks. For this purpose, characteristics particular to communication networks other than the degree distribution are important. First, we investigate the link bandwidth distribution of an actual ISP network in Japan using the capacity information presented in Ref. [63]. Results show that the link bandwidth distribution also obeys the power-law with exponent -1.11. Second, to reveal the reason why the link bandwidth distribution obeys the power-law, we investigate the network performance under the constraint of link bandwidth. We generate three distributions of link bandwidth; power-law, exponential and identical. Then we assign the link bandwidth to ISP topologies, and compare the network performance. Results show that topologies with power-law link bandwidth distribution can accommodate more traffic than that with other two distributions. We also reveal that it is insufficient to consider only the constraint of node processing capacities, even though it is discussed in Ref. [21], in evaluating the network performance.

This chapter is organized as follows. In Section 3.3, we investigate link bandwidth distribution of an actual ISP network in Japan. In Section 3.4, we explain about topology generation models and topologies which are used for evaluation. In Section 3.5, we investigate the difference of network performance by three distributions of link bandwidth. Section 3.6 concludes this chapter.

3.2 Related Work

There are relatively few studies on modeling router-level Internet topology [19–23]. Heckmann et. al. [22] present parameter settings for topology generator tools such as BRITE, TIERS, and GT-ITM to construct POP (point-of-presence)-level ISP topologies. Topology generation methods that consider technical constraints of router-level topologies are discussed in [20, 21]. Reference [20] presents a topology generation

model where a newly added node connects with existing nodes that minimize the sum of Euclidian distance between the nodes and logical distance from the existing node. The authors demonstrate that the degree distribution obeys the power-law with appropriate parameter settings. However, the topology generated by the model has too many nodes whose degree is one [64], and thus the topology is much different from the actual ISP topologies [27]. This is because ISP topologies are designed with technological, geographic, and business constraints [56, 65]. In Ref. [21], the authors enumerate several topologies having the same degree distribution. Then, they evaluate the amount of traffic that the network can accommodate in each topology under the constraint of node processing capacities. The results indicate that the network throughput highly depends on the structure of topologies even though they have the same degree distribution. With the constraint of node processing capacities, a product of node degree and its corresponding link bandwidth is bounded. Thus, for maximize the traffic that the network can accommodate, non-hub nodes having small number of links connects with higher-bandwidth links, and hub-nodes having many links connects with lower-bandwidth links.

Moreover, many researchers have discussed about capacity distribution in power-law topologies [28, 30, 66]. Zhang et al. mentioned that node capacity allocation based on node betweenness centrality is appropriate to enhance network capacity [28]. Betweenness centrality is defined as the number of shortest paths that pass through the nodes or links. In Ref. [66], Xia et al. assigned optimal node capacity with an optimization problem to maximize the system performance and the utilization ratio of capacity. This method provides a similar capacity distribution to that is allocated to be proportional to betweenness centrality of nodes in BA topologies. These researchers assumed that link bandwidth is unlimited. Actual communication systems like the Internet have heterogeneous bandwidth. In Ref. [30], Ling et al. discussed about bandwidth allocation strategy in BA topologies. They introduced an allocation strategy based on the product of end nodes' degree of links to maximize network throughput. In these papers, model-based topologies are concerned. Model-based topologies have different properties from ISP topologies'. So, we evaluate the relationship between link bandwidth distribution and network throughput in ISP topologies.

3.3 Link Bandwidth Distribution in ISP Router-level Topology

Ref. [63] investigates node processing capacities of a commercial ISP network in Japan. The author uses a disclosed information of link bandwidth in IIJ (Internet Initiative Japan) [67]. In the paper, a node processing capacity is defined as the sum of link bandwidth connected to the node. Figure 3.1, which was also presented in Ref. [63], shows the rank of node processing capacities in 2002. The vertical axis represents the node processing capacity and the horizontal axis represents the rank of that node. We observe that node processing capacity distribution obeys the power-law; its exponent is nearly -2.6.

The result is interesting in that the power-law relationship again appears in the node processing capacity. However, one question arises whether the link bandwidth distribution follows the power-law or not. To answer this, we show the link bandwidth distribution in Figure 3.2; the result shows that the exponent of the link bandwidth distribution (for 40 higher-ranking) is -1, i.e., the distribution obeys the zipf's law. The figure is obtained by using the information on IIJ backbone network in 2002, thus it is uncertain that other ISP topologies have the observation that the link bandwidth distribution obeys the zipf's law. However, we believe that the observation is general because we can easily derive it by the non-blocking configuration of routers. Supposing that there is one uplink port with the bandwidth of X in a router and there are α numbers of links, then if the router to be non-blocking, the bandwidth of each link should be X/α. With this case, the exponent of link bandwidth distribution becomes -1. Base of this discussion, in the next section, we investigate the performance of networks when the link bandwidth distribution obeys the power-law with exponent -1.

3.4 Topologies for Evaluation

In this section, we explain topologies used for evaluation.

3.4.1 ISP Topologies

We prepare two ISP topologies; the AT&T and the Sprint [34]. The AT&T topology has 523 nodes and 1304 links, and the Sprint topology has 467 nodes and 1280 links. These topologies measured using the

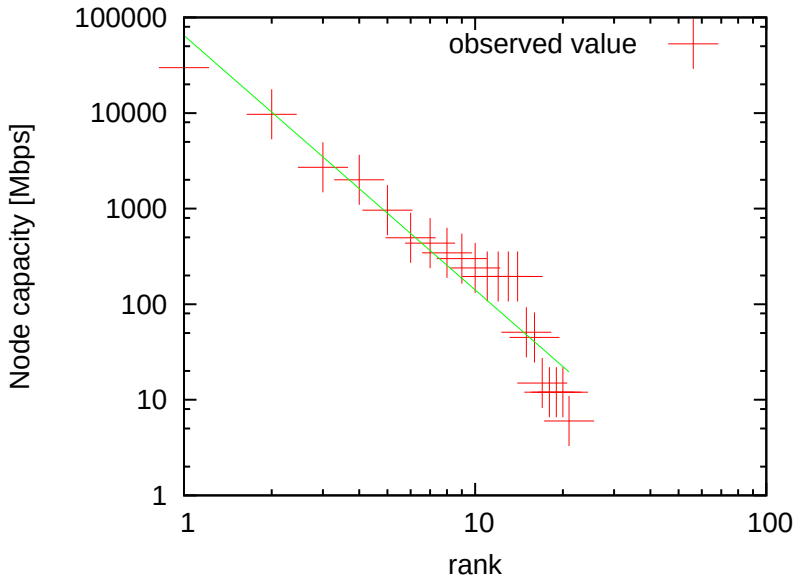


Figure 3.1: Node processing capacity distribution in the IIJ network

Rocketfuel tool [32].

3.4.2 The Abilene-inspired topology

As a router-level topology, we prepare the Abilene-inspired topology that is also evaluated in Ref. [21]. This topology is based on Abilene Network used for higher education.The Abilene-inspired topology has 869 nodes and 877 links.

3.4.3 The BA Topology

Barabasi and Albert proposed a BA model to generate topologies having a power-law degree distribution [19]. We call the BA topology as the topology generated by BA model. The BA model is characterized by two features: Incremental Growth and Preferential Attachment, The model starts with a topology with a small number of nodes and works as follows:

Step1 Make an initial topology that has m_0 nodes.

Step2 Incremental Growth: Add a new node at each time step.

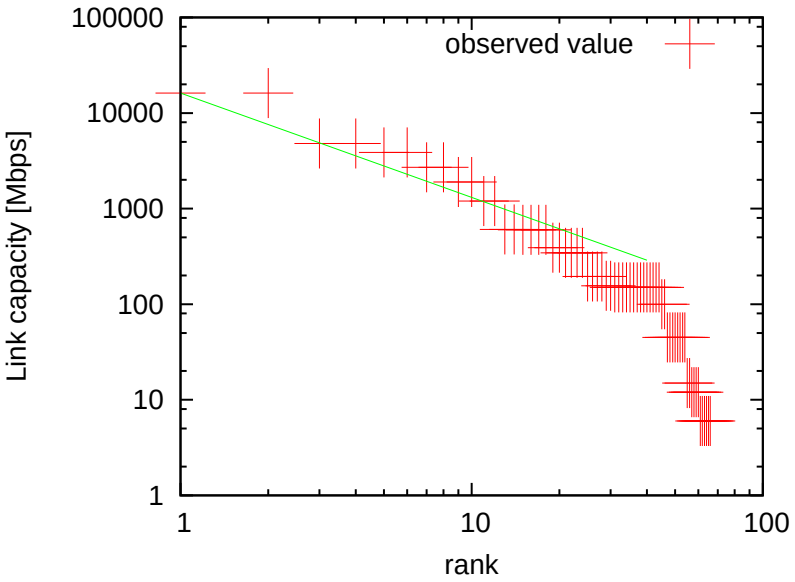


Figure 3.2: Link bandwidth distribution in the IIJ network

Step3 Preferential Attachment: Connect the new node to m different nodes chosen with probability Π:

$$\Pi(k_i) = \frac{k_i}{\sum_j k_j} \tag{3.1}$$

Where k_i is the degree of node i.

To follow this process, the topology with its degree distribution obeys the power-law is generated. In this chapter, we generate BA topology to be the same number of nodes and links in the Abilene-inspired topology. At first we use BA model with $m_0 = 1$, and generate 869 nodes and 868 links topology. Then, nine links are added based on rule of Step3. The topology with 869 nodes and 877 links is generated.

3.4.4 The FKP Topology

The FKP model proposed by Fabrikant et al. [20] revealed that the power-law property of degree distribution can still be obtained by minimizing “distance” metrics. This model does not use preferential attachment to add links, and instead uses minimization-based link attachment. More specifically, the FKP model works as follows. Each new node arrives at randomly in the Euclidean space $\{0, 1\}^2$. After arriving

at new node i, the FKP model calculates the following equation for each node, j, already existing in the network: $\beta \cdot w_{ij} + l_{0j}$, where w_{ij} is the Euclidean distance (i.e., physical distance) between nodes i and j, and l_{0j} is the hop-counts distance between node j and a pre-specified "root" node (node 0). β is a parameter that weights the importance of physical distance. If β has a lower value, each node tries to connect to higher degree nodes; $\beta = 0$ is an extreme scenario that creates a star-topology. If β has a higher value, each node tries to connect their nearest nodes. A topology with high a β is shown to behaves like an ER topology. The power-law property of the degree distribution appears at a moderate value of β value. Here, there are several hub-nodes in each region, and the hub-nodes form a power-law. However, it is mentioned that the number of node whose degree is one is much more than that of ISP topology [21] and the degree distribution is different from that of ISP topology [34].

In this chapter, we generate FKP topology to be the same number of nodes and links in the Abilene-inspired topology. For this purpose, we modify the FKP model; we first select nine nodes randomly before starting the topology generation and when each of these nine nodes is added, we generate two links between itself and a node satisfying the equation ($\beta \cdot w_{ij} + l_{0j}$), and another next node satisfying the equation. At last, a topology which has 869 nodes and 877 links is generated.

3.5 Performance Evaluation on Link Bandwidth Distribution

In Section 3.3, we show that the link bandwidth distribution of Japanese ISP obeys the power-law. For revealing the reason that link bandwidth distribution obeys power-law, we investigate network performance under the constraint of link bandwidth. In this chapter, we define the network performance as the amount of traffic that the topology can accommodate. For comparison purpose, we generate three distributions of link bandwidth; power-law, exponential, and identical where bandwidth of all links are identical. Then, the amount of traffic that the topology can accommodate for each distribution is compared.

3.5.1 Link Bandwidth Distributions

According to the observation in Section 3.3, we generate link bandwidth distribution that obeys power-law with exponent the -1. We first prepare m links with the bandwidth X [bps]. Then, we prepare $\alpha \times m$ links each of which has X/α link bandwidth. By setting $m \leftarrow \alpha \times m$ and $X \leftarrow X/\alpha$, we repeatedly prepare links.

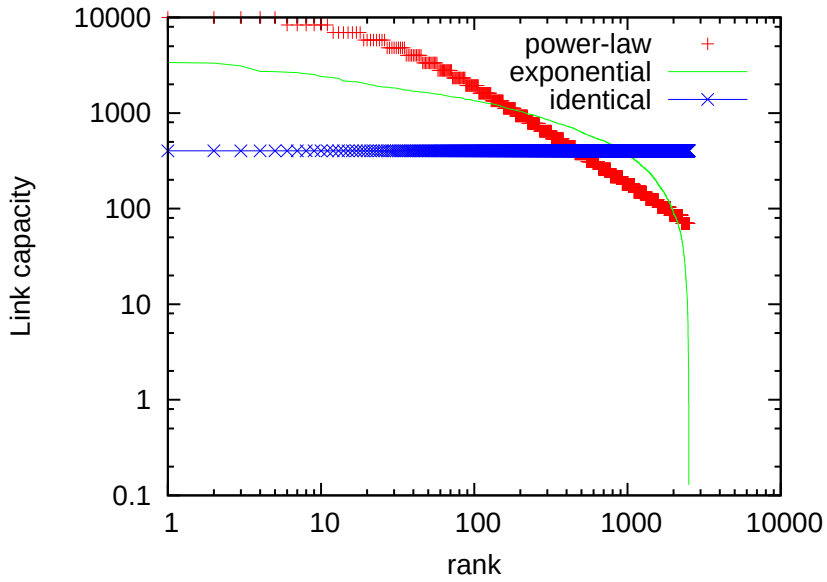


Figure 3.3: Normalized link bandwidth distributions

With this case, the exponent of link bandwidth distribution becomes −1.

For comparison purpose, we generate other two link bandwidth distributions that obey exponential and identical. Figure 3.3 shows obtained distributions. The vertical axis represents the link bandwidth and horizontal axis represents the rank of that link. Note that we normalize the exponential and the identical distributions such that the sum of link bandwidth in each distribution is equaled.

3.5.2 Assigning Link Bandwidth

We then assign link bandwidth to each link in the topology. In this chapter, we assume that the link bandwidth is assigned based on the edge betweenness centrality. More exactly, we assign link bandwidth in a descending-order to link with the edge betweenness centrality. The edge betweenness centrality is defined as the number of minimum hop paths that pass through the link. That is, the edge betweenness centrality represents the amount of contribution of a link to the communication in the network.

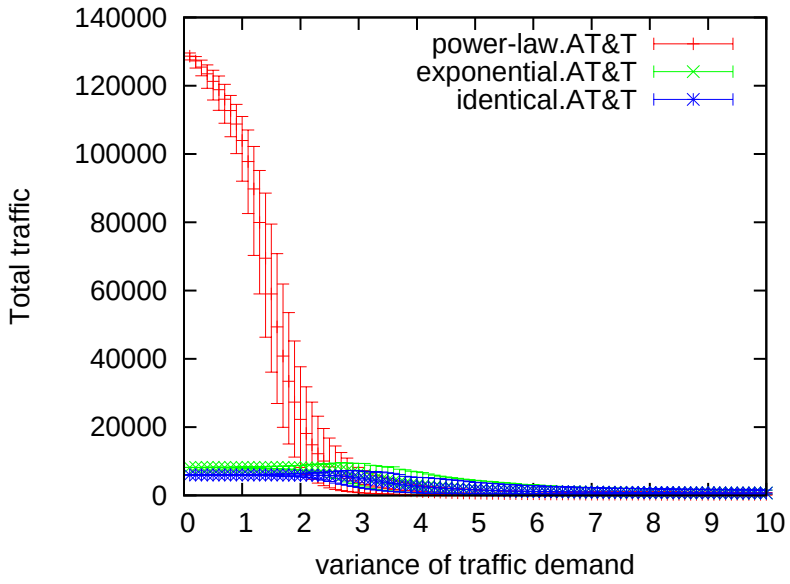


Figure 3.4: The amount of traffic that the AT&T topology can accommodate

3.5.3 Evaluation

We evaluate the amount of traffic that each topology can accommodate under the constraint of link bandwidth. The traffic demand is generated by log-normal distribution since the distribution is known to be suitable for modeling the traffic demand [68]. We change the variance of the log-normal distribution from 0.1 to 10.0 with the step 0.1. The traffic demand is generated for each variance and sum of the traffic demands is normalized to 1. A minimum hop routing is applied to select the route of the traffic. Note that if there are several paths that have the same number of hops, one of them is selected randomly. The result is shown in Figure 3.4-3.8. The vertical axis represents the amount of traffic that the topology can accommodate and horizontal axis represents the variance of traffic demand. In these figures, 95% confidence interval is also presented.

In ISP topologies (Figure 3.4 and Figure 3.5), when the variance of traffic demand is less than 2, the topology with the power-law link bandwidth distribution can accommodate more traffic than topologies

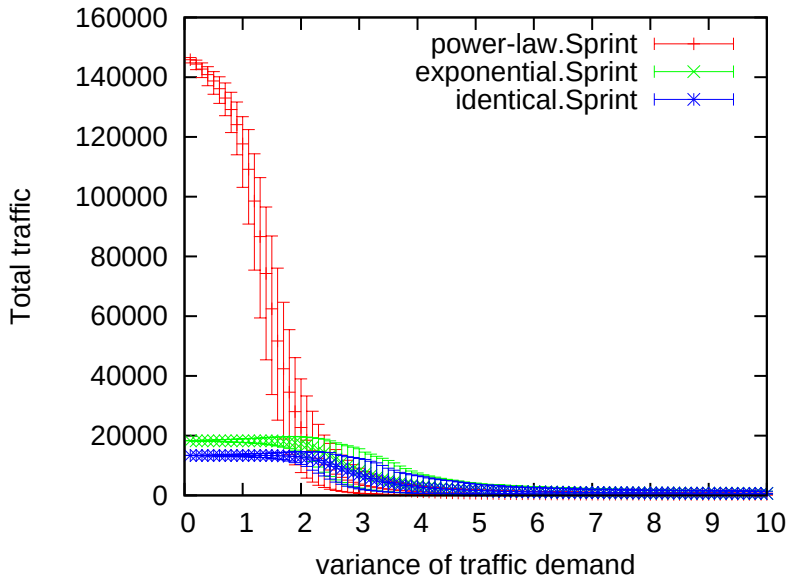


Figure 3.5: The amount of traffic that the Sprint topology can accommodate

with other distributions. In Ref. [68], the variance of log-normal distribution is nearly 1 for fitting the observed traffic in real ISPs. Thus, in the realistic situation, the power-law link bandwidth distribution makes the topology more efficient. The topology with the power-law distribution accommodates the largest traffic among three distributions even though the variance of traffic demand changes around 1.0. However, the advantage of the power-law distribution decreases and diminishes when the variance is too large since an elephant traffic appears in the traffic matrix. The same results are also observed in Figure 3.6, 3.7, 3.8. From the results of these figures, the power-law link bandwidth distribution can accommodates larger amount of the traffic than that of other distributions.

