PROPOSAL AND EVALUATION OF A COOPERATIVE MECHANISM FOR HYBRID P2P FILE-SHARING NETWORKS

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

ABSTRACT
Overlay networks, such as P2P, Grid, and CDN, have been widely deployed over physical IP networks. Since simultaneous overlay networks compete for network resources and their selfish behaviors disrupt each other, we consider cooperative mechanisms for overlay networks to enhance the collective performance and improve the QoS at the application level. In this paper, we proposed a cooperative mechanism for hybrid P2P file-sharing networks, where peers can find more files and exchange files with more peers. Through simulation experiments, we verified the effectiveness of cooperation. On the other hand, however, our results indicated that the system load would be increased by cooperation of networks.

KEY WORDS
Overlay Networks, P2P (Peer-to-Peer), Cooperative Networks, File-Sharing

1 Introduction

With emerging needs for application-oriented network services, overlay networks, such as P2P, Grid, and CDN, have been widely deployed over physical IP networks. To satisfy their own application-level QoS, these overlay networks often probe physical networks, for example using *ping* and *traceroute*, to learn and monitor underlaying physical topologies and dynamic attributes such as available bandwidth and delay. Based on those observations, overlay networks dynamically construct more efficient overlay topologies, select more efficient paths, and conduct traffic engineering independently. Since those behaviors are selfish and greedy, these overlay networks compete for limited physical resources and disrupt each other. To enhance the collective performance of competing overlay networks and efficiently utilize network resources, several research papers on cooperation among overlay networks have been published in recent years [1] [2] [3] [4]. In [4], they investigated a spectrum of cooperation among coexisting overlay networks. As an example, they proposed an architecture where overlay networks cooperated with each other in inter-overlay routing where a message from one overlay network was forwarded to another which provided a shorter path to the destination.

The analysis on coexistence of competitors has been investigated in the field of biology. In the ecosystem, organisms in the same environment live together with direct and/or indirect interactions with each other. In [5], they established the mathematical model of the metabolic pathways of bacterial strains to elucidate mechanisms of coexistence of living organisms of closely related species. They revealed that the coexistence emerged not only from interactions among competitors, but also from changes of their internal states.

Taking inspirations from biology, our research group consider the symbiosis among competing overlay networks [6]. We regard an overlay network as an organism. In the model of symbiotic overlay networks, overlay networks in a system evolve, interact with each other, and dynamically change internal structures, as living organisms in the same environment do. Overlay networks meet and communicate with each other in a probabilistic way. Overlay networks that benefit from each other reinforce their relationship, eventually have many inter-overlay links, and become one. Otherwise, they part from each other. All of evolutions, interactions, and internal changes are performed in a self-organizing way. Each node independently decides its behavior based on locally available information. Symbiosis among overlay network emerges as a consequence of independent and autonomous behaviors of nodes and networks.

In this paper, we focus on the cooperation among hybrid P2P file-sharing networks. Peer-to-Peer (P2P) file-sharing applications are still attracting users. Both in terms of the number of users and traffic volume, a P2P file-sharing network is one of the most important overlay networks in the Internet today. P2P networks are categorized into two, those are, a hybrid P2P network with central entities such as meta-servers, and a pure P2P network without any servers [7]. Whereas a pure P2P network consists of only peers, a typical hybrid P2P file-sharing network involves peers and one or more meta-servers which maintain a directory of available files on a P2P network to assist peers in finding files. As shown in Fig. 1, a hybrid P2P file-sharing network has a two-tiers topology, where meta-servers compose a core network and peers are connected to meta-servers to form star-shaped networks. When a client joins a hybrid P2P network, it connects with a meta-server and then registers meta-information about files to

