

Estimation of Traffic Amounts on all Links by Using the Information From a Subset of Nodes

Yuya Tarutani*, Yuichi Ohsita[†], Shin'ichi Arakawa*, and Masayuki Murata*

*Graduate School of Information Science and Technology, Osaka University, Osaka, Japan
{y-tarutn, arakawa, murata}@ist.osaka-u.ac.jp

[†]Graduate School of Economics, Osaka University, Osaka, Japan
y-ohsita@econ.osaka-u.ac.jp

Abstract—Traffic information is required to perform traffic engineering. However, as the network that require traffic engineering becomes large, the overhead to collect the traffic amount information required for traffic engineering becomes large. In this paper, we propose a method to reduce the overhead to collect the traffic amount information. In our method, we select a subset of nodes and collect the traffic amount information only from the selected nodes. Then, we estimate the traffic amount on each link by using the information collected from the selected nodes. According to simulation results, our method can estimate the traffic amount on each link required for traffic engineering accurately by monitoring 30% of all nodes.

Keywords-Estimation; Selection of Monitoring Nodes; Traffic Matrix; Traffic Engineering.

I. INTRODUCTION

In recent years, various applications are deployed and their traffic is carried over the Internet. The traffic in the Internet still doubles each year, the network providers are required to accommodate their traffic in a cost effective way.

Traffic engineering is a method to optimize the performance of networks by dynamically changing the topology and/or route of traffic [1-5]. The topology and route of traffic are calculated and controlled with a server called Path Computation Element (PCE). To perform the traffic engineering at the PCE, we may need to know the information of traffic in the network. Depending on the granularity of traffic engineering, we may require different degree of traffic information. For example, when we apply some optimization technique to determine the topology and route of traffic, we need the traffic matrix that expresses the traffic amount for each edge-to-edge traffic in the network.

However, collecting all the edge-to-edge traffic requires much overhead: one reason is the monitoring overhead at each router. Header inspection is necessary at the router to identify the edge-to-edge traffic that monitored packets belong to. However, the header inspection caused the overhead at each router; another reason is collecting overhead at the PCE server. To collect all the edge-to-edge traffic, the PCE server has to query the nodes monitoring edge-to-edge traffic and obtain the information of the traffic amounts. Especially, as the number of routers increases, the number of nodes the PCE server has to query and the size of information

to be collected become large, which causes the significant overhead at the PCE server.

To overcome these overheads, traffic engineering using only the information of traffic amount on each link has been investigated. Juva [5] calculates the range of each edge-to-edge traffic by using the information of traffic amount on each link, and optimizes the traffic routes that minimize the worst-case link utilization. Roughan et al. [2] and Ohsita et al. [3, 4] use the traffic matrices estimated from the information of traffic amount on each link. The traffic amount on a link can be easily counted at the node connected to the link, and the PCE can obtain the information of traffic amount on the link by querying the node.

However, the granularity of traffic information depends on the application of them. Recently, network virtualization [6] to support deployments of various network services, such as P2P services and cloud computing services, has been investigated. Virtual networks are prepared and reconfigured for each service. In this case, traffic engineering is required for each virtual networks and the PCE has to collect the information of traffic amount on each virtual link of each virtual network. It will be thought that diversification of the service in the network advances more and the number of virtual networks that require traffic engineering remarkably increases in future. That is, the number of links whose information must be collected by the PCE increases in future as the number of service increases. This may cause the heavy collecting overhead at the PCE.

In this paper, we propose a method to reduce the overhead to collect traffic information necessary for traffic engineering by estimating traffic amounts on all links from the traffic information collected from a subset of nodes. In our method, we first select the nodes we collect the traffic information from, and collect the information of traffic amount on each link from the selected nodes. Then, we estimate the traffic amounts of all links by using only the information collected from the selected nodes. Throughout this paper, we call the selected node *monitoring node*.

The rest of this paper is organized as follows. Section II explains the existing methods to estimate traffic matrices. In Section III, we propose a method to select monitoring nodes and estimate traffic amount on each link by using

the information collected from the monitoring nodes. In Section IV, we evaluate our method by simulation and clarify that our method can estimate traffic amount on each link accurately by selecting the monitoring node properly. Finally, Section V provides a conclusion.

