Evaluation Indices of Many-to-Many Multicast Routing Tree to Represent Delay Performance

Hajime KARASAWA[†], Shinsuke TERADA[††], Takumi MIYOSHI[‡,†],

Kyoko YAMORI[†], and Yoshiaki TANAKA[†,‡‡]

[†]Global Information and Telecommunication Institute, Waseda University
1-3-10 Nishi-Waseda, Shinjuku-ku, Tokyo, 169-0051 Japan
[††]Graduate School of Information Science and Technology, The University of Tokyo
[‡]College of Systems Engineering, Shibaura Institute of Technology
[‡‡]Advanced Research Institute for Science and Engineering, Waseda University
E-mail: hajimekarasawa@toki.waseda.jp, terada@mcl.iis.u-tokyo.ac.jp,
miyoshi@se.shibaura-it.ac.jp, kyamori@waseda.jp, ytanaka@waseda.jp

Abstract — The optimization of multicast routing tree is an important issue for efficient multicast communication. Some previous papers have tried to analyze the one-to-many multicast tree and have obtained that the most efficient tree depends on the characteristics of the receivers such as the number of receivers. For many-to-many multicast, on the other hand, the efficiency of the tree has not been analyzed. This paper analyzes how various components affect the efficiency of the many-to-many multicast tree. As a result, it is shown that the most efficient many-to-many multicast tree does not depend on the number of multicast members. The result also tells that the diameter of the tree, the variance of node degrees, and the number of leaf nodes mostly decide the delay performance of the tree.

Index terms — multicast, many-to-many, multiple regression analysis

A. INTRODUCTION

Recently, many-to-many communication such as tele-conferencing or network game begins to spread. In such applications, they have to send and receive large data on real time. Therefore, a much wider bandwidth and a much lower delay are required. Multicasting is well suited for these applications since it makes efficient use of the network resources [1].

To realize an efficient multicast, the optimization of the multicast tree is an important issue. To invent a multicast routing algorithm, it is necessary to understand the characteristics of the multicast communication. Some previous researches have focused on such the analysis [2]. If a static multicast tree is used for one-to-many multicast, the most efficient tree depends on the characteristics of the receivers such as the number of receivers. On the other hand, as for many-to-many multicast, the efficiency of the tree has not been analyzed. This paper therefore analyzes how various components affect the delay performance of the many-to-many multicast tree using multiple regression analysis.

B. MANY-TO-MANY MULTICAST ROUTING

Various multicast routing algorithms have been suggested [1]. Multicast routing problems are categorized to either static or dynamic. In static routing problems, the members of the multicast group remain unchanged during the lifetime of the multicast connection. In dynamic routing problems, members can join or leave the group during the lifetime of connection. The multicast tree cost can be minimized by reconstruct tree dynamically [3],[4]. However, the dynamic routing causes a large processing overhead in the reconstruction of the multicast tree. Therefore, even the dynamic routing uses a static tree generally [5].

When using a static tree in one-to-many multicast, it is known that efficient trees are different depending on the characteristics of the receivers. If the number of receivers is small, shortest-path tree (SPT) is efficient. On the other hand, if the number of receiver is large, minimum-spanning tree (MST) is efficient [2],[4]. However, as for many-to-many multicast, the efficiency of the trees has not been analyzed.

In many-to-many multicast, the member of a multicast group has to send/receive large data to/from each other in real time. Therefore, the members have to send data efficiently with low delay. Two types of strategies to construct multicast trees have been proposed in many-to-many multicast routing. One uses different one-to-many multicast tree for every source. Since it can construct the most efficient tree for every source, the sources can send data with low delay. As the members increase, however, the number of the trees increases and many control messages are required. Thus, this method is not scalable.

The other strategy uses only one shared tree connected with all members. Though it has to filter unnecessary data on some application to avoid sending unnecessary data to all members, it needs only one shared tree, and it is scalable [6],[7]. Scalability is a very important factor for the spread of multicast communication [8], and accordingly it is desirable to use the shared tree strategy in many-to-many multicast routing.

