

Partial Traffic Engineering over Cost Metric Network and Analysis on Its Oversubscription Capability

Masashi Hashimoto and Aki Fukuda

Abstract—Traffic engineering (TE) is promising to provide a more efficient traffic accommodation while cost metric routing such as OSPF, which sets routes so as to minimize link cost, is widely used. However, a path setting by TE does not offer cost minimum paths, which may be the best paths from the viewpoint of the quality of service (QoS). Here, we study a partial traffic engineering (Partial TE) scheme. First in this scheme, a part of the traffic matrix is subjected to cost metric routing and its bandwidth is reserved. Then, we perform path calculation for the remainder of the traffic matrix which realizes TE. We also consider a special situation of Partial TE, called an oversubscription. Oversubscription is the traffic accommodation that can be expected and really possible, even if the bottleneck link(s) is caused by QoS traffic accommodated in advance. The simulation results confirm that Partial TE allows the TE to set paths that coexist with the cost minimum paths, which can be used to increase traffic accommodation when non-uniform metric or traffic is used. And good oversubscription properties are obtained, too, in many examples in which non-uniform metric or traffic is used.

Index Terms—Traffic engineering, cost metric, shortest path, metric.

I. INTRODUCTION

THE traffic volume imposed on core networks has been rapidly increasing because of the increasing number of broadband subscribers as well as advances in broadband access technologies. The explosive spread of the Internet in recent years has led to various applications emerging on IP (Internet Protocol) networks. Such applications require assured levels of quality of service (QoS) and high network reliability. One of the major concerns is how to construct a backbone network that offers the very economical transport of large-scale IP-based packet traffic streams while achieving high reliability and offering QoS guarantees. One of the key technologies to achieve the economical transport of high-capacity IP traffic while satisfying the above requirements lies in traffic

engineering(TE)[1]-[4].

One merit of applying cost minimum routing, which is usually used in OSPF[5] is that QoS-specific requirements can be satisfied. These requirements, which are indicated by the metrics such as delay, reliability, and bandwidth, can be represented as cost. Those metric-based routing schemes tend to allow traffic to aggregate on particular routes, especially if their metrics are QoS-oriented [6]. There are some studies that realize TE while metric-based routing is carried on by setting metrics appropriately [4][7]. In the former case, the shortest path is to be determined via link costs which are acquired by taking advantage of TE's dual problem, and with the shortest path, traffic can be distributed at the same level on which TE is implemented. In the latter case, costs are acquired heuristically. Although the congestion rate becomes slightly higher, TE can be realized by using the shortest path calculated from the acquired costs. However, both methods employ link cost setting for TE, and so cannot satisfy QoS-oriented requirements for traffic routing. There usually are differences between TE paths and paths in a cost-based metric network, which carries out routing by using path that minimizes link costs given from outside to the link.

This paper studies partial traffic engineering (Partial TE)[8] to realize the two key goals of meeting network quality requirements and realizing traffic grooming. TE calculation is done under the condition that only some part of traffic are available for traffic grooming, while for the other traffic flows cost minimum paths are guaranteed. Some studies have shown the effectiveness of TE for metric-based traffic under certain conditions [9][10]. However, traffic that is routed according to a metric impacts TE, which may respond with path changes. This makes it difficult to maintain network quality. In the latter case, traffic dispersion within a network is not considered in routing QoS traffic, so a situation may arise where QoS traffic between particular nodes is zero, which means that there is no guarantee that realistic traffic requirements can be satisfied. On the other hand, as for Partial TE explained here, paths that carry metric-based traffic are fixed and their quality is guaranteed. Reference [8] proposed Partial TE to achieve a certain effect, but failed to detailed under what range of conditions the effect could be realized. We assume TE application would yield different results for the condition in which traffic converges onto a particular path and the condition in which traffic is

Manuscript received December 9, 2011.

Masashi Hashimoto is with Dept. of Computer Science and Engineering, Faculty of Engineering and Resource Science, Akita University, Akita, 010-8502 Japan. (e-mail: hashimoto@ie.akita-u.ac.jp).

