On the Stability of Autonomous Decentralized Flow Control in High-Speed Networks with Asymmetric Configurations

Chisa Takano†, Masaki Aida‡ and Shin-ichi Kuribayashi‡

† Traffic Engineering Division, NTT Advanced Technology Corporation
2-4-15, Naka-cho, Musashino-shi 180-0006, Japan
chisa@m.ieice.org

‡ NTT Information Sharing Platform Laboratories, NTT Corporation
Musashino-shi 180-8585, Japan
{aida.masaki,kuribayashi.shinichi}@lab.ntt.co.jp

Abstract

In high-speed network, the principle adopted for a time-sensitive flow control mechanism should be autonomous decentralized control. In this mechanism, each node in a network handles its local traffic flow only on the basis of the information it is aware of, although it is desirable that the decision-making at each node leads to high performance of the network as a whole. To implement this mechanism, we need first to investigate the relationship between the flow control mechanism of each node and network performance. In particular, the stability of the network performance is important. Our previous studies have proposed a simple mechanism for autonomous decentralized flow control and have shown that it has desirable stability. However, since we used a simple, symmetric, and homogeneous network model having a single bottleneck in the evaluation, the results cannot be generalized. In this paper, by using more complex network configurations in evaluations, we show that the principle of autonomous decentralized flow control can be applied to high-speed network with asymmetric configurations.

1 Introduction

In a high-speed network, propagation delay becomes the dominant factor in the transmission delay because the speed of light is an absolute constraint. Therefore, at any given time, a large amount of data is being propagated on links in the network (Fig. 1). The amount of such data is characterized by the *bandwidth-delay product*, *i.e.*, the propagation distance multiplied by the transmission rate. Therefore, in high-speed and/or long-distance transmission, there is more data in transit on the links than there is in the nodes.

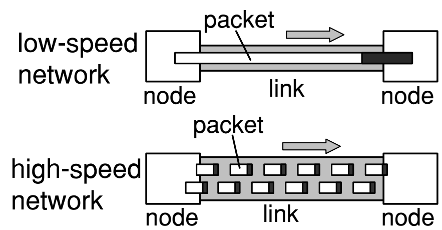

Figure 1. Effect of large bandwidth-delay product.

Since the state of the nodes changes rapidly in accordance with their clock speed, but the speed of light is constant, it is impossible to exert time-sensitive control based on collecting global information about the network. If we allow sufficient time to collect global information, the data so gathered is too old to apply to time-sensitive control. So, in a high-speed network, the principles adopted for time-sensitive control are inevitably autonomous decentralized systems [1, 2, 3].

This paper focuses on flow control realized as an autonomous decentralized system. In our model, nodes in networks handle their local traffic flow themselves based only on the information they are aware of. Since time-sensitive control in high-speed networks cannot collect global information about the network, nodes can use only restricted local information. We assume that each node can be aware of the following information: the distance between the node and adjacent nodes, the number of packets stored in the node at the present moment, and the feedback information that is received from the adjacent nodes. It is, of course, desirable that the decision making of each node should lead to high performance of the whole network. In flow control, we use the total throughput of a network as a global perfor-

Proceedings of the 23 rd International Conference on Distributed Computing Systems Workshops (ICDCSW'03)
0-7695-1921-0/03 $17.00 © 2003 IEEE

mance measure.

In our previous studies, we have investigated the behaviors of local packet flow and the global performance measure when a node is congested, and demonstrated an appropriate flow control model through simulation results [3]. In addition, we investigated the stability and adaptability of the network performance when the capacity of a link is changed [4].

However, since we used a simple, symmetric, and homogeneous network model having a single bottleneck in the evaluation, the results cannot be generalized. In this paper, by using more realistic network models in evaluations, we show that the framework of autonomous decentralized flow control can be applied to high-speed network with asymmetric configurations.

