Multiple Paths Protocol for a Cluster Type Network

Hiroshi Mineno, Lu Junshu, Ken Ohta, Masahiro Aono, Tetsuo Ideguchi† and Tadanori Mizuno

†Faculty of Information Science and Technology
Aichi Prefectural University
Nagakute, aichi, 480-1198 Japan

Graduate School of Science and Engeering
Shizuoka University
3-5-1 Johoku, Hamamatsu, 432 Japan

Abstract

We can access Internet by carrying a portable computer and using the wireless communication. The wireless network with PHS(Personal Handy phone System) and portable cellular telephone has only rates of tens of Kbps to a few Mbps. Compared with the cable network, the transfer rate cannot generally satisfy a highly developed communication services such as large file transfer and real-time communications. This paper proposes a protocol, SHAKE, for sharing multiple paths in cluster type network that is a kind of LAN in which some mobile hosts temporarily connect mutually. SHAKE provides the functions for composing cluster type network, and dispersing traffic efficiently by measuring transfer rate and round-trip time. As a mobile host has only low transfer capacity in individual to communicate with outside, if whole capacities of other hosts which compose cluster type network are shared, we can get larger transfer capacity and satisfy the required communication services.

1 Introduction

With the rapid development of the network technology and the improvement of the computer's performance, we can access Internet by carrying a portable computer and using the wireless communication. In the present status, the wireless network with PHS(Personal Handy phone System), portable cellular telephone or in-house wireless LAN has only rates of tens of Kbps to a few Mbps. Compared with the wired network, the transfer rate cannot generally satisfy a highly developed communication services such as transfer of large files or multimedia data, and real-time communication which should guarantee the delay time. This means that the demands for capacity may sometimes exceed what is available, and packets may be lost.

In this paper, we propose a protocol SHAKE (SHAed multiple paths protocol for a cluster type networK Environment), for sharing multiple paths in a

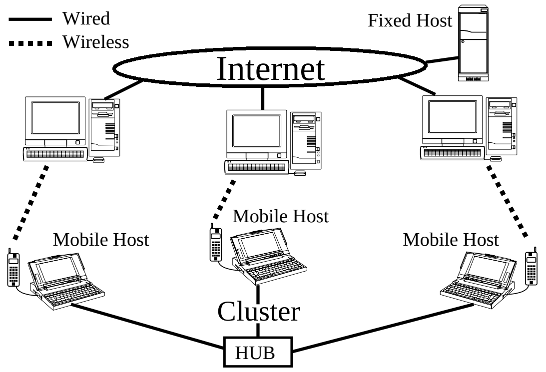

Figure 1: An Example of the System.

cluster type network. Cluster type network is a kind of LAN in which some mobile hosts temporarily connect mutually. In mobile units such as trains, ships and aircrafts where some persons gather, there are two ways to get larger capacity by composing the cluster type network. One way is that the host which has larger or more stable path is made to represent and communicate with outside. No mobile host which connects cluster type network needs to have its own outside path. The other is that as a mobile host has only low transfer capacity in individual to communicate with outside, if whole capacities of other hosts which compose cluster network are shared, we can get larger transfer capacity and spread burst traffic equally.

We focus the latter(Fig. 1). We can select several kinds of paths to communicate with outside. So the communication media such as PHS or portable cellular telephone can be freely selected by each mobile host. But using different paths to communicate, the quality of each paths such as transfer rate, delay time, and the rate of packet loss are different, and for this reason, it is important to consider how to distribute the burst traffic and how to reconstruct the packets.

This paper is organized as follows. The next section describes the background of this research work and the related research works; Section 3 describes the shared multiple paths protocol, SHAKE and its functions; Section 4 talks about the implementation of the prototype; And Section 5 concludes this paper.

2 Related Works

2.1 Traffic Dispersion

Generally, aggregating the resource can improve the performance and efficiency in shared system or communication network[1]. One of the resource aggregation methods is that the traffic is transmitted in parallel over multiple paths.

