

# LLR: A Construction Scheme of a Low-Diameter, Location-Aware, and Resilient P2P Network

Masahiro Sasabe  
Cybermedia Center  
Osaka University  
1-32 Machikaneyamacho, Toyonaka-shi  
Osaka 560-0043, Japan  
Email: m-sasabe@cmc.osaka-u.ac.jp

Naoki Wakamiya and Masayuki Murata  
Graduate School  
of Information Science and Technology  
Osaka University  
1-5 Yamadaoka, Suita-shi  
Osaka 565-0871, Japan  
Email: {wakamiya, murata}@ist.osaka-u.ac.jp

**Abstract**—Since a peer searches for its desired file in a P2P file sharing system, the structure of an overlay network determines the effectiveness of search. In this paper, based on the Barabási-Albert (BA) model, we propose a novel scheme (LLR) to construct a low-diameter and location-aware overlay network where peers can easily find physically-close file holders. LLR has a rewiring method to improve the structure of an overlay network and a recovery method to cope with disappearance of peers. Through several simulation experiments using real physical topologies, we found that LLR could construct an overlay network that had the higher reachability than BA and the higher correlation between physical and logical distances.

## I. INTRODUCTION

In recent years, P2P file sharing systems have become widely deployed. There are three kinds of architectures for P2P systems: centralized, decentralized-unstructured, and decentralized-structured. In centralized P2P systems, such as Napster [1], a peer searches for its desired file by emitting a query to a meta server that maintains information on peers and file locations. Since a query message is relayed only among meta servers, response time and network load can be kept low. However, query messages concentrating on meta servers cause serious congestion as the number of peers increases. Furthermore, a meta server is a single point of failures.

To tackle these problems, many researchers have focused on decentralized P2P systems. Decentralized-unstructured P2P systems, such as Gnutella [2] and KaZaA [3], are the most popular in the current Internet because of its simplicity. In decentralized-unstructured P2P systems, a peer tries to find its desired file by flooding a query in a P2P overlay network. When a peer has a desired file, it returns a response message to a querying peer. Among peers, i.e., file holders from which response messages are received, a querying peer determines a provider peer from which it retrieves a file. Although there is no-centralized control, such unstructured systems are often criticized for their search inefficiency in flooding.

Decentralized-structured P2P systems are also decentralized, but introduce a rule of locating files. They typically employ Distributed Hash Table (DHT) to place files on designated

peers [4-6]. In a DHT scheme, each peer is assigned a key by a hash function and becomes responsible for files of the same or similar key. A peer can easily find a file by inquiring a file of a peer of the corresponding key. This scheme reduces the load of search, but it has been pointed out that managing the structure of a P2P network against join and leave of peers causes much control overhead and deteriorates system performance [7].

The structure of an overlay network determines the effectiveness of search in terms of network load and user QoS. Especially in a decentralized-unstructured P2P system, the unorganized structure of an overlay network emerging from independent behavior of peers leads to the waste of network resources and spoils the usefulness of the system. If an overlay network is randomly constructed without taking into account the topology of an underlying physical network, a logical link may be established between physically-distant peers. Consequently, passing a message from one peer to another takes much time and consumes more network resources. Furthermore, a peer has to spend time and network resources on each search trial to identify physically-close file holders from which it can quickly retrieve its desired file, because file holders found by a search are not necessarily physically-close. In addition, it is desirable to construct a low-diameter overlay network to diffuse query messages efficiently. As the diameter of an overlay network decreases, more peers query messages can reach with a smaller TTL value.

From the above observations, we can conclude that an overlay network should reflect characteristics of an underlying physical network, e.g., the degree distribution and physical proximity [8]. Figures 1(a) and 1(b) illustrate examples of efficient and inefficient overlay networks constructed on the same physical network, respectively. The physical network consists of six hosts ( $H1 \sim H6$ ) and three routers ( $R1 \sim R3$ ). Peers ( $P1 \sim P6$ ) are on the hosts. There are two significant characteristics in an efficient overlay network: low-diameter and location-awareness. The diameter of an overlay network is defined as the average number of logical hops among arbitrary two peers. The diameter of the overlay network in Fig. 1(a)



network based on Gnutella 0.6 protocol [8]. In LTM, peer  $P$  collects delay information on peers within two logical hops by periodically sending probing packets. Based on the delay information,  $P$  conjectures the topology of an overlay network within two logical hops and finds peer  $S$  that has two or more logical paths to  $P$ . If the longest connection among the logical paths is established between  $P$  and its neighbor,  $P$  cuts off the connection and obtains a new neighbor based on the Gnutella protocol. Although LTM can construct a location-aware overlay network, it does not contribute to reduction of the diameter of an overlay network. Furthermore, Gnutella-based random selection of neighbors can not necessarily find physically-close peers. This not only induces unnecessary traffic into an underlying physical network but also requires multiple times of rewiring to obtain physically-close neighbors.