II. OVERVIEW OF TRAFFIC MATRIX ESTIMATION

Traffic matrix is the matrix of $T_{s,d}$ that represents the traffic amount from node s to node d . Let N be the number of nodes in the network. Then, the traffic matrix is represented as,

$$T = \begin{bmatrix} T_{1,1} \\ T_{1,2} \\ \vdots \\ T_{N,N} \end{bmatrix}. \quad (1)$$

As this equation indicates, obtaining the traffic matrix requires the traffic information between all nodes and requires more overhead as the number of nodes increases. Therefore methods to estimate a traffic matrix from the traffic amount on each link have been investigated. The traffic amount on each link is determined from the routing information A , which is known to the network administrator, and traffic matrix T , which is unknown to the network administrator. That is, the following equation is hold;

$$AT = X, \quad (2)$$

where X is a matrix of X_i that represents the traffic amount that pass through the link i . That is,

$$X = \begin{bmatrix} X_1 \\ \vdots \\ X_L \end{bmatrix}. \quad (3)$$

In the above equation, L is the number of links in the networks. A is a matrix that has an element $A_{s,d,l}$ that represents the route of flow between node s and d and when the flow passes through the link l , $A_{s,d,l}$ takes one, otherwise takes zero. Note that when we consider the splittable flow, $A_{s,d,l}$ is the rate of end-to-end traffic $T_{s,d}$ flows the link l . A is called *routing matrix*.

$$A = \begin{bmatrix} A_{1,1,1} & A_{1,2,1} & \cdots & A_{N,N,1} \\ A_{1,1,2} & A_{1,2,2} & \cdots & A_{N,N,2} \\ \vdots & \vdots & \ddots & \vdots \\ A_{1,1,L} & A_{1,2,L} & \cdots & A_{N,N,L} \end{bmatrix} \quad (4)$$

Traffic matrix estimation is an approach to estimate T that satisfies the Eq. 2, based on the monitored traffic amount X and the routing matrix A . However, we cannot obtain the unique traffic matrix that satisfies Eq. 2 since the number of equations in Eq. 2 is usually less than the number of elements in T . That is, there are several candidates for the traffic matrix to satisfy the Eq. 2.

Many approaches have been considered to obtain the true traffic matrix from the candidates. One of approaches is to

use the model of traffic matrix [7-12]. Zhang et al. [11] proposed the estimation method called *tomogravity method* that estimates traffic matrix so as to follow the gravity model where the traffic amount between two nodes is proportional to the product of the traffic of the two nodes. The tomogravity method works as follows. At first, the method estimates the edge-to-edge traffic $T_{s,d}^{grav}$ based on monitored traffic in the ingress and egress links to follow the gravity model by the following equations;

$$T_{s,d}^{grav} = X_{i_s^{in}} \frac{X_d^{out}}{\sum_k X_k^{out}}, \quad (5)$$

where i_s^{in} is the ingress link at node s and j_d^{out} is the egress link at node d . We denote T^{grav} as the matrix in which each entry is $T_{s,d}^{grav}$. Then, the tomogravity method estimates a traffic matrix \hat{T} by following equations;

$$\begin{aligned} \min \|\hat{T} - T^{grav}\|, \\ \text{s.t. } A\hat{T} = X. \end{aligned} \quad (6)$$

That is, the tomogravity method calculates \hat{T} that satisfies Eq. 2 and minimize the difference between \hat{T} and T^{grav} . Although the traffic information required for the tomogravity method is much smaller than the case of collecting traffic matrix information directly, L numbers of traffic information are still required to be collected to estimate traffic matrices.

One approach to reduce the collecting overhead is to collect traffic amount information only from a subset of nodes and estimate the uncollected traffic amount information from the collected information. Zhang et al. [12] proposed a method to estimate the uncollected traffic amount information from the traffic amount information that was monitored and collected before at the same point or currently at the different points. In this method, we calculate the correlation between each traffic amount information monitored at different times or different points by using the traffic amount information collected before. Then, we estimate the uncollected traffic amount information by using the correlation. However, this method cannot estimate the uncollected traffic amount information accurately when the traffic change that is different from a past tendency occurs.

In this paper, we investigate the method to estimate the uncollected traffic amount information of each link from the information collected from a subset of monitoring nodes without using the past information. In addition, because the accuracy of the estimation depends on selection of the monitoring nodes, we also propose a method to select the monitoring nodes.