This paper is organized as follows. In Section 2, we discuss related works, comparing them with our work by categorizing flow control mechanisms. In Section 3, we describe a performance measure for the whole network and our flow control model. In Section 4, we show the way to set the values of the parameters used in our flow control model. This is required for applying to asymmetric network models. In Section 5, we describe two simulation models and the corresponding results. One simulation model has two bottlenecks and the other model includes links which differ in length. From the evaluations, our flow control model is shown also to be effective in high-speed networks with complex and asymmetric configurations. Finally, we conclude this paper in Section 6.

2 Related Works

In general, the mechanism used for flow control for high-speed networks should satisfy the following requirements:

1. With regard to the collection of information: it must be possible to collect the information used in the control.

2. With regard to the delay in applying control: the control should take effect immediately.

There are many other papers that study optimization of flow control problems in a framework of solving linear programs [5, 6, 7, 8, 9]. Their works assume the collection of global information about the network, but it is impossible to realize such a centralized control mechanism in high-speed networks. In addition, solving these optimization problems requires enough time to be available for calculation, and so it is difficult to apply them to decision-making in a very short time-scale.

Decentralized flow control by end hosts including TCP is widely used in the current networks, and there is much research in this area [8, 9]. However, since the end-to-end or the end-to-node control does not apply to decision-making in a time-scale shorter than the round-trip delay, it is insufficient to apply to decision-making in very short time-scale.

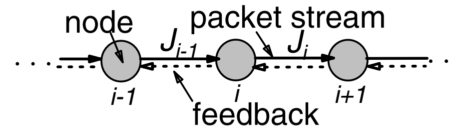

Figure 2. Interaction between nodes.

In our previous studies, we have investigated the characteristics of autonomous decentralized flow control in a high-speed network, considering the two above requirements [1, 2, 3]. We proposed a simple and effective method of flow control in [3] and investigated its stability and adaptability in [4].

3 Flow Control Model

3.1 Performance Measure

Each packet in a network is either in a node or on a link. Since the packets currently stored in nodes are not being transmitted over the network, it is natural to define the total throughput of the network as a global performance measure as follows. We define the total throughput of a network at time t as the amount of data being propagated on the network [1, 2, 3, 10]. In other words, it is the number of packets being propagated on all links in the network at time t.

On the other hand, the only packets we can control are those stored in nodes, and not those being propagated. Thus, higher performance of the whole network involves many uncontrollable packets being propagated on links. Therefore, inappropriate flow control cannot produce a state that has high performance and stability.

3.2 Node Model

Figure 2 shows the interaction of our flow control between nodes using the network model with a simple 1-dimensional configuration. All nodes have two incoming links and two outgoing ones for a one-way packet stream and for feedback information, that is, node i ($i = 0, 1, 2, \dots$) transfers packets to node $i+1$ and node $i+1$ sends feedback information (node information) to node i. For simplicity, we assume that packets have a fixed length in bits.

All nodes are capable of receiving and sending node information from/to adjacent downstream and upstream nodes, respectively. Each node i can receive node information sent from the downstream node $i+1$, and can send the node information about node i itself to the upstream node $i-1$. When node i receives node information from downstream node $i+1$, it determines the transmission rate for packets to the downstream node $i+1$ using the received node information and adjusts its transmission rate towards the downstream node $i+1$. The framework of node behavior and flow control is summarized as follows:

- Each node i autonomously determines the transmission rate J_i based only on information it is aware of, that is, the node information obtained from the downstream node $i+1$ and its own node information.
- The rule for determining the transmission rate is the same for all nodes.
- Each node i adjusts its transmission rate towards the downstream node $i+1$ to J_i . (If there are no packets in node i, the packet transmission rate is 0.)
- Each node i autonomously creates node information according to a predefined rule and sends it to the upstream node $i-1$.
- The rule for creating the node information is the same for all nodes.
- Packets and node information both experience the same propagation delay.

As mentioned above, the framework of our flow control model involves both autonomous decision-making by each node and interaction between adjacent nodes. There is no centralized control mechanism in the network. More precisely, it is impossible to realize centralized control in a high-speed network environment.