On the other hand, R. Albert, et al. [13] proposed a modified BA with rewiring of links and studied a condition where the degree distribution of a constructed overlay network followed a power-law. However, the model randomly selects nodes to disconnect and does not consider the topology of an underlying physical network.

Rewiring of links also contributes to resilience to peer disappearances. A peer randomly leaves a P2P overlay network due to user's intention or a node failure. In addition, a malicious user may attack to some of peers highly connected to other peers to fragment the network [14]. It has been pointed out that a power-law network is vulnerable to such attacks because high-degree peers are centrally located in the network. When a high-degree peer suddenly disappears for some reasons, all logical links connected to the peer are lost and the network may be broken into small components. Peers neighboring the removed peer should establish one or more new connections to others to maintain the network connectivity.

To summarize, we consider that a low-diameter and location-aware network is a network where the degree distribution follows a power-law and logical links are established among physically-close peers. In LLR, a peer obtains information on peers that its neighboring peers know by periodically exchanging ping-pong messages as in the Gnutella protocol. Based on the constructed peer list, a peer first finds peers that are physically closer and with a higher-degree than the current neighbors and then tries rewiring. In addition, a peer chooses a physically-close and high-degree peer and establishes a logical link when it detects a link failure due to peer disappearances. To investigate the effectiveness of LLR, we conduct several simulation experiments using real physical topologies that also follow power-law distributions. We evaluate the diameter and physical distance between neighbors of overlay networks constructed by existing schemes and LLR.

The rest of the paper is organized as follows. We propose LLR that constructs a low-diameter, location-aware, and resilient overlay network in section II. Next, in section III we evaluate LLR through several simulation experiments. Finally, we conclude the paper and describe future works in section IV.

## II. LLR: A CONSTRUCTION SCHEME OF A LOW-DIAMETER, LOCATION-AWARE, AND RESILIENT P2P NETWORK

In this section, we consider methods to construct an overlay network that satisfies both low-diameter and location-awareness. We first propose a BA-based construction method with a modification in the node selection for PA in order to consider the physical proximity of neighbor nodes. Then, we further propose a rewiring method to improve the structure of an overlay network. Finally, we also discuss a failure recovery method.

### A. Construction Method

Our method puts restrictions on nodes considered in PA so that a new peer connects high-degree and physically-close peers. Our method is based on the modified BA [11] in which they introduce 'affinity' to restrict target nodes for PA. In Ref. [11], the affinity means user preferences such as bookmarks in WWW systems and references in papers. A new node with a random affinity first selects nodes with a similar affinity in the network. Then, nodes to connect are chosen among them according to PA. Through several simulation experiments, they showed that the power-law feature was not lost by introducing the restriction. Our method considers the physical distance instead of the affinity. In this paper, we regard the number of physical hops of the shortest path as the physical distance. Note that we can also use other definition of the physical distance. By using delay or bandwidth as the physical distance, we expect that the structure of an overlay network can adapt to changes in network conditions. However, it also requires much traffic to estimate network conditions.

When new peer  $i$  joins to an overlay network, it chooses  $m$  peers to connect among peers it knows according to the following algorithm.

- 1) Obtain set  $S_p$  of  $x$  peers from a bootstrapping server.
- 2) Calculate the physical distance to each peer in  $S_p$  by using the existing measurement tools, e.g., traceroute.
- 3) Obtain set  $S_c$  of  $\mu x$  peers in  $S_p$  in an ascending order of the physical distance.  $\mu$  is a control parameter that ranges (0,1]. If  $\mu$  is set to one, this method is equivalent to the original BA with the limitation of the size of candidate nodes. On the other hand, the smaller value of  $\mu$  intends to construct an overlay network which emphasizes the physical distance rather than the degree.
- 4) According to PA, select  $m$  peers in  $S_c$ . The probability  $P_n(d_j)$  that peer  $j$  with degree  $d_j$  is chosen is given by:

$$P_n(d_j) = \frac{d_j}{\sum_{k \in S_c} d_k}. \quad (1)$$

We should note here that bandwidth consumed by a new peer in traceroute can be justified by the fact that a location-aware network frees peers from evaluation of distance to file holders per search. The file-holder who returns a response fastest is the closest.

### B. Rewiring Method

In realistic situations, a newly added peer knows only some of peers from a bootstrapping node. This means that a peer can not necessarily find a peer that is physically close enough and with a sufficient degree at the join phase. Therefore, there is a possibility to refine the structure of a constructed overlay network by finding the other peers that are unknown and changing the structure. For this purpose, we propose a rewiring method where a peer first disconnects an inefficient connection and then establishes a connection to a physically-closer and higher-degree peer. This approach is inspired by BA with a rewiring method described in Ref. [13]. Our rewiring method differs from the model in selection of peers to disconnect. While the model randomly selects a peer to disconnect, our rewiring method chooses the connection to the most physically-distant neighbor to disconnect.