III. ESTIMATION OF THE TRAFFIC AMOUNTS ON ALL LINKS FROM THE INFORMATION OF A SUBSET OF NODES

In this section, we propose a method to estimate traffic amounts on all links by using the traffic amounts monitored at a subset of nodes. In addition, since the accuracy of the estimated traffic amounts depends on the selection of the

monitoring nodes, we also propose a method to select the monitoring nodes. Our method works as the following steps.

- Step. 1 Select monitoring nodes and collect the information of traffic amount from the selected monitoring nodes.
- Step. 2 Estimate the traffic amount on each link X' roughly by using the number of edge-to-edge traffic passing the link.
- Step. 3 Estimate the traffic matrix \hat{T} from the traffic amount on each link X' and the routing matrix A .
- Step. 4 Estimate the traffic amount on each link \hat{X} from the estimated traffic matrix \hat{T} and the routing matrix A .

After performing the above steps, we designate \hat{X} as the final estimation results for the traffic amount on each link. The details of the above steps are described below.

A. Selecting monitoring nodes

In this subsection, we propose a method to select monitoring nodes so as to estimate the traffic amounts on all links accurately. The edge-to-edge traffic whose amount is not monitored at any monitoring nodes is difficult to estimate and may also cause large estimation errors on the traffic amount on each link. Thus, in our method, we select monitoring nodes so as to cover as many edge-to-edge traffic as possible. In addition, when no nodes can increase the number of edge-to-edge traffic covered by the selected monitoring nodes, we select the nodes where the number of edge-to-edge traffic passing the node is the largest so as to increase the accuracy of as many edge-to-edge traffic as possible.

In our method, initially we regard all nodes as the candidates for the monitoring nodes. Then, we eliminate the selected nodes from the candidate until the number of remaining candidates becomes the target number of monitoring nodes H .

To select the nodes eliminated from the candidates for the monitoring nodes, we use the number of edge-to-edge traffic monitored by node i (Q_i), the number of edge-to-edge traffic that cannot be monitored at any other candidates than node i (P_i), and the number of candidates passed by the edge-to-edge traffic from node n to node m ($R_{n,m}$). Our method selects the monitoring nodes by the following steps.

- Step. 1.1 Select all nodes as candidates for the monitoring nodes .
- Step. 1.2 Initialize P_i to 0, Q_i to the number of edge-to-edge traffic passing the node i and $R_{n,m}$ to the number of nodes passed by the edge-to-edge traffic from node n to node m .
- Step. 1.3 If there exists the node whose P_i is 0, eliminate the node whose Q_i is the smallest among the candidates whose P_i is 0 from the candidates, and then go to Step 1.5. Otherwise, go to step 1.4.
- Step. 1.4 If $P_i > 0$ for all nodes, eliminate the node whose P_i is the smallest from the candidates.

Step. 1.5 If the number of candidates is larger than the threshold H , update $R_{n,m}$ and P_i for all candidates and go back to Step 1.3. Otherwise, go to step 1.6.

Step. 1.6 Designate the remaining candidates as the monitoring nodes.

In the Step. 1.5 of the above steps, $R_{n,m}$ is updated by decrementing its value if the edge-to-edge traffic from node n to node m passes the node eliminated from the candidates. Then, P_i is updated by counting the elements of $R_{n,m}$ where the edge-to-edge traffic from node n to node m passes through the node i and $R_{n,m} = 1$.

B. Estimation of traffic amounts by using the number of edge-to-edge traffic

In our method, we use only traffic amount information monitored at the selected monitoring nodes. However, the lack of traffic amount information causes the difficulty in estimating traffic matrices. Thus, we estimate the uncollected traffic amount information before estimating the traffic matrix.

To estimate the traffic amounts, we use the relation between the number of edge-to-edge traffic passing a link and the traffic amounts on the link. We investigate this relation by simulation. In this simulation, we use AT&T's router-level topology (523 nodes and 1304 links) measured in Ref. [13]. We add one ingress link and one egress link for all nodes in the AT&T topology, and generate traffic between each pair of ingress and egress links.