3.5.4 Effects of Node Processing Constraint

In Ref. [21], it is mentioned that the number of links connected to a router affects bandwidth of link. Figure 3.9 shows a relationship between node degree and node processing capacity. Due to the constraint of router technology, a product of node degree and its link bandwidth is bounded by the node processing capacity. That is, a router having few links can accommodate large capacity links, and a router having

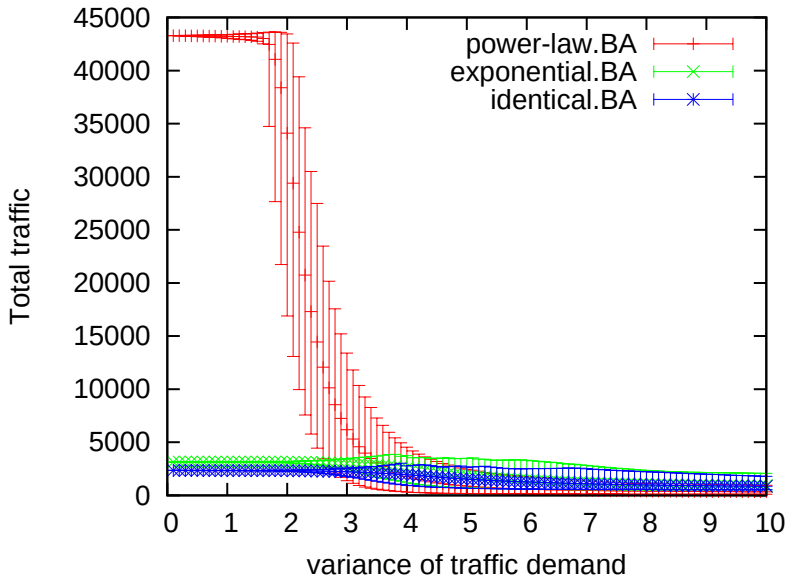


Figure 3.6: The amount of traffic that the BA topology can accommodate

small bandwidth of links can accommodate many links. This section examines the difference of network performance between two constraints for the link bandwidth and the node processing capacities. We apply the relation in Figure 3.9 to the Sprint topology and the Abilene-inspired topology. Figure 3.10 and Figure 3.11 show the amount of traffic that the topology can accommodate under the constraint of node processing capacities. For the comparison purpose, we again show the result of previous section in (a) of each figure and the result under the constraint of node processing capacities in (b) of each figure.

Comparison between (a) and (b) of Figure 3.10 and Figure 3.11 shows the difference of amount of the traffic that the topology can accommodate with each constraint; the link bandwidth and the node processing capacities. These figures have a common tendency that amount of the traffic that the topology can accommodate gets smaller as the variance of traffic demand becomes larger. In other three topologies (BA topology, FKP topology and the Abilene-inspired topology), the same tendency is observed but are not shown here for the lack of space. However, the amount of traffic that the topology can accommodate under the constraint of link bandwidth is strongly smaller than that under the constraint of node processing capacities. Thus, evaluation of network performance only the constraint of node processing capacities by

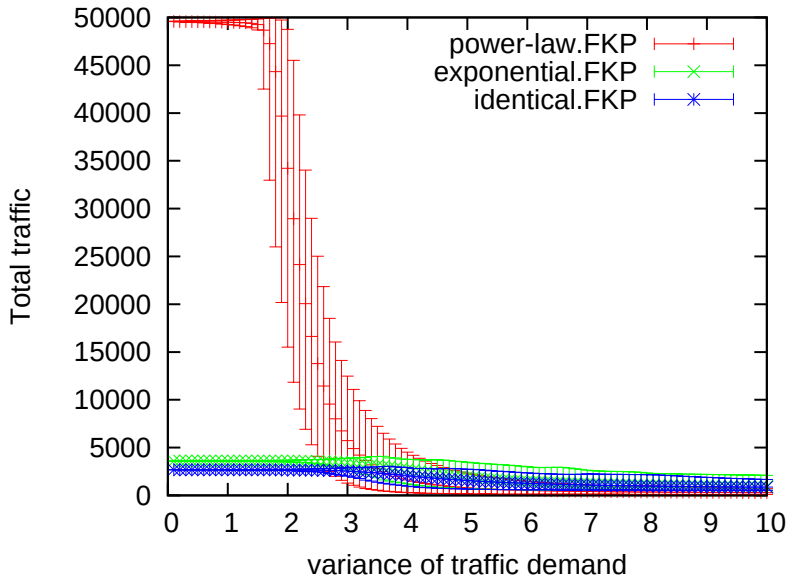


Figure 3.7: The amount of traffic that FKP topology can accommodate

its degree is inadequate for network evaluations.

3.6 Conclusion

Evaluation of the network control method needs proper model of the communication network. In this chapter, we focused on link bandwidth in ISP router-level topology. First, we investigate link bandwidth distribution in Japanese ISP backbone network. Results show that link bandwidth distribution in ISP obeys the power-law. Next, for revealing the reason that link bandwidth distribution obeys power-law, we compared network performance with three link bandwidth distributions; power-law, exponential and identical. Results show that power-law link bandwidth distribution can accommodate large amount of the traffic. Moreover, under the constraint of only node processing capacities by its degree, amount of the traffic that the network can accommodate is over estimated. Thus to take characteristic that link bandwidth distribution obeys power-law into the modeling of router-level topology is important to evaluate the network control method properly.

In this chapter, we assigned link bandwidth distribution to the AT&T topology and the Sprint topology.

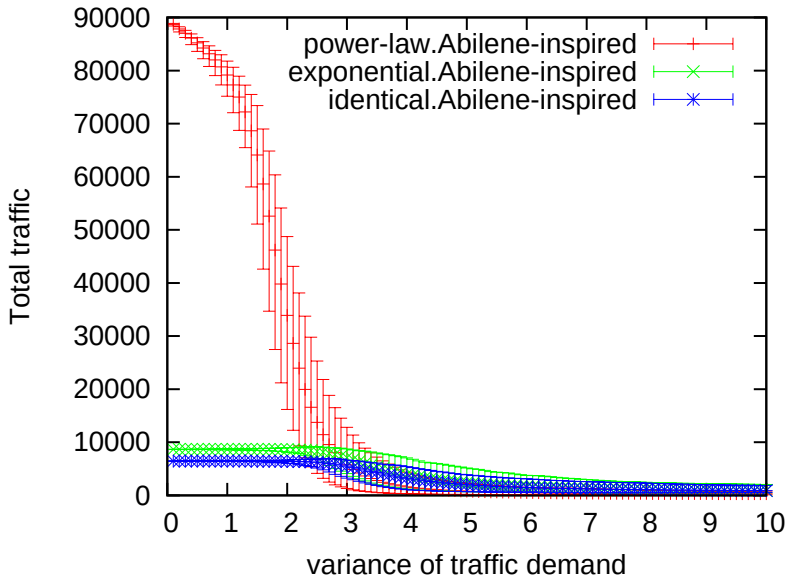


Figure 3.8: The amount of traffic that the Abilene-inspired topology can accommodate

From the comparison results of network performance under the constraint of link bandwidth, we showed validity of power-law link bandwidth distribution in router-level topology. However, we consider that there would be a close relationship between link bandwidth and structure of topology. For future work, we will reveal the relationship between link bandwidth and structure in ISP router-level topology, and then we develop the modeling method for the router-level topology.

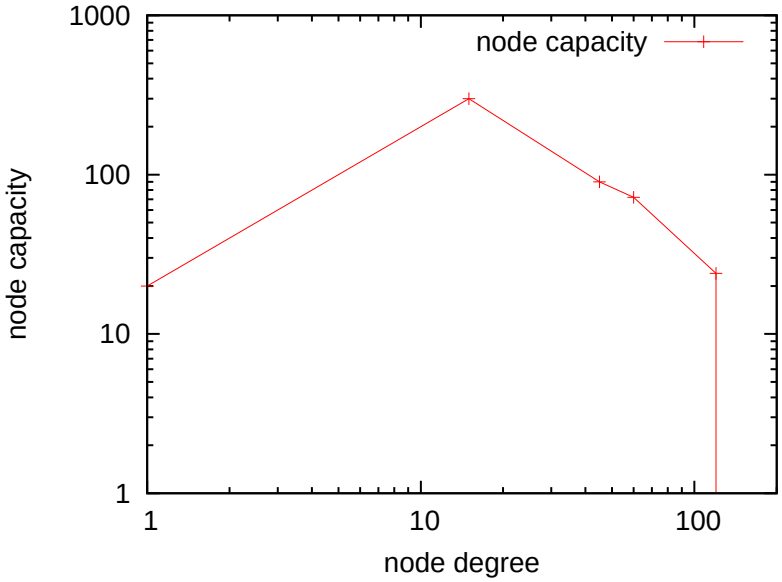
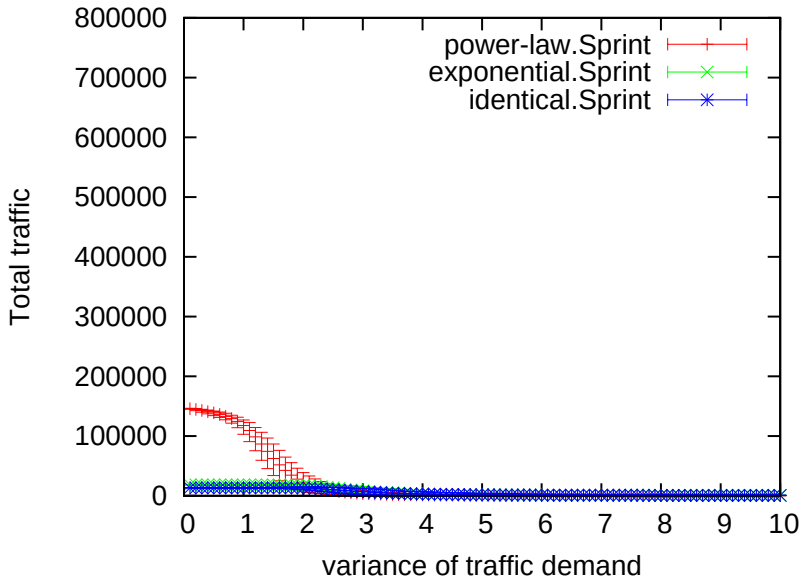
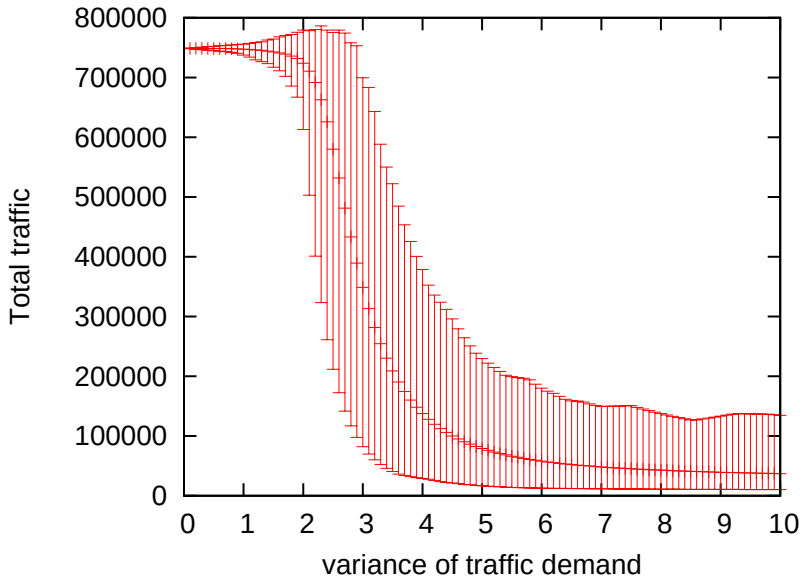


Figure 3.9: Relationship between node degree and node processing capacity

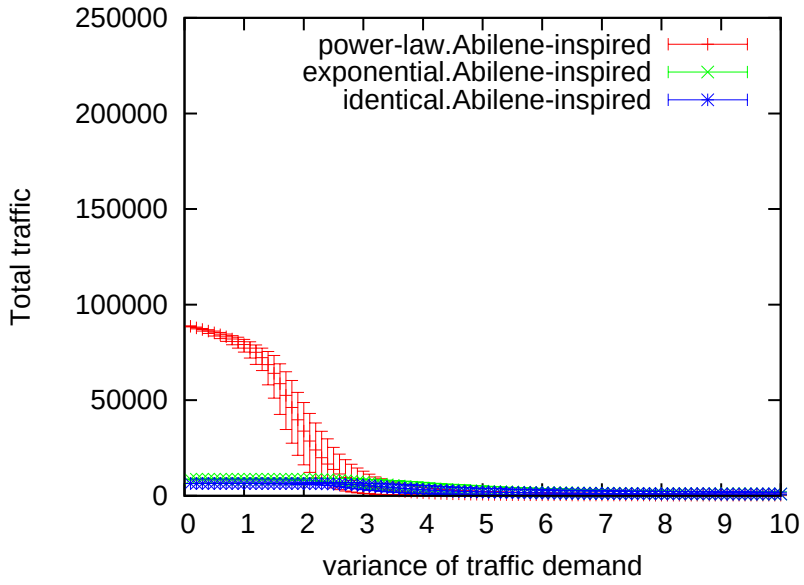


(a) The case of link bandwidth constraints

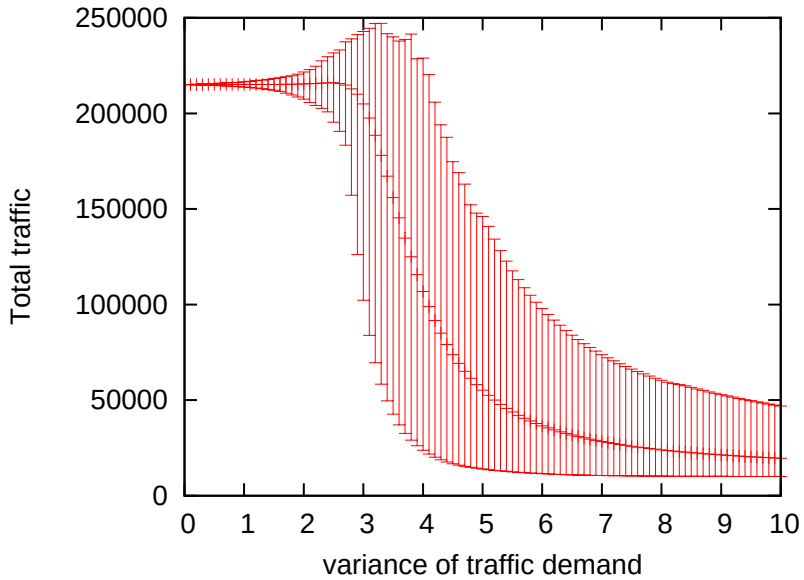


(b) The case of node processing capacity constraints

Figure 3.10: The network performance of the Sprint topology under constraints



(a) The case of link bandwidth constraints



(b) The case of node processing capacity constraints

Figure 3.11: The network performance of Abilene-inspired topology under constraints

Chapter 4

Modeling Link Bandwidth Distribution in ISP Router-level Topologies

4.1 Introduction

The Internet has been advanced through improvements to various network-related protocols based on functional partitioning. However, as new networking technologies are developed, designing the Internet according to such partitioning becomes increasingly difficult. For instance, conflicts between selfish behavior in overlay networks and traffic engineering in underlay networks lead to performance degradation [5]. Lately, dynamic interactions between network protocols have increased the complexity of Internet traffic behavior. Modeling router-level networks while considering the traffic dynamics induced through these interactions is critical to establishing new network control methods.

Previous research has revealed that Internet service provider (ISP) router-level topologies ("ISP topologies" hereinafter) have power-law degree distributions; that is, the probability that nodes with k links exist is proportional to $k^{-\gamma}$ (for constant γ). Barabási et al. [19] proposed the Barabási–Albert (BA) model as a method for generating a power-law topology based on two rules: incremental growth and preferential attachment. The BA model is widely employed in studies examining network performance in power-law topologies. However, the structural properties of actual ISP topologies are different from those of model-based topologies. Li et al. [21] presented several topologies with the same degree distribution but different

structures. They pointed out that the structural differences lead to changes in network performance, such as the traffic volume that a network can accommodate. Hence, we should not only focus on power-law characteristics when discussing the performance of information networks.

One of the more notable characteristics is the link bandwidth. We previously demonstrated that power-law bandwidth distributions in ISP topologies could accommodate large traffic volumes in 3. That result indicated that the bandwidth distribution affects the performance of communication protocols or routing strategies. Recently, the relations between node capacity or link bandwidth distributions and power-law topologies have been considered [28, 30, 66]. However, these studies utilized topologies generated by the BA model or its variants to investigate the network throughput performance in power-law topologies. Indeed, modeling of the link bandwidth distribution in ISP topologies is barely addressed.

ISP networks are designed to accommodate dynamic traffic. ISPs monitor the traffic, and forecast the peak demand in the near future. On the basis of this prediction, they then enhance the bandwidth of a bottleneck link, either periodically or occasionally [40]. Because bandwidth enhancement of a certain link can trigger unexpected traffic changes, new bottlenecks emerge at a variety of points. For example, bandwidth enhancement of access links can create new bottlenecks in core links [69]. The current bandwidth distributions of ISP topologies thus result from the accumulation of link bandwidth enhancements. As ISPs monitor the traffic demand as an indicator of required bandwidth enhancement, the resulting bandwidth distribution is affected by the traffic dynamics induced by a dynamic flow control, such as the transmission control protocol (TCP) [35–39]. In actuality, we revealed that the dynamic interaction between the TCP flow control and modularity structure of ISP topologies affects the queue-length fluctuation of routers in Chapter 3. In the present study, we investigate the evolution process of bandwidth distributions to accommodate TCP traffic in ISP topologies, and examine the resulting bandwidth distributions through simulations. Bandwidth distributions are, therefore, a consequence of periodic enhancements in the ISP topologies. The results indicate that bandwidth distributions of ISP and BA topologies follow a power law. We then propose a bandwidth allocation model by considering the TCP traffic dynamics, and show that our model generates a bandwidth distribution with similar properties to that obtained by periodic enhancement.