Aki Fukuda was with Dept. of Computer Science and Engineering, Faculty of Engineering and Resource Science, Akita University, Akita, 010-8502 Japan.

distributed by ECMP (Equal Cost Multi-Path). Thus, this paper focuses on metric-based routing and traffic dispersion to elucidate the conditions under which that Partial TE can be expected. Specifically, we examine the effectiveness of Partial TE from the two points of metric and traffic non-uniformity. The effectiveness of Partial TE is readily confirmed by calculating the percentage of overall traffic that is impacted by TE and the routes selected.

We also studies traffic accommodation by Partial TE from the point of oversubscription. Oversubscription means the inputted traffic consists of the bottleneck traffic and oversubscription traffic. This means that the maximum traffic which can be accommodated using the sure cost minimized routes, and oversubscription traffic means further possible traffic in the network. In the oversubscription, congestion ratio is one and best effort traffic does not raise congestion ratio. Simulations are conducted to obtain the traffic gain possible with Oversubscription. The impact of topology is also considered.

In this paper, Section II describes the consideration on the metric-base network and its LP formulation, while the the Partial TE method and its evaluation procedure are explained in Section III. We show simulation results in Section IV. Section V reviews the results of oversubscription and our conclusion is stated in Section VI.

II. EVALUATION OF TRAFFIC CONCENTRATION IN THE COST METRIC NETWORK

In the cost metric network, traffic tends to concentrate on particular, cost minimum paths. We first examine this feature.

A. LP Formulations for Cost Metric Network

Figure 1 shows a network model with flexible route setting such as MPLS. We can assume that traffic between nodes is transferred through a route set that can be set explicitly. It is described below as a linear programming network issue. Network $G(V,E)$ is defined with node set V and link set E as shown follow.

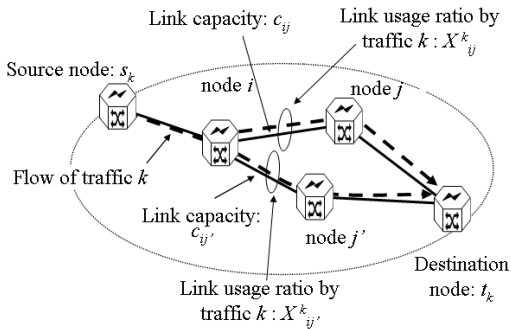


Fig.1 Network model.

- i, j nodes
- (i, j) directional link between node i and node j
- c_{ij} capacity of link (i, j)

- k, K each traffic demand and a set of traffic ($k \in K$)
- d_k bandwidth of traffic demand k
- s_k, t_k start node and destination node of traffic k
- X_{ij}^k ratio of traffic k transmitting through link (i, j)
- $cost_{ij}$ cost defined on link (i, j) , here, delay or hop

The formula for cost minimum routing is given as the eqn.(1). Eqn.(2),(3) are constraints on flow conservation[11][12].

$$\min \sum_{k \in K} \sum_{(i,j) \in E} cost_{ij} X_{ij}^k \quad (1)$$

subject to:

$$\sum_{j:(i,j) \in E} X_{ij}^k - \sum_{j:(i,j) \in E} X_{ji}^k = \begin{cases} 1, & i = s_k \\ 0, & i \neq s_k, t_k \end{cases} \quad (2)$$

$$\forall k \in K$$

$$0 \leq X_{ij}^k \leq 1, \forall (i,j) \in E, \forall k \in K \quad (3)$$

As shown in eqn.(3), the flow between any node pair is allowed to be split into one or more paths with arbitrary ratio. If there is unique cost minimum path, each flow transmits through that path, which means X_{ij}^k takes a value of 0 or 1.

Another important objective commonly used for path calculation is minimizing congestion ratio, which is the maximum link utilization ratio among all links. Its purpose is traffic grooming namely, TE, over the whole network. The objective of the TE is given in eqn.(4). The minimized α is a congestion ratio.

$$\min \alpha \quad (4)$$

$$\sum_{k \in K} d_k X_{ij}^k \leq \alpha c_{ij}, \quad \forall (i,j) \in E \quad (5)$$

$$\alpha \geq 0 \quad (6)$$

This optimization is very effective with regard to traffic grooming because it exploits all possible paths. However, cost metric routes are not generally same as the routes defined by TE.