Maxemchuk presented that for load balancing and fault handling in packet-switched networks, it is better in space rather than in time to disperse the traffic[2, 3, 4]. According to this method, a message from a source is distributed into several submessages, which are transmitted in parallel over different paths in network. And as a transmission error on one path is independent from errors on the other paths, forward error correcting codes can be successfully used. Furthermore, in an algorithm proposed by Lee and Liew[5], the traffic from a source is partitioned into submessages, each consisting of K packets. This algorithm encodes the K packets into N>K packets, which are transmitted in parallel over N separate paths. When any K of the N packets are correctly received, the original message can be reconstructed. And if the bit error detection is used for every packet, the algorithm is capable of correcting N−K packet errors in each submessage.

IETF standardizes MP(Multilink Point-to-Point Protocol) with the technique which consists of a number of different switched WAN services such as ISDN and ATM to form a virtual path[6, 7]. It can strengthen the functions of PPP(Point-to-Point Protocol) data communications between two points, and a number of virtual connections can be set up between equipments by taking advantages of characters of switched WAN service. If necessary, the bandwidth which a user requested can be divided, and a number of physical lines and virtual lines can be used as a single logical line.

The technology of aggregating multiple communication paths and dispersing traffic into each communication path are being developed to improve the performance and efficiency of transmission. And they are generally used to communicate with different kinds of WAN lines which are aggregated among routers.

2.2 Mobile Distributed Co-operation System

In mobile computing environment, a mobile host works as a terminal for personal operation. Furthermore it is also considered to form a group to co-operate with. There is a mobile groupware which supports co-operation for a number of mobile hosts. In the area where a number of mobile hosts can communicate mutually, the number of mobile hosts can construct a temporary network to support co-operation for people gathered in the area. The network is called cluster type network or ad hoc network, and some research works are being developed [8, 9, 10].

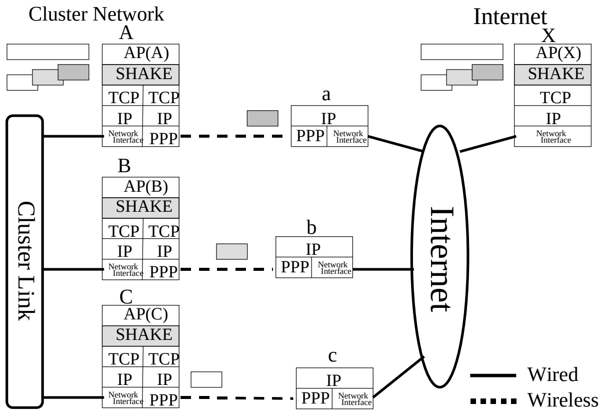

Figure 2: SHAKE Protocol Stack.

In addition, it was also considered how to communicate with outside networks. The one is that in some kinds of transportations such as trains, ships or aircrafts, if there is a communication system with high communication capacity which works as a representative of all mobile hosts to communicate with outside networks, the transfer time and cost can be reduced. And the other is that as a mobile host has only low transfer rate in individual to communicate with outside networks, but if whole capacities of other hosts which compose cluster network are shared, we can get larger transfer capacity.

So the cluster has some advantages which a single mobile host hasn't. And the latter communication protocol in which other hosts' communication capacity can be shared is called SHAKE, which we are developing. By using this, the communication paths of other mobile hosts which compose the cluster are shared and they are considered logically one path.

3 Shared Multiple Paths

3.1 SHAKE

We think there are severaral advantages of the sharing multiple paths: The end-to-end transfer rate of data can be increased and the delay time be reduced. The bursty traffic is dispersed into several paths and network load is equalized. Transmission errors on each path are independent from errors on the other paths.