The remainder of this chapter is organized as follows. In Section 4.3, we outline the network model employed in the simulations and demonstrate its validity. Weintroduce related work in Section 4.2.Then, in

Section 4.4, we determine the bandwidth distribution and elucidate its properties in large-scale topologies. The proposed bandwidth allocation model is introduced in Section 4.5, and we verify that this model produces similar properties to those found in the simulation results. Finally, in Section 4.6, we conclude this chapter and discuss future work.

4.2 Related Work

Many researchers have formulated resource allocation methods for power-law topologies [28, 30, 66]. Zhang et al. [28] stated that node capacity allocation based on each node's betweenness centrality enhances the network capacity. Here, betweenness centrality is defined as the number of shortest paths that pass through a node or link. Xia et al. [66] assigned capacities to nodes by solving an optimization problem that maximized the system performance and the utilization ratio of the capacity. This method provides a similar capacity distribution to that when the allocation in BA topologies is in proportion to each node's betweenness centrality. Both of the above studies assumed that link bandwidth was unlimited; however, actual communication systems like the Internet have heterogeneous bandwidths. Our group demonstrated that differences in bandwidth distributions led to different network throughputs, and that power-law bandwidth distributions can accommodate larger amounts of traffic than other types of distributions in Chapter 3. Ling et al. [30] also discussed a bandwidth allocation method for BA topologies. To maximize the network throughput, they introduced an allocation method based on the product of the degree of both end nodes. However, previous research has evaluated network performance using Poisson, or flow-level, traffic. As mentioned in Chapter 2, the interaction between topological structure and end-to-end flow control impacts on traffic dynamics. Hence, the traffic dynamics induced by TCP flow control must be considered when modeling the Internet.

4.3 Modeling and Validation

4.3.1 Simulation Model

Packet Processing Model at Nodes

Each node maintains limited buffers for each outgoing link. When a packet arrives at its destination node, the packet is removed from the network. Otherwise, the packet is forwarded by the current node to the buffer of an outgoing link, which is connected to a node according to a minimum hop routing algorithm. Each outgoing link delivers C packets per unit time to the next nodes based on first-in-first-out and drop-tail queuing principles. In addition, we do not employ dynamic routing; instead, each packet traverses the shortest path, which is calculated a priori. When multiple shortest paths to the destination are found, the next node is determined by the packet's source node. We apply this static routing strategy because, according to [58], traffic fluctuations caused by TCP-like flow control appear on short timescales, such as those of the order of round trip times (RTTs).

Flow Control between End Hosts

In the simulations, the source and destination nodes are randomly selected, a pre-specified number of sessions are created between them, and each session arrives at a node pair according to the Poisson process. Each session always has data to send during the simulations; that is, once a session is generated, data is continually sent to destination nodes until the simulation ends. Of course, this model is unrealistic as only one session exists between each node pair. However, it is sufficient to capture the average behavior of TCP traffic.

In the model, source nodes control the number of DATA packets based on slow-start and congestion avoidance algorithms, which are basic flow-control functions of TCP. If the congestion window size (*cwnd*) is smaller than the slow-start threshold ($cwnd < ssthresh$), then the source node uses the slow-start algorithm. When the source node receives an acknowledgment (ACK) packet, it increases *cwnd* by one packet size ($= smss$, the segment size) and sends two new DATA packets to the destination node. If $cwnd > ssthresh$, the source node uses the congestion avoidance algorithm. In this case, when the source node receives an ACK packet, it increases *cwnd* by $1/cwnd$ and sends an adequate number of

DATA packets to the destination node. In our model, *cwnd* does not exceed a predetermined maximum size. If the source node does not receive an ACK packet within the retransmission timeout (RTO) period, it recognizes that serious congestion has occurred, and resends the lost DATA packet while reducing *cwnd* by *smss*. The RTO period is defined according to the estimated RTT, and is doubled for every timeout.

In addition, we use the fast retransmit and fast recovery algorithms outlined by RFC 2581 [59]. The source node utilizes the fast retransmit algorithm when it detects packet loss and light congestion through the arrival of three duplicate ACKs. When the source node receives the third ACK, it halves *cwmd* and resends the lost DATA packet. After retransmission, the source node then increases *cwmd* based on the fast recovery algorithm. The source node continues to increase *cwmd* by *smss* until it no longer receives the duplicate ACKs.

Bandwidth Allocation Algorithm

ISPs appear to monitor traffic conditions and offset any bandwidth shortages. Here, we assume that the current bandwidth distributions of ISP topologies are the result of accumulated periodic enhancements. Therefore, we allocate bandwidths according to the following algorithm.

Initial condition: Set the default bandwidth of all links to C_0.

Bandwidth increase: At each time step $t = 1, 2, \cdots$ in the simulation, find the bottleneck link with the largest packet drop rate and add $C_b(t)$ (the bottleneck link bandwidth on the tth step) to the bottleneck link bandwidth; thus, $C_b(t+1) = 2C_b(t)$. Furthermore, the buffer size of the bottleneck link is doubled, and the next step of the simulation is performed with the updated bandwidth distribution. If another bottleneck link is found, increase the bandwidth of the new bottleneck, and continue iteratively. Note that the total bandwidth increases with t.

By repeatedly adding extra bandwidth, we generate a bandwidth distribution that can accommodate the TCP traffic.

4.3.2 Effect of TCP Traffic on the Bandwidth Distribution

To show the effect of TCP traffic, let us examine the simulation results of the parking lot topology shown in Fig. 4.1, which has seven source nodes (S1, S2, $\cdots$, S7) sending packets to a single destination node (Dst).

Table 4.1: Simulation parameter values

Number of sessions	100,000
Maximum *cwnd*	64 kByte
DATA packet size	1.5 kByte
ACK packet size	40 Byte
Initial bandwidth (C_0)	96 Mbps
Initial buffer size	10 MByte
Simulation time	1 s

The number of sessions between the nodes is seven. We set the initial bandwidth $C_0 = 9.6$ Mbps and the initial buffer size to 1 MByte. As described in 4.3.1, the bandwidth and buffer size of the bottleneck link are then doubled during each time step. All other parameter values are as in Table 4.1. The bottleneck links at each time step are listed in Table 4.2. These results show the potential for links to become bottlenecks, even when they have a small number of sessions passing through them. The link connecting S7 and E became a bottleneck at $t = 3$, induced by the enhancement of the link connecting E and Dst at $t = 2$. Because the session from S7 to Dst has no competition, S7 sends the greatest number of packets to Dst. As a result, the traffic increase caused by TCP makes the link between S7 and E a bottleneck. Under the TCP model, source nodes dynamically change their sending rate to utilize the available bandwidth. As a result, packet drops also occur at links that conduct only a small number of sessions.

The results also indicate the possibility that a new bottleneck link could have a higher drop rate than the previous one due to dynamic flow control. For example, when $t = 1$, the link connecting D and E was found to be a bottleneck with a drop rate of 0.14. The bandwidth of this link was thus increased and, because of this enhancement, the link connecting E and Dst became a bottleneck with a higher drop rate of 0.18. In the same way, the maximum drop rate can increase due to this enhancement. When the bandwidth of a bottleneck link is improved, the source nodes of sessions passing through the bottleneck link send more packets to their destinations. However, if a low-bandwidth link that is unable to accommodate the increased traffic still exists on the path, a new bottleneck appears and the link starts to drop packets. Thus, the maximum drop rate increases.

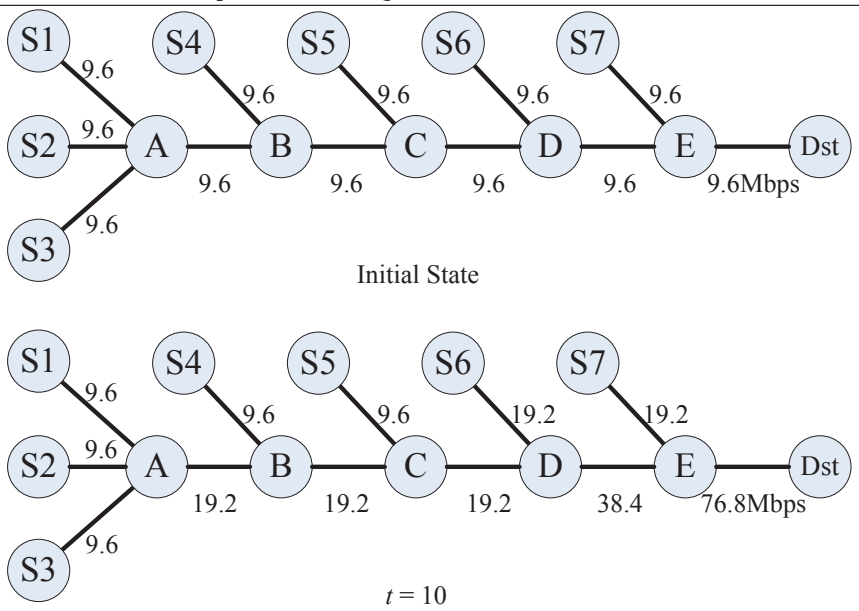


Figure 4.1: The Parking lot topology

4.3.3 Validation of the Bandwidth Allocation Algorithm with an ISP Topology

We evaluate the IIJ (Internet Initiative Japan) [67] topologies to confirm that our bandwidth allocation reproduces the characteristics of the observed bandwidth distribution. Reference [63] investigated the node processing capacities of a commercial ISP network in Japan. The author used disclosed information of the link bandwidth in IIJ. In the paper, the node processing capacity is defined as the sum of bandwidth of links connected to a node. We observe that the node processing capacity distribution obeys the power-law; its exponent is nearly -2.6. This result is interesting in that the power-law relationship again appears in the node processing capacity. However, one question arises—does the link bandwidth distribution follow a power-law? To answer this, we evaluate the bandwidth distribution of the IIJ topology [51]. Figure 4.2 shows the bandwidth distributions under IIJ. The x-axis represents the bandwidth and the y-axis represents the rank of that link. This shows that the exponent of the bandwidth distributions (for the 40 highest rankings) is nearly -1.

Figure 4.3 shows the bandwidth distributions obtained through these simulations. Here, the y-axis represents the link bandwidth and the x-axis represents the rank in descending order. We use the parameters

Table 4.2: Bottleneck links during simulation of parking lot topology

t	Bottleneck	Drop Rate
1	D-E	0.14
2	E-Dst	0.18
3	S7-E	0.21
4	C-D	0.13
5	E-Dst	0.19
6	B-C	0.22
7	D-E	0.18
8	S6-D	0.21
9	A-B	0.11
10	E-Dst	0.018

summarized in Table 4.1 for the simulations. The IIJ topology has 23 nodes and 29 bidirectional links. Note that the number of sessions in this case is 506 because of the upper bound on the number of node pairs. From this figure, we see that our bandwidth allocation algorithm reproduces a similar distribution in that it has a power-law decay with an exponent of around -1. This result shows the validity of our bandwidth allocation algorithm.

4.4 Effects of Bandwidth Allocation Considering TCP Traffic Dynamics in Large-scale Topologies

4.4.1 Evaluation of Bandwidth Distribution and Network Performance

The IIJ topology is one of many ISP topologies, and it is very small in scale. It is uncertain whether other large-scale ISP topologies have the same characteristics as the IIJ topology, and so it is not clear that the above results are applicable to ISP topologies in general. Therefore, we evaluate the bandwidth distributions of other large-scale ISP topologies. If not otherwise specified, we use the parameters summarized in Table 4.1 for these simulations.

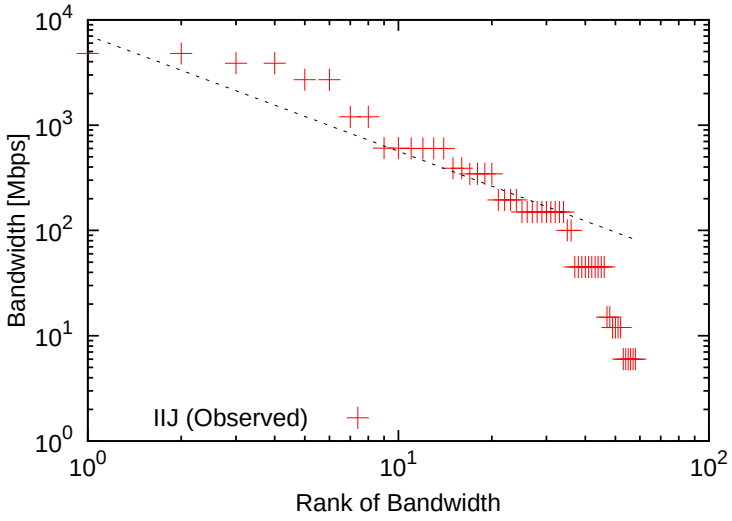


Figure 4.2: Observed bandwidth distribution of the IIJ topology: The slope of the line is -1

We simulate several ISP topologies and measure their performance using Rocketfuel [32]. BA topologies are then generated by the BA model such that the number of nodes and links are the same as in the corresponding ISP topologies. In [21] and [57], it was shown that the BA model is insufficient for modeling ISP topologies. However, BA topologies are still utilized in the current paper, because one of our main purposes is to clarify the fundamental structure that characterizes bandwidth distributions in ISP topologies. Rocketfuel is a topology measurement tool based on `traceroute`. Note that actual ISP networks may have some routers that are invisible to `traceroute`. We believe that such routers are used for backup purposes, because they are placed on paths that are inactive for Internet users. Hence, Rocketfuel may not cover the entirety of the ISP topologies. However, even when Rocketfuel misses some of the routers, adequate investigation of traffic dynamics can be achieved because the active paths between routers have been included.

First of all, we compare the AT&T topology with a corresponding BA topology to identify the fundamental structure that characterizes the bandwidth distribution of large-scale ISP topologies. These two networks have 523 nodes and $1,304$ bidirectional links. We then confirm and discuss the characteristics of the AT&T topology by comparing these with the results for other ISP topologies.

Figure 4.4 shows the bandwidth distributions of the two topologies at $t = 500$ and $t = 1,500$. Both

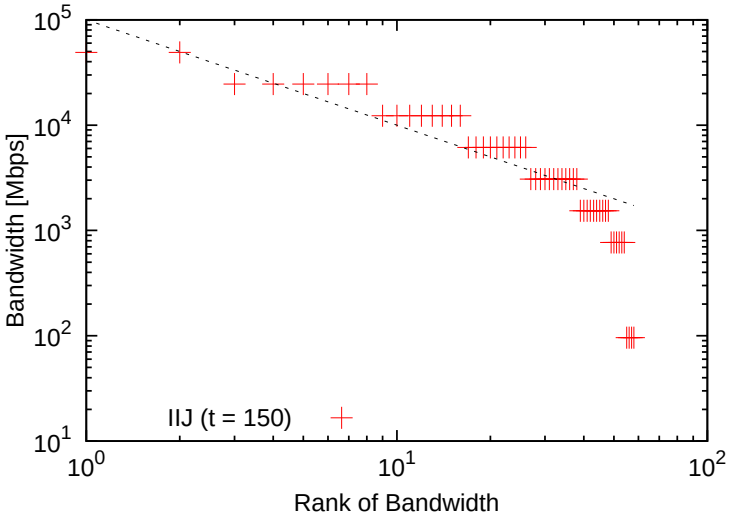


Figure 4.3: Bandwidth distributions of the IIJ topology: The slope of the dashed line is −1

topologies are shown to have a power-law bandwidth distribution. In particular, as t increases, links with high bandwidth appear, and the bandwidth distribution of the AT&T topology approximates a power-law distribution with an exponent of around −1. The power-law characteristic of the distribution in the AT&T topology is more apparent than that of the IIJ topology. We regard this difference is caused by the size of each topology. Unlike the IIJ topology, the AT&T topology has many nodes and requires a large number of enhancements in order to accommodate a large amount of traffic between end hosts. Because of a large number of enhancements, the bandwidth distribution of the AT&T topology shows the power-law characteristic more clearly. In addition, the exponent of the BA topology is about −0.6. Because the two networks have the same number of nodes and links, the difference between the bandwidth distributions is considered to be induced by their different topological structures.

Figure 4.5 shows the correlation between the bandwidth and the betweenness centrality rank of each topology. The x-axis represents the number of passing sessions of each link and the y-axis represents the link's bandwidth. From Figs. 4.4 and 4.5, we can see that the AT&T topology has a greater number of high-bandwidth links than the BA topology. One reason for this difference is the large number of sessions traversing core links in the AT&T topology. Specifically, additional bandwidth tends to be allocated to core links in order to cope with a large number of packets, and so frequently congested links are iteratively

enhanced. Indeed, the AT&T topology has a smaller number of enhanced links than the BA topology—on completion of the $1,500$th step, 827 links are enhanced in the AT&T topology compared to $1,060$ in the BA topology.

The number of passing sessions for a link is approximately proportional to its betweenness centrality when source and destination node pairs are randomly selected. From this point of view, a bandwidth allocation method based on the betweenness centrality of links seems appropriate for accommodating TCP traffic. When dynamic flow control is considered, however, centrality-based allocation may not be suitable for these topologies. Figure 4.5 shows that high bandwidth is also allocated in both topologies for links with a small number of passing sessions. As described in Section 4.3.2, the flow control functionality of TCP causes packet loss at links that conduct only a small number of sessions. These results indicate that bandwidth allocation based on the number of passing sessions may be inappropriate when TCP flow control is considered.

In addition, even though additional bandwidth is allocated to a given link, the maximum drop rate may increase because of the dynamic flow control. Figure 4.6 shows the transition of the maximum drop rate. The x-axis represents time t and the y-axis represents the maximum drop rate at each step. As t increases, the maximum drop rate does not always decrease, in spite of the bandwidth growth. Even if the overall trend is decreasing, the drop rate intermittently increases. This also occurs because of the dynamic flow control of TCP (see Section 4.3.2).

4.4.2 Effect of Topological Structure on Performance

In this subsection, we discuss the network performance induced by the topological structure. Figure 4.7(a) shows the network throughput (y-axis) at each time step (x-axis) for the AT&T and BA topologies. We define the network throughput as the total number of DATA packets received by a destination node. For a small number of enhancements, the throughput increases to a greater extent in the AT&T topology than in the BA topology.

However, the AT&T topology increase comes at a higher cost. Figure 4.7(b) shows the correlation between the throughput (y-axis) and the total amount of bandwidth (x-axis). In the simulation, the initial total bandwidth is about 250 Gbps; that is, $1,304$ (links) $\times$ 2 (directions) $\times$ 96 (Mbps). Increases in

throughput for the BA topology are sharper than those for the AT&T topology for the same increase in total bandwidth. Therefore, the BA topology requires a lower cost to achieve a high throughput.