3.3 Diffusion-Type Flow Control Mechanism

This subsection briefly reviews our flow control model described in [3, 4]. In this paper, we focus on the stability of flow control in the congested state, and we consider packet flow in a heavy-traffic environment. The packet flow is defined as the number of sent packets per unit of time, and it is the same as the transmission rate toward the downstream node in a heavy-traffic environment. That is, we let the packet flow be $J_i(t)$ if the transmission rate specified by node i is $J_i(t)$. This is because node i has sufficient packets to transfer. Hereafter, we identify the packet flow with the transmission rate specified by the node.

The packet flow $J_i(t)$ should be controlled by the behavior of node i in the framework described in Sec. 3.2. This means the packet flow can be expressed using the node information obtained from the downstream node $i+1$ and its own node information. We define the packet flow as

$$J_i(t) := \alpha\, r_i(t-d_i) - D\,(n_{i+1}(t-d_i) - n_i(t)), \quad (1)$$

where $n_i(t)$ denotes the number of packets in node i at time t, $r_i(t-d_i)$ is the target transmission rate specified by the downstream node $i+1$ as node information, $\alpha\ (>0)$ and $D\ (>0)$ are constants, and d_i denotes the propagation delay between node i and node $i+1$. In addition, $(r_i(t-d_i), n_{i+1}(t-d_i))$ is notified from the downstream node $i+1$ with the propagation delay d_i. The first term on the right hand side of Eq. (1) reflects the target rate specified by the downstream node, and the second term is proportional to the gradient of the packet density. We call α and D the flow intensity multiplier and the diffusion coefficient, respectively.

If there is no packet loss in the network, the temporal variation of $n_i(t)$ is expressed as

$$n_i(t+\epsilon) - n_i(t) = \epsilon\,[J_{i-1}(t-d_{i-1}) - J_i(t)]\,, \quad (2)$$

where $\epsilon > 0$ is a small number.

To estimate the temporal variation roughly, we replace i with x and apply continuous approximation. Then the propagation delay becomes $d_i \to 0$ for all i and the packet flow is expressed as

$$J(x,t) = \alpha\, r(x,t) - D\,\frac{\partial n(x,t)}{\partial x}, \quad (3)$$

and the temporal variation of the number of packets at x is expressed as a diffusion type equation,

$$\frac{\partial n(x,t)}{\partial t} = -\alpha\,\frac{\partial r(x,t)}{\partial x} + D\,\frac{\partial^2 n(x,t)}{\partial x^2}, \quad (4)$$

by using the continuous equation

$$\frac{\partial n(x,t)}{\partial t} + \frac{\partial J(x,t)}{\partial x} = 0. \quad (5)$$

That is, our method aims to perform flow control using the analogy of a diffusion phenomenon. We can expect that packets in the congested node to be distributed to the whole network and normal network conditions to be restored after some time.

In our diffusion-type flow control, node i's packet transmission rate to the downstream node $i+1$ is determined as

$$J_i(t) = \alpha\, r_i(t-d_i) - D_i\,(n_{i+1}(t-d_i) - n_i(t)), \quad (6)$$

where the diffusion coefficient D in Eq. (1) has been replaced with D_i, which depends on i. The meaning of this replacement is revealed in the next section.

In addition, the node information of node i sent to the upstream node $i-1$ is determined as

$$r_{i-1}(t) = J_i(t). \quad (7)$$

In the framework of Eqs. (6) and (7), the node information of i specified to the upstream node $i-1$ is a pair of values $(r_{i-1}(t),\, n_i(t))$.

4 Parameter Tuning for the Diffusion-Type Flow Control

The diffusion-type flow control includes two parameters: the flow intensity multiplier α and the diffusion coefficient D_i. These parameters should be appropriately determined based on only the local information that node i is aware of.

In [3, 4], the values of α and D_i are chosen to be independent of i. Diffusion-type flow control with these parameters exhibits appropriate performance for simple and symmetric network. However, since all the links in such networks are the same lengths, it is an unrealistic situation. Therefore, we consider an appropriate way to determine the values of the parameters for networks with links of different lengths.