According to Ref. [11], actual traffic matrices follow the gravity model. In addition, according to Ref. [14], each element of actual traffic matrices obeys a lognormal distribution. Thus, in this simulation, we generate traffic matrix T indicating traffic amounts between each ingress and egress links so as to follow both the gravity model and a lognormal distribution. The traffic matrix T used in this simulation is generated as

$$T = T^{grav} + \Delta, \quad (7)$$

where T^{grav} is a traffic matrix generated so as to follow both the gravity model and a lognormal distribution, and Δ is a matrix indicating the white Gaussian noise with the mean of 0 and the variance of 1. We generate $T_{i,j}^{grav}$ as

$$T_{i,j}^{grav} = G_i * G_j, \quad (8)$$

where G_i is the weight for node i . We generate G_i based on the lognormal distribution with a mean of $e^{4.8}$ and the variance $e^{9.7}$ so as to match the results described in Ref. [14]. In this simulation, the unit of the traffic amount of the edge-to-edge traffic generated the above steps is Mbps.

Fig. 1 shows the relation between the number of edge-to-edge traffic passing a link and the traffic amount on the link obtained by our simulation. According to Fig. 1, we can model the relation as

$$W_i = \alpha Z_i + \beta, \quad (9)$$

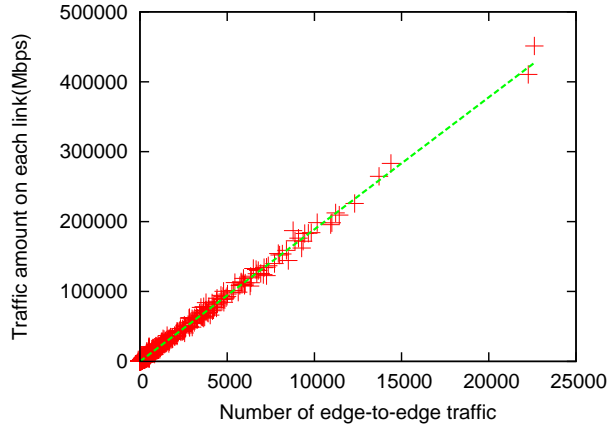


Figure 1. Relations of the number of edge-to-edge traffic and the traffic amount on each link

where W_i is the traffic amount of the link i , Z_i is the number of edge-to-edge traffic passing the link i , and α and β are the constant parameters. Z_i for any node i can be calculated from the routing matrix.

By using this relation, we estimate the traffic amount on each link as following steps. First, we calculate the constant parameters, α and β by using traffic amount on each link collected from the selected monitoring nodes. To calculate α and β , we use the least-square method. That is,

$$\alpha = \frac{|S| \sum_{i \in S} Z_i W_i - \sum_{i \in S} Z_i \sum_{i \in S} W_i}{|S| \sum_{i \in S} Z_i^2 - (\sum_{i \in S} Z_i)^2}, \quad (10)$$

$$\beta = \frac{\sum_{i \in S} Z_i \sum_{i \in S} W_i - \sum_{i \in S} Z_i W_i \sum_{i \in S} Z_i}{|S| \sum_{i \in S} Z_i^2 - (\sum_{i \in S} Z_i)^2}, \quad (11)$$

where S is the set of links connected to the monitoring nodes. Then, we estimate the traffic amount U_j on the link j that is not collected from the monitoring nodes as

$$U_j = \alpha Z_j + \beta. \quad (12)$$

Finally, we define the matrix X' which is a matrix indicating the roughly estimated traffic amount on each link as

$$X' = \begin{bmatrix} X'_1 \\ \vdots \\ X'_L \end{bmatrix}, \quad (13)$$

where

$$X'_l = \begin{cases} X_l & \text{if } l \text{ is the link connected to the monitored nodes,} \\ U_l & \text{otherwise.} \end{cases} \quad (14)$$

C. Estimating traffic matrices

We estimate the traffic matrix from the roughly estimated traffic amount on each link. If we apply the tomography method to estimate traffic matrix from the estimated traffic amount on each link, the estimation errors may become

large, because the estimation errors included in the traffic amounts on ingress and egress links cause the inaccurate estimation of T^{grav} and large estimation errors of the tomography method even when traffic amounts on other links are estimated accurately. Therefore, we need a traffic matrix estimation method where estimation errors included in the traffic amounts on particular links do not affect the estimation results significantly.