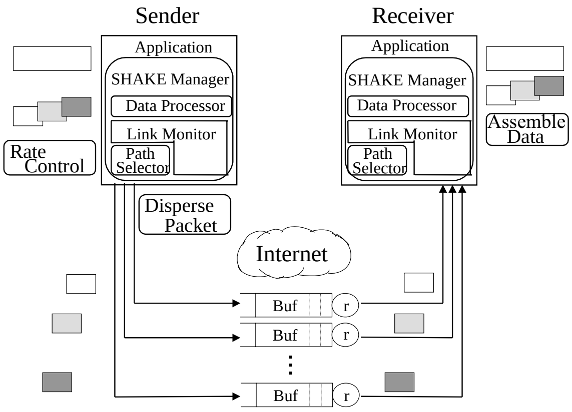

Figure 3: Illustration of SHAKE Modules.

To transfer packets redundantly, the characteristics of delay time and packet loss on different paths can be reduced, and reliability can be improved. It can form a cluster in which each mobile host can gather anywhere. Any mobile host in a cluster can freely select any kinds of medium for communication between the cluster and the outside network. If one of mobile hosts in cluster has not a path to communicate with outside, it can transfer data by relaying other's path.

We are now developing the protocol SHAKE to realize the shared multiple paths. SHAKE is implemented at all mobile hosts which form a cluster (cluster means cluster type network) and at fixed hosts which are connected to Internet and communicate with one of mobile hosts in a cluster (Fig. 2). And for easy to implement SHAKE, it is built between the Transport Layer and the Application Layer. So the mobile host which isn't implemented SHAKE can't share multiple paths.

In the cluster, we call a mobile host that has a communication path with outside network (like Internet) as a relay host. Each relay host in the cluster has two network interfaces: one is used to communicate with outside network, and the other is assigned the local private IP address before forming a cluster. Cluster is formed by using Ethernet, serial cable, wireless LAN, IrDA and so on, and the shape of the network can be considered as ring type network or star type network. A mobile host can freely select any communication path and any kind of medium to communicate with outside. For those hosts which have no connection with outside, the communication can be also done through the relay hosts. The logical link which connects each mobile host to form the cluster is called "cluster link". A mobile host became a part of the cluster if it is connected by cluster link.

3.2 Functions of SHAKE

As a protocol of Transport Layer, SHAKE uses TCP to transmit data. TCP connection is formed between mobile hosts in the cluster and fixed hosts in the Internet. There are three kinds of SHAKE packets which go through this connection, each is identified by special SHAKE header; the first is data packet which contains data from the Application Layer; the second is control packet to form a cluster, and the last is report packet which is used to report communication status. The protocol SHAKE is realized from the following 4 modules.

- **Link monitoring module**:
- **Path selecting module**:
- **Data processing module**:
- **SHAKE management module**:

The relationship among these modules is shown in Fig.3. From these modules, we can realize the SHAKE for sharing multiple paths in cluster type network.

3.2.1 Link Monitoring Module

This module monitors the status of communication paths with outside networks and estimate the sending rate. Here the status of communication paths means the theoretical bandwidth of used medium, throughput which changes actively, transit delay, cut of communication paths, etc.. And if a notable change of the status is happened, it changes the sending rate to use each path's bandwidth effectively.

Here we will talk about the estimation of sending rate for transferring data. The method is to control the definite amount of data which pass the network. We consider the network as a large buffer (we call it a network buffer) which has it's own queue and sending rate. For using each path effectively, it is better to leave network buffer in constant value to avoid that the receiving buffer in the object is often empty. But the problem will be happened if too much data is sent, which can cause the transit delay largely, so it is important to control the amount of data in the network buffer. Like in Fig. 3, it seems that each communication path has a buffer and all buffers are managed by Link Monitoring module. The SHAKE evaluates how many packets have been sent to the object and how many packets still overstock in each network buffer. The SHAKE can estimate sending rate to keep amount of data in all network buffers.