Let us consider again the packet flow by continuous approximation at a certain point $x = 0$,

$$J(0,t) = \alpha\, r(0,t) - D\,\frac{\partial n(0,t)}{\partial x}. \tag{8}$$

If we change the scale of x as $x \to \xi = a\,x$, $(a > 0)$, then we have

$$J(0,t) = \alpha\, r(0,t) - D\,\frac{\partial n(0,t)}{\partial \xi}. \tag{9}$$

This operation corresponds to changing the distance between nodes by using a scale parameter a. Since the rule for determining the transmission rate should be the same for all nodes, the form of Eq. (9) is unchanged under $x \to \xi$. Therefore, by choosing $D' = D/a$, we obtain

$$J(0,t) = \alpha\, r(0,t) - D'\,\frac{\partial n(0,t)}{\partial x}. \tag{10}$$

This implies that, for the propagation delay d_i between node i and node $i+1$, we should choose the diffusion coefficient as

$$D_i \propto \frac{1}{d_i}, \tag{11}$$

and the flow intensity multiplier α is independent of d_i.

For the flow intensity multiplier α, [4] shows that the value of $\alpha \geq 1.0$ enables stability of the performance.

5 Simulation

In our simulation study, we consider the case where the capacity of a link in the network is suddenly reduced to a narrow bandwidth. This situation occurs when there is an accident to links, an occupancy of bandwidth caused by background traffic, and so on. For example, Fig. 3 illustrates a change of available bandwidth influenced by cross traffic. In high-speed networks, no node is aware of the change of the link state and the new capacity of the link. We investigate the performance and stability of the diffusion-type flow control model through observations of the total throughput of the network.

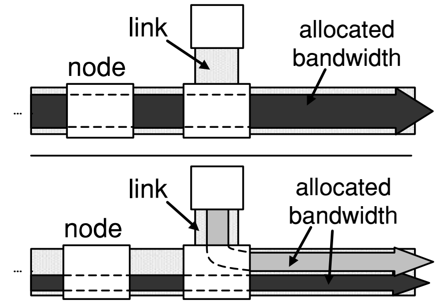

Figure 3. Example of a bottleneck link.

Although the diffusion-type flow control mechanism described in the previous sections achieves both desirable performance and stability, those results are obtained from simulations using the simple, symmetric, and homogeneous network model, as shown in [3, 4]. The network models used in [3, 4] have just one bottleneck link and the lengths of all links are the same. In this section, we investigate the effectiveness of the diffusion-type flow control mechanism through simulation using two different types of network model, which are more complicated than the model used in [3, 4]. The first set of models include two bottleneck links and the second set of models include links which differ in length (in other words, different propagation delays).

5.1 Evaluation for Network with Multiple Bottlenecks

In this subsection, we consider network models containing two narrow bandwidth bottleneck links, and investigate the performance of the diffusion-type flow control mechanism.

5.1.1 Network Model and Simulation Conditions

Figures 4 and 5 show our network models, which are closed networks with 1-dimensional configurations and toroidal boundaries. Both models have two congested nodes and corresponding two bottleneck links in the network. In network *model 1*, the congested nodes are located at opposite sides from each other as shown in Fig. 4, and in network *model 2*, they are configured at adjacent positions as shown in Fig. 5. All the other nodes and links are in the same condition as each other. These models simulate the situations in which congestion occurs at the congested nodes.

Detailed conditions of our network models are listed below:

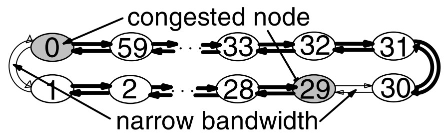

Figure 4. Network model (*model 1*) with two bottleneck links ($i = 0$ and 29).

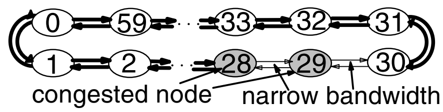

Figure 5. Network model (*model 2*) with two bottleneck links ($i = 29$ and 30).