This result is attributed to the difference between the topological structures of the networks. We previously compared the structural differences of the AT&T topology and the BA topology [34]. Our results indicated that the design principles used in networks greatly affect the modularity structure of the ISP topology: the design principles determine the node functionality, which in turn determines the node connectivity. To confirm that result, let us now divide each network into multiple modules and examine the node connectivity. Observe from the schematic in Fig. 4.8(b) that the BA topology has many connector hub nodes that connect modules through a number of inter-module links. Conversely, the AT&T topology has numerous provincial hub nodes that connect the nodes within each module (Fig. 4.8(a)). The AT&T topology thus has a highly modular structure, with a dense concentration of node connections within a module and a sparsity of node connections between modules. We have observed that other ISP topologies have similar, highly modular structures (in Chapter 2).

Packets forwarded between different modules in the AT&T topology pass through a number of inter-module links. If link bandwidths are uniformly allocated, these inter-module links tend to be congested and packet drop occurs. In fact, inter-module packets that traverse certain inter-module links rarely reach their destination nodes. Source nodes of inter-module sessions send packets with low frequency, and so the throughput of these sessions is low. In contrast, intra-module sessions have high throughput because packets forwarded between two nodes in the same module arrive at their destinations easily. Hence, the throughput of the AT&T topology is higher than that of the BA topology in the initial state. As t increases, however, the inter-module links in the AT&T topology are iteratively allocated additional bandwidth. Owing to this enhancement, inter-module packets reach their destinations more easily, and the source nodes of inter-module sessions send a greater number of packets. Consequently, the source nodes of intra-module sessions have to reduce their sending rate because of the increased competition, resulting in a decrease in their throughput. Because the growth in throughput of inter-module sessions is relatively moderate, the total throughput increases only slightly.

Conversely, throughput increases substantially in the BA topology, and surpasses that in the AT&T topology when the total bandwidth reaches about 300 Gbps. The BA topology has many connector hub nodes that ensure that modules are only connected over short distances. When an inter-module link of a

connector hub node is allocated additional bandwidth, the many sessions passing through the node also attain higher throughput. In addition, the BA topology has numerous inter-module links. By adding any unused extra bandwidth to these inter-module links, the total throughput is increased.

4.4.3 Relations between Bandwidth Distribution and ISP Topological Structure

We next conduct simulations using other ISP topologies. Here, we use the Sprint (owned by Sprint Nextel Corporation) and Telstra Corporation Limited topologies. The Sprint topology has 467 nodes and $1,280$ bidirectional links, and the Telstra topology has 296 nodes and 594 bidirectional links. Simulation parameter values are as in Table 4.1, except that the number of sessions is $87,320$ because of the upper bound on the number of node pairs. Figure 4.9 shows the bandwidth distributions of the two ISP topologies. These also have power-law degree distributions, although the exponents of the bandwidth distributions for all three ISP topologies differ from one another. This result indicates that bandwidth allocation through consideration of TCP traffic dynamics may lead to power-law distributions, but the exponents of the distributions vary depending on the topological structure.

4.5 Modeling Bandwidth Distribution with Consideration of TCP Traffic Dynamics

In the previous section, we determined the bandwidth distributions accommodating TCP traffic dynamics using simulations with a large number of iterations. However, considerable time was needed to perform each simulation. In this section, we propose a model for generating bandwidth distributions with similar properties to those obtained through simulations.

The basic concept of this model is to add a weight to each link, and then repeatedly enhance the bandwidth of the bottleneck link. The initial weight of each link is defined as the number of passing sessions, and the link with the largest weight is regarded as the bottleneck. When a link gains additional bandwidth, its weight is reduced and the surplus is distributed to its peripheral links. To account for the

effects of TCP flow control, the surplus weight is divided between links with larger hop counts to the connector node than those at the ends of the enhanced link. Explicitly, the weight is distributed between the links close to edge nodes. If a link located near other modules is enhanced, the end hosts at edge nodes in the module send more traffic, because sessions between intra-module links rarely compete.

4.5.1 Modeling Bandwidth Allocation and Evaluation

We allocate bandwidth according to the following process.

Initial condition:

- Uniformly set the default bandwidth of all links to C_0. Set the default weight of each link, which is defined as the number of sessions that pass through the link in both directions.
- Set the hierarchical level of each node, which is determined by the hop count to the nearest connector node with inter-module links. The level of a connector node is 0. The levels of other nodes then increase with increasing distance from the nearest connector node.

Bandwidth addition:

1. Add extra bandwidth to the bottleneck link with the largest weight. We denote the weight of the bottleneck link at the tth step as $w_b(t)$, and the end nodes of the bottleneck link as i and j. To enhance the bottleneck link, double its bandwidth (i.e., $C_b(t+1) = 2C_b(t)$), and then reduce the weight of the link by $\alpha w_b(t)$ (i.e., $w_b(t+1) = (1-\alpha)w_b(t)$).
2. Uniformly distribute $\alpha w_b(t)$ between the links connecting node i and its neighbors, except for the links that connect to nodes above level i. The neighboring links of j are selected in the same manner as for i. We denote the set of target links as S_1. Thus, all target links in S_1 gain the additional weight $\alpha(1-\alpha)w_b(t)/n_1$, where n_1 is the number of links contained in S_1.
3. Distribute weight to the neighbors of each link $l_s \in S_1$, except for links with greater bandwidth than $C_b(t+1)$ and those connecting to nodes with the same or a higher level than both ends of l_s. We denote the set of target links as S_2. Apart from links already included in S_1, each link contained in S_2 gains the additional weight $\alpha^2 w_b(t)/(n_1 \cdot n_2)$, where n_2 is the total number of links in S_2. Note that n_2 varies with l_s.

4. Repeat the above steps t times.

A schematic of the proposed model is shown in Fig. 4.10. By performing a number of iterations, we obtain a similar bandwidth distribution to the simulation results for the AT&T topology. Here, as in Section 4.4.2, we divide the topology into several modules by Newman's method [15]. A bandwidth distribution generated by our model is shown in Fig. 4.11. This distribution is highly similar to that resulting from the simulation. In particular, the exponent of this distribution is approximately -1. To construct this distribution, we first set up $100,000$ sessions between randomly chosen node pairs, following Section 4.3.1. Then, with α set to 0.5, we perform iterations while the total bandwidth is less than 1.25 Tbps.

To establish that our model produces a bandwidth distribution with similar properties to that generated through the simulations, we compare their throughput distributions. Figure 4.12 shows the throughput of three types of bandwidth distributions. The simulated and modeled distributions are created by repeating bandwidth addition while the total bandwidth is less than 1.25 Tbps. We also evaluate the bandwidth distribution constructed through allocation based on the number of flows traversing each link (termed the "flow-proportional" allocation). The flow-based bandwidth distribution is, of course, unrealistic, but we utilize it for comparative purposes. The flow-proportional bandwidth of link l is defined as:

$$B_l = \frac{f_l}{\sum_l f_l} \times C_s, \tag{4.1}$$

where $C_s = 1.25$ Tbps is the total bandwidth and f_l is the number of flows that pass through l. The flow-based distribution is simulated using the parameter values in Table 4.1. Each session in Fig. 4.12 is classified based on the hop count between its source and destination nodes. The trends in throughput distributions obtained by the modeled bandwidth distribution are somewhat comparable to those of the distributions obtained through the simulations with TCP traffic, especially in the 2–7 hop range. We have also verified that our model produces comparable results for the Sprint and the Telstra topologies (not shown). Hence, our proposed model can explain the effects of complex TCP traffic dynamics on link bandwidths.

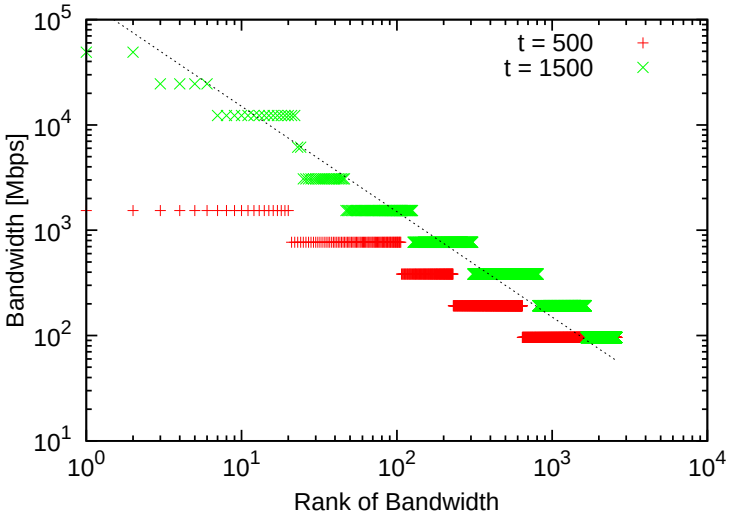
4.5.2 Modeling Bandwidth Distribution with a Simple Allocation Approach

In the previous subsection, we showed that our bandwidth allocation model produces similar properties to those generated through simulations. However, these results raise an obvious question: What is the difference when bandwidth is allocated according to the number of passing sessions for each link? To answer this, we evaluate a simple bandwidth allocation. That is, we find the bottleneck link with the minimum flow rate based on the assumption that bandwidth is shared evenly among all sessions using that link. As is the case in other simulations, we double the bandwidth of the bottleneck link. By repeatedly finding and enhancing the bottleneck link in the AT&T topology, we obtain the bandwidth distribution shown in Fig. 4.13 (denoted as "fair-share based"). The number of iterations is $1,500$. Unlike the model in the previous subsection, bandwidth allocation based on the flow rate of each link does not produce the characteristics obtained through simulations, as the links with high bandwidth do not appear in the bandwidth distribution generated by the simple bandwidth allocation. That is, the simple allocation scheme based only on the flow rate of each link is inadequate.

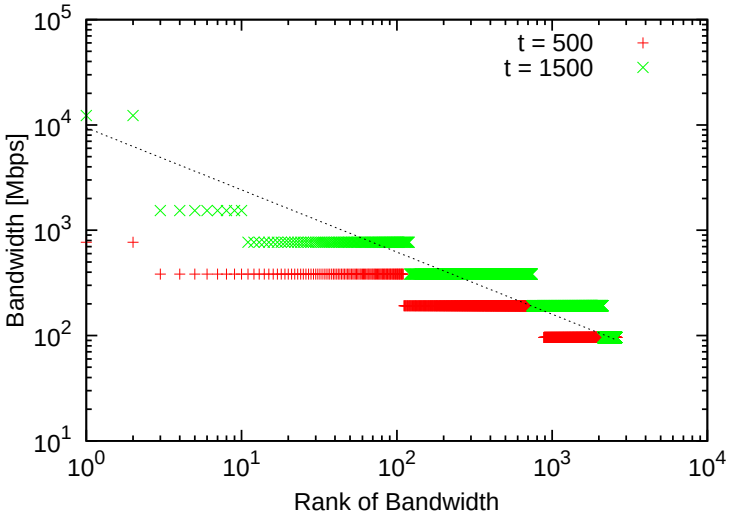
4.6 Conclusion

In this chapter, we evaluated the bandwidth distributions obtained by periodic bandwidth enhancements to accommodate TCP traffic in ISP topologies. As a consequence of periodic bandwidth enhancement, we revealed that the bandwidth distributions of these topologies follow a power-law whose exponent varies according to the topology. A comparison of simulation results for the AT&T topology and the BA topology showed that the highly modular structure of the AT&T topology characterizes its bandwidth distribution. The AT&T topology has a greater number of high-bandwidth links due to its highly congested inter-module links. Our results indicate the possibility that a power-law characteristic is induced in the bandwidth distribution of ISP topologies by the periodic bandwidth enhancement to accommodate dynamic TCP traffic. Hence, we have proposed a bandwidth allocation model that accounts for TCP traffic dynamics. By considering the impact of a given link's bandwidth enhancement on its peripheral links, our model generated a bandwidth distribution with similar properties to the distribution obtained by TCP traffic simulations.

In future work, we will explore a network design method that includes the interaction between TCP flow control and the structural properties of topologies.

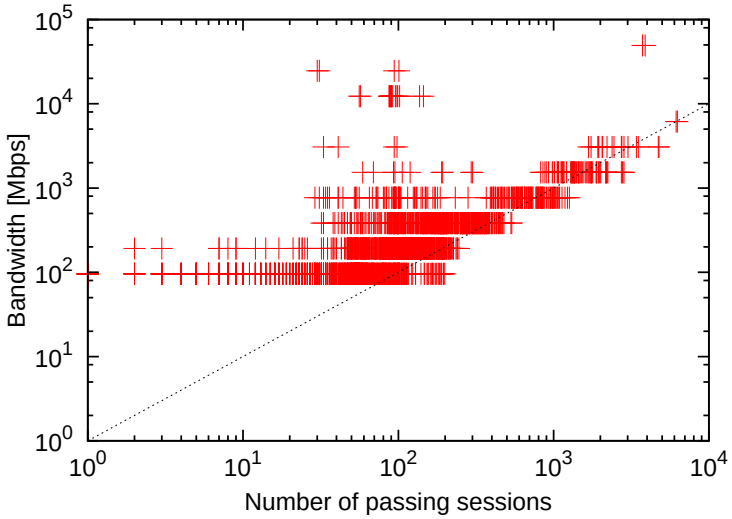


(a) The AT&T topology: The dashed line has a slope of −1

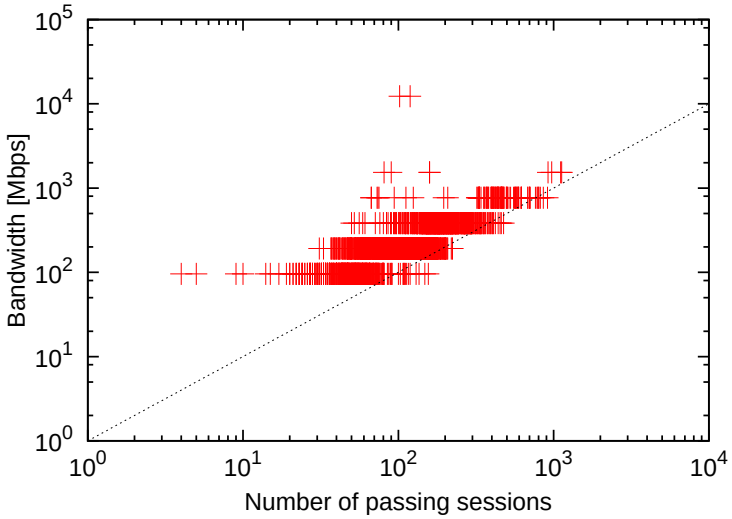


(b) The BA topology: The dashed line has a slope of −0.59

Figure 4.4: Bandwidth distribution of each topology



(a) The AT&T topology ($t = 1,500$)



(b) The BA topology ($t = 1,500$)

Figure 4.5: Correlation between bandwidth and number of passing sessions: The slope of the dashed line is 1

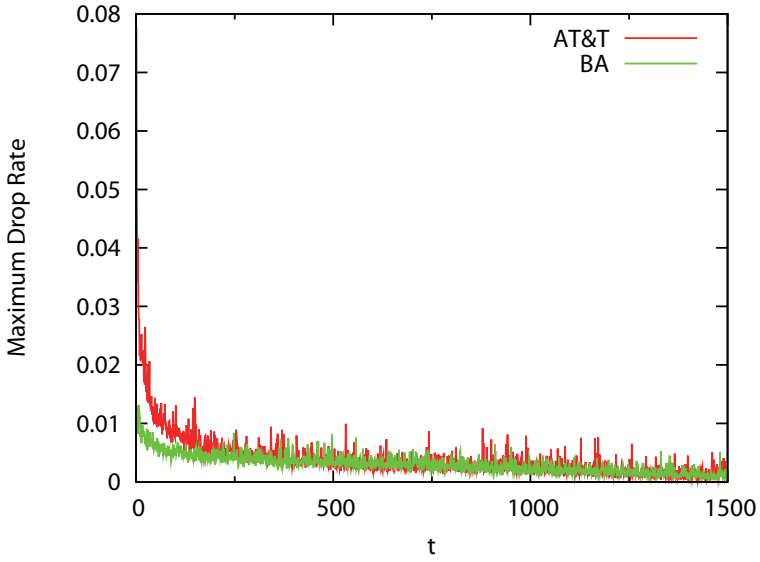
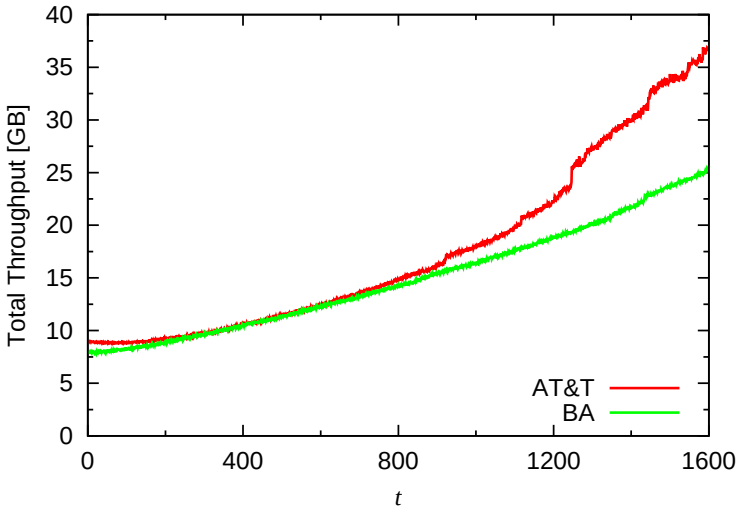
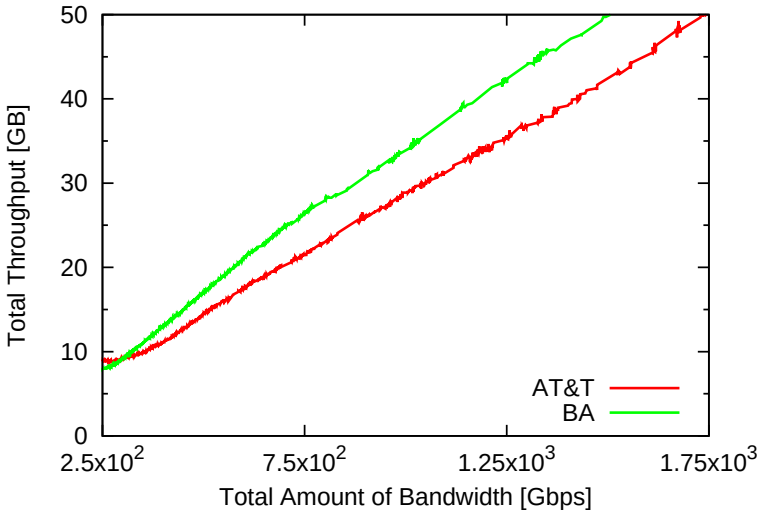