The remained amount of data in network buffers can be evaluated as follows. In sending host, each of the dispersed packets is added a sending time stamp

ST_1 in SHAKE header and sent to the receiving host. On the other side, at regular intervals, the receiving host sends back a receiving report to the sending host, which contains the newest received packet's ST_1, RT_1 which is the time it received, RT_2 which is the time this report was sent, and R_{recv} which is rate of receiving. So when the sending host received the report at the time ST_2, the sending host can estimate the RTT by the following expression. And to avoid the wide fluctuation of RTT_{avg} which means average of RTT, it is estimated as follows by using low-pass filter $(0 < \alpha \leq 1)$.

$$RTT = (ST_2 - ST_1) - (RT_2 - RT_1) \quad (1)$$

$$RTT_{avg} = (1 - \alpha)RTT_{avg} + \alpha RTT \quad (2)$$

And then Buf_{cur} which means current network buffer is estimated as follows by using the R_{rcv}, RTT_{avg}, and RTT_{min} which is minimum RTT value till the time:

$$Buf_{cur} = R_{rcv}(RTT_{avg} - RTT_{min}) \quad (3)$$

Comparing the expected value Buf_{des} with current amount of data in network buffer Buf_{cur}, the sending rate R_{snd} is increased if expected value is small, or reduced if expected value is large. And using the receiving rate R_{rcv} in the receiving report, we can calculate the accurate sending rate:

$$R_{snd} = R_{rcv} + \frac{(Buf_{des} - Buf_{cur})}{interval} \quad (4)$$

3.2.2 Path Selecting Module

A sending host gets information about communication paths with outside from the link monitoring module, and then distributes the data packets which come from the application layer. If each communication path has same length, type, throughput and transit delay, the packets can be distributed equally to each of communication paths in order. But actually, in each path, characters of network are different since each path can be selected freely by mobile hosts. So the distribution rate of each path must be changed according to the status of each communication path. This is done by analyzing the information about the sending rate R_{snd}, and disconnection report from link monitoring module.

The distribution procedure is to distribute packets according to the rate of each path's R_{snd}, and adjust the amount of data in each network buffer. If there are n paths in cluster and the sending rate is R_i in path $i (i \leq n)$, the packet distribution rate P_i of path i can be calculated as follows. And according to the packet distribution rate P_i, the packets can be distributed for each communication path.

$$P_i = \frac{R_i}{\sum_{k=1}^{n} R_k} \quad (5)$$

3.2.3 Data Processing Module

The received data from application layer are divided into a number of packets. Each packet is transmitted to transport layer after it is added a special header of SHAKE.This packet is a data packet, and the header includes: packet identifier (Data), link number, packet length, sequence number, ST1, and data. And there is a report packet which includes: packet identifier (Report), link number, packet length, ST1, RT1, RT2, and receiving rate of the link. Fig. 4, Fig. 5 show the format of each SHAKE packet. And last, there is a control packet which is only used to form a cluster and make a link table.

When each host receives a SHAKE packet, it identifies the packet whether it's a data packet or a report packet or a control packet by checking the identifier in the header, and according to the link number in the header, it judges where to relay the packet. Each host has a table which relates host with link number. So if the packet is destined the host itself, the packet is given to the application and the others are relayed to the next host.

Here we explain how to put in order the received data packets with irregular sequence since they were transmitted from multiple paths. Although the path selecting module selects paths from which packets be received by destination in right sequence, actually it's difficult to do it completely because of different characters like transit delay and length of each path. If a packet with irregular sequence is received by SHAKE, it cannot be given to the application directly. SHAKE will store it to the buffer temporarily, and then if any packet which can be given to the application is received, SHAKE will check the buffer if there is any packet with the following sequence number. If there is, of course it will be given to the application, and SHAKE will continue to search whether there is any packet else in the buffer. If no one else, SHAKE receives the next packet and repeats the above processing.

3.2.4 SHAKE Management Module

The work of this module includes controlling the other modules, forming the cluster, ensuring communication paths and so on. The application of the upper layer

0		32bit
Packet ID	Link Number	Unused
Packet Length		
Sequence Number		
ST1		
Data		
....		

Figure 4: Data Packet Format.

0		32bit
Packet ID	Link Number	Unused
Packet Length		
ST1		
RT1		
RT2		
Receiving Rate		

Figure 5: Report Packet Format.

needn't to consider how and on which paths are selected to transmit data. It realizes the shared multi-link procedure by using multimedia through multiple paths to form a logical path.