- Number of nodes: $m = 60$. Each node is specified by i (mod 60).
- Propagation delay between adjacent nodes for all links: 1.0 (unit time).
- Index of the congested nodes:

  model 1 : $i = 0$ and 29.

  model 2 : $i = 29$ and 30.
- Total number of packets in the network: $N = 6000$.
- Maximum number of packets on a link (except the bottleneck link): $L_c = 100$.
- Maximum number of packets on the bottleneck link between adjacent nodes ($i = 0$ and 29 for *model 1* or $i = 29$ and 30 for *model 2*): $L_b = 25$, 50, or 75 (that is, 1/4, 1/2, or 3/4 of the bandwidth of other links, and having the same length).

To investigate the stability under congestion, in addition to the above conditions, we set the initial condition for congested nodes as follows:

- Number of packets in each of two congested nodes at time $t = 0$: 400.
- The other 5200 packets are randomly distributed amongst the other nodes and the other links.

As a model for the diffusion-type flow control, Eqs. (6) and (7), we set $D_i = 0.1$ and use the following flow control model. Since the packet flow is restricted by the link capacity, the diffusion-type flow control is expressed as follows:

$$J_i(t) = \min(\max(\tilde{J}_i(t), 0), L_i(t)), \tag{12}$$

$$r_{i-1} = J_i, \tag{13}$$

where

$$\tilde{J}_i = \alpha\, r_i - D_i\,(n_{i+1} - n_i) \tag{14}$$

$$L_i = \begin{cases} L_b, & (i = 0 \text{ and } 29,\ 29 \text{ and } 30), \\ L_c, & (\text{otherwise}), \end{cases} \tag{15}$$

$$\alpha = 1.0. \tag{16}$$

5.1.2 Simulation Results: Total Throughput and Its Stability

From the simulation results, we discuss the performance and stability of the diffusion-type flow control model, through observations of the total throughput.

Figure 6 shows the total throughput for *model 1* and *model 2* where the capacities of the two bottleneck links are the same, and take three different values, $L_b = 25$, 50, or 75. The horizontal axis denotes the simulation time and the vertical axis denotes the total throughput (*i.e.*, the total number of packets being propagated on links). The results show that, diffusion-type flow control achieves stable total throughput of the network, irrespective of the locations of the two bottleneck links and of the value of L_b. From the quantitative point of view, for the case where the link capacity of the bottleneck links $L_b = 50$, the maximum value of the sustainable total throughput (the maximum number of packets propagated stably on the links) is 3000, *i.e.*, 50 packets/link $\times$ 60 links. Thus, the diffusion-type flow control achieves 77% of the maximum value of the total throughput and its value is stable for both *model 1* and *model 2*.

Similarly, Fig. 7 shows the total throughput for *model 1* and *model 2* where the capacities of the two bottleneck links are different. For both *model 1* and *model 2*, we choose the pair of L_bs as (25,50), (25,75) or (50,75). Both graphs in this figure show that the diffusion-type flow control also achieves high performance and stability, in cases where the bottleneck links have different capacities. The value of total throughput is seen to be determined by the lower capacity in the pair of L_bs and it is almost the same as the value shown in Fig. 6

In our flow control model, although no node is aware of the bandwidth of the bottleneck link, high performance and stability are achieved without dependence on the locations of the bottleneck links.

5.2 Evaluation for Network with Different Link Lengths

In this subsection, we consider network models with different propagation delays between adjacent nodes and investigate the performance of the flow control using these models.

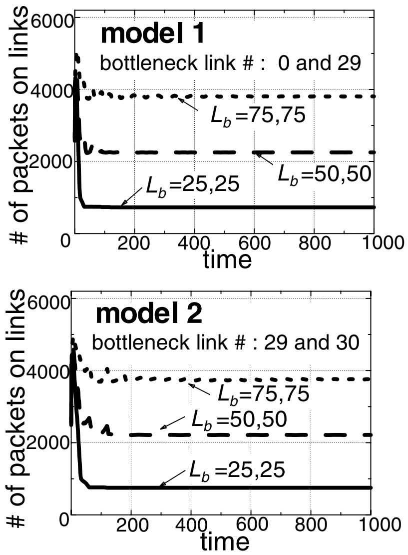

Figure 6. Total throughput of the network in *model 1* and *model 2* (in the case where the two values of L_b are equal).