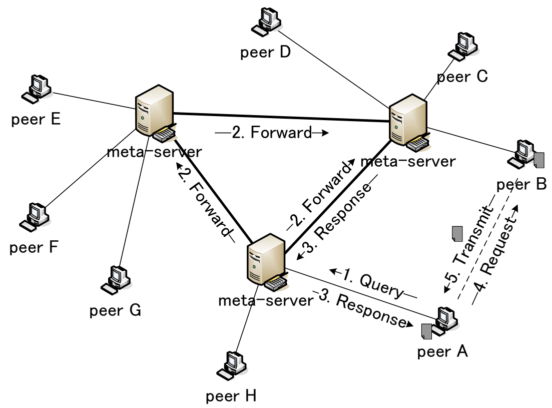

Figure 1. A Hybrid P2P File-Sharing Network

share with the other peers. In a hybrid P2P application, a peer sends a query message to a connected meta-server to find a file. If meta-information about the desired file exists in its directory, the meta-server sends a response message to the querying peer. Otherwise, the meta-server forwards the query message to other neighboring meta-servers. On receiving a query message, a meta-server investigates its directory. If it does not have any meta-information matching a query message, it forwards the query message to its neighboring meta-servers except that one from which it received the query message. This method is generally called flooding. If a meta-server has any matching meta-information, it generates a response message and sends it to the neighboring meta-server from which it received the query message. A response message is relayed among meta-servers following a reverse path of the corresponding query message and finally it reaches the querying peer. Once a peer receives a response message, it can directly retrieve a file from a provider peer, which appears in the response message.

There are several benefits from cooperative P2P file-sharing networks. In a P2P file-sharing application, when a peer failed in finding a file, it often repeatedly retries searches by changing keywords. Such repetitive and redundant query messages increase the load on meta-servers and waste physical network resources. When a meta-server halts for the overload or links among meta-servers are disconnected, meta-information on a failed or isolated meta-server is lost or becomes inaccessible. As a result, the possibility of successful search decreases and the application-level QoS deteriorates. If multiple P2P file-sharing networks cooperate with each other and share their files among them by exchanging query and response messages, a peer can find more files at more peers and the possibility of successful search increases. Consequently, the number of repetitive and redundant query messages decreases and resultant extra load also decreases. In addition, a peer can choose the best, i.e., the fastest or the most reliable peer among many provider peers found in a search. Furthermore, even when a P2P network is disconnected due to failures of meta-servers or links, meta-servers and peers in a P2P network are still able to communicate with each other since their message are relayed through cooperating P2P networks.

In this paper, we discuss cooperation among hybrid P2P file-sharing networks to improve their collective application-level QoS. We propose two cooperation approaches. One is called a *Shared-Peer-Based Approach*, where peers participating in multiple P2P networks play a role of cooperative points or gateway nodes, and the other is called a *Server-Chain-Based Approach*, in which hybrid P2P networks cooperate through inter-meta-server connections. Due to space limitation, we focus on the former and describe detailed mechanisms.

The rest of the paper is organized as follows. We describe our proposed cooperation approaches for hybrid P2P file-sharing networks, i.e., shared-peer-based approach and server-chain-based approach, in Section 2. The details of cooperative mechanisms for a shared-peer-based approach are presented in Section 3. Next, in Section 4, we evaluate our proposed mechanisms for a shared-peer-based approach through simulation experiments. Finally, we conclude the paper and describe future works in Section 5.

2 Cooperation among Hybrid P2P File-Sharing Networks

In this paper, we assume that, intending to improve its application-level QoS, a peer or a meta-server introduces a cooperative program to achieve the cooperation among hybrid P2P file-sharing networks. By a cooperative program, a P2P network, i.e., a peer or a meta-server can discover other P2P networks, decide whether P2P networks cooperate with each other or not, and cooperate.

In this section, we propose two cooperative approaches, that is, a Shared-Peer-Based (SPB) approach where a cooperative program is introduced into those peers that participate in multiple P2P networks, and a Server-Chain-Based (SCB) approach where a cooperative program is introduced into meta-servers. In the following subsections, we will briefly describe them.