We assume that each peer can obtain information on other peers that are not current neighbors by exchanging ping-pong messages with current neighbors in the same manner as Gnutella. In the case of Gnutella, an interval of sending ping messages is a few minutes. A pong message sent by a neighbor includes information on peers that the neighbor knows. When a peer finds a new peer, it conducts rewiring. At first, by using a similar way to the construction method, it examines the physical distance to new peers. Then, it conducts PA for a set of the most distant neighboring peers and closer non-neighboring peers. Detailed algorithm is described as follows.

- 1) Calculate set  $S_w$  of peers that are the most physically distant among the current neighbors. We denote the physical distance as  $h_w$ . Note that neighbors with degrees of one are not included in  $S_w$  to prevent the fragmentation of an overlay network.
- 2) Calculate set  $S_m$  of peers whose physical distance does not exceed  $h_w$  among non-neighboring peers which it newly knows from pong messages.
- 3) Select a peer with probability  $P_r(d_j)$  among peers in  $S_w \cup S_m$  according to PA as follows:

$$P_r(d_j) = \frac{d_j}{\sum_{k \in S_w \cup S_m} d_k}. \quad (2)$$

If the selected peer is not the current neighbor, namely a member of  $S_m$ , it is replaced with one of distant neighbors randomly chosen from  $S_w$ .

### C. Failure Recovery Method

A peer conducts failure recovery when it detects a link failure due to peer disappearances. At first, it finds a new candidate for its neighbor from a peer list it has. Then, it determines a neighboring peer by starting from Step 3 in the algorithm described in subsection II-A except for the following settings. We set  $m = 1$  and  $S_p$  as a set of known peers excluding neighboring peers.

## III. SIMULATION EXPERIMENTS

To investigate the effectiveness of LLR, we conducted several simulation experiments. We evaluate LLR from a view

point of the structure of a constructed overlay network. We use the reachability to evaluate the diameter of an overlay network. The reachability is defined as the ratio of the number of peers to which a query message emitted by a peer reaches to the number of peers in the overlay network. The higher reachability means that a query message is effectively disseminated over an overlay network with a small TTL and it leads to the higher probability of successful searches and the larger number of discovered file holders. The neighbor distance is used to evaluate to what extent an overlay network considers an underlying physical network. The neighbor distance is defined as the number of physical hops between peers connected with a logical link. If neighbors on an overlay network are physically close, a peer can quickly obtain a desired file by choosing a logically-close file holder as a provider peer.

### A. Simulation Model

We used topological data of the real physical networks: Abilene [15] and Sprint [16]. The Abilene network is an Internet backbone network and a part of the Internet2. It has a power-law degree distribution and forms a hierarchical structure. It is comprised of sparsely meshed core routers and many edge routers each of which is highly-connected to end users. This structure considers the constraints of router technology where a router can have a few high bandwidth connections or many low bandwidth connections. The Sprint network describes the topology of a major ISP in the USA. The router level topology of the Sprint network was obtained by using a measurement tool called Rocketfuel [16]. Figure 2 illustrates physical topologies used in simulation experiments. In Fig. 3, we also present the characteristics of them, namely degree distribution and leaf-leaf distance that is the number of physical hops between leaf nodes whose degree is one. In both physical topologies, we assumed that peers were on leaf nodes. The number of peers in the Abilene and Sprint networks were 698 and 6478, respectively. We assumed that the latency of each physical link was identical so that we could compare LLR with LTM.

Since LTM is based on a Gnutella protocol, a peer periodically tries to establish connections until the number of neighbors reaches a pre-determined degree limit. We set the degree limit of each peer at 8 [8]. On the other hand, BA and LLR have no degree limitation. Instead, they restricted the number of connections established at the join phase to  $m$ . We used two values of  $m$ , they are  $\lceil \frac{d_l}{2} \rceil$  and  $\lfloor \frac{d_l}{2} \rfloor$ . Here,  $d_l$  corresponded to the average degree of an overlay network constructed by LTM. In the following simulation experiments,  $d_l$  of the Abilene and Sprint networks were 5 and 4.8, respectively. The inter-arrival time between two successive peer participations followed an exponential distribution whose average was 120 seconds. We set the interval that a peer sent ping messages to 120 seconds. Since we also set the interval of rewiring in LTM to 120 seconds, there was no difference in the messaging overhead between LLR and LTM. Depending on the limitation on the number  $x$  of peers initially obtained from a bootstrapping node, we focused on three