C. ANALYTICAL MODEL

In this paper, we assume that one static shared tree is used for a many-to-many multicast communication. Many-to-many communication has to send and receive data in real time. Therefore, we assume that the lower-delay tree is a higher-performance tree. As a measure of performance, the maximum end-to-end hop counts are used. The networks are modeled as power-low networks. BRITE [9] is used for generating the network models. Table 1 shows the numbers of nodes and links in the network models and the methods of analyses. In the networks that have 8 to 15 nodes, we make all patterns of the shared trees that span all the nodes in the analyses. In the network with more than 15 nodes, it is hard to make all patterns of the shared trees and complete survey is difficult. We thus use a random sampling rather than all possible regression in the analysis. To be concrete, we pick out 100 patterns of the shared trees from the network that consists of 50 nodes and 97 links. The number of necessary samples in random sampling is clarified by testing population.

We calculate the following parameters:

(1) Average of max hop counts: The max hop count is the maximum end-to-end hop count along the shared tree among all node pairs in a multicast group. When the multicast group size is 3 to $|V|$-1, where $|V|$ represents the number of nodes, we calculate the average of max hop counts for each group size. For example, when the multicast group size is 3 in a 15-node network, we calculate the max hop counts from all patterns of choosing 3 nodes out of 15 nodes, and then calculate the average of them.

(2) Diameter of multicast tree: The maximum value of end-to-end hop counts of the shared tree.

(3) Variance of node degrees: A node degree is the number of links connected with each node. The variance of the node degrees shows the deflection of the node degrees in the shared tree.

(4) Number of leaf nodes: Leaf node is the node that has only one link. This index will show the degree of the spread of the shared tree.

Table 1. Network Model.

Number of nodes	Number of links	Method of analysis
8	13	Complete survey
9	15	Complete survey
10	17	Complete survey
11	19	Complete survey
12	21	Complete survey
13	23	Complete survey
14	25	Complete survey
15	27	Complete survey
50	97	Sampling survey

Though we analyzed three different networks for each network size, the obtained results were very similar. Therefore, we will show one of the obtained results for each network size. We also performed the same analyses in random networks [3] generated by BRITE, and obtained the similar results as those shown in the following sections.

D. CORRELATION ANALYSIS

Figure 1 shows the diameter and the average of max hop counts of every multicast tree, and Figure 2 shows the diameter, the variance of node degrees, and the number of leaf nodes of every multicast tree in an 8-node network. Figures 3 and 4 show the above parameters in case of a 50-node network. In Figures 1 to 4, sequential numbers are assigned to the multicast trees, and the tree number is sorted in ascending order with the average of max hop counts when multicast group size is 3.

From Figures 1 and 2, we can observe obvious characteristics: As the average of max hop counts increases, the tree diameter increases, and the variance of node degrees and the number of leaf nodes decrease. Even when the network size becomes large, from Figures 3 and 4, the same characteristics can be found. Thus, it is sure that those four parameters, the average of max hop counts, the diameter, the variance of node degrees, and the number of leaf nodes, have correlation. In addition, Figures 1 and 3 show the averages of max hop counts only in cases where the multicast group sizes are 3 and $|V|$-1 (7 and 49, respectively). Even in cases of other group sizes, the similar characteristics were obtained. Namely, the correlation coefficient between the average of max hop counts and the tree diameter is independent of the multicast group size. In Figures 3 and 4, we can find some fluctuations in observed values. These fluctuations occur due to random sampling analyses, and will be smaller if the number of analytical network models increases.

Figure 5 shows the correlation coefficients between the average of max hop counts and other parameters in an 8-node network: the diameter, the variance of node degrees, and the number of leaf nodes. Figures 6 and 7 show the above parameters in 15- and 50-node networks, respectively. In Figures 5 to 7, as the multicast group size becomes larger, the correlation coefficient with the tree diameter becomes stronger, the correlation coefficient with the variance of node degrees remains constant, and the correlation coefficient with the number of leaf nodes becomes slightly weak.