2.1 Shared-Peer-Based Approach

In the SPB approach, a peer which participates in two or more P2P networks is called a shared peer. It becomes a point of cooperation, called a cooperative peer, by introducing a cooperative program (Fig. 2). Through logical links established between cooperative peers, multiple P2P networks exchange their query and response messages among them. Then, peers in a P2P network can obtain files from other P2P networks. We hereafter call a P2P network from which a query message is originated as a guest network and a P2P network to which a query message is forwarded as a host network.

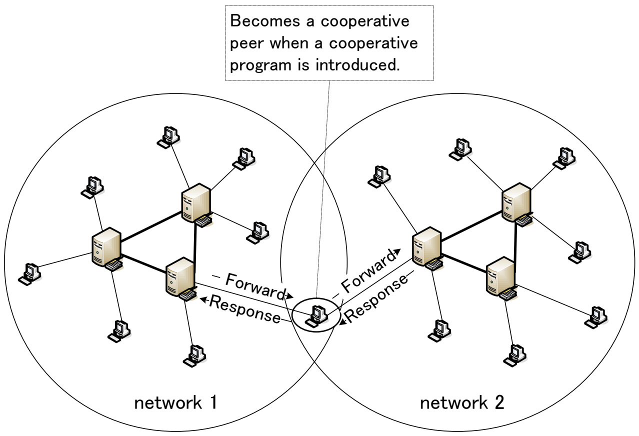

Figure 2. Shared-Peer-Based (SPB) Approach

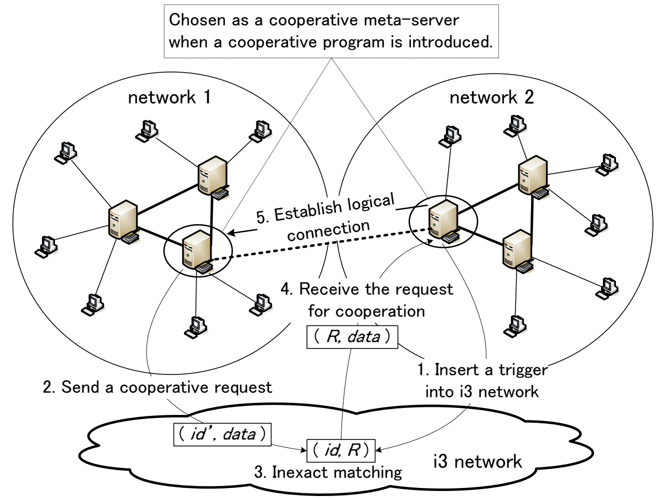

Figure 3. Server-Chain-Based (SCB) Approach

A new cooperative peer, which newly introduced a cooperative program, first decides whether or not to initiate cooperation among P2P networks to which it belongs. The decision is made according to some criteria such as the compatibility of protocols, the size of P2P network and categories of files that peers are interested in and shared in P2P networks. If P2P networks offer files of the same or similar categories, peers can find and obtain more files by cooperation. When P2P networks use different protocols for searching or retrieving files, the load on a cooperative peer increases for protocol conversion and relaying files. Therefore, it is desirable that P2P networks with the same or similar files and protocols cooperate with each other.

In the SPB approach, a cooperative peer behaves not only as a peer, but also as a meta-server. When a meta-server disseminates a query message by flooding, the query message is also forwarded to a cooperative peer since it is regarded as one of neighboring meta-servers. After applying protocol conversion to the query message if needed, a cooperative peer transmits it to a meta-server in host P2P networks as shown in Fig. 2. A query message is disseminated in a host network as usual to find a desired file. When two P2P networks are connected by more than two cooperative peers, the same query message will be forwarded to a host network. To detect the duplication, each query message has a unique identifier, and a meta-server discards redundant and duplicated query messages with the same identifier.

If a desired file is found in a host network, a response message is generated by a meta-server in a host network and it is sent back to the corresponding cooperative peer through a reverse path of the query message. Preparing for future query messages for the same or similar file, a cooperative peer deposits the meta-information in a response message into its local cache. After protocol conversion on necessity, a cooperative peer forwards a response message to its neighboring meta-server as a normal meta-server does. A response message eventually arrives at the peer which emitted the cooresponding query message. It retrieves a desired file directly from a provider peer in a host network. If a host network employs a different protocol for file retrieval, a cooperative peer replaces information about a provider peer with itself in a response message. Then, a querying peer establishes a connection to a cooperative peer to retrieve a file. A cooperative peer obtains a file in place of a querying peer on request. Consequently, a peer can benefit from the cooperation without recognizing it. We will describe details of cooperative mechanisms for this approach in Section 3.