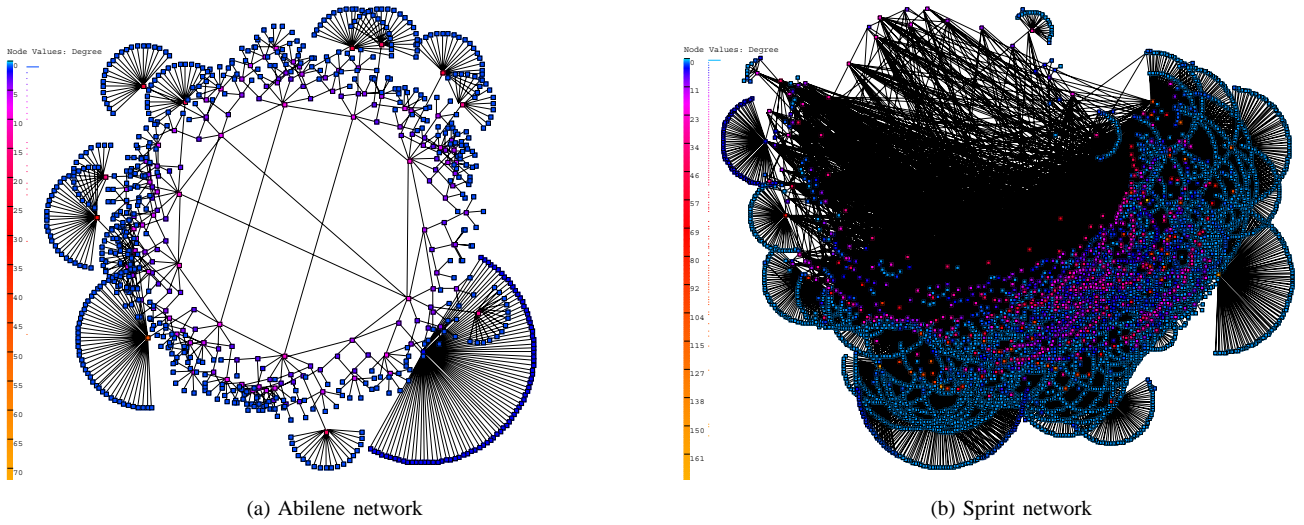


Fig. 2. Physical topologies used for simulation experiments

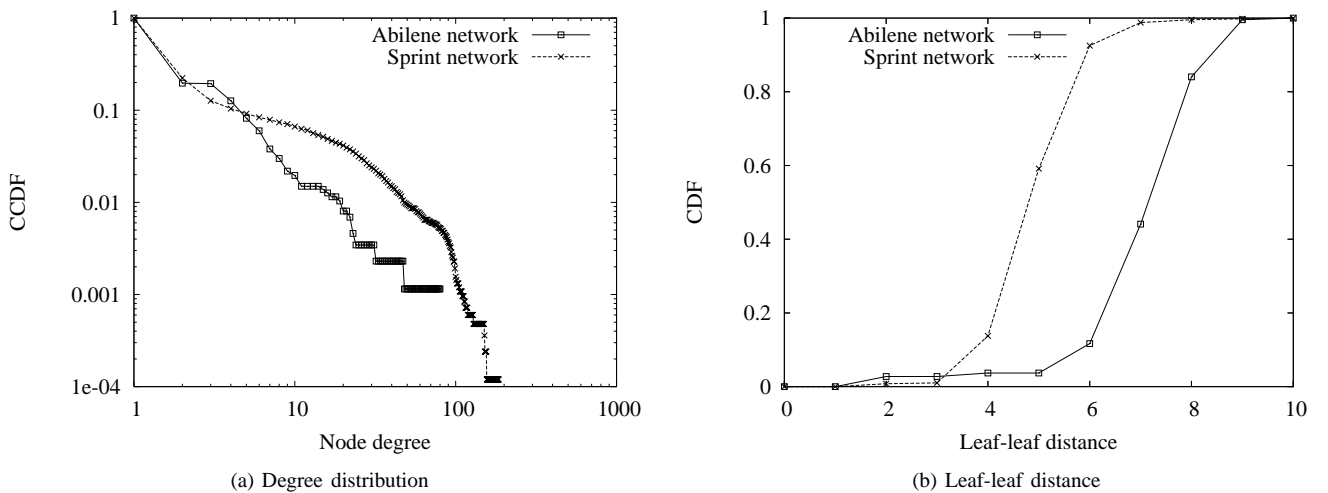


Fig. 3. Characteristics of physical topologies used for simulation experiments

cases of LLR: (1) no limitation on  $x$ , (2)  $x = 20$ , and (3)  $x = 20$  with the rewiring method. Although we conducted several simulation experiments by changing  $\mu$ , i.e., the ratio of peers to be considered in PA, we only show the results of  $\mu = 0.2$  and  $m = 3$  in the following figures. We evaluated characteristics of overlay networks when the last peer, namely 698th peer in Abilene and 6478th peer in Sprint, joined the overlay networks.