4 Implementation

In this section, we present the implementation environment and the evaluation of the prototype which we are developing now.

4.1 The Implementation Environment

The cluster network is formed by mobile hosts (Toshiba SatellitePro420, CPU: Pentium100MHz Memory:16.0MB, and Mitsubishi Amity SP, CPU: 486DX4-75MHz Memory:16.0MB) which installed Windows95 and with two PCMCIA card slots. For sending packets within the cluster, the network interface which is connected to the cluster is assigned the private IP address like 192.168.X.X, and each host connected with about 1m's 10Base-T and 4 port HUB to form a cluster. The other network interface is assigned a normal IP address to connect the Internet, which may be assigned by DHCP or decided value.

For the implementation of the SHAKE prototype, we developed a simply special data transfer application with Visual C++, and the SHAKE is used in the communication section of the application. It is implemented at all mobile hosts which form a cluster, and the host in the Internet which communicates with one of the host in cluster network. Then it can divide an application data into a number of determined size SHAKE packets and distributes to the destination. And in destination, the received packets is reconstructed and given to the application in order.

4.2 The Status of Prototype

The present link monitoring module forms a path information table of all end-to-end path information when the cluster is formed, and the information is updated by receiving the report packets. When some new hosts are added to the cluster, new sections are added to this table after new connecting. It performs a role as a routing table indicated among link number and relay host and destination host, and it contains various information like theoretical bandwidth, round-trip delay, and receiving rate. These values are estimated by receiving report packet from the recieving host at regular intervals, and each link's suitable sending rate is also estimated by this report.

The path selecting module searches if the destination host is in the information table when the application requests for sending data. And then it checks the path, and simply sends through these paths in order. That is to say, the function which can change the packet distribution rate according to the each path's sending rate is being developed.

The data processing module is almost completely implemented. It divides application data into a number of SHAKE packets, and adds a special SHAKE header to each packet. When the packet is received, it is processed to determine if it is relayed or be given to the application according to the information of header. If the packet is to be given to the application, it should be put back in order since they may be transmitted in different paths and received in irregular sequence. In the present method, the sequence number in SHAKE header will be checked to determine the action. If the packet can't be given to the application, it is stored in buffer and its sequence number will be used as a key for searching. When a new packet is received and given to the application, the SHAKE checks if the next packet is stored in buffer. If it is, the packet will be took out and haded to the application, then the receiver repeats the above processing.

The tasks of SHAKE management module are checking among hosts and managing any other modules. It keeps communications with the application directly, and an additional window of SHAKE is opened for the request of forming a cluster.

4.3 Experiments

In our experiments, we used two mobile hosts to form a cluster network, and used wireless LAN links (2.4GHz frequency hopping spread spectrum) for communication with outside networks. We did the following three kinds of experiments.

1. In experiment one, we used one access point for wireless LAN, and the mobile host communicates with the fixed host which connected with our laboratory's LAN by using a wireless LAN (Fig.6).

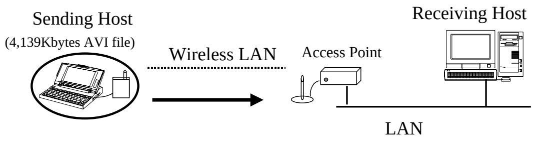

Figure 6: Experiment 1.

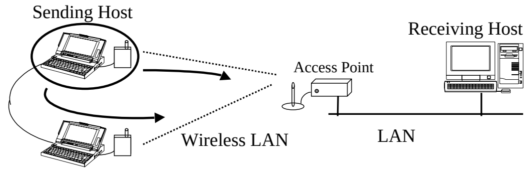

Figure 7: Experiment 2.

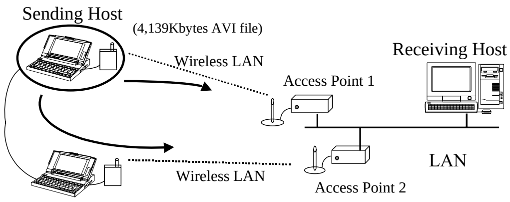

Figure 8: Experiment 3.