2.2 Server-Chain-Based Approach

In the SCB approach, P2P networks cooperate with each other through logical connections established between meta-servers as shown in Fig. 3.

A meta-server which introduces a cooperative program becomes a candidate of cooperative meta-servers. One among them is chosen as a cooperative meta-server taking into account several criteria such as the number of peers connecting with it, the number of meta-information in its directory, and the distance to the other meta-servers in a P2P network. Then, a cooperative meta-server finds cooperative meta-servers of other P2P networks by using, for example, i3 [3]. A logical connection is established between cooperative meta-servers to exchange query and response messages with each other.

In the SCB approach, a cooperative meta-server behaves not only as a meta-server in its own P2P network, but also as a peer against cooperative meta-servers in host networks. When a cooperative meta-server receives a query message and has no corresponding meta-information in its local directory, it forwards the query message to all of its neighboring meta-servers in a P2P network by flooding. In addition, after applying protocol conversion if needed, it

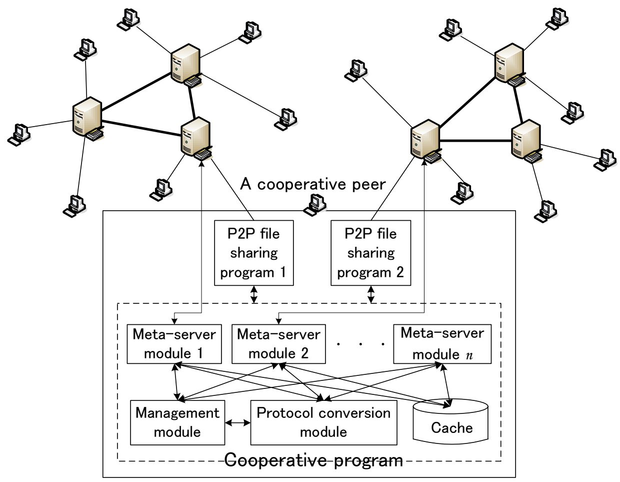

Figure 4. Modules Constituting a Cooperative Program for the SPB Approach

also sends the query message to cooperative meta-servers in host networks via logical links as a peer. The query message is treated as a normal query message in host P2P networks to find a desired file. When the desired file is found in a host P2P network, a response message is generated by a meta-server in the host network and it is sent back to the cooperative meta-server in the guest network. Preparing for future query messages, a cooperative meta-server deposits the meta-information in a response message into its local cache. After applying protocol conversion and replacing information about a provider peer if needed, a response message is sent back to the querying peer following a reverse path of the query message. The peer obtains the desired file directly from a provider peer in a host network or with the mediation of a cooperative meta-server. Thus, also in the SCB approach, the cooperation among hybrid P2P networks is achieved in a transparent way where other meta-servers and peers are unaware of the cooperation.

3 Cooperative Mechanisms for a Shared-Peer-Based Approach

In this section, we describe details about cooperative mechanisms for the SPB approach. Figure 4 illustrates components constituting a cooperative program: a management module, a protocol conversion module, and meta-server modules.

A management module allows a cooperative program to communicate with P2P file-sharing programs, manages the other modules, and decides whether or not to make P2P networks cooperate with each other. Once a cooperative program is initiated, a management module detects P2P file-sharing programs working in a cooperative peer. Then it decides whether or not to perform cooperation taking into account several criteria described in 2.1. To start cooperation, a management module initiates meta-server modules corresponding to each of cooperative P2P networks.