### B. Evaluation of Reachability

Figure 4 illustrates the relationship between the reachability and the range of search defined by a TTL value on a query message. The higher reachability means that a diameter of a constructed overlay network is low. At first, LLR (1)-(3) and BA can construct lower-diameter overlay networks than LTM because the degree distribution of constructed overlay networks follows the power-law. When the number of peers initially obtained from a bootstrapping node is limited, the

reachability of LLR without rewiring, i.e., case (2), becomes lower than that of BA. By introducing the rewiring method, i.e., case (3), the reachability is significantly improved. Specifically, the reachability obtained by LLR with rewiring becomes higher than BA by 0-60%. This is because that rewiring leads to a heavier tail of the degree distribution than that of an overlay network constructed by BA.

We also find that differences in the obtained reachability among methods are larger in the Sprint network than in the Abilene network. This is because the number of rewiring in the Sprint network is about ten times larger than that in the Abilene network. The number of rewiring is proportional to the size of the network.

### C. Evaluation of Neighbor Distance

Figure 5 depicts the cumulative distribution function (CDF) of the neighbor distance. LTM can shorten the neighbor distance than BA by disconnecting physically-distant neighbors.

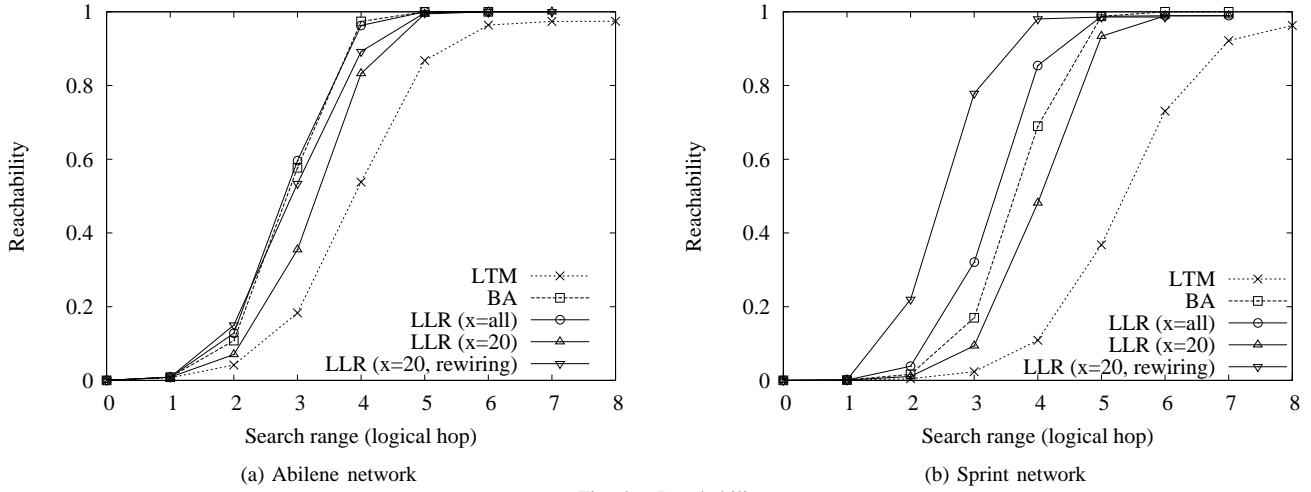


Fig. 4. Reachability

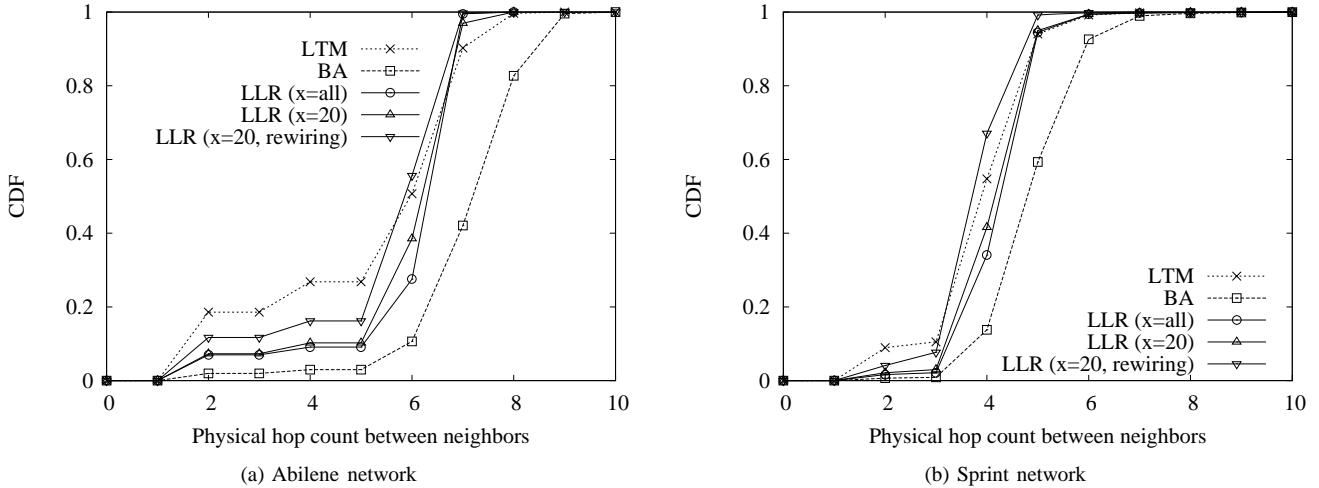


Fig. 5. Neighbor distance

LLR also shows better performance than BA by taking into account the physical distance in choosing neighbors. Furthermore, by introducing the rewiring method, LLR can construct an overlay network where logical neighbors are physically close with each other as with LTM.