Though there may be more sophisticated estimation method, in our evaluation described in Section IV, we use the simple approach to estimate the traffic matrix by minimizing the following equation;

$$\min \|X' - A\hat{T}\|. \quad (15)$$

The results shown in Section IV clarifies that we can estimate the traffic amount on each link accurately even when we use this simple approach to estimate the traffic matrix.

D. Estimating traffic amount from estimated traffic matrices

Once we obtain the estimated traffic matrix \hat{T} , we calculate the matrix \hat{X} that represents the traffic amount on each link as

$$\hat{X} = A\hat{T}. \quad (16)$$

Then, we designate \hat{X} as the final estimation results for the traffic amount on each link. By estimating \hat{X} as Eq. 16, even when significant traffic changes on a small number of edge-to-edge traffic occurs and the changes are not captured by X' , \hat{X} may follow the traffic changes since \hat{X} is estimated so as to fit the current traffic amount information collected from the monitored nodes.

IV. NUMERICAL EVALUATIONS

In this section, we evaluate our method by simulation. In this evaluation, we use the same topology and traffic matrix as Section III-B

In this evaluation, we investigate the accuracy of the estimation of the traffic amount on each link, because the traffic amount on each link is important information for traffic engineering and the estimation errors of the traffic amount on each link may cause the misidentification of the congested links.

To evaluate the accuracy of the estimation of the traffic amount on each link, we use the Root Mean Squared Error (RMSE) and the Root Mean Squared Relative Error (RMSRE). The RMSRE (X_{RMSRE}) and the RMSE (X_{RMSE}) are defined as

$$X_{RMSRE} = \sqrt{\frac{1}{L} \sum_{k=1}^L \left(\frac{\hat{X}_k - X_k}{X_k} \right)^2}, \quad (17)$$

$$X_{RMSE} = \sqrt{\frac{1}{L} \sum_{k=1}^L (\hat{X}_k - X_k)^2}, \quad (18)$$

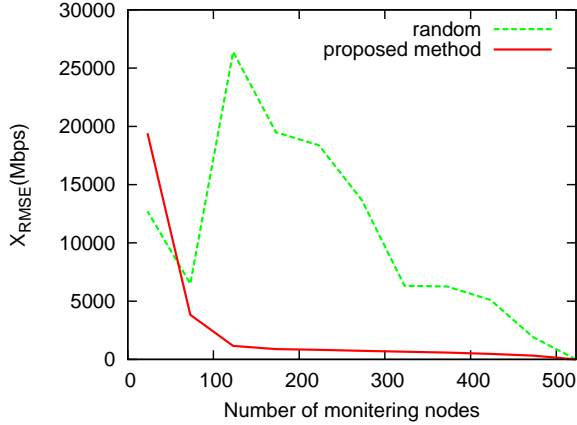


Figure 2. RMSE of traffic amount on each link

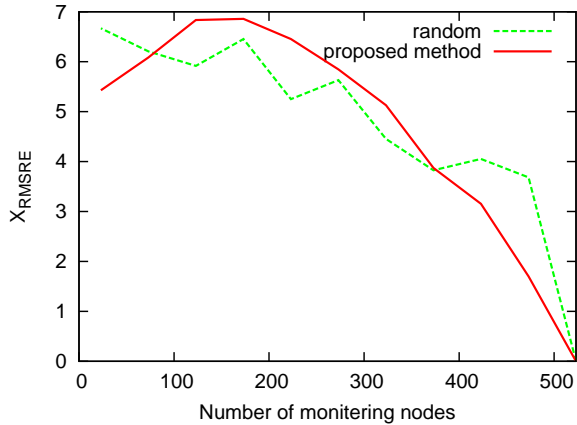


Figure 3. RMSRE of traffic amount on each link

where L is the number of links in the network, \hat{X}_k is the estimated traffic amount of link k , and X_k is the actual traffic amount of the link k .

Figures 2 and 3 show X_{RMSE} and X_{RMSRE} respectively when we change the number of monitoring nodes. In these figures, the vertical axis is X_{RMSE} or X_{RMSRE} , and the horizontal axis is the number of monitoring nodes. In these figures, “proposed method” indicates the case that we select the monitoring nodes by our method and “random” indicates the case that we select the monitoring nodes randomly.