When there are ECMPs and cost minimum routing having small congestion ratio is preferable, following optimization provides suitable solution.

$$\min \left(\sum_{k \in K} \sum_{(i,j) \in E} cost_{ij} X_{ij}^k + \alpha \right) \quad (7)$$

Here, the order of each term in eqn.(7) can be different, so we multiplied a certain constant number to make the first term larger than the second term when we actually calculated. We use eqn.(7) for calculation using hop metric¹.

¹ In calculation using delay metric, there is no ECMP. So, eqn.(1) and eqn.(7) give same routes.

B. Evaluation Conditions

We examine concentration of traffic feature using two real-world topologies of cost239[13] and nsfnet[14] shown in Fig.2. It is assumed all links have equal capacity and two cost metric types of hop count and delay are considered. For the delay on each link (i,j) , we refer to the distance between real existing cities for the real network model. Hop count as metric yields many ECMPs and is considered as a metric more even than delay. We use full mesh traffic matrix as the traffic demand. As traffic distribution among nodes, uniform traffic distribution and distribution by the gravity model are considered. In the traffic distribution with gravity model, traffic between each node is assumed to be proportional to population of each node[15]. We denote this model by gravity1. Gravity may include feature that is inversely proportional to square of the distance between each node[16]. We also consider this model denoted by gravity2. Distributions of gravity1 and gravity2 are nonuniform models and gravity2 has more nonuniformity than gravity1. We assume that this distribution is fixed in each evaluation by changing traffic volume and consider the average value of traffic matrix components.

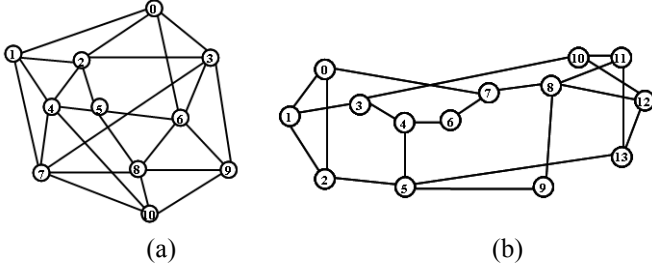


Fig.2 Network topologies, (a) cost239 and (b) nsfnet.

C. Evaluation Results

Table I shows slope of increase of congestion ratio α , when average traffic increases. Here, the set path accommodates the traffic given by formulas (3) or (7). We denoted this slope by β . Here, the β is normalized by TE calculation. So, β should be equal to or higher than one and value near one means that traffic is evenly distributed in the network like using TE. From table I, we can observe that for uniformly distributed traffic with HOP count as cost metric routing is very efficient since β in table I is close to one. These characteristics were outstanding for both uniform traffic dispersion (uniform) and non-uniform traffic dispersion (only gravity 1). When hop count is used as a metric, there are many ECMPs. That is a characteristic that is also attributed to topologies, but regarding cost239 in Figure 2 (a), for example, ECMP appears between cities with the minimum HOP count of 2. Traffic is branched off between ECMPs. However, this traffic uses the paths identified by minimization with Formula (7) and the branching is not uniform, unlike the OSPF case. In these cases, ECMPs having the shortest hop are fully utilized for TE purpose.

On the other hands, we can find β is much larger than one,

especially for β using delay as cost metric. In these cases, α becomes large more rapidly versus traffic increase. When the metric of delay, which is set by referring to the actual distance between cities, is used, there is only one shortest route between each pair of cities, and so traffic becomes concentrated on that path. Even when HOP count is used as the metric, when gravity 2, which has the most non-uniform traffic dispersion, is used, β takes a greater value. This indicates that effective traffic grooming observed in the case of hop count as cost metric for uniform traffic is lost with noneven metric and/or nonuniform traffic. So, some kind of the introduction of TE would be promising to increase network efficiency, especially when there is a large non uniformity in the network.

Table I Traffic Concentration Coefficient as Comparing to TE Calculation, β normalized slope by TE.

traffic distribution	cost239		nsfnet	
	metric		metric	
	hop count	delay	hop count	delay
uniform	1.01	3.22	1.06	1.88
gravity1	1.00	3.33	1.12	2.22
gravity2	1.71	2.29	1.92	2.55

III. PARTIAL TRAFFIC ENGINEERING OVER COST METRIC NETWORK

A. The Partial TE (Partial Traffic Engineering)

With the Partial TE, traffic is accommodated as described below[10].