Table I summarizes the correlation coefficient between the logical distance and the physical distance. LLR and LTM lead to the higher correlation between the overlay and underlying physical networks, that is, logically-close peers are also physically close.

Next, we discuss what extent LLR considers an underlying physical topology by comparing the neighbor distance of LLR with the leaf-leaf distance. Figure 6 presents the results of the case of the Abilene network. In LLR, a new peer first selects its neighbors from  $\mu x$  physically close peers initially obtained from a bootstrapping node. After that, it conducts the rewiring method when it finds a new candidate peer for its neighbor. The candidate peer must not be more distant than any of current neighbors. Therefore, the CDF of the neighbor

distance of LLR can be regarded as the CDF of the leaf-leaf distance whose range is restricted from 0 to  $\mu$ . The latter CDF can be derived by dividing the CDF of the leaf-leaf distance by  $\mu$ . We find that there is almost no difference between them except for the case in which the neighbor distance is greater than six. Thus, we can conclude LLR attains the upper bound of the neighbor distance that is determined by  $\mu$ . On the other hand, the CDF of the neighbor distance of BA is almost the same as the CDF of the leaf-leaf distance because BA does not consider an underlying physical topology. We also obtained similar results for the case of the Sprint network.

#### D. Evaluation of Resilience to Failures

It is desirable that a system can recover from a failure and shows the similar characteristics as before the failure. In this paper, we focus on two kinds of system failures: random disappearance of peers and an attack. In the case of random disappearance, a peer leaves an overlay network by termination of a P2P application by a user or a failure. On the other hand,

	LTM	BA	LLR with rewiring
Abilene	0.19	-0.01	0.45
Sprint	0.46	-0.17	0.42

TABLE I  
CORRELATION COEFFICIENT BETWEEN LOGICAL DISTANCE AND  
PHYSICAL DISTANCE

an attack considers the case where a malicious user attacks high-degree peers to break an overlay network into pieces and stop the service. We only show results of the case of the Abilene network, but we obtained similar results for the case of the Sprint network.

Figure 7 depicts results of the case of random disappearance. Every time a new peer joins, we conduct a disappearance event at probability  $P_d$ . When the disappearance event occurs, a peer is randomly selected and removed from an overlay network. We change  $P_d$  as 0, 0.1, 0.2, and 0.3. As shown in Fig. 7(a), the reachability becomes higher by recovery than the case without failures, because the number of peers decreases and consequently the diameter of the overlay network decreases. On the other hand, Fig. 7(b) shows that the neighbor distance does not change much. Therefore, we can conclude that LLR is resilient to random disappearance and keeps the characteristics of an overlay network.

Next, we consider the case of attack. First, we removed  $N_d$  peers in a descending order of degree after all peers had joined. Then, we conducted two types of simulation experiments. In scenario 1, an overlay network tried to recover from the attack by our recovery method. For the comparison purpose, an overlay network was rebuilt from the initial condition by adding peers that had remained after the attack one by one (scenario 2). Time for recovery in scenario 1 and time for reconstruction in scenario 2 were the same. We can consider that LLR is resilient to an attack if those two networks have similar properties.

We evaluate the resilience to attacks by changing the scale of attacks. Figure 8 illustrates results when we set the number  $N_d$  of removed peers at 10 and 125.  $N_d = 125$  corresponds to the critical point, 18%, at which percentage of node disappearance a scale-free network is completely fragmented [14]. As shown in Fig. 8(a), the reachability of scenario 1 is lower than that of scenario 2, independently of  $N_d$ . In scenario 1, when a high-degree peer disappears, its neighboring peers try to establish a connection to a high-degree peer among physically-close peers. It means that a high-degree peer is