Figure 4.6: Transition of maximum drop rate

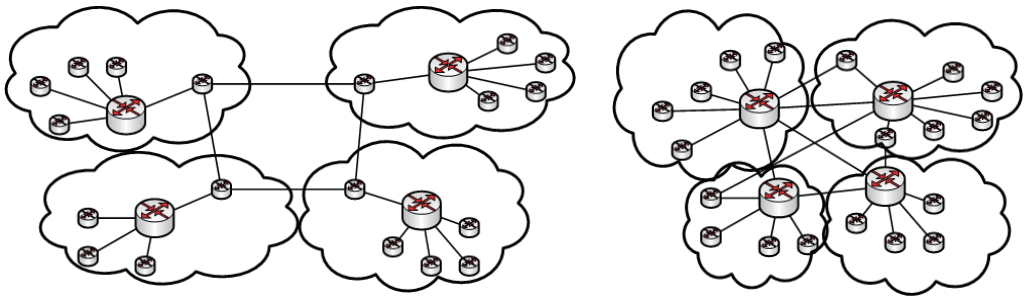


(a) Correlation between throughput and t



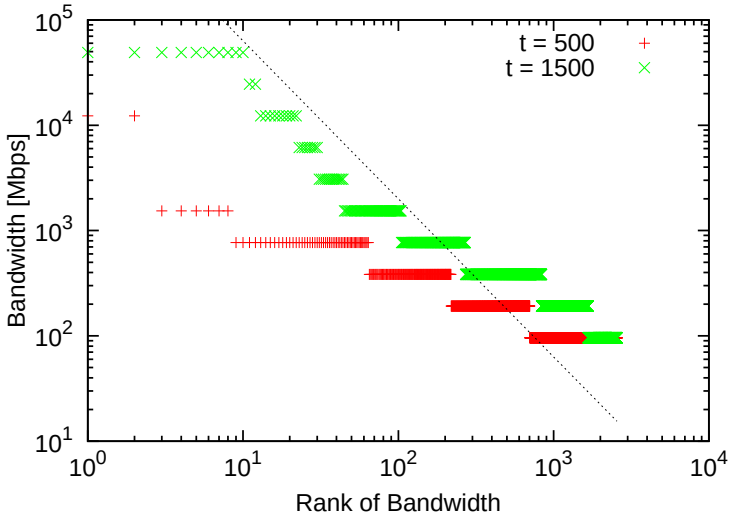
(b) Correlation between throughput and total bandwidth

Figure 4.7: Transition of network throughput

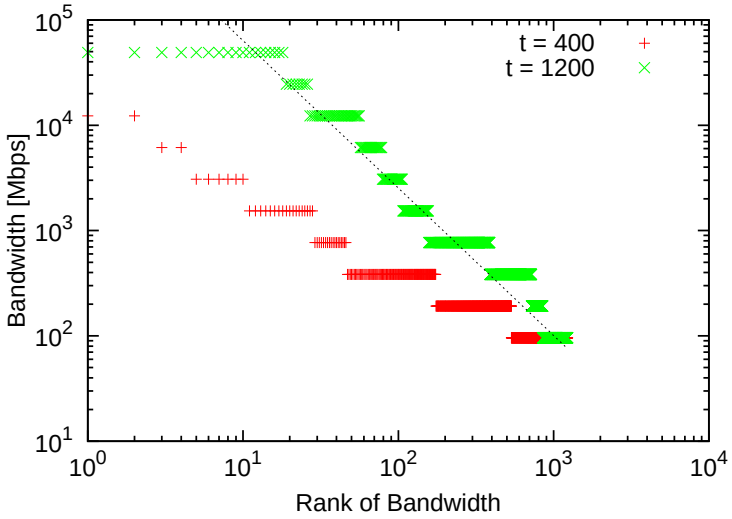


(a) The AT&T topology (b) The BA topology

Figure 4.8: Schematic of the structure of each topology



(a) The Sprint topology: The slope of the dashed line is −1.5



(b) The Telstra topology: The slope of the dashed line is −1.4

Figure 4.9: Bandwidth distributions of ISP topologies

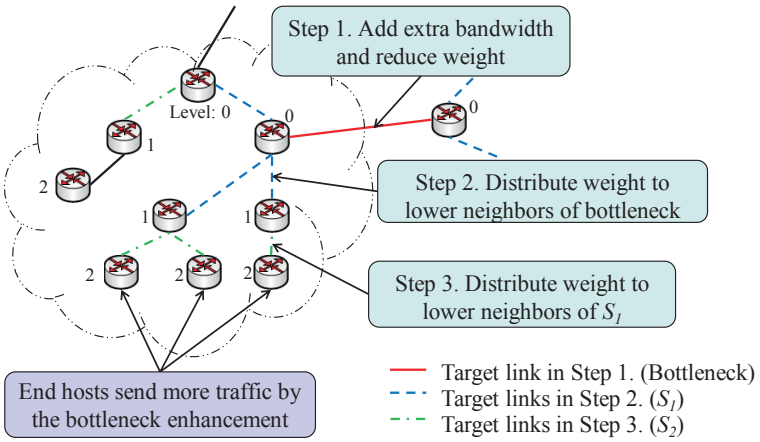


Figure 4.10: Schematic of proposed model

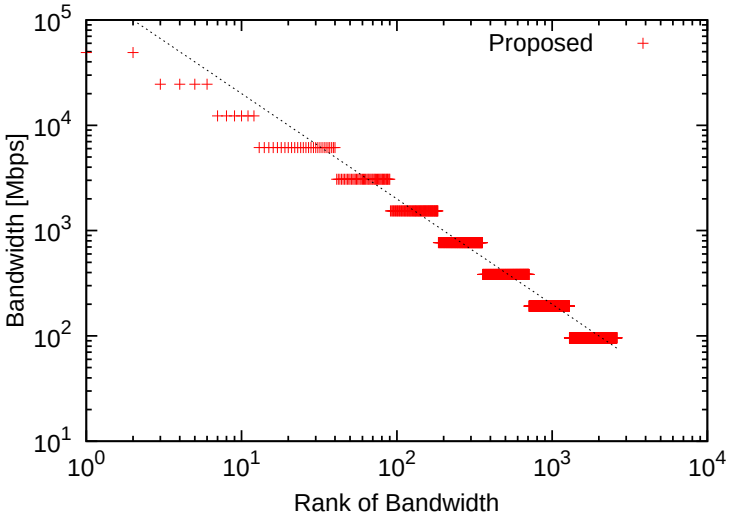


Figure 4.11: Bandwidth distribution of the AT&T topology generated by the proposed model ($\alpha = 0.5$, $t = 1,346$): The slope of the dashed line is -1 and the total bandwidth is about 1.25 Tbps

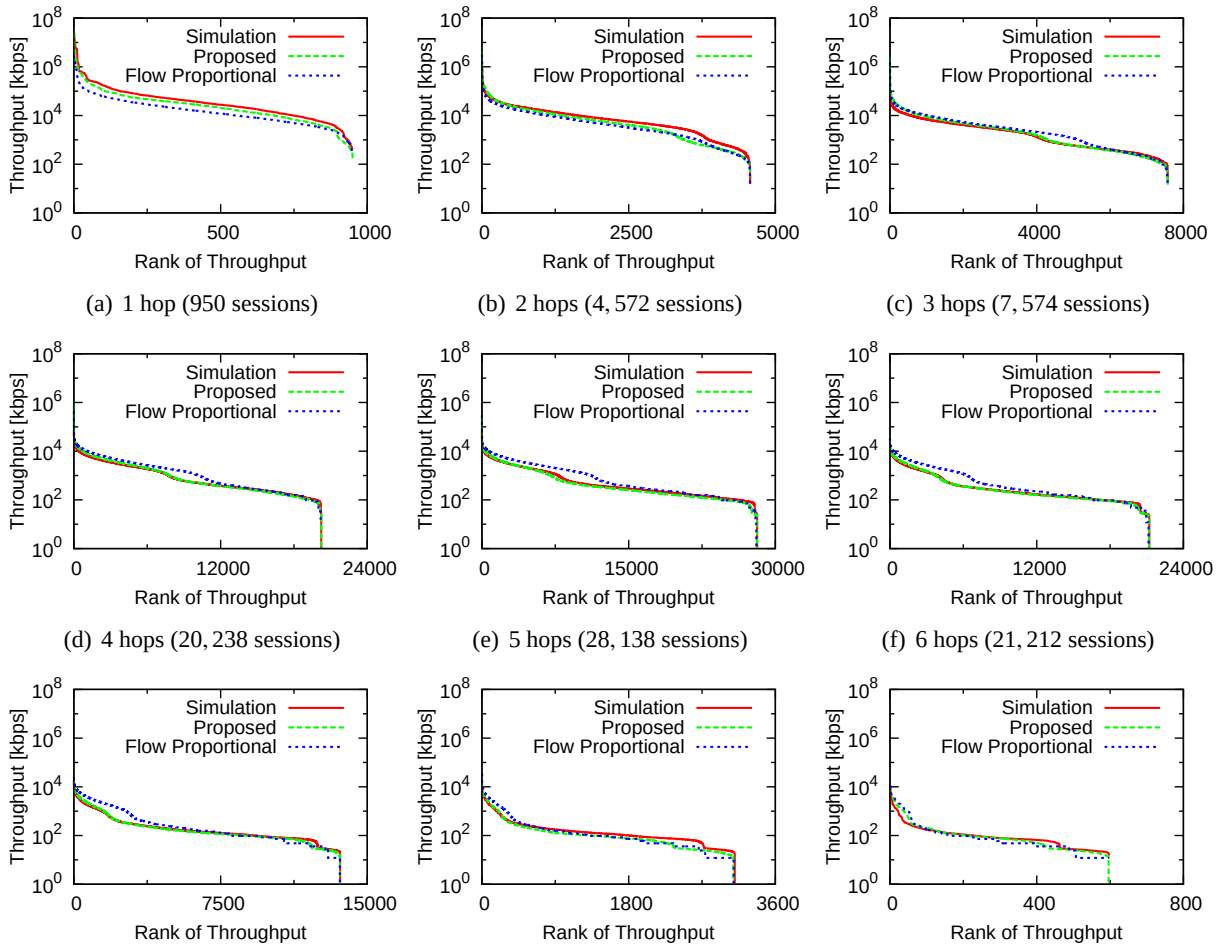


Figure 4.12: Throughput comparison for the AT&T topology: The comparison is between the simulation result, the proposed model, and the flow-proportional allocation. Each session is distinguished by the hop count between its source and destination nodes.

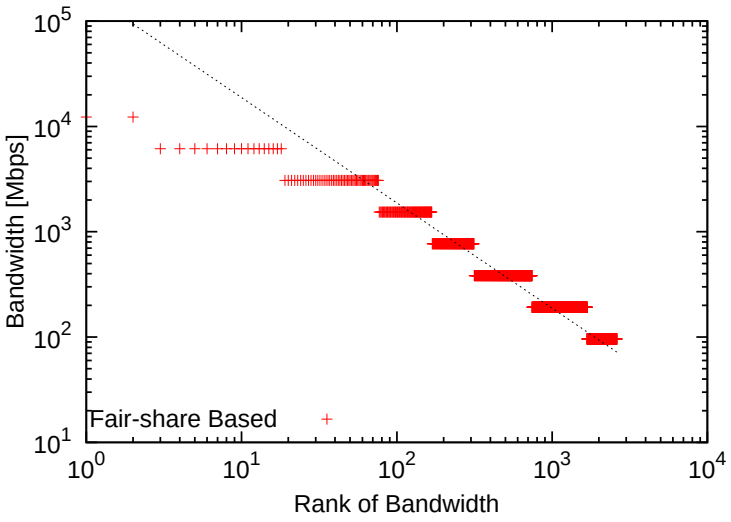


Figure 4.13: Bandwidth distribution obtained by simple bandwidth allocation

Chapter 5

Design of Interconnected Networks between ISP Topologies

5.1 Introduction

The Internet is composed by connected a large number of autonomous systems (ASs), such as ISPs. Nowadays, Internet traffic volume keeps increasing. According to Cisco VNI forecast [1], the total volume of global IP traffic is expected to grow from a volume of 31 Exabytes per month in 2011 to a volume of 110 Exabytes per month in 2016. The authors of Ref. [70] predicted that total Internet inter-domain traffic in the mid of 2012 at around 120Tbps, using findings from [41] and traffic growth forecast in Ref. [71]. ISPs should consider how to accommodate external traffic from neighboring ISPs. In the past, ISPs have constructed peering and transit connections with other ISPs at their own policies in order to increase their profit. However, communication efficiency obtained by the discretion of each ISP does not lead to globally optimized efficiency. Therefore, interconnection topology between ISPs is important to design networks.

In previous studies, profit distribution mechanism between ISPs is proposed to obtain a better engineered Internet from global perspective [72]. However, this mechanism includes global optimal routing strategies on the assumption that traffic and topology information gathered in router-level topologies and AS level topologies are applicable. Furthermore, even if all required information is available, computational complexity increases in large-scale network. Unlike this study, we discuss about how to design

interconnection topology between ISPs with graph-theoretical approach.

Recently, the concept of "networks of networks" formed by subnetworks have been concerned [11–13]. We call a network constructed by some subnetworks "interconnected network" and call a topology composed by some links connecting between subnetworks "interconnection topology." The authors introduced "inter-similarity" as a property measure of interconnected networks. Inter-similarity is defined by the two coefficients, inter degree-degree coefficient (IDDC) and inter clustering coefficient (ICC). They claimed that interconnected networks with high inter-similarity, where both IDDC and ICC value are high, are robust against cascade failure. In these studies, it is assumed that all nodes in each subnetwork have only one link to a certain node in the other subnetwork. By contrast, the authors of Ref. [13] evaluated the relationships between robustness and the fraction of links connecting between subnetworks. Starting from two isolated networks, adding a few connecting links suppresses the largest cascades in each subnetwork. However, too many connecting links cause large cascades. They also demonstrated that asymmetric node capacity among institutive subnetworks may lead to an arms race for greater capacity of nodes. These studies indicate that interconnectivity affects robustness of interconnected networks.

However, these studies are insufficient from the point of view of network performance such as total amount of bandwidth required to accommodating traffic. In addition, it is expected that topological structure of interacting subnetworks must have some effect on performance of interconnected network. We therefore focus on relationships between topological structures of interconnection topology formed by subnetworks. Ref. [16] categorized structures of tree-based topologies into five classes. They pointed out the multiscale structure, where subnetworks are connected by links of various distance and link density changes depending on the hierarchy level, suppresses load of node and possess high robustness against node failure. In the track of this, we investigate performance of interconnection topologies having each structure. Comparing the total amount of bandwidth to accommodate traffic in the network, we show that interconnection topologies constructed in order to have the multiscale structure requires the least cost regardless of structures of subnetworks. We then demonstrate heterogeneous degree distribution of ISP topologies affects interconnection topology. Nevertheless, interconnection topology having the multiscale structure suppresses the sum of bandwidth. These results show possibility of a solution of the Internet design.

This chapter is organized as follows. We introduce related work in Section 5.2. In Section 5.3, we

explain the construction methods of interconnection topologies. And then, we performance of interconnected network constructed by each class of subnetworks and ISP topologies in 5.4. Finally, in Section 5.5, we conclude this chapter and mention future work.

5.2 Related Work

5.2.1 Robustness against Cascade Failure of Interconnected Networks

Robustness of interconnected networks has been concerned in Refs. [11–13]. Interconnected networks, as it called "networks of networks," is formed by institutive small-scale subnetworks. They pointed out that interconnected networks having high "inter-similarity" are robust against cascade failure. They introduced two measures to assess the level of inter-similarity. One is inter degree-degree coefficient (IDDC) and the other is inter clustering coefficient (ICC). IDDC is measured by the joint probability e_{jk} that a link is connected to a node in subnetwork A with degree j and to a node in subnetwork B with degree k. IDDC displays high value when high degree nodes from network A tend to couple with high degree nodes from network B and vice versa. ICC evaluates for a pairs of interconnected nodes a in a subnetwork A and b in the other subnetwork B how many the neighbors of a connect to the neighbors of b and vice versa. For increasing values of ICC more of the neighbors of a connect to the neighbors of b. They claimed that interconnected networks having both high IDDC and high ICC shows high robustness against cascade failure.

In Ref. [11, 12] assumed that all nodes in subnetwork A connect to nodes belonging to the other subnetwork B. Unlike these studies, the correlation between robustness and fraction of connecting links is investigated [13]. Starting from two isolated networks, adding some interconnecting links suppresses the largest cascades in each subnetwork. However, too many connecting links cause large cascades. They also demonstrated that asymmetric capacity among interacting subnetworks may lead to an arms race for greater capacity. In these studies, cascade failure is mainly focused to evaluate properties of interconnected networks. It seems insufficient from the view point of performance of communication network. We therefore investigate network performance of interconnected networks, such as the total amount of bandwidth in order to accommodate traffic demand.

5.2.2 Robustness of the Multiscale Structure

Dodds et. al. proposed a network construction algorithm that constructs five classes of networks [16]. The algorithm starts from a tree topology with branching ratio b and hierarchical level L and adds m links chosen stochastically. The probability $P(i, j)$ that two nodes i and j will be connected depends on the depth D_{ij} of their lowest common ancestor a_{ij} and also their own depths d_i and d_j (in Fig. 5.1, for example, $D_{ij} = 2$, $d_i = 2$, and $d_j = 3$). The probability $P(i, j)$ is defined as follows,

$$P(i, j) \quad \propto \quad e^{-D_{ij}/\lambda} e^{-x_{ij}/\zeta}, \tag{5.1}$$

where λ and ζ are tunable parameters. x_{ij} represents organizational distance between two nodes i and j to be $x_{ij} = (d_i^2 + d_j^2 - 2)^{1/2}$ (valid for $d_i + d_j \geq 2$). By changing the values of λ and ζ, this algorithm generates various topological structures. The authors categorized resulted networks into the following five classes.

- Random (R): For $(\lambda, \zeta) \to (\infty, \infty)$, links are added uniformly at random.
- Random interdivisional (RID): For $(\lambda, \zeta) \to (\infty, 0)$, links are added exclusively between pairs of nodes that share the same ancestor in high level.
- Local Team (LT): $(\lambda, \zeta) \to (\infty, 0)$, links are allocated exclusively between pairs of nodes that have short organizational distance, regardless of their rank. The resulting topology has teams form at all levels of the hierarchy.
- Core-periphery (CP): $(\lambda, \zeta) \to (0, 0)$, links are added exclusively between subordinates of the node in top level. The resulting topology has dense-connected core network.
- Multiscale (MS): The values of (λ, ζ) have intermediate values in the inside among other four classes. The resulting topology has connectivity dominated by the range from a small x_{ij} to a large x_{ij}. In multiscale networks, link density decreases as the hierarchical level decreases, such that the high level nodes exhibit the high density. The multiscale structure has small-world property.