Figure 8 shows the correlation coefficients between the average of max hop counts and other parameters when multicast group size is 3 in every network size. Looking into Figures 5 to 8 simultaneously, the correlation coefficient between the average of max hop counts and the tree diameter is very strong in every network size. On the other hand, the correlation coefficients with the variance of node degrees and the number of leaf nodes become weaker
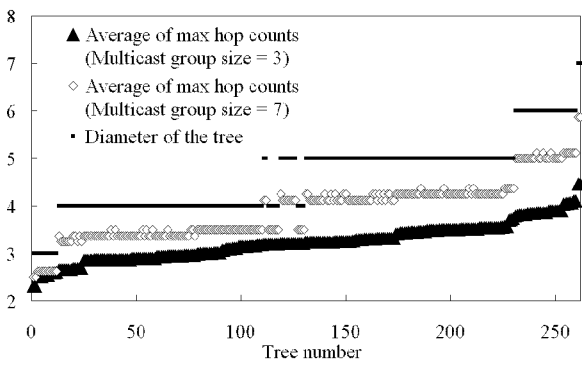
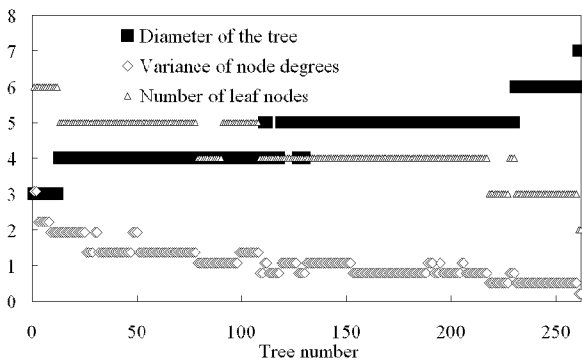Figure 1. Diameter and average of max hop counts of every tree in 8-node network.

Figure 2. Diameter, variance of node degrees, and number of leaf nodes of every tree in 8-node network.
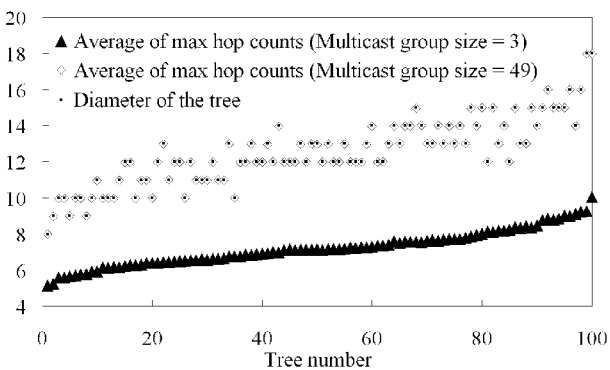

Figure 3. Diameter and average of max hop counts of every tree in 50-node network.
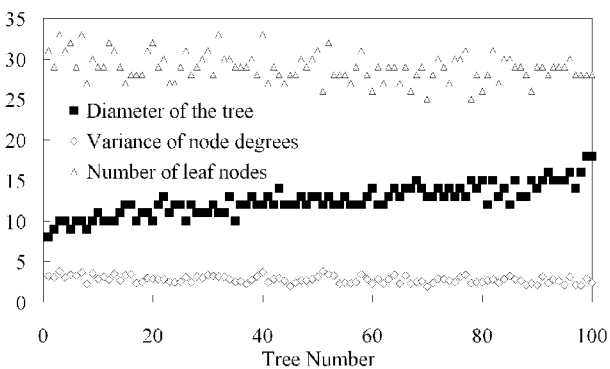

Figure 4. Diameter, variance of node degrees, and number of leaf nodes of every tree in 50-node network.
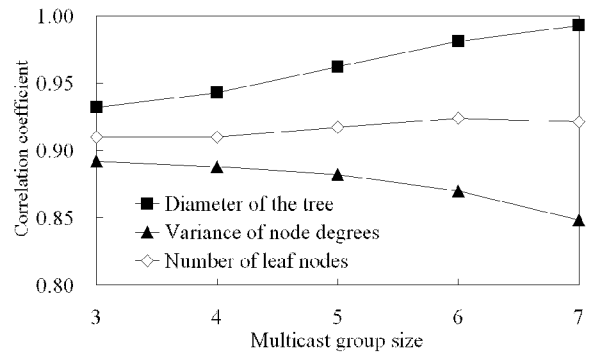