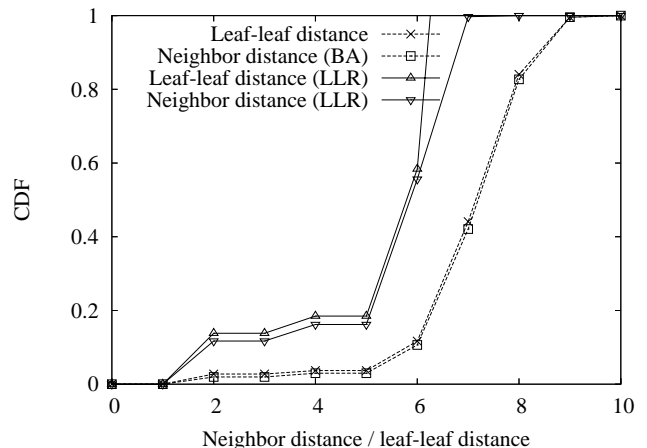


Fig. 6. Upper bound of neighbor distance (Abilene)

not necessarily chosen as a neighbor if it is physically apart. As a result, the degree of a high-degree peer grows slower than in scenario 2. Because of the same reason in the random disappearance, the reachability of  $N_d = 125$  is slightly higher than that of  $N_d = 10$ . On the other hand, the neighbor distance is almost the same when the scale of attacks is relatively small ( $N_d = 10$ ) as shown in Fig. 8(a). Furthermore, the neighbor distance deteriorates at most 0.2 even when the overlay network is suffered from massive attacks ( $N_d = 125$ ). Consequently, we can conclude that LLR has the resilience to attacks too.

#### IV. CONCLUSIONS

In this paper, we proposed LLR to construct a low-diameter, location-aware, and resilient overlay network. Through several simulation experiments using real physical topologies, we showed that LLR could construct an overlay network that led to the higher reachability than BA and the physically close neighbors as LTM. Furthermore, we also demonstrated that LLR had the resilience to both random disappearance of peers and an attack of a malicious user.

As future research works, we further consider load balancing among peers. In a constructed overlay network, query messages tend to concentrate on high-degree peers. We expect to reduce the load on a high-degree peer by introducing a cache mechanism. If information about popular files is cached at peers around high-degree peers, most of query messages are blocked before reaching high-degree peers. High-degree peers then serve as a repository of information about unpopular files.

#### ACKNOWLEDGEMENT

This research was supported by a Grant-in-Aid for Young Scientists (B) 17700058 and “New Information Technologies for Building a Networked Symbiosis Environment” of The 21st Century Center of Excellence Program of the Ministry of Education, Culture, Sports, Science and Technology in Japan.

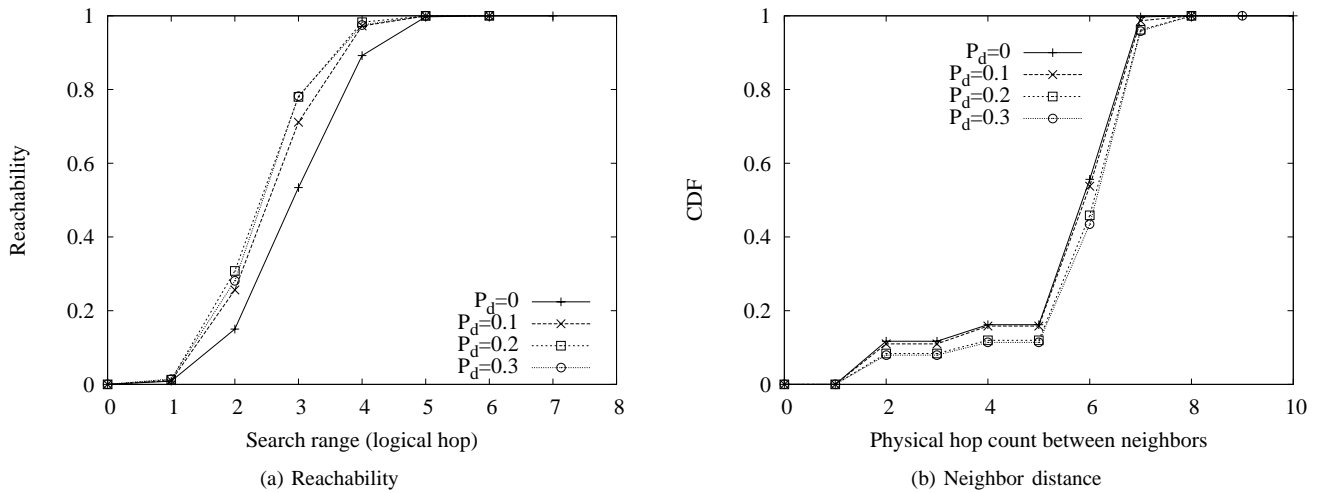


Fig. 7. Random disappearances (Abilene)