2. In experiment two, we used one access point for wireless LAN, and two wireless LAN adapters for sharing the paths. The cluster is formed by two mobile hosts, only one mobile host communicates with the fixed host which connected with the laboratory's LAN (Fig. 7).

3. In experiment three, we used two access points. We set up two channels for them to share the communication paths under the situation that there is no affection between them. The cluster is formed by two mobile hosts, and done the same experiment (Fig. 8).

The used file is about 4Mbytes (4139kbytes) AVI file. It was sent from the client host in the cluster to server host in fixed network. Then in the server host, we investigated the amount of received data and the receiving time of data (which means the time beginning the first data packet received to the time that all data packets were received). And the throuput is estimated at the server host. The relationship is shown in graphs (Fig. 9).

Next we did above experiments with PHS terminals on PIAFS mode. PIAFS (the abbreviation for PHS Internet Access Forum Standards) is the data transmission standards to operate a high speed data communication service with 32Kbps access via PHS. This environment is shown in Fig. 10.

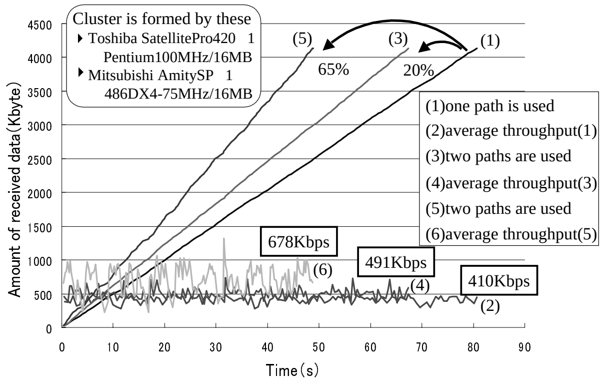

Figure 9: Results of Wireless LAN experiments.

4.4 Results

There are six graphs in the Figure 9. Graph (1)(3)(5) show the relationship between the received time and the amount of received data, and graph (2)(4)(6) show the average of the throughput which is calculated when five data packets are received. Graph (1)(2) are the results of experiment 1, graph (3)(4) are experiment 2, and graph (5)(6) are experiment 3.

The results shows that to form a cluster network and share multiple paths can reduce the transfer time and improve the throughput comparing with using a single wireless LAN adapter. When one wireless path is used, the average of throughput is about 410Kbps, and when two wireless paths are used with same access point, the average is about 491Kbps. In spite of sharing the same channel with two wireless LAN adapters, the average of throughput is improved a little (20%). And if the using channels are differ from, the throughput is better (65%).

The communication speed is changed according to different conditions, but the average of throughput which is guaranteed by the manufacture company is about 400 Kbps. So both of throughputs we calculated from graphs (2)(4) of Fig. 9 are the results we expected. And we know that if sharing the same channel, we can use the channel effectively and the average of throughput nearly approaches the limit value of using one channel.

From the results of experiment 3, graph (5)(6), we could get the 65% higher throughput than using one wireless path. The value is not yet twofold, but if the prototype is more developed, we may improve the throughput better. And in graph (6), the average of throughput is unstable. We think that the reason why the graph is so unstable is the function of resequencing packets. Some functions, such as the sending rate control and the distribution rate control, were not im-

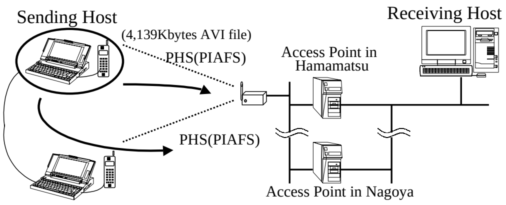

Figure 10: Experiment with PHS terminals.

plemented in the present prototype, so many packets are received out of order. After we implement the functions, the performance could be better.