According to Fig. 2, we can estimate the traffic amount of each link accurately by selecting more than 173 monitoring nodes, while the RMSE of the traffic amount on each link become significantly large if the number of monitoring nodes is less than 173 since the number of traffic amount information is too small to estimate the parameters of Eq. 12.

Fig. 2 also shows that we can estimate traffic amounts much more accurately in the case of selecting monitoring nodes by our method, compared with the case of selecting

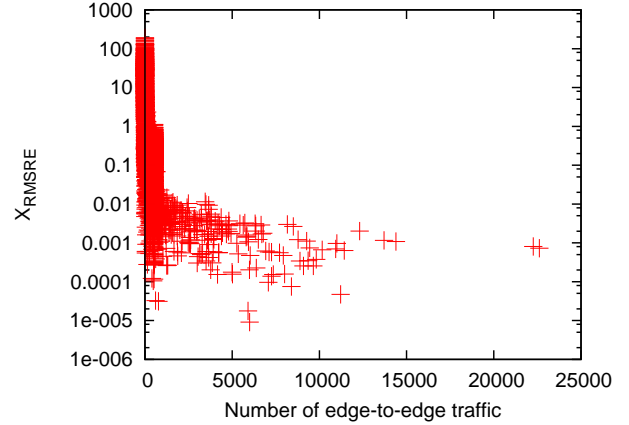


Figure 4. Relations of the number of edge-to-edge traffic and RMSRE when our method selects 173 nodes

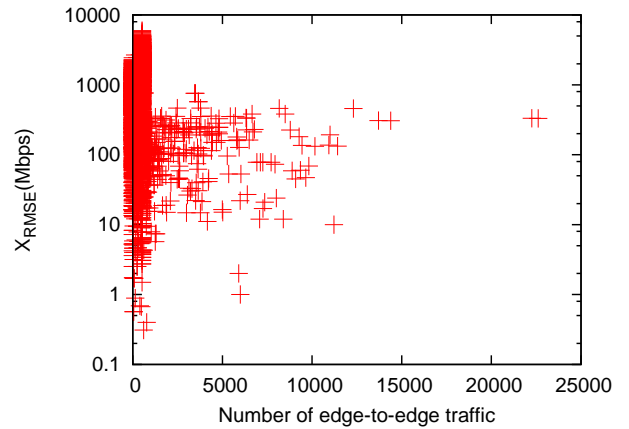


Figure 5. Relations of the number of edge-to-edge traffic and RMSE when our method selects 173 nodes

monitoring nodes randomly. This is because our method selects the monitoring nodes so as to cover as many edge-to-edge traffic as possible. Therefore, most of edge-to-edge traffic pass at least one of monitoring nodes selected in our method and can be estimated from the information of the traffic amounts collected from the monitoring nodes. On the other hand, in the case of selecting monitoring nodes randomly, several edge-to-edge traffics pass no monitoring nodes. Since we cannot obtain the traffic amount information corresponding to such edge-to-edge traffics from the monitoring nodes, such edge-to-edge traffics cannot be estimated accurately. As a result, the estimation errors of traffic amounts on links passed by such edge-to-edge traffic whose estimation error is large become also large.

However, according to Fig. 3, unlike the RMSE, the RMSRE of our method is still large and close to the RMSRE of the case of selecting monitoring nodes randomly. To investigate this in more detail, we show the estimation error of traffic amount on each link and discuss whose estimation

error is large.

Figs. 4 and 5 show the relations between the number of edge-to-edge traffic passing a link and RMSRE or RMSE of the traffic amount on the link when we select 173 monitoring nodes by our method. According to Fig. 4, the RMSREs only for the traffic amounts on the links where the number of edge-to-edge traffic is small become large. The actual traffic amount on the link where the number of edge-to-edge traffic is small may be small. The small actual traffic amount makes the value of relative error quite large even when the estimation error is not large.

In addition, according to Fig. 5, the estimation errors for the traffic amount on the links where the number of edge-to-edge traffic is small also become large. This is because one-hop traffics whose source and destination nodes are both ends of a link cannot be estimated from the traffic amounts monitored at any other links and their estimation errors become large. The ratio of one-hop traffic among the total traffic on the link increases as the number of edge-to-edge traffic passing the link becomes small. Thus, the estimation errors of one-hop traffic cause the large estimation errors of traffic amounts on the links where the number of edge-to-edge traffic is small.