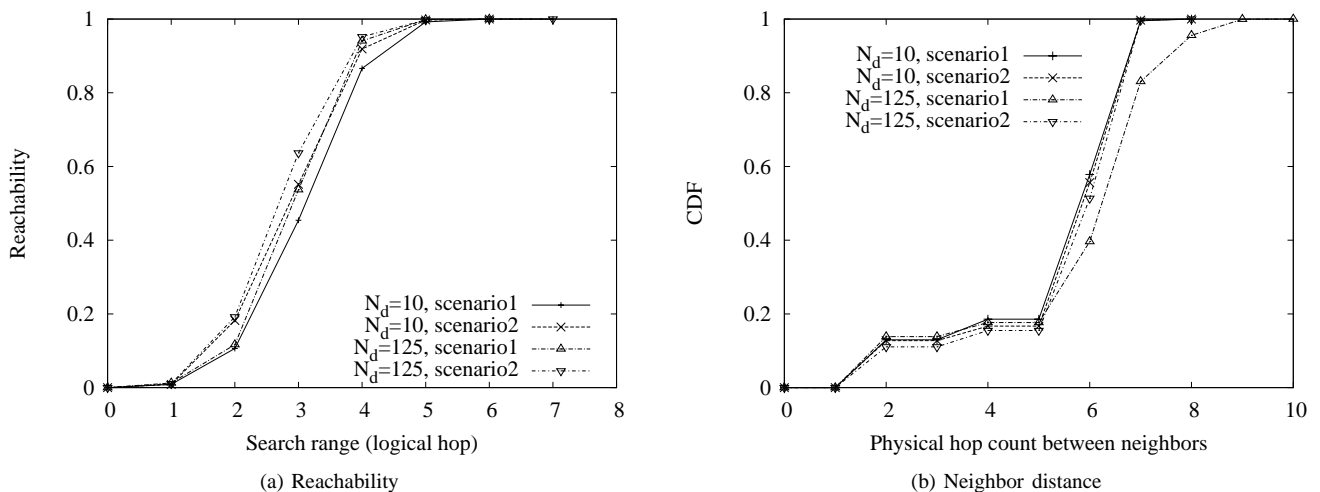


Fig. 8. Attacks from malicious users (Abilene)

## REFERENCES

- [1] Napster, available at <http://www.napster.com>.
- [2] Gnutella, available at <http://gnutella.wego.com>.
- [3] KaZaA, available at <http://www.kazaa.com>.
- [4] S. Ratnasamy, P. Francis, M. Handley, R. Karp, and S. Schenker, "A Scalable Content-Addressable Network," in *Proceedings of ACM SIGCOMM 2001*, New York, NY, USA, 2001, pp. 161–172.
- [5] I. Stoica, R. Morris, D. Karger, M. F. Kaashoek, and H. Balakrishnan, "Chord: A Scalable Peer-to-Peer Lookup Service for Internet Applications," in *Proceedings of ACM SIGCOMM 2001*, New York, 2001, pp. 149–160.
- [6] A. Rowstron and P. Druschel, "Pastry: Scalable, Distributed Object Location and Routing for Large-Scale Peer-To-Peer," in *Proceedings of Middleware 2001*, London, 2001, pp. 329–350.
- [7] R. H. Wouhaybi and A. T. Campbell, "Phenix: Supporting Resilient Low-Diameter Peer-to-Peer Topologies," in *Proceedings of INFOCOM 2004*, Hong Kong, Mar. 2004.
- [8] Y. Liu, L. Xiao, X. Liu, L. M. Ni, and X. Zhang, "Location Awareness in Unstructured Peer-to-Peer Systems," *IEEE Transactions on Parallel and Distributed Systems*, vol. 16, no. 2, pp. 163–174, 2005.
- [9] A.-L. Barabási and R. Albert, "Emergence of Scaling in Random Networks," *Science*, vol. 286, 1999.
- [10] P. Karbhari, M. Ammar, A. Dhamdhere, H. Raj, G. Riley, and E. Ien Zegura, "Bootstrapping in Gnutella: A Measurement Study," in *Proceedings of PAM 2004*, Antibes Juan-les-Pins, Apr. 2004.
- [11] J. Gómez-Gardeñes and Y. Moreno, "Local versus Global Knowledge in the Barabási-Albert Scale-Free Network Model," *Physical Review E*, vol. E69, no. 037103, pp. 1–4, 2004.
- [12] A. Beygelzimer, G. Grinstein, R. Linsker, and I. Rish, "Improving Network Robustness," in *Proceedings of ICAC 2004*, 2004, pp. 322–323.
- [13] R. Albert and A.-L. Barabási, "Topology of Evolving Networks: Local Events and Universality," *Physical Review Letter*, vol. 85, no. 24, Dec. 2000.
- [14] R. Albert, H. Jeong, and A.-L. Barabási, "Error and Attack Tolerance of Complex Networks," *Nature*, vol. 406, 2000.
- [15] L. Li, D. Alderson, W. Willinger, and J. Doyle, "A First-Principles Approach to Understanding the Internet's Router-Level Topology," in *Proceedings of ACM SIGCOMM 2004*, Portland, Aug. 2004, pp. 3–14.
- [16] N. Spring, R. Mahajan, D. Wetherall, and T. Anderson, "Measuring ISP Topologies with Rocketfuel," *IEEE/ACM Transactions on Networking*, vol. 12, no. 1, pp. 2–16, 2004.