- 1) Accommodate metric-base traffic using the shortest routes.
- 2) Use the remaining bandwidth to accommodate best-effort traffic with TE.

In this scheme, we assume each traffic (d_k) consists of metric-base traffic (denoted by d_k^{metric}) which uses only shortest route and TE-base traffic (denoted by d_k^{TE}) which is best-effort traffic. And we also assume that the ratio d_k^{metric} and d_k^{TE} to d_k is same among all traffic. And this ratio is represented by ρ and $(1-\rho)$. So, d_k is expressed as follows.

$$d_k = \rho d_k^{metric} + (1 - \rho) d_k^{TE} \quad (8)$$

B. Partial TE Evaluation

First, we calculate shortest routes for d_k^{metric} using eqns.(1) -(3), or (7) instead of (1), and denote obtained variable by Y_{ij}^k instead of X_{ij}^k . TE calculation is conducted under following condition (9) which is rewritten form of eqn.(5). Here, Y_{ij}^k is constant parameter and Z_{ij}^k is variable for routes of d_k^{TE} .

$$\sum_{k \in K} (1 - \rho) d_k Y_{ij}^k + \sum_{k \in K} (1 - \rho) Z_{ij}^k \leq \alpha c_{ij} \quad (9)$$

C. Evaluation with Average Traffic

We fix the traffic distribution and evaluate congestion rate α by changing QoS traffic proportion, ρ , against all traffic. The

average traffic volume, described as T_{total} , consists of metric-base traffic and best-effort traffic. T_{total} is an average value of each element of traffic matrix and traffic distribution is fixed when evaluation is done. So each traffic demand is determined by setting T_{total} . We assume that metric-base traffic volume and best-effort traffic volume are described as T_{metric} ($=\rho T_{total}$) and T_{TE} ($=(1-\rho)T_{total}$), respectively. Hereafter, those values show average values within the distribution when each traffic volume is considered.

We remark traffic for which $\alpha=1$ is true when $\rho=1$ as T_{metric}^{max} and traffic for which $\alpha=1$ is true when $\rho=0$ as T_{TE}^{max} , respectively. T_{metric}^{max} is the maximum value of traffic volume with the condition of using only the shortest routes, while T_{TE}^{max} is the maximum value of traffic volume that the network can accommodate. α and T_{total} have a linear relationship if ρ is fixed. When the relationship between α and ρ is evaluated, we only need to perform one evaluation using one specified T_{total} . In the evaluation of the relation between α and ρ , we use T_{metric}^{max} as T_{total} . That means we fix traffic T_{total} as T_{metric}^{max} and change the included QoS traffic, T_{metric} , by changing ρ , and evaluate the effectiveness as congestion rate, α , decreases.

To consider oversubscription, we denote T_{total} as T_{metric}^{max} plus T_{osub} . The oversubscription capability is estimated using the following optimization. When traffic of T_{metric}^{max} is accommodated in the minimum cost path, $\alpha=1$, which means that one or more bottleneck link(s) will appear. T_{osub} is the traffic amount that can be accommodated after maximizing QoS traffic.

$$\max T_{osub} \quad (10)$$

IV. SIMULATION RESULTS

A. Simulation Conditions

Simulation conditions are same as shown in evaluation conditions of II-B.

B. The Relation between α and ρ

Fig.3 (a) and (b) show the results using delay as cost metric for two topologies. These figures show the Partial TE method is very effective in reducing α especially for the case of delay as cost. In this case, for example, by using the traffic half of d_k as d_{metric} , α becomes 0.5. So, the congestion ratio becomes half even if half of the total traffic flows is guaranteed to transmit cost minimum paths. Furthermore, α is approximate to α in the value which obtained condition that ρ is 0. On the case of delay as cost with traffic of approximately uniform distribution, Reference [18] analyzed above characteristics into detail. In the case of hop count, the same trend is also observed for gravity2. The condition of $\rho=0$ means whole traffic is accommodated using TE calculation. And α has always the value of ρ or greater. With this meaning, in many examples, it can be said that we acquired fairly good results in terms of lowering the value of α to its limit. Fig.4 (a) and (b) show the results using hop count as cost metric for two topologies. In the case of hop count, but for traffic distribution of gravity2, significant effect by Partial

TE is not obtained. These results show that using hop count as metric for these distributions is enough effective from the point of traffic accommodation. Given that the effectiveness of Partial TE as shown in Figures 3 and 4, setting $\rho \leq 0.5$ or 0.6 yields a value of α similar to that obtained when all traffic is accommodated by TE calculation. This suggests that even if we fix 50% or 60 % of the traffic, as QoS traffic, to the minimum cost paths, there is small degradation in the congestion rate and accommodating efficiency rate is comparable to the case in which TE calculation is used.