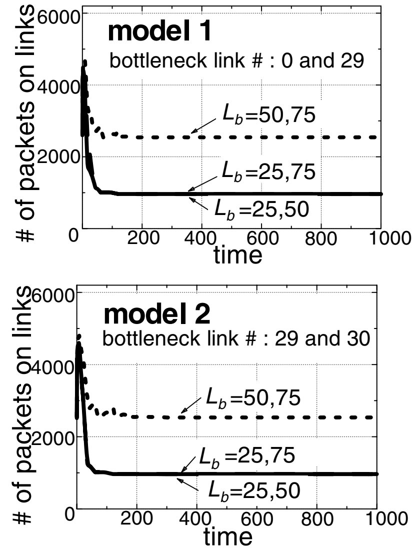

Figure 7. Total throughput of the network in *model 1* and *model 2* (in the case where the two values of L_b are different).

5.2.1 Network Models and Simulation Conditions

The simulation conditions are almost the same as those in the previous Sec. 5.1.1 except for the length of each link and the number of bottleneck links. The differences are as follows:

- Propagation delays between adjacent nodes:

  model 3: Each distance between the node and the adjacent node is short/long (that is, small/large propagation delay), where the ratio of the length of the short links to that of the long links is 1:50. Each link state appears with the probability $1/2$. So, we choose 0.04 or 1.96 for the propagation delays of the 60 links so that the mean propagation delay is 1.0 and the variance is 0.94.

  model 4: The length of each link follows a lognormal distribution, where the mean delay in the network model is 1.0 and the variance is 5.2.

- Index of the (single) congested node: $i = 29$.

- Number of packets in two congested nodes at time $t = 0$: 400.

- The other 5600 packets are randomly distributed amongst the other nodes and the other links.

This model is shown in Fig. 8.

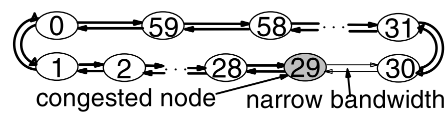

Figure 8. Network models (*model 3* and *model 4*) with different propagation delays between adjacent nodes.

As a model for the diffusion-type flow control, we use the flow control model Eqs. (12) and (13), where

$$\tilde{J}_i = \alpha\, r_i - D_i\,(n_{i+1} - n_i) \tag{17}$$

$$L_i = \begin{cases} L_b, & (i = 29), \\ L_c, & (\text{otherwise}), \end{cases} \tag{18}$$

$$\alpha = 1.0. \tag{19}$$

The value of the diffusion coefficient D_i is explained in the following subsection.

5.2.2 Simulation Results: Total Throughput and Stability

From the simulation results, we discuss the performance and stability of the diffusion-type flow control model for *model 3* and *model 4*, through observations of the total throughput. In this discussion, the value of the diffusion coefficient D_i shown in Sec. 4 is justified.

Figure 9 shows the total throughput for *model 3* and *model 4*, where the horizontal axis denotes the simulation time and the vertical axis denotes the total throughput. The

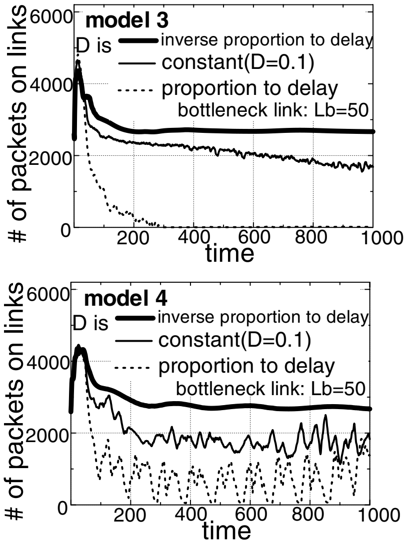

Figure 9. Total throughput of the network (*model 3* and *model 4*) (in the case where the delays between the adjacent nodes are different).

three lines in this figure show the results from the diffusion-type flow control models with

$$D_i = 0.1 \times \frac{1}{d_i}, \tag{20}$$

$$D_i = 0.1, \quad \text{and} \tag{21}$$

$$D_i = 0.1 \times d_i. \tag{22}$$

We can see from this figure that the total throughput becomes stable on a higher level of performance in the case when D_i is inversely proportional to the propagation delay. The other cases, $D_i = \text{const.}$ and D_i is proportional to the propagation delay, fail in flow control. This result implies that Eq. (11) is appropriate for realizing high performance and stability in networks with asymmetric configurations.