Figure 5. Correlation coefficients between average of max hop counts and other parameters in 8-node network.
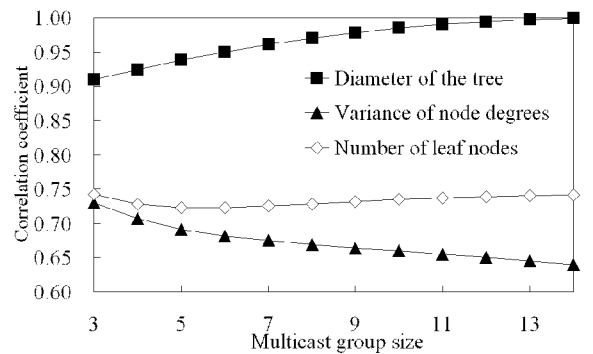

Figure 6. Correlation coefficients between average of max hop counts and other parameters in 15-node network.
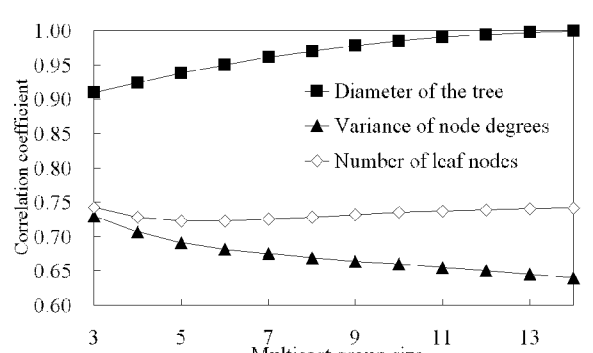

Figure 7. Correlation coefficients between average of max hop counts and other parameters in 50-node network.
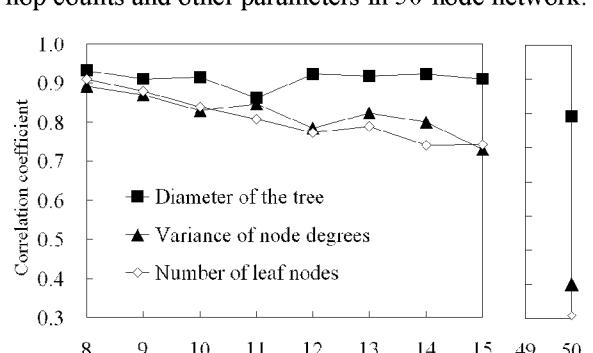

Figure 8. Correlation coefficients between average of max hop counts and other parameters in every network size when multicast group size is 3.
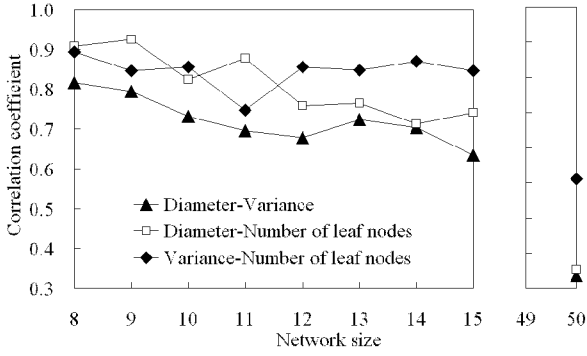
Figure 9. Correlation coefficients between each parameter pair in every network size.

as the network size is larger.

Figure 9 shows the correlation coefficients among the diameter of the tree, the variance of node degrees and the number of leaf nodes. From Figure 9, the correlation coefficients are strong in a small network, and those become weaker as the network become larger.

In Figure 9, we can find some fluctuations in observed values. These fluctuations will be smaller if the number of analytical network models increases.

E. MULTIPLE REGRESSION ANALYSIS

In multiple regression analysis, regarding the result as a criterion variable, and the factors as explanatory variables, find a regression equation shown in Eq. (1), which can clarify the criterion variable is affected by the explanatory variables.

$$y = a_1 x_1 + a_2 x_2 + \cdots + a_p x_p + a_0, \qquad (1)$$

where

y : criterion variable,

$x_1, x_2, \cdots, x_p$: explanatory variables,

p: the number of explanatory variables,

$a_1, a_2, \cdots, a_p$: partial regression coefficient, and

a_0 : residual.