However, routes of one-hop traffic are rarely changed by traffic engineering. In addition, the links where the number of edge-to-edge traffic is small are on the edge of the network. Thus, the estimation errors on the traffic amount on such links may have only little impact on the traffic engineering. According to Figs. 4 and 5, most of the traffic amount on each link passed by many edge-to-edge traffic can be estimated accurately. That is, our method can estimate the traffic amount on each link required for traffic engineering accurately by using only the information from a subset of nodes.

V. CONCLUSION

In this paper, we proposed a method to select the monitoring nodes and estimate the traffic amounts on all links from the traffic information collected from the selected monitoring nodes. Through the simulation, we clarified that our method can estimate the traffic amount on each link required for traffic engineering accurately by monitoring 30% of all nodes.

One of our future research topics is to evaluate the performance of traffic engineering using the traffic amount on each link estimated by our method.

ACKNOWLEDGMENTS

This work was partly supported by Grant-in-Aid for Scientific Research (B) 22300023 and Grant-in-Aid for Young Scientists (B) 21700074 of the Ministry of Education, Culture, Sports, Science and Technology in Japan.

REFERENCES

- [1] B. Fortz and M. Thorup, "Internet traffic engineering by optimizing OSPF weights," in *Proceedings of IEEE INFOCOM*, vol. 2, pp. 519–528, Mar. 2000.
- [2] M. Roughan, M. Thorup, and Y. Zhang, "Traffic engineering with estimated traffic matrices," in *Proceedings of the 3rd ACM SIGCOMM Conference on Internet Measurement*, pp. 248–258, Nov. 2003.
- [3] Y. Ohsita, T. Miyamura, S. Arakawa, S. Ata, E. Oki, K. Shiimoto, and M. Murata, "Gradually reconfiguring virtual network topologies based on estimated traffic matrices," *IEEE/ACM Transactions on Networking*, vol. 18, pp. 177–189, Feb. 2010.
- [4] Y. Ohsita, T. Miyamura, S. Awakawa, E. Oki, K. Shiimoto, and M. Murata, "Estimation of current traffic matrices from long-term traffic variations," *IEICE Transactions on Communications*, vol. E92-B, pp. 171–183, Jan. 2009.
- [5] I. Juva, "Robust load balancing," in *Proceedings of GLOBECOM*, pp. 2708–2713, Nov. 2007.
- [6] T. Miyamura, Y. Ohsita, E. Oki, S. Arakawa, Y. Koizumi, A. Masuda, K. Shiimoto, and M. Murata, "Network virtualization server for adaptive network control," in *Proceedings of 20th ITC Specialist Seminar on Network Virtualization - Concept and Performance Aspects*, May 2009.
- [7] Y. Vardi, "Network tomography: Estimating source-destination traffic intensities from link data.," *Journal of the American Statistical Association*, vol. 91, pp. 365–377, Mar. 1996.
- [8] J. Cao, D. Davis, S. Wiel, and B. Yu, "Time-varying network tomography: Router link data," *Journal of the American Statistical Association*, vol. 95, Feb. 2000.
- [9] I. Juva, S. Vaton, and J. Virtamo, "Quick traffic matrix estimation based on link count covariances," in *Proceedings of IEEE ICC*, vol. 2, pp. 603–608, June 2006.
- [10] A. Soule, A. Nucci, R. Cruz, E. Leonardi, and N. Taft, "Estimating dynamic traffic matrices by using viable routing changes," *IEEE/ACM Transactions on Networking*, vol. 15, pp. 485–498, June 2007.
- [11] Y. Zhang, M. Roughan, N. Duffield, and A. Greenberg, "Fast accurate computation of large-scale IP traffic matrices from link loads," *ACM SIGMETRICS Performance Evaluation Review*, vol. 31, pp. 206–217, June 2003.
- [12] Y. Zhang, M. Roughan, W. Willinger, and L. Qiu, "Spatio-temporal compressive sensing and Internet traffic matrices," *ACM SIGCOMM Computer Communication Review*, vol. 39, pp. 267–278, Aug. 2009.
- [13] 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.
- [14] 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.