The authors evaluated two types of robustness. One is congestion robustness and the other is connectivity robustness. Congestion robustness is measured by the maximum probability that any given message

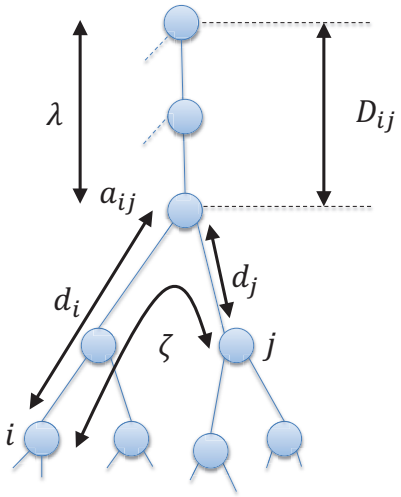


Figure 5.1: Illustrative example of the network construction algorithm

will be processed by a given node. Connectivity robustness defined by the fractional size of the largest connected component remaining after the removal of some number of nodes. In this study, the multiscale structure has high two types of robustness. The authors called this property “ultrarobustness.”

5.3 Network Connecting Model

We propose a model connecting two subnetworks inspired by the algorithm introduced in Ref. [16]. For simplify, we connect two subnetworks having exactly the same topology and links connecting between two topologies are selected stochastically. Of course, this situation is quite unlikely to occur in real information networks. However, we use this model to clarify stochastic property of interconnection topology between two subnetworks.

Before connecting two subnetworks, we prepare subnetwork A and subnetwork A' having the same topology. Then, we calculate connection probability $P(i, j)$ of all node pairs (i, j) in one subnetwork A. After calculating connection probability, we connect two nodes belonging to different subnetworks. We start with selecting connected node pair i and j according to $P(i, j)$. We then connect between node i and node j' in subnetwork A'. Node j' is the corresponding node of j. We repeat adding links between two

subnetworks m times.

Our model is based on the assumption that hierarchical level of nodes is preliminarily defined. The probability is calculated by using Eq. 5.1 in the same way in Ref. [16]. However, we arrange this calculation method for use of connecting between two subnetworks. One of the extensions is for use in non-tree-based topology such as the Abilene-inspired topology [21] (in Fig. 5.2). The Abilene-inspired topology is based on Abilene Network used for higher education. As shown this figure, The Abilene-inspired topology has meshed cores and tree-like structure hanging down from the nodes in core. The algorithm in Ref. [16] does not consider existence of horizontal links connecting nodes in the same level. We therefore extend the algorithm to calculate probability in subnetwork where horizontal links exist. We redefine organizational distance x_{ij} by using three values d_i, d_j, and d_{ij}. In our model, d_i is defined as the numbers of upstream links in the shortest path from source node i to destination node j. In the same way, d_j is the number of downstream links and d_{ij} is the number of horizontal links (shown in Fig. 5.3 as an example, $d_i = 1$, $d_j = 1$, and $d_{ij} = 1$). The redefined organizational distance of nodes i and j is $x_{ij} = (d_i^2 + d_j^2 + d_{ij}^2)^{1/2}$ (valid for $d_i + d_j + d_{ij} \geq 0$). Another extension is ignoring nodes in the lowest level of a subnetwork. It seemed that nodes on the edge of ISP topologies concentrate to aggregate traffic from customers and do not forward traffic from external networks [21]. From this point of view, we consider nodes in the lowest level L as end hosts and does not have connection to the other subnetwork. The other extension is calculation in the case when node i and j are located in the same place, i.e., $i = j$. When two subnetworks are connected, the pair of node i and node i' in the subnetwork A' has possibility to be connected each other like Fig. 5.4. For this reason, we include the case $i = j$ in probability calculation process. In that case, organizational distance x_{ij} equals 0 and D_{ij} is defined by level of the node i (shown in Fig. 5.4 for example, $D_{ij} = 2$).

By tuning parameters λ and ζ, we generate five classes of interconnection topology between two subnetworks. In this chapter, we use the same definition of five classes in the same way in Ref. [16].

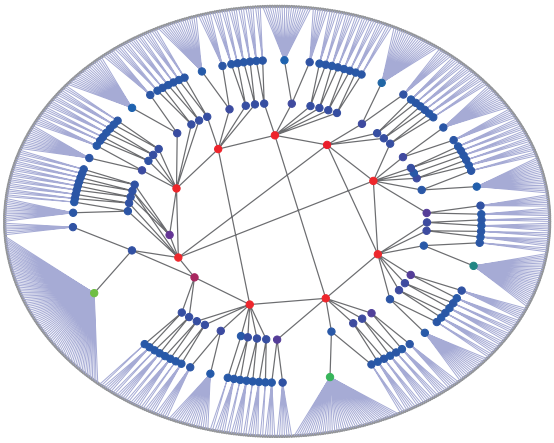


Figure 5.2: The Abilene-inspired topology referenced from Ref. [21]. The number of nodes is 869 and the number of links is 877.

5.4 Performance Evaluation of Interconnection Topology

In the previous section, we explain about how to connect between two the same subnetworks. In this subsection, we investigate network performance of interconnection topology having each structure.

5.4.1 Evaluation of Interconnected Networks Formed by Subnetworks Having Five Classes of Structure

First, we compare when subnetworks have five classes of structure. We generate five classes of structures starting with branching ratio $b = 9$ and hierarchical level $L = 4$. And then, we add $m = 50$ links according to the algorithm in Ref. [16]. The number of nodes is 820 and the number of links is 869. We also evaluated subnetworks generated by other m values, from $m = 5$ to $m = 200$. However, the results denote the same tendency of the following results. Resulting structures in this way have the similar number of nodes and links to the Abilene-inspired topology. We set values of (λ, ζ) into five patterns; $(\lambda, \zeta) = (\infty, \infty)$ (random, R), $(\lambda, \zeta) = (\infty, 0.05)$ (local team, LT), $(\lambda, \zeta) = (0.05, \infty)$ (random interdivisional, RID), $(\lambda, \zeta) = (0.05, 0.05)$ (core-periphery, CP), and $(\lambda, \zeta) = (0.5, 0.5)$ (multiscale, MS).

After generating five classes of subnetworks, we connect between subnetworks according to probability calculated by the algorithm in Section 5.3. We use the same values of (λ, ζ) to generate five classes

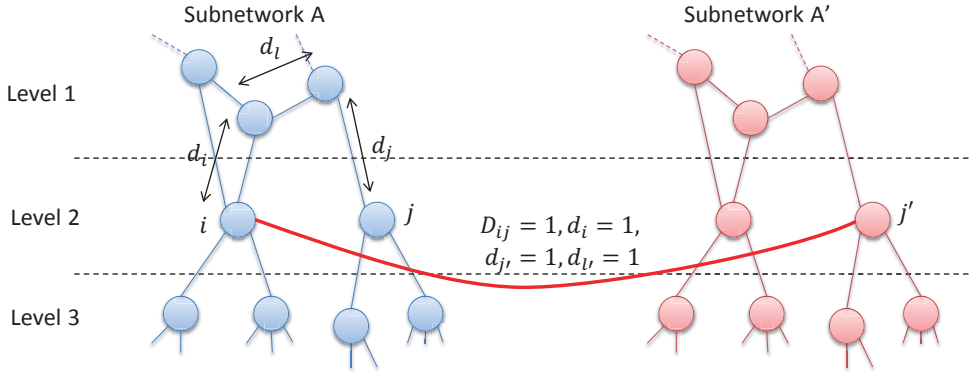


Figure 5.3: Construction of interconnected network: fundamental form

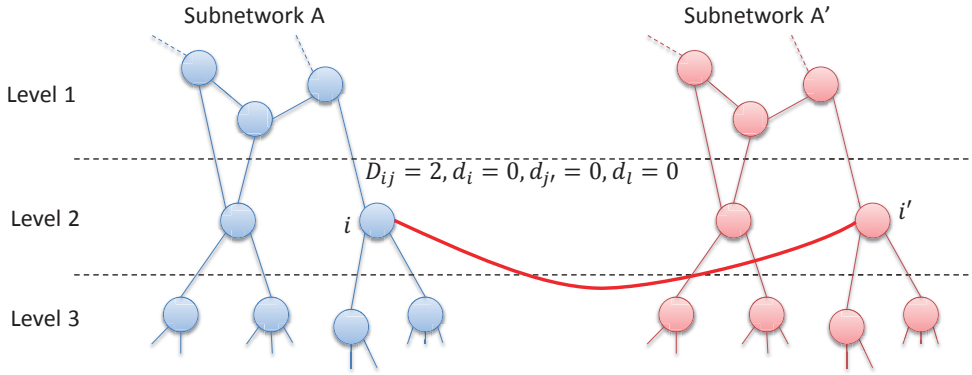


Figure 5.4: Construction of interconnected network when two nodes are located on the same point

of interconnection topologies. In the term of performance evaluation, we assign traffic flows between all pairs of end hosts in the lowest level ($L = 4$). The number of end hosts is $729 \times 2 = 1,458$. To simplify, values of traffic demand are the same (demand equals 1). We evaluate network performance with the maximum load and the total amount of link bandwidth required to accommodate all flows. The maximum load is defined by the maximum fraction of passing flows to the number of total flows. Bandwidth of the link is defined as the number of passing flows of the link. In this evaluation, flows go through the shortest path to their destinations. Like realistic flow between ISP topologies, we use hot-potato routing, i.e., flows from end hosts in subnetwork A to end hosts in subnetwork A' select the path that minimizes the distance through subnetwork A in the shortest paths to their destinations.

The Figures 5.5, 5.6, 5.7, 5.8, and 5.9 show the results of performance evaluation of interconnection topology between five classes of subnetworks. All plotted points represent average value of the results obtained by 10 times simulations. The *x*-axis in these figures represents the number of connecting links between two subnetworks. In each figure (a), the *y*-axis shows the maximum load of node in the interconnected network and *y*-axis represents the total amount of bandwidth in each figure (b). In these figures, the structure of interconnection topologies is compared.

The common characteristics of these results, the total amount of bandwidth shows the lowest value when interconnection topology has the multiscale structure. In the multiscale structure, the connecting links between two subnetworks cover a wide range of hierarchy and organizational distance. These features shorten average path length of the connected networks and contribute to suppress the total bandwidth. As proof of this inference, the maximum load rises when interconnection topology has the multiscale structure as shown in each figure (a). This is because flows concentrate on the node contributing to shorten path length in top level. The results of interconnection topology having the core-periphery structure exhibit the similar tendency.

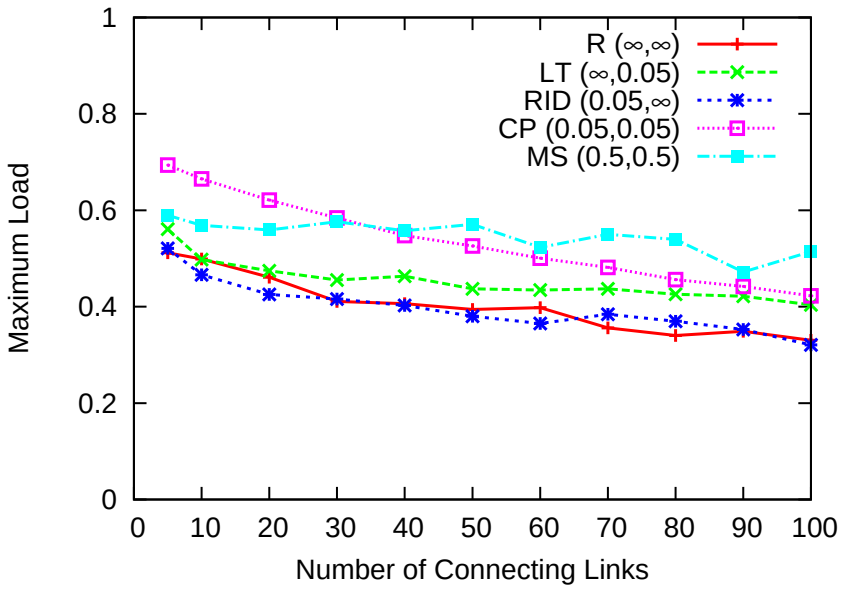
Another point to note is the total amount of bandwidth of the core-periphery structure is lower than the multiscale structure when the number of connecting links is low. These results indicate that it is important to connect between nodes in higher level when the number of links between two subnetworks is small. However, as the number of connecting links increases, the total amount of bandwidth of the multiscale structure falls below that of the core-periphery structure. In addition, the core-periphery structure requires the links having high-bandwidth compared to the multiscale structure. Figure 5.10 shows the bandwidth distributions of the core-periphery structure and the multiscale structure. In this figure, the *y*-axis represents the link bandwidth and the *x*-axis represents the rank in descending order. This figure indicates that the core-periphery structure needs high-bandwidth links because a large amount of the link connecting between two subnetworks. From the view point of technological constraint, it is so hard to construct the core-periphery structure due to high-bandwidth links. We conclude that connecting only between the nodes in high level is not expected to performance improvement as the number of connecting link increases.

We finally mention about effects of differences in structure of subnetworks. The maximum load displays low value when subnetworks have the multiscale or the core-periphery structure. There is consistency with Ref. [16], where two classes of structure show high congestion robustness, i.e., the maximum load keeps lower than that of other classes.

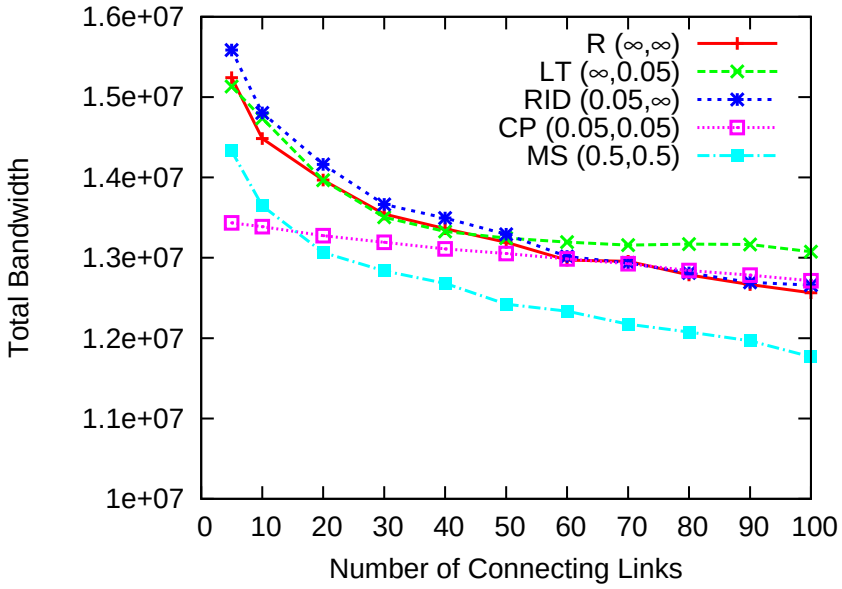
5.4.2 Evaluation of Interconnected Networks Formed by ISP Topologies

We also evaluate performance of interconnection topologies when two the Abilene-inspired topologies are connected. The Abilene-inspired topology has hierarchical structure with $L = 4$ and random mesh structure in top level as shown in Fig. 5.2. Having core network is the difference from topologies generated by method in Ref. [16]. We connect between the two Abilene-inspired topologies according to our algorithm. We then assign traffic flows between all pairs of end hosts. The number of end hosts is $698 \times 2 = 1,396$. Figure 5.11 shows the results of interconnection topology constructed by two the Abilene-inspired topologies. Horizontal axis and vertical axis represent in the same manner as Fig. 5.5. Fig. 5.11 shows the similar characteristics to the results of structures generated by the model in Ref. [16].