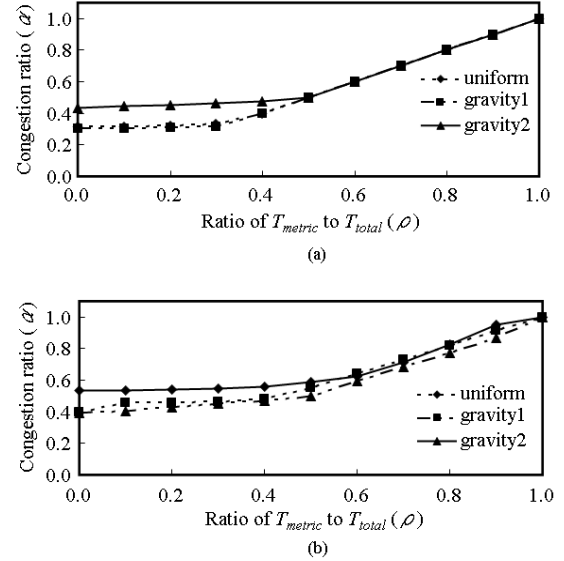


Fig.3 Congestion ratio (α) vs. a ratio of metric-base traffic (ρ) when cost is delay, (a) cost239, (b) nsfnct.

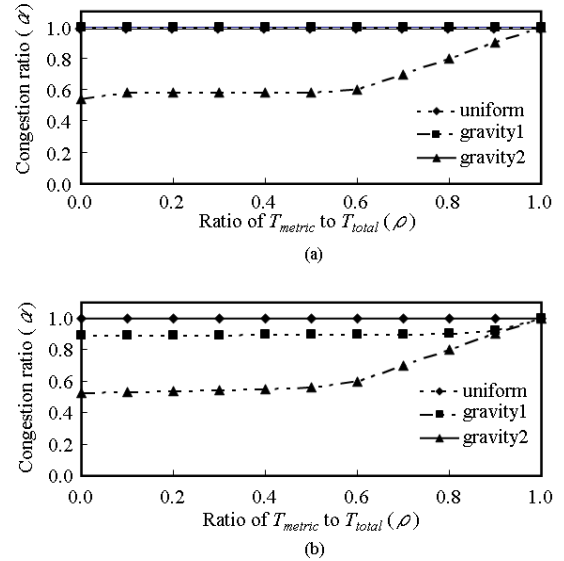


Fig.4 Congestion ratio (α) vs. a ratio of metric-base traffic (ρ) when cost is hop count, (a) cost239, (b) nsfnct.

C. Oversubscription Characteristics

Table II shows numerical results of capable T_{osub} and it is

normalized against T_{metric}^{max} . So, for example, value 0.71 means totally 1.71(=1+0.71) by T_{metric}^{max} can be accommodated.

For the cost239 topology, we can find that large additional traffic can be accommodated for delay metric. Using hop metric, only for gravity2 traffic, additional traffic can be also accommodated, but it is smaller than that using delay metric. For the nsfnets topology, expect the combination of hop count and gravity1, additional traffic can be accommodated but it is smaller than cost239 as a whole.