A meta-server module can perform some of meta-server's functions, including relaying query and response messages, caching meta-information into its local cache, and generating a response message. A meta-server module connects to a meta-server in a guest network to receive query messages as a neighboring meta-server. On receiving a query message, a meta-server module forwards it to a protocol conversion module. A protocol conversion module manipulates a query message to fit to a protocol used in host networks and sends it to meta-server modules of host networks. A query message is sent to a meta-server in a host network and disseminated by flooding.

If the corresponding meta-information for a desired file is found on a meta-server in a host network, a response message is generated and it reaches the corresponding meta-server module through the same path, but in the reversed direction, that the query message traversed. A meta-server module deposits meta-information in a response message into its local cache, and then relays the response message to a protocol conversion module. If needed, a protocol conversion module manipulates a response message, replaces a provider peer with itself, and then forwards it to a meta-server module for a guest network. Finally, a meta-server module sends a response message to a neighboring meta-server in a guest network so that it is relayed to a querying peer.

On receiving a response message, a querying peer establishes a logical link to a provider peer recorded in the response message. If the provider peer is a cooperative peer for the need of protocol conversion, the cooperative peer obtains the information about an original provider peer from the corresponding response message in its local cache and retrieves the file in place of the querying peer.

4 Experimental Results

In this section, we conduct simulation experiments to evaluate our proposed cooperative mechanisms for the SPB approach in terms of application-level and system-level performance.

4.1 Simulation Model

Referring to KaZaA's topology [8] [9], we simulate a hybrid P2P file-sharing network by following steps. First, m meta-servers and n peers are randomly placed in a two-dimensional region. Next, a randomly chosen meta-server is connected to the closest meta-server to construct an initial seed of a meta-server network. Then, a meta-server is randomly chosen one by one and is connected to the closest meta-server in a meta-server network. Finally, peers are connected to the closest meta-server. An example of a generated hybrid P2P network is shown in Fig. 5. In our simulation, two hybrid P2P networks are generated in the

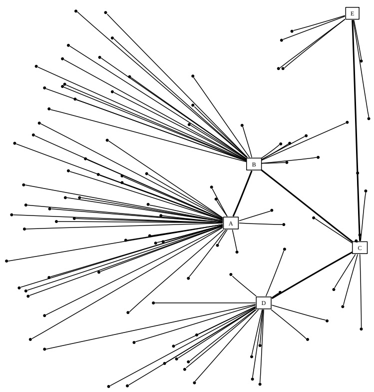

Figure 5. An Example for Hybrid P2P File-Sharing Network Topology ($m = 5, n = 100$)

same manner.

When we consider cooperative P2P networks, c cooperative peers are randomly placed in the two-dimensional region. Each cooperative peer is connected to two meta-servers each of which is the closest to the peer in each of P2P networks. In order to keep the number of peers, c peers are randomly chosen and removed in each network.

F kinds of files are available in two networks. Each file has an identifier f_r ($0 \leq r \leq F-1$), where f_0 is for the most popular file and f_{F-1} is for the least popular one. The popularity of files follows a Zipf distribution with $\alpha = 1.0$. The number of each file exsiting in networks also follows a Zipf distribution with $\alpha = 1.0$, where the number of the least popular file is 1 and the number of the most popular file is F. Files are assigned to peers at random. Peers register meta-information about assigned files to its designated meta-server.

We conducted simulation experiments based on the query-cycle model [10]. Peers generate query messages following the poisson process whose rate λ is randomly chosen from 0 to 0.5 at the uniform distribution. It means that the probability that a peer issues x query messages in one query cycle becomes $p(x) = \frac{e^{-\lambda} \cdot \lambda^x}{x!}$. The probability that a peer generates query message q_r for file f_r is given by the popularity of file f_r. A peer does not issue a query message for a file that it already owns.

In our simulation experiments, we assume that two networks use the same protocol and a cooperative peer does not deposit meta-information into its local cache. We conducted 100 set of simulations of 10 query cycles and show averaged values in the following results. We should note here that a TTL value for a query message is not defined as in a pure P2P file-sharing application, since the most of hybrid P2P file-sharing application do not use a TTL mechanism.