However, there is one of the notable points of these results that the maximum load shows lowest value when interconnection topology has the core-periphery structure. There is the other notable point that core-periphery structure approaches the performance of the multiscale structure even if the number of connecting links is large. The reason comes from heterogeneous degree distribution of the Abilene-inspired topology. Degree distribution of the Abilene-inspired topology follows a power law [21]. That is, the nodes having large degree, so-called “hub nodes,” exist in the Abilene-inspired topology. By contrast, most nodes of model-based subnetworks have the same degree at the initial state. So, resulting degree distributions of generated topologies do not vary widely. In the Abilene-inspired topology, hub nodes contribute to shorten average path length in the network. However, if hub nodes earn connecting links, they bear a burden, i.e., heavy traffic. In the Abilene-inspired topology, hub nodes tend to be located in $L-1$ level because hub nodes play the role of aggregating traffic from end hosts in the lowest level L. If the

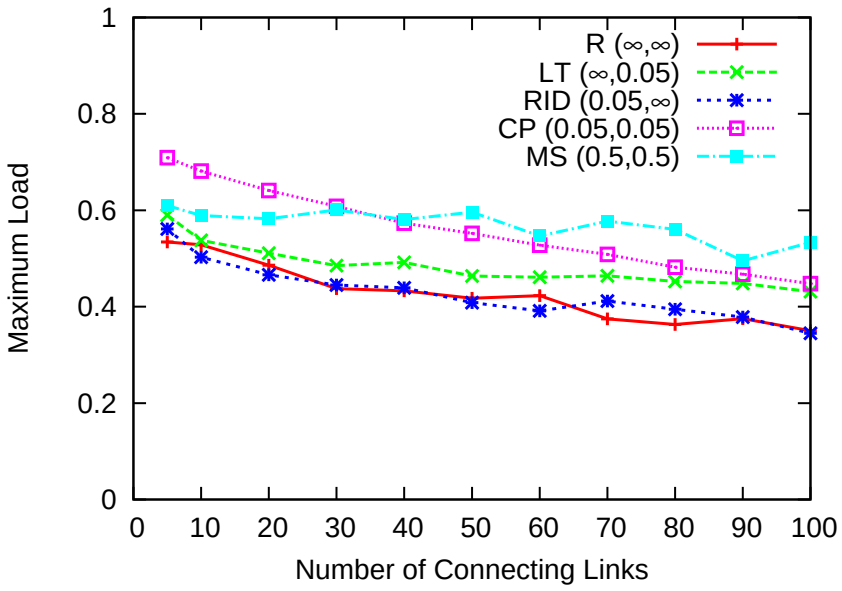


(a) Maximum load

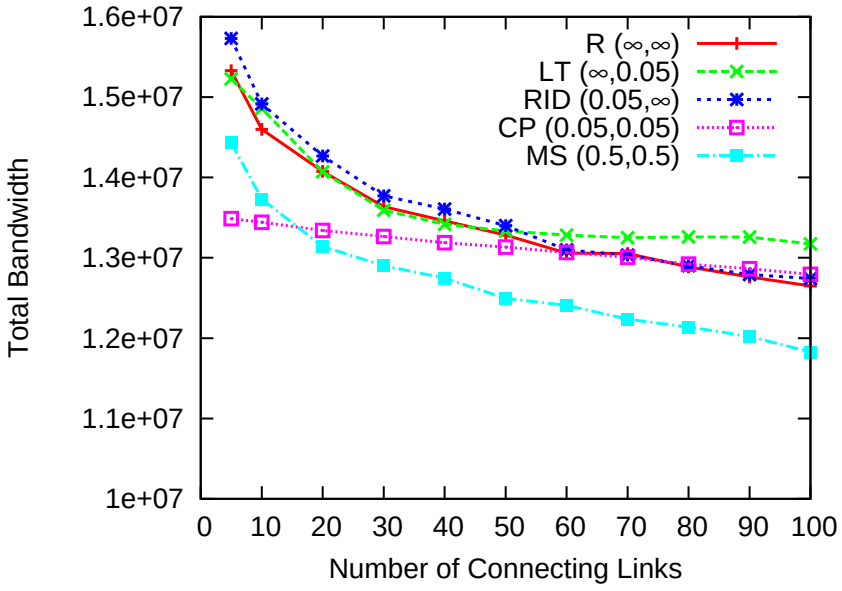


(b) Total amount of bandwidth

Figure 5.5: Evaluation of interconnected networks formed by subnetworks having the random structure

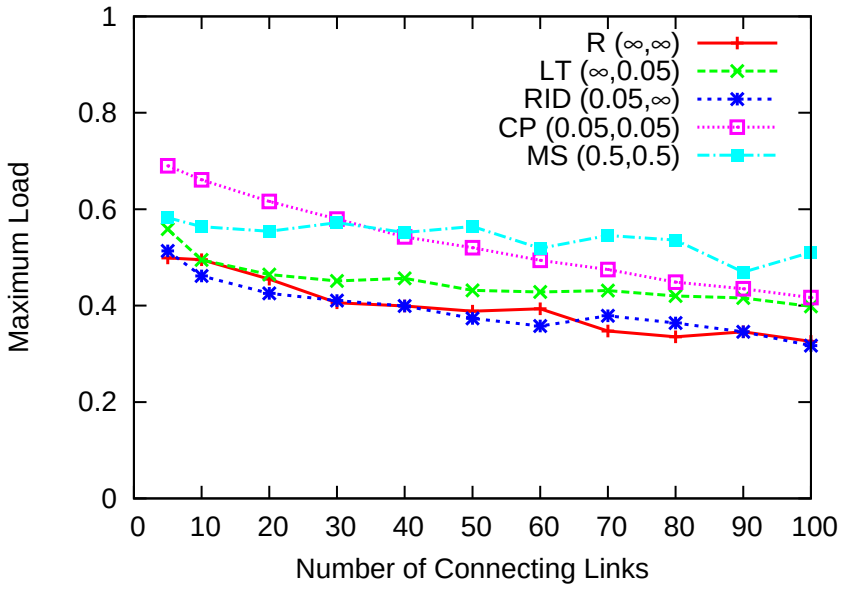


(a) Maximum load

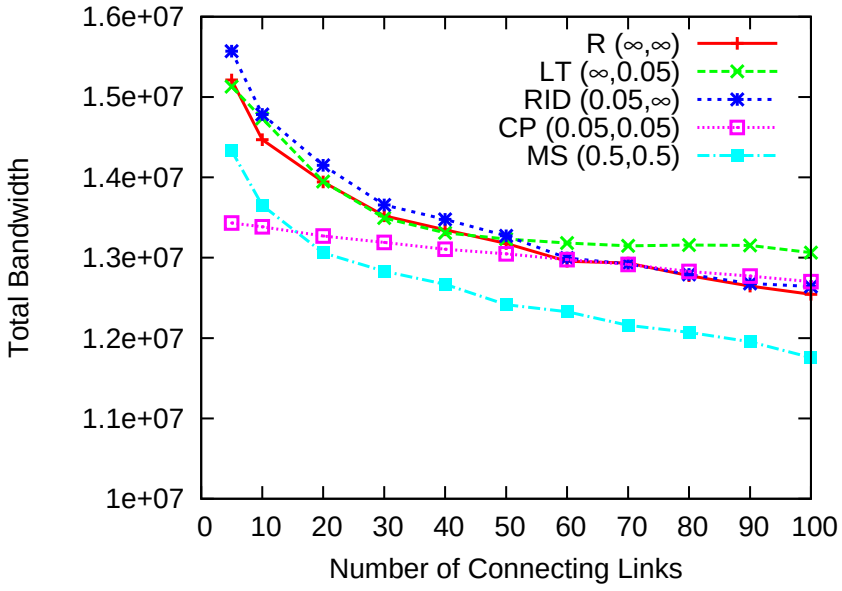


(b) Total amount of bandwidth

Figure 5.6: Evaluation of interconnected networks formed by subnetworks having the local team structure

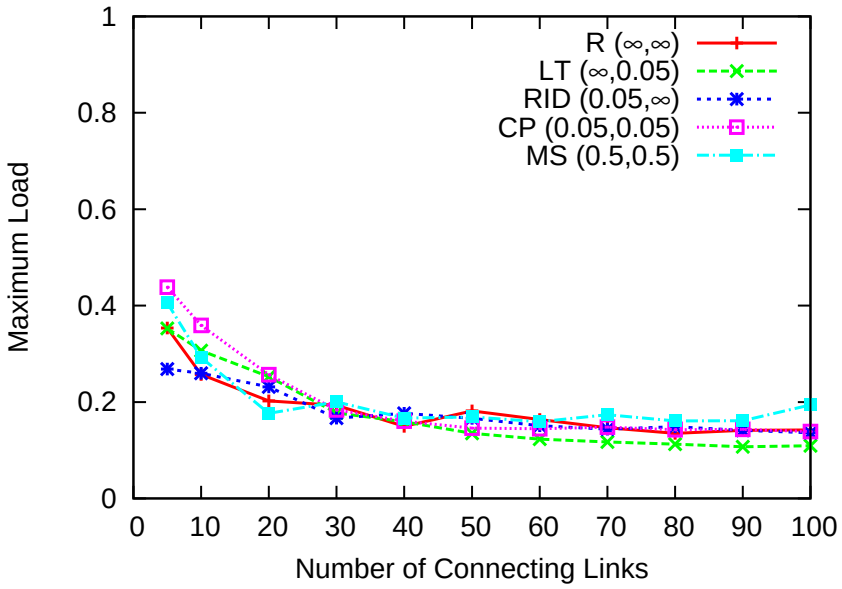


(a) Maximum load

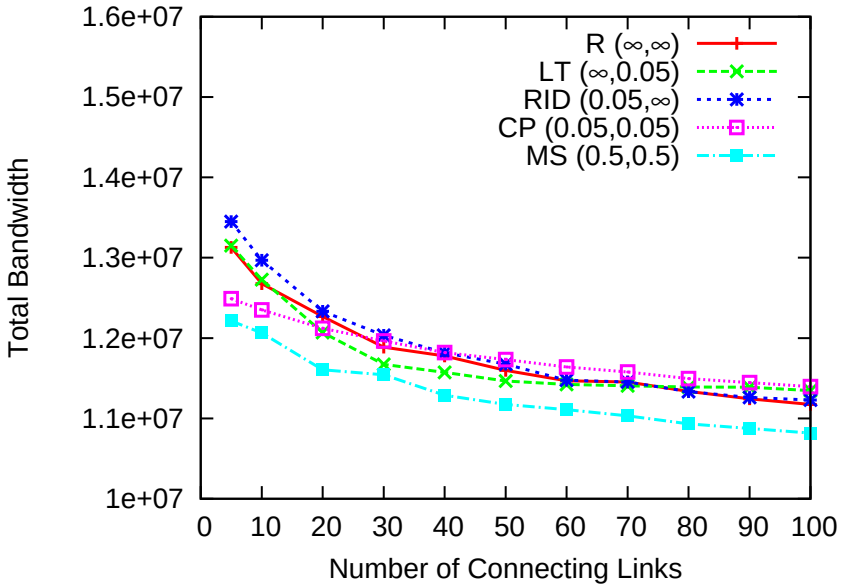


(b) Total amount of bandwidth

Figure 5.7: Evaluation of interconnected networks formed by subnetworks having the random interdivisional structure

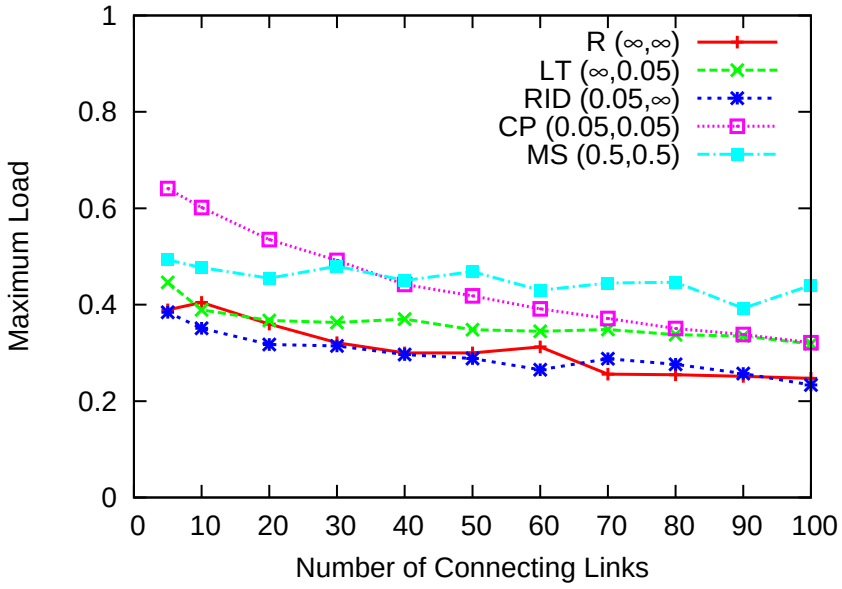


(a) Maximum load

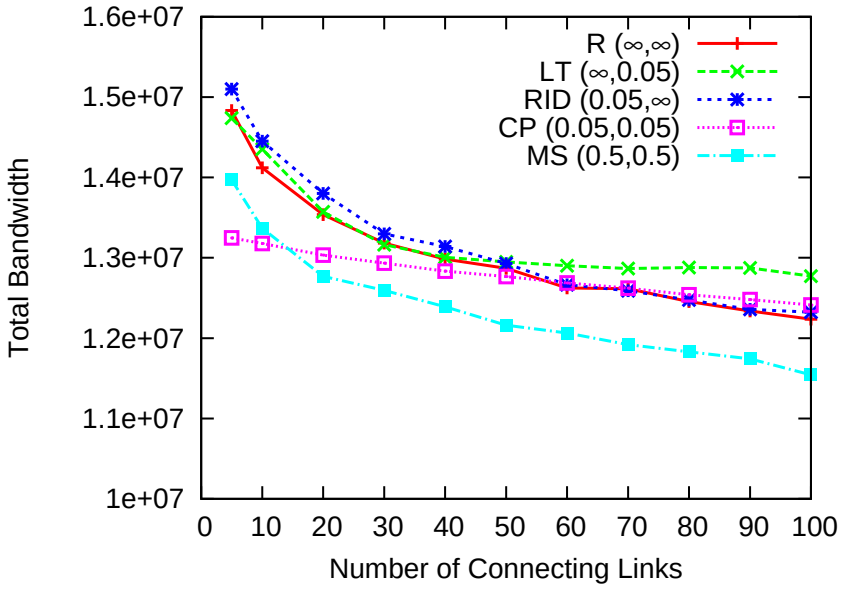


(b) Total amount of bandwidth

Figure 5.8: Evaluation of interconnected networks formed by subnetworks having the core-periphery structure



(a) Maximum load



(b) Total amount of bandwidth

Figure 5.9: Evaluation of interconnected networks formed by subnetworks having the multiscale structure

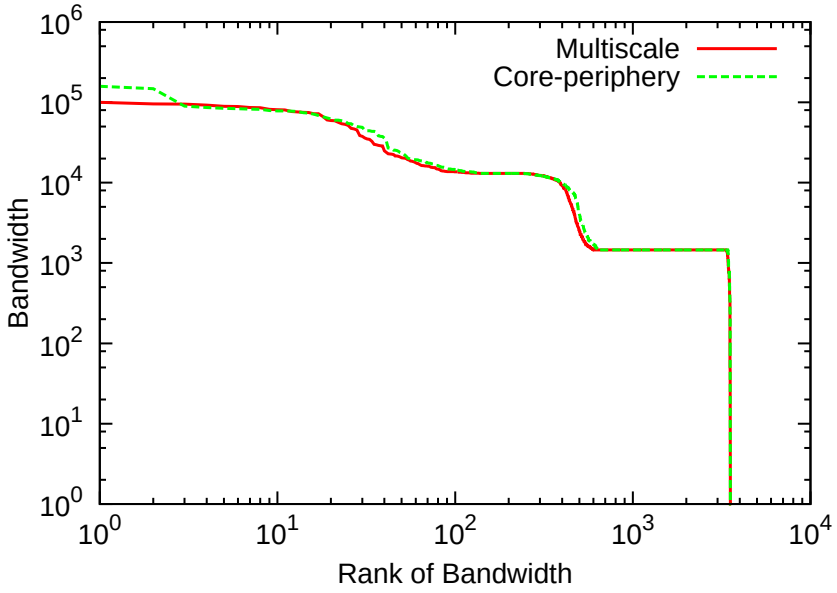


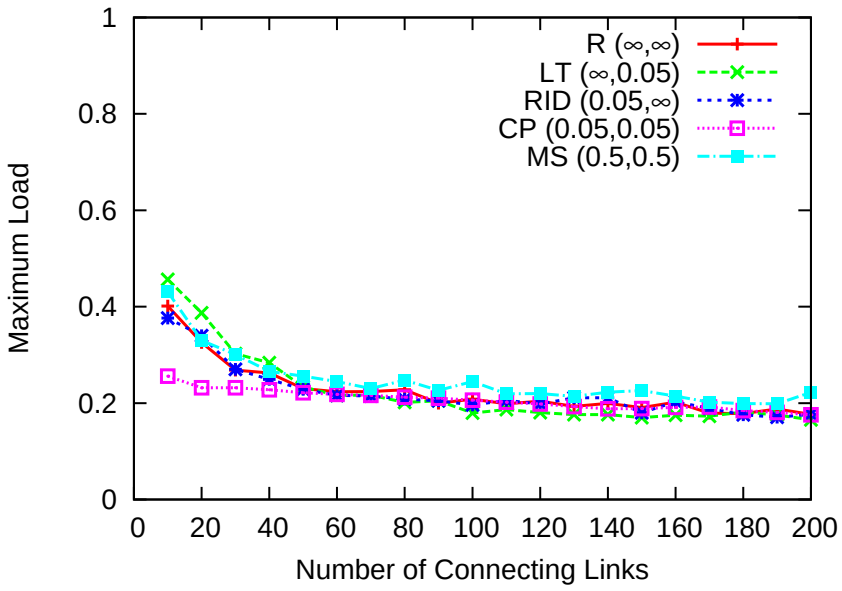
Figure 5.10: Bandwidth distributions of the core-periphery structure and the multiscale structure: The number of connecting links is 50.

Abilene-inspired topologies are connected in order to have the core-periphery structure, hub nodes do not have connecting links. Therefore, the maximum load is suppressed. In addition, core network constructed by nodes in top level also contribute to suppress the load. Because flows passing through top level nodes are distributed in core network, load of top level nodes is absorbed.

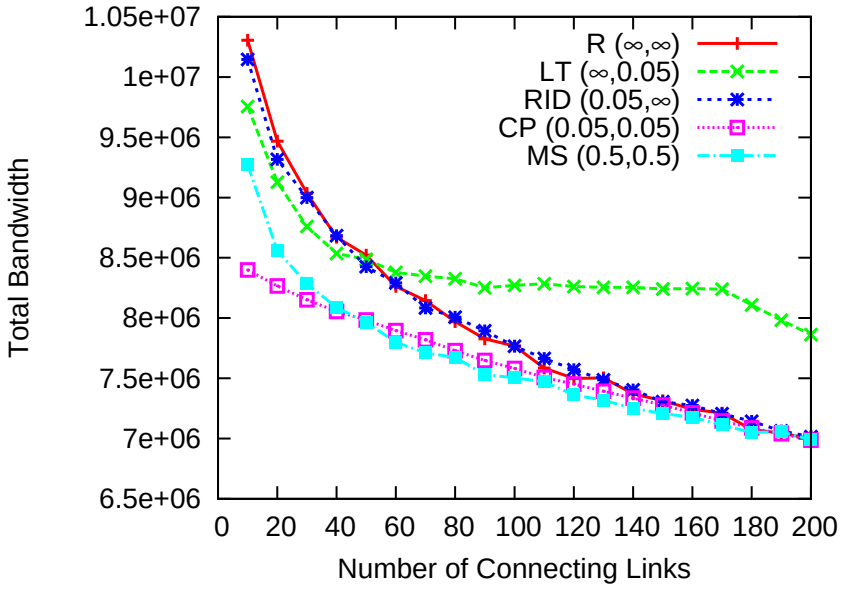
The results of interconnected the Abilene-inspired topologies indicate that performance of interconnected networks is affected by not only structures of subnetworks but also degree distributions of subnetworks. Nevertheless, the multiscale structure suppresses the total amount of bandwidth in spite of the heterogeneous degree distribution.

5.5 Conclusion

In this chapter, we discussed about performance of interconnection topology between ISP topologies towards establishment of network design methods. The Internet is composed by a large number of ISPs. Traffic coming from neighboring ISPs is considerable for each ISP. We focused on network performance



(a) Maximum load



(b) Total amount of bandwidth

Figure 5.11: Evaluation of interconnected networks formed by the two Abilene-inspired topologies

of interconnected networks constructed by institutive subnetworks. By reference to Ref. [16], we construct five classes of interconnection topology between two subnetworks. We compared network performance of five classes of interconnection topologies. The results show that interconnection topology having the multiscale structure suppresses the total amount of bandwidth. We also investigated the case when the two Abilene-inspired topologies are connected. Because of heterogeneous degree distribution of the Abilene-inspired topology, the results are different from those of model-based topologies in point of the maximum load of nodes. Nevertheless, the multiscale structure suppresses the total amount of bandwidth. Our work sheds the light on the problem of designing interconnection topology between ISP topologies.