Table II Possible Oversubscription, T_{osub} .

	cost239		Nsfnet	
traffic distribution	Metric		Metric	
	hop count	delay	hop count	delay
uniform	0.00	1.97	0.22	0.69
gravity1	0.00	2.17	0.01	0.16
gravity2	0.71	1.09	0.76	1.01

V. DISCUSSION: POSSIBILITY OF OVERSUBSCRIPTION

A. From the Point of β and Early Link Bottlenecking

Comparing to TE calculation, congestion ratio increases and becomes 1 rapidly in many cases. Such inefficient traffic accommodation due to cost metric routing is considered to allow oversubscription. Figure 5 shows a relationship between α and traffic. There are two regions in traffic. In one region, whole traffic can use cost minimum routes, while in another region, traffic, which is additional traffic to T_{real}^{max} , is accommodated as oversubscription. In the case that oversubscription region in Fig.5 is large, expected T_{osub} is large, and in the case that oversubscription region is small, expected T_{osub} is small by nature, respectively. It is true for most cases. For example, in the calculations for cost239 with delay metric, we found above oversubscription region is large. In these cases, α becomes 1 more rapidly versus traffic increase, which means that at least one of the links has become fully occupied. So, we consider oversubscription possibility can be estimated by this early link bottlenecking characteristics.

But we found it was not true for the combination of gravity1 and delay metric in nsfnets topology. For this combination, although oversubscription region is large, whose β is 2.22 as shown in table I, but little oversubscription is allowed.

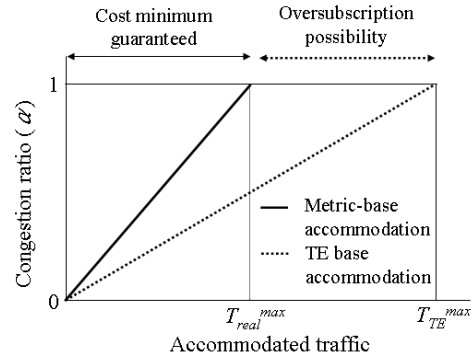


Fig. 5 Relation between congestion ratio vs. traffic having possible oversubscription region.

B. From the Point of Early Node Bottlenecking

We examined this case in detail, the reason is found to be node bottlenecking. Here, node bottlenecking means all of incoming/outgoing links to at least one node is fulfilled. We examine the maximum usage ratio of links from/to node in bottlenecking condition. This usage ratio is defined as total traffic from/to node divided by total capacity of the links from/to node. In the Tab.III, the maximum values of node bottlenecking among all nodes are shown. Here, the value of 1 means at least one node is isolated incoming or outgoing direction. From this table, for the combination of gravity1 and delay metric in nsfnets topology, T_{metric}^{max} makes the condition near the node bottlenecking and it is understood as reason for small oversubscription capability. In this case, one node is almost isolated because it is connected to network through only bottlenecking and approximately bottlenecking links. Figure 6 shows this node isolation. So, from β , we can roughly expect oversubscription, especially when topology has large node degree like cost239. But in the case that node degree is small, considering node bottlenecking provides more rigid estimation.

Table III Maximum Usage of Links from/to Node in Bottlenecking Condition.

traffic distribution	cost239	nsfnets
Uniform	0.50	0.71
Gravity1	0.35	0.99
Gravity2	0.53	0.76

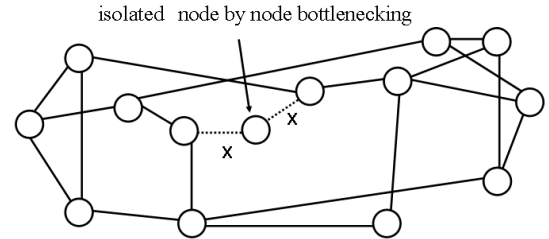


Fig. 6 Isolated node by node bottlenecking.

VI. CONCLUSION

In this paper, we clarified the efficiency of Partial TE from the two points of acceptable traffic range and acceptable metric range. We introduced an LP method, Partial TE, and evaluated its effectiveness in lowering the congestion rate under uniform and non-uniform traffic patterns with routing based on a cost metric. We demonstrated that its effectiveness strengthens with non-uniform traffic patterns. We obtained a satisfactory decrease in α without making ρ , the ratio of QoS traffic to total traffic, zero. Simulations confirmed its good performance when ρ is 0.5 or 0.6. Its effectiveness is limited under uniform traffic.

We also evaluated the case of oversubscription, which is highly likely with non-uniform traffic or metrics. Characteristics of the oversubscription expected can be roughly estimated from the level of concentration of metric-based traffic. However, the topology also influences the characteristics.

This paper confirmed that Partial TE (1) can guarantee the satisfaction of QoS traffic requirements through the use of metric-based routing as well as adequate TE effect, (2) can increase traffic capacity while maximizing support of QoS traffic, and (3) is most effective under non-uniform traffic and/or non-uniform metrics.