The values of total throughput realized by the model with $D_i \propto 1/d_i$ shown in Fig. 9 are almost the same as the cases where the smaller capacity of the bottleneck links is 50 as shown in Figs. 6 and 7. This means the setting of the diffusion coefficient Eq. (11) absorbs the complexity of the network model, and realizes high performance and stability even if the configuration of the network becomes complex.

6 Conclusions

This paper has presented the performance and stability of the diffusion-type flow control mechanism. The framework of this flow control is an autonomous decentralized system in high-speed networks. In the diffusion-type flow control, nodes handle their local traffic flow themselves based on only the information they are aware of.

To apply this method to networks with complex and asymmetric configurations, we investigated the appropriate values of the flow intensity multiplier and the diffusion coefficient, α and D_i, in our flow control model, and we found the conditions $\alpha = \text{constant}$ and $D_i \propto 1/d_i$, from observations of the packet flow.

We have shown simulation results for two cases: multiple bottlenecks and different propagation delays. Both results show that diffusion type control achieves high performance and is stable even if the network is congested.

In particular, if we choose the diffusion coefficient as $D_i \propto 1/d_i$, the diffusion-type flow control absorbs the complexity of the network model, and realizes high performance and stability even if the link delays in the network are different.

We are interested in the application of the diffusion-type flow control to networks with more complex topology. These issues will be the subject of further study.

References

[1] M. Aida and K. Horikawa, "Stability analysis for global performance of flow control in high-speed networks based on statistical physics," *IEICE Transactions on Communications*, vol. E82-B, no. 12, pp. 2095–2106, December 1999.

[2] K. Horikawa, M. Aida, and T. Sugawara, "Traffic control scheme under the communication delay of high-speed networks," In *Proc. of Int. Conf. on Multi-Agent Systems (ICMAS) '96*, pp. 111–117, 1996.

[3] M. Aida and C. Takano, "Stability of autonomous decentralized flow control schemes in high-speed networks," In *Proc. of IEEE ICDCS 2002 Workshop (ADSN 2002)*, pp. 63–68, 2002.

[4] C. Takano and M. Aida, "Stability and adaptability of autonomous decentralized flow control," In *Proc. of the Sixth International Symposium on Autonomous Decentralized Systems (IEEE ISADS 2003)*, 2003.

[5] Y. Bartal, J. Byers, and D. Raz, "Global optimization using local information with applications to flow control," In *Proc. of the 38th Ann. IEEE Symp. on Foundations of Computer Science*, October 1997.

[6] S. H. Low and D. E. Lapsley, "Optimization flow control-I: basic algorithm and convergence," *IEEE/ACM Transactions on Networking*, vol. 7, no. 6, pp. 861–874, 1999.

[7] K. Kar, S. Sarkar, and L. Tassiulas, "A simple rate control algorithm for maximizing total user utility," In *Proc. of IEEE INFOCOM 2001*, pp. 133–141, 2001.

[8] J. Mo and J. Walrand, "Fair end-to-end window based congestion control," *IEEE/ACM Trans. Networking*, vol. 8, no. 5, pp. 556–567, October 1999.

[9] S. K. R. Srikant, "A decentralized adaptive ECN marking algorithm," `citeseer.nj.nec.com/487389.html`

[10] D. Mitra, "Optimal design of windows for high speed data networks," In *Proc. of IEEE INFOCOM '90*, pp. 1156–1163, 1990.

0-7695-1921-0/03 $17.00 © 2003 IEEE