In this study, let the average of max hop counts to be criterion variable, and let the diameter of the tree, the variance of node degrees, and the number of leaf nodes to be explanatory variables, as shown in Table 2. In the analyses, every explanatory variable is normalized by its maximum value.

Table 2. Regression equation.

Regression equation	$y = a_1 x_1 + a_2 x_2 + a_3 x_3 + a_0$
Criterion variable	y : Average of max hop counts
Explanatory variables	x_1 : Diameter of the tree
	x_2 : Variance of node degrees
	x_3 : Number of leaf nodes

Table 3. Result of analyses. (8-node network)

Group size	Proportion (%)	a_1	a_2	a_3	a_0
3	92.0	0.596	-0.380	-0.029	2.268
4	93.1	0.658	-0.355	0.004	2.088
5	95.2	0.725	-0.290	0.000	1.812
6	97.6	0.803	-0.201	-0.015	1.401
7	99.0	0.892	-0.103	-0.019	0.784

Table 4. Result of analyses. (15-node network)

Group size	Proportion (%)	a_1	a_2	a_3	a_0
3	87.1	0.792	-0.340	0.131	2.071
6	91.8	0.916	-0.222	0.144	0.851
9	96.2	0.964	-0.135	0.097	0.355
12	99.0	0.989	-0.062	0.047	0.130
14	99.9	0.998	-0.019	0.015	0.040

Table 5. Result of analyses. (50-node network)

Group size	Proportion (%)	a_1	a_2	a_3	a_0
3	80.6	0.869	-0.174	0.105	0.957
10	89.9	0.947	-0.107	0.107	-1.852
20	95.4	0.976	-0.070	0.066	-1.334
30	98.2	0.990	-0.042	0.038	-0.785
40	99.5	0.997	-0.018	0.017	-0.345
49	100.0	1.000	-0.002	0.002	-0.028

Table 3, 4, and 5 show the results of multiple regression analyses in an 8-, 15-, and 50-node networks, respectively. Every regression equation is estimated to be positive with a 5% significant level. Since the explanatory variables are normalized, the partial regression coefficients directly show how much accurately the explanatory variables explain the criterion variable. A negative partial regression coefficient shows the negative correlation coefficient between the criterion variable and its explanatory variable. In other words, the larger the explanatory variable is, the smaller the criterion variable is. A proportion shows how much accurately the regression equation explains the criterion variable.

Figure 10 shows the proportion of each size of multicast group in 8-, 12-, 15-, and 50-node networks. Looking into Tables 3, 4, and 5 and Figure 10, the regression equations always explain the criterion variable appropriately since every proportion is over 80%. Moreover, a larger multicast group size brings higher proportion. Thus, the average of max hop counts can be explained by three elements, the tree diameter, the variance of node degrees, and the number of leaf nodes. Every table leads us to the conclusion that the element that most strongly impacts the average of max hop counts is the tree diameter, since its partial regression coefficient is large. Furthermore, it is
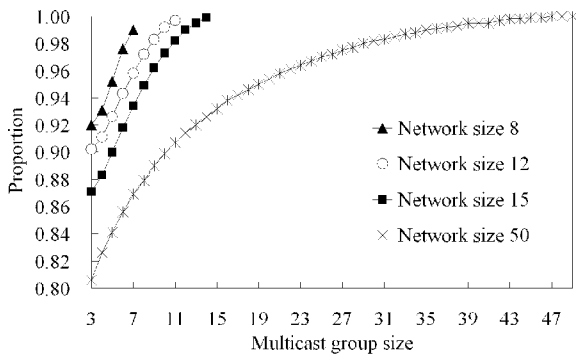
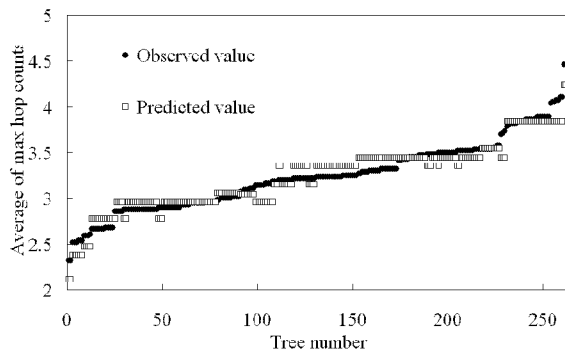Figure 10. Proportion vs. multicast group size in every network size.