- [17] M. Hashimoto, E. Oki, "A Simplified Delay-Guaranteed Traffic Engineering Method," *Cyberjournals, Journal of Selected Areas in Telecommunications (JSAT)*, pp.1-8, Oct. 2011.

Masashi Hashimoto is an Associate Professor at Akita University, Akita, Japan. He received the B.E. degree in electrical engineering from Waseda University, Tokyo, Japan in 1982 and M.E. degree in electrical engineering from Kyoto University, Kyoto, Japan in 1984, respectively. He joined Nippon Telegram and Telephone Public Corporation (presently, NTT), 1984. From then, he has been engaged in the study and development of optical subscriber transmission system, and conducted research in optical parallel signal processing from 1987. He received a doctoral degree from Waseda University in 1995. He joined the Akita University, Akita, Japan, in April 2005. His present research interests involve traffic engineering and application flow control which suits to current and future Internet age. He is a member of IEICE, IPSJ of Japan and IEEE.

Aki Fukuda was received B.E. and M.E. degrees from Akita University, Japan, in 2007 and 2009, respectively.

REFERENCES

- [1] E. Rosen, A. Viswanathan, and R. Callon, "Multiprotocol Label Switching Architecture," IETF, RFC3031, January 2001.
- [2] D. Awduche, J. Malcolm, J. Agogbua, M. O'Dell, and J. McManus, "Requirements for Traffic Engineering Over MPLS," IETF, RFC2702, September 1999.
- [3] Y. Wang, and Z. Wang, "Explicit Routing Algorithms for Internet Traffic engineering," IEEE ICCCN 1999, pp.582-588, 1999.
- [4] Y. Wang, Z. Wang and L. Zhang, "Internet Traffic engineering without Full Mesh Overlaying," IEEE INFOCOM 2001, pp.565-571, 2001.
- [5] J. Moy, "OSPF Version 2", IETF RFC 2328.
- [6] D. Awduche, A. Chiu, A. Elwalid, I. Widjaja, and X. Xiao, "Overview and Principles of Internet Traffic Engineering," IETF, RFC3272, May 2002.
- [7] B. Thorup and M. Fortz, "Optimizing OSPF/IS-IS weights in a changing world," IEEE JSAC, Vol. 20, pp.756-767, 2002.
- [8] Masashi Hashimoto, Ken-ichi Yukimatsu, Takashi Miyamura, and Kohei Shiimoto, "Partial Traffic Engineering for Guaranteeing Network Quality", APSITT2008, April, pp.48-52, 2008.
- [9] Y. Lee, Y. Seok, and Y. Choi, "Traffic Engineering with Constrained Multipath Routing in MPLS Networks," IEICE TRANS., COMMUN., Vol.E87-B, No.5, pp.1346-1356, 2004.
- [10] D. Mitra and K. G. Ramakrishnan, "A Case Study of Multiservice, Multipriority Traffic Engineering Design for Data Networks," IEEE Globecom 1999, Part B, pp.1077-1083, 1999.
- [11] R. K. Ahuja, T. L. Magnanti, and J. B. Orlin, "Network Flows: Theory, Algorithm, and Applications," PRENTICE HALL, 1993.
- [12] E. Oki and A. Iwaki, "Performance of Optimal Routing by Pipe, Hose, and Intermediate Models," IEICE Trans. Commun., vol. E93-B, no. 5, pp.1180-1189, May 2010.
- [13] M.J. OMahony, M.C. Sinclair, and B. Mikac, "Ultra-high capacity optical transmission network: European research project cost 239," Information, Telecommunications, Automata Journal, vol.12, pp.33-45, June 1993.
- [14] X. Niu, W. D. Zhong, G. Shen, and T. H. Cheng, "Connection establishment of label switched paths in IP/MPLS over optical networks," Photonic Network Communications, vol.6, pp.33-41, July 2003.
- [15] M. Roughan, M. Thorup, and Y. Zhang, "Traffic Engineering with Estimated Traffic Matrices," IMC 2003, Oct. 2003.
- [16] W. J. Reilly, *Reilly's Law of Retail Gravitation*, Knickerbocker Press, 1931.