In the future work, we evaluate network performance with consideration of traffic dynamics. And then, we investigate performance of connected network composed by more than three topologies.

Chapter 6

Conclusion and Future Work

In this chapter, we summarize the discussion in each of the previous chapters and describe future research works.

Chapter 1 mentioned research background and overview of our studies. Development of new networking applications contributes to an increase in complexity of Internet traffic behavior. Thus, understanding the complex behavior of the Internet is important for designing cost effective networks.

In Chapter 2, we investigated the interaction between the structures 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 improves network throughput in the ISP topology. On the other hand, the functionality of TCP makes the queue length fluctuate. Even in this case, the highly-modularity structure of ISP topologies reduces the number of highly fluctuating links compared with BA topologies. We also evaluated the queue length on other topologies and confirmed that the modularity structure can reduce the number of highly fluctuating links. Our results suggest that reproducing the modularity structure is important when evaluating the performance of transport protocols.

Chapter 3 focused on link bandwidth in ISP topologies. First, we investigate the link bandwidth distribution of an actual ISP backbone network in Japan. Results show that link bandwidth distribution in the actual ISP network obeys a power law. To see the reason of the power-law characteristic, we compared network performance with three link bandwidth distributions; power-law, exponential and identical. Results show that power-law link bandwidth distributions accommodate large amount of the traffic. Results

also show that, under the constraint of only node processing capacities, amount of the traffic that the network can accommodate is over estimated. Our conclusion of this chapter is that the link bandwidth with power-law distribution is essential to characterize the network performance and is important to evaluate the network control method properly.

In Chapter 4, we proposed a model of bandwidth distribution in ISP topologies. We first evaluated the bandwidth distributions obtained by periodic bandwidth enhancements to accommodate TCP traffic in ISP topologies. As a consequence of periodic bandwidth enhancement, we revealed that the bandwidth distributions of these topologies follow a power-law whose exponent varies according to the topology. We confirm that periodic enhancement reproduces the similar power-law bandwidth distribution of the actual ISP topology in Japan. A comparison of simulation results for the ISP topology and the BA topology showed that the high-modularity structure of the ISP topology characterizes its bandwidth distribution. The ISP topology has a greater number of high-bandwidth links due to its highly congested inter-module links. Our results indicate the possibility that a power-law characteristic of bandwidth distributions of ISP topologies is induced by the periodic bandwidth enhancement to accommodate dynamic TCP traffic. Then, we proposed a bandwidth allocation model that accounts for TCP traffic dynamics. By considering the impact of a given link's bandwidth enhancement on its tributary links, our model generated a bandwidth distribution with the similar properties to the distribution obtained by periodic enhancement.

Chapter 5 discussed about interconnected networks between two subnetworks to design cost effective networks. ISPs should consider how to accommodate external traffic from neighboring ISPs because each ISP is affected by traffic from the other neighboring ISPs. We investigate network performance of interconnection topology for the purpose of designing interconnected networks formed by ISP topologies. Comparing the total amount of bandwidth to accommodate traffic, we showed that the multiscale structure reduces the sum of bandwidth regardless of the difference in structures of subnetworks and the difference in degree distribution of subnetworks.

In this thesis, we first analyzed traffic dynamics in ISP topologies. In large scale topologies, bottlenecks emerge at a variety of points and show various size of fluctuation because of the difference in traffic load of each link. Analyzing traffic fluctuation of all links in large scale topologies revealed that the high-modularity structure of ISP topologies has significant effect on traffic dynamics. This knowledge is not obtained by investigation in small scale topologies having a single bottleneck link. Traffic dynamics

induced by the high-modularity structure effects on network design such as bandwidth allocation. We then discussed and modeled link bandwidth distribution including traffic dynamics in ISP topologies. The results showed that inter-module links and their peripheral links tend to be enhanced at many times. To avoid a large number of enhancements at inter-module links, we focused on structure of interconnection topology formed by institutive modules (subnetworks). Traffic load distribution is affected not only by the number of interconnecting links but also by the structure of interconnection topology between subnetworks. We then investigated performance of interconnection topology and showed that the multiscale structure reduces the sum of bandwidth. These results show the possibility of a solution of the Internet design.

Finally, we discussed about design of interconnected networks as a part of cost effective network design. We conclude this thesis by summarizing several challenging tasks in future work. In Chapter 2, we showed the high-modularity structure is one of the important characteristics of ISP topologies. From this finding, it is expected to develop a topology generation method that reproduces the high-modularity structure for evaluating the network control method properly. In Chapter 3 and 4, we revealed that the power-law bandwidth distribution of the actual ISP is obtained by periodic bandwidth enhancement considering TCP traffic dynamics and proposed a model reproducing the power-law bandwidth distribution. As for future works, a network design method that includes the interaction between TCP flow control and the structural properties of topologies should be explored. In Chapter 5, we investigate network performance of interconnection topology between two subnetworks. However, traffic dynamics was not considered in the simulations unlike Chapters 2, 3, and 4. We should evaluate network performance of interconnected networks in more realistic situations. For example, we investigate performance of interconnected networks composed by more than three topologies, with consideration for traffic dynamics. We believe that discussions in this thesis contribute to establish new design methods for ISP networks.

Bibliography

[1] Cisco Systems, “Cisco visual networking index: The zettabyte era.” Cisco White Paper, 2012.

[2] R. Bruschi, F. Davoli, and F. Cucchietti, “Energy efficiency in the future Internet: A survey of existing approaches and trends in energy-aware fixed network infrastructures,” *IEEE Communications Surveys & Tutorials*, vol. 13, pp. 223–244, Second Quarter 2011.

[3] G.-Q. Zhang, Q.-F. Yang, S.-Q. Cheng, and T. Zhou, “Evolution of the Internet and its cores,” *New Journal of Physics*, vol. 10, pp. 1–11, Dec. 2008.

[4] D. Neilson, “Photonics for switching and routing,” *IEEE Journal of Selected Topics in Quantum Electronics (JSTQE)*, vol. 12, pp. 669–678, July 2006.

[5] Y. Koizumi, T. Miyamura, S. Arakawa, E. Oki, K. Shiomoto, and M. Murata, “Stability of virtual network topology control for overlay routing services,” *OSA Journal of Optical Networking*, vol. 7, pp. 704–719, July 2008.

[6] M. Armbrust, A. Fox, R. Griffith, A. D. Joseph, R. Katz, A. Konwinski, G. Lee, D. Patterson, A. Rabkin, I. Stoica, and M. Zaharia, “A view of cloud computing,” *Communications of the ACM*, vol. 53, pp. 50–58, Apr. 2010.

[7] N. Gupta, J. Kataria, and A. Bansal, “A survey on cloud providers and migration issues,” *International Journal of Computer Applications*, vol. 56, pp. 38–43, Oct. 2012.

[8] Cisco Systems, “Cisco visual networking index: Global mobile data traffic forecast update, 2011?2016.” Cisco White Paper, 2012.

[9] G. M. Luigi Atzori, Antonio Iera, "The Internet of things: A survey," *Computer Networks*, vol. 54, pp. 2787–2805, Oct. 2010.

[10] M. Brede, "Coordinated and uncoordinated optimization of networks," *Physical Review E*, vol. 81, 066104, June 2010.

[11] R. Parshani, C. Rozenblat, D. Ietri, C. Ducruet, and S. Havlin, "Inter-similarity between coupled networks," *Europhysics Letters*, vol. 92, pp. 68002–p1–68002–p5, Jan. 2011.

[12] J. Gao, S. V. Buldyrev, H. E. Stanley, and S. Havlin, "Networks formed from interdependent networks," *Nature Physics*, vol. 8, pp. 40–48, Dec. 2011.

[13] C. D. Brummitt, R. M. D'Souza, and E. A. Leicht, "Suppressing cascades of load in interdependent networks," *Proceedings of the National Academy of Sciences (PNAS)*, vol. 109, pp. E681–E689, Mar. 2012.

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

[15] M. E. J. Newman, "Modularity and community structure in networks," *Proceedings of the National Academy of Sciences (PNAS)*, vol. 103, pp. 8577–8582, Apr. 2006.

[16] P. S. Dodds, D. J. Watts, and C. F. Sabel, "Information exchange and the robustness of organizational networks," *Proceedings of the National Academy of Sciences (PNAS)*, vol. 100, pp. 12516–12521, Oct. 2003.

[17] M. Faloutsos, P. Faloutsos, and C. Faloutsos, "On power-law relationships of the Internet topology," *ACM SIGCOMM Computer Communication Review*, vol. 29, pp. 251–262, Oct. 1999.

[18] B. Zhang, R. Liu, D. Massey, and L. Zhang, "Collecting the Internet AS-level topology," *ACM SIGCOMM Computer Communication Review*, vol. 35, pp. 53–61, Jan. 2005.

[19] A. L. Barabási and R. Albert, "Emergence of scaling in random networks," *Science*, vol. 286, pp. 509–512, Oct. 1999.

[20] A. Fabrikant, E. Koutsoupias, and C. H. Papadimitriou, "Heuristically optimized trade-offs: A new paradigm for power laws in the Internet," in *Proceedings of International Colloquium on Automata, Languages and Programming (ICALP)*, pp. 110–122, July 2002.

[21] 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.

[22] O. Heckmann, M. Piringer, J. Schmitt, and R. Steinmetz, "Generating realistic ISP-level network topologies," *IEEE Communications Letters*, pp. 335–336, July 2003.

[23] S. Zhou, "Characterising and modelling the Internet topology – the rich-club phenomenon and the PFP model," *BT Technology Journal*, vol. 24, pp. 108–115, July 2006.

[24] 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, 36102, Mar. 2004.

[25] Z. Q. F. Ming-yang Zhou, Shi-min Cai, "Traffic dynamics in scale-free networks with tunable strength of community structure," *Physica A: Statistical Mechanics and its Applications*, vol. 391, pp. 1887–1893, Feb. 2012.

[26] Y. Zhang, S. Zhou, Z. Zhang, J. Guan, S. Zhou, and G. Chen, "Traffic fluctuations on weighted network," *arXiv:0902.3160*, Nov. 2011.

[27] J. Alvarez-Hamelin and N. Schabanel, "An Internet graph model based on trade-off optimization," *European Physical Journal B*, pp. 231–237, Mar. 2004.

[28] G.-Q. Zhang, S. Zhou, D. Wang, G. Yan, and G.-Q. Zhang, "Enhancing network transmission capacity by efficiently allocating node capability," *Physica A*, vol. 390, pp. 387–391, Jan. 2011.

[29] M. Woolf, D. Arrowsmith, R. Mondragon, J. Pitts, and S. Zhou, "Dynamical modelling of TCP packet traffic on scale-free networks," *Institut Mittag-Leffler*, Oct. 2004.

[30] X. Ling, M.-B. Hu, W.-B. Du, R. Jiang, Y.-H. Wu, and Q.-S. Wu, "Bandwidth allocation strategy for traffic systems of scale-free network," *Physics Letters A*, vol. 374, pp. 4825–4830, Nov. 2010.

[31] Y. Zhuo, Y. Peng, C. Liu, Y. Liu, and K. Long, "Traffic dynamics on layered complex networks," *Physica A: Statistical Mechanics and its Applications*, vol. 390, pp. 2401 – 2407, June 2011.

[32] N. Spring, R. Mahajan, D. Wetherall, and T. Anderson, "Measuring ISP topologies with Rocketfuel," *IEEE/ACM Transactions on Networking*, vol. 12, pp. 2–16, Feb. 2004.

[33] R. Fukumoto, S. Arakawa, and M. Murata, "On routing controls in ISP topologies: A structural perspective," in *Proceedings of Chinacom*, pp. 1–5, Oct. 2006.

[34] S. Arakawa, T. Takine, and M. Murata, "Analyzing and modeling router-level internet topology and application to routing control," *Computer Communications*, vol. 35, pp. 980–992, May 2012.

[35] 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.

[36] 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.

[37] 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.

[38] 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.

[39] 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.

[40] W. L. S. Bauer, D. Clark, "The evolution of Internet congestion," in *Proceedings of The Research Conference on Communication, Information and Internet Policy*, pp. 1–34, Aug. 2009.

[41] C. Labovitz, S. Iekel-Johnson, D. McPherson, J. Oberheide, and F. Jahanian, "Internet inter-domain traffic," *ACM SIGCOMM Computer Communication Review*, vol. 41, pp. 75–86, Aug. 2011.

[42] T. Hirayama, S. Arakawa, K. Arai, and M. Murata, "Dynamics of feedback-induced packet delay in ISP router-level topologies," *IEICE Transactions on Communications*, vol. E95-B, pp. 2785–2793, Sept. 2012.

[43] T. Hirayama, S. Arakawa, K. Arai, and M. Murata, "On the packet delay distribution in power-law networks," in *Proceedings of International Conference on Evolving Internet (INTERNET)*, Aug. 2009.

[44] T. Hirayama, S. Arakawa, K. Arai, and M. Murata, "Queue dynamics in power-law networks," in *Proceedings of International Workshop on Sensor Networks and Ambient Intelligence (SeNAmI)*, Dec. 2009.

[45] T. Hirayama, S. Arakawa, K. Arai, and M. Murata, "Dynamics of feedback-induced packet delay in power-law networks," in *Proceedings of International Conference on Systems and Networks Communications (ICSNC)*, Aug. 2010.

[46] T. Hirayama, S. Arakawa, K. Arai, and M. Murata, "Modularity structure and traffic dynamics of ISP router-level topologies," in *Proceedings of International Symposium on Nonlinear Theory and its Applications (NOLTA)*, Oct. 2012.

[47] T. Hirayama, S. Arakawa, K. Arai, and M. Murata, "Congestion propagation in ISP topologies having powr-law degree distribution," *Technical Report of IEICE (NS2008-68)*, vol. 108, pp. 1–6, Oct. 2008. (in Japanese).

[48] T. Hirayama, S. Arakawa, K. Arai, and M. Murata, "Characteristics of packet transfer delay in power-law networks," *Technical Report of IEICE (NS2009-82)*, vol. 109, pp. 31–36, Oct. 2009. (in Japanese).

[49] T. Hirayama, S. Arakawa, K. Arai, and M. Murata, "Dynamics of feedback-induced packet transfer delay in power-law networks," *Technical Report of IEICE (NS2010-38)*, vol. 110, pp. 1–6, July 2010.

[50] T. Hirayama, S. Arakawa, K. Arai, and M. Murata, "Effect of modularity structure on traffic dynamics," *Technical Report of IEICE (NwGN2010-39)*, pp. 15–20, Jan. 2011.

[51] T. Hirayama, S. Arakawa, S. Hosoki, and M. Murata, "Models of link capacity distribution in ISP's router-level topologies," *International Journal of Computer Networks and Communications (IJCNC)*, vol. 3, pp. 205–216, Sept. 2011.

[52] T. Hirayama, S. Arakawa, and M. Murata, "Modeling link capacity distribution with consideration of TCP traffic dynamics in router-level networks," *submitted to Computer Communications*, 2012.

[53] T. Hirayama, S. Arakawa, and M. Murata, "Analysis of connection structure between ISP router-level topologies," *submitted to Physical Review E*, 2012.

[54] 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.

[55] M. Cusumaano, "Cloud computing and saas as new computing platforms," *Communications of the ACM*, vol. 53, pp. 27–29, Apr. 2010.

[56] D. Alderson, L. Li, W. Willinger, and J. C. Doyle, "Understanding Internet topology: principles, models, and validation," *IEEE/ACM Transactions on Networking*, vol. 13, pp. 1205–1218, Dec. 2005.

[57] P. Mahadevan, D. Krioukov, K. Fall, and A. Vahdat, "Systematic topology analysis and generation using degree correlations," in *ACM SIGCOMM Computer Communication Review*, vol. 36, pp. 135–146, Aug. 2006.

[58] 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," *ACM SIGCOMM Computer Communication Review*, vol. 29, pp. 301–313, Oct. 1999.

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

[60] 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.

[61] M. E. J. Newman and M. Girvan, "Finding and evaluating community structure in networks," *Physical Review E*, vol. 69, 026113, Feb. 2004.

[62] F. Sally and P. Vern, "Difficulties in simulating the Internet," *IEEE/ACM Transactions on Networking*, vol. 9, pp. 392–403, Feb. 2001.

[63] H. Maruta, "Power law structure of the Internet backbone network," *GLOCOM Review*, vol. 9, pp. 1–16, Sept. 2003. (in Japanese).

[64] N. Berger, B. Bollobás, C. Borgs, J. Chayes, and O. Riordan, "Degree distribution of the FKP network model," in *Proceedings of International Colloquium on Automata, Languages and Programming (ICALP*, pp. 725–738, July 2003.

[65] D. Alderson, J. Doyle, R. Govindan, and W. Willinger, "Toward an optimization-driven framework for designing and generating realistic Internet topologies," *ACM SIGCOMM Computer Communication Review*, vol. 33, pp. 41–46, Jan. 2003.

[66] Y. Xia and D. Hill, "Optimal capacity distribution on complex networks," *Europhysics Letters*, vol. 89, pp. 58004–p1–p6, Mar. 2010.

[67] "Internet Initiative Japan." `http://www.iij.ad.jp/`.

[68] A. Nucci, A. Sridharan, and N. Taft, "The problem of synthetically generating IP traffic matrices: Initial recommendations," *ACM SIGCOMM Computer Communication Review*, vol. 35, pp. 19–32, July 2005.

[69] A. Akella, S. Seshan, and A. Shaikh, "An empirical evaluation of wide-area Internet bottlenecks," in *Proceedings of ACM SIGCOMM conference on Internet measurement*, pp. 101–114, ACM, 2003.

[70] J. C. C. Restrepo and R. Stanojevic, "IXP traffic: A macroscopic view," in *Proceedings of Latin American Networking Conference (LANC)*, pp. 1–8, Oct. 2012.

[71] Cisco Systems, "Cisco visual networking index: Forecast and methodology." Cisco White Paper, 2008.

[72] R. T. B. Ma, D. M. Chiu, J. C. S. Lui, V. Misra, and D. Rubenstein, "Internet economics: The use of shapley value for ISP settlement," *IEEE/ACM Transactions on Networking*, vol. 18, pp. 775–787, June 2010.