Next we will talk the results of the experiments with PHS terminals on PIAFS mode. The results are shown in the graph (Fig. 11). Graph (1)(2) are the results of experiment 1 via PIAFS access point in Nagoya, and the average of throughput is about 19Kbps. And graph (3)(4) are the results of experiment 1 via PIAFS access point in Hamamatsu, and the average of throughput is about 26Kbps. Nagoya is about 100kms away from Hamamatsu, so the defference of delay between them was about 80ms (calculated from the measured RTT). Graph (5)(6) are the results of using two PIAFS links. Compared with the case of using one PIAFS link, the average of throughput is improved 93% better than the sum of the former two cases.

5 Conclusions

We proposed a protocol, SHAKE, for sharing multiple paths in a cluster type network in which a number of mobile hosts temporarily connect mutually. Each host in the cluster can share other hosts' communication links. And SHAKE consists of four parts: link monitoring module, path selecting module, data processing module and SHAKE management module. Till now we implemented the prototype of SHAKE and did communication tests with wireless LAN systems and PHS terminals.

From now on, we plan to continue the implementation of the protocol, to simulate the affection of a cluster's size, to evaluate the efficiency of control of transmission rate and data transmission with distributed packets, and so on.

Acknowledgments

The authors gratefully acknowledge the help of Junshu Lu who assisted us to write in English, and Nicholas F. Maxemchuk who supported several references.

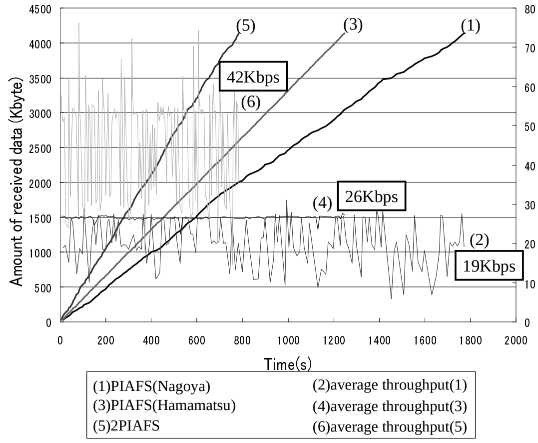

Figure 11: Results of PHS experiments.

References

[1] E. Gustafsson, and G. Karlsson, "A Literature Survey on Traffic Dispersion," *IEEE Network March/April*, Vol. 11, No. 2, pp. 28–36, 1997.

[2] N. F. Maxemchuk, "Dispersity Routing", *ICC'75*, pp. 41–10–41–12, June 1975.

[3] N. F. Maxemchuk, "Dispersity Routing in High-Speed Networks", *Computer Networks and ISDN Systems*, Vol. 25, No. 6, pp. 645–661, January 1993.

[4] N. F. Maxemchuk, "Dispersity Routing in ATM Network", *IEEE INFOCOM'93*, pp. 347–357, March 1993.

[5] T. T. Lee and S. C. Liew, "Parallel Communications for ATM Network Control and Management", *IEEE GLOBECOM'93*, Vol. 1, pp. 442–446, December 1993.

[6] K. Sklower, B. Lloyd, G. McGregor, and D. Carr, "The PPP Multilink Protocol (MP)", November 1994. RFC1717.

[7] K. Sklower, B. Lloyd, G. McGregor, D. Carr and T. Coradetti, "The PPP Multilink Protocol (MP)", August 1996. RFC1990.

[8] P. Gupta and P. R. Kumar, "A system and traffic dependent adaptive routing algorithm for ad hoc networks", in Proc. IEEE 36th Conf. on Decision and Control, pp. 2375–2380, 1997.

[9] B. Das, R. Sivakumar, and V. Bharghavan, "Routing in Ad-Hoc Networks Using a Spine", in Proc. 6th International Conference on Computer Communications and Networks, September, 1997.

[10] "Mobile Ad Hoc Networking (MANet)", http://tonnant.itd.nrl.navy.mil/manet/manet_home.html.