Figure 11. Observed value and predicted value in 8-node network when multicast group size is 3.
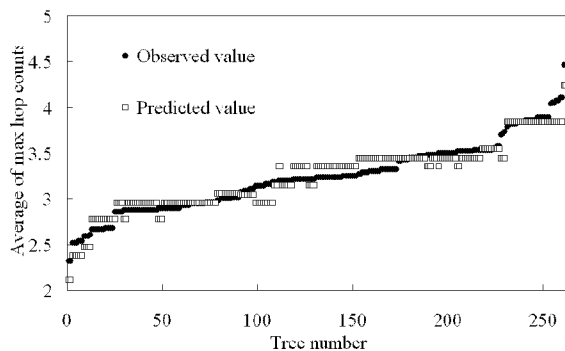

Figure 12. Observed value and predicted value in 50-node network when multicast group size is 3.

observed that the partial regression coefficients of both the variance of node degrees and the number of leaf nodes are slightly large when the multicast group size is small. This fact brings that these two elements also affect the average of max hop counts when multicast group size is small.

Figure 11 shows both the observed values and the predicted values when multicast size is 3 in an 8-node network, and Figure 12 shows them in a 50-node network in the same fashion. Looking into Figures 10, 11 and 12 at the same time, it is observed that the proportion becomes higher as the network size is smaller. Thus, the proposed regression equation more accurately explains the average of max hop counts when network size is smaller. Though the proportion is the lowest when the multicast group size

is 3, the predicted values are similar to the observed values. Therefore, we can conclude that the proposed regression equation well explains the average of max hop counts.

F. CONCLUSION

In this paper, we first clarified that the performance of a multicast tree shows the similar tendency in every multicast group size, in case where a shared multicast tree is used for many-to-many multicast communication. Then, we showed that the delay performance of the multicast tree can be accurately explained by three elements: the tree diameter, the variance of node degrees, and the number of leaf nodes. Moreover, we proposed a regression equation that explains the performance of each multicast tree, and evaluated it using multiple regression analyses. The obtained results showed that the proposed equation can accurately explain the average of max hop counts, and can apply to all size of networks.

It is left for further study to propose a novel multicast routing algorithm that suits many-to-many multicast communication using the result of the analyses in this study.

REFERENCES

[1] K. C. Almerth, "The evolution of multicast: from the MBone to interdomain multicast to Internet2 deployment," IEEE Network, pp.10-20, Jan./Feb. 2000.

[2] T. Miyoshi and M. Morimoto, "Performance estimation method for multicast routing trees and its evaluation," 5th Asia-Pacific Symposium on Information and Telecommun. Tech. (APSITT 2003), pp. 313-318, Nov. 2003.

[3] B. M. Waxman, "Routing of multipoint connections," IEEE J. Select. Areas Commun., vol.6, no.9, pp. 1617-1622, Dec. 1988.

[4] T. Asaka, T. Miyoshi, and Y. Tanaka, "Multicast routing algorithm using predetermined path search," IEICE Trans. Commun., vol.E83-B, no.5, pp.1128- 1135, May 2000.

[5] D. Estrin, D. Farinacci, A. Helmy, D. Thaler, S. Deering, M. Handley, V. Jacobson, C. Liu, P. Sharma, and L.Wei, "Protocol independent multicast – sparse mode (PIM-SM) : Protocol specification," RFC 2362, IETF, June 1998.

[6] A. Fei, "Constructing shared-tree for group multicast with QoS constraints," IEEE GLOBECOM 2001, vol.4, pp. 2389-2394, Nov. 2001.

[7] S. J. Koh, "Enhancement of the CBT multicast routing protocol," 8th Int. Conf. Parallel and Distributed Systems (ICPADS), pp.209-213, June 2001.

[8] C. Diot, B. N. Levine, B. Lyles, H. Kassem, and D. Balenseifen, "Deployment issues for the IP multicast service and architecture," IEEE Network, vol.1, no. 14, pp.78-88, Jan./Feb. 2000.

[9] "BRITE: Boston university Representative Internet Topology gEnerator," http://cs.bu.edu/brite/, 2005.