Table 1. Ratio of Available Files and Hit Rate

		Ratio of Available Files	Hit Rate
100:100	network1	0.69	0.90
	network2	0.69	0.89
1000:1000	network1	0.69	0.93
	network2	0.69	0.93
10000:10000	network1	0.69	0.95
	network2	0.70	0.95

4.2 Evaluation of Application-Level Behaviors

In this section, we evaluate our proposed cooperative mechanisms for the SPB approach in terms of the ratio of available files and the hit rate. When two P2P networks cooperate with each other, peers can find files not only in their own network but also in the other. The ratio of available files is defined as the ratio of the number of kinds of files in one network to the total number of kinds of files available in two networks, that is, F. We also define the hit rate as the ratio of the number of query messages whose desired files are found to the total number of query messages. By the definition, the hit rate is one when two P2P networks cooperate.

Table 1 summarizes the ratio of available files and the hit rate for two P2P networks, i.e., network1 and network2, of the same number of peers. In the table, $n1 : n2$ stands for the number of peers in network1 and network2, respectively. We find that the ratio of available files of an independent network is only about $69\% \sim 70\%$, and it increases by about 30% when they cooperate with each other. Furthermore, the hit rate also increases by cooperation regardless of the network size and the degree of increase is higher with smaller networks.

In Fig. 6, we set the number of peers in network2 at 1000 while changing the number of peers in network1 from 100 to 1000. It is shown that the effect is higher for cooperation among imbalance networks, and a network with less number of peers benefits more. For example, the ratio of available files of network1 of 100 peers increases by about 77% when it cooperates with network2, i.e., a ten times larger network.

4.3 Evaluation of System-Level Behaviors

To evaluate the load introduced to cooperative peers and meta-servers by cooperation, we define the load on cooperative peers as the average number of messages that a cooperative peer received and relayed. Analogously, the load on meta-servers is defined as the average number of query messages that a meta-server received and relayed and response messages that a meta-server generated, received, and relayed.

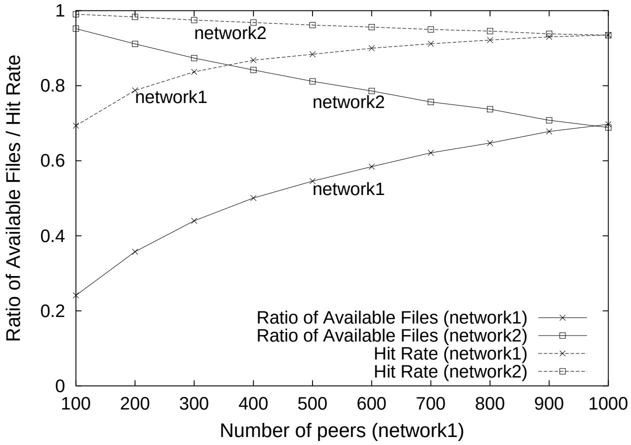

Figure 6. Ratio of Available Files and Hit Rate (100∼1000:1000)

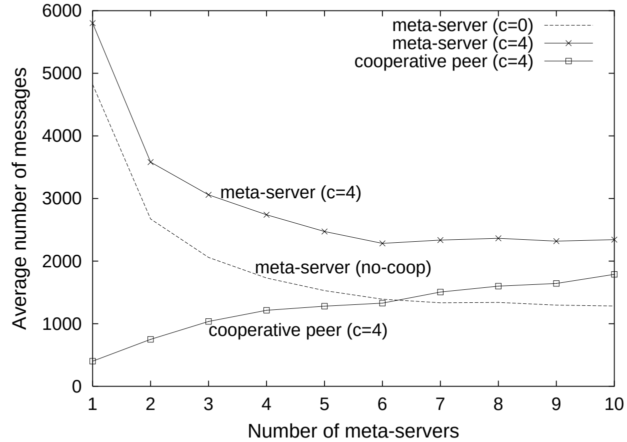

Figure 8. Influence of Number of Meta-servers on Load (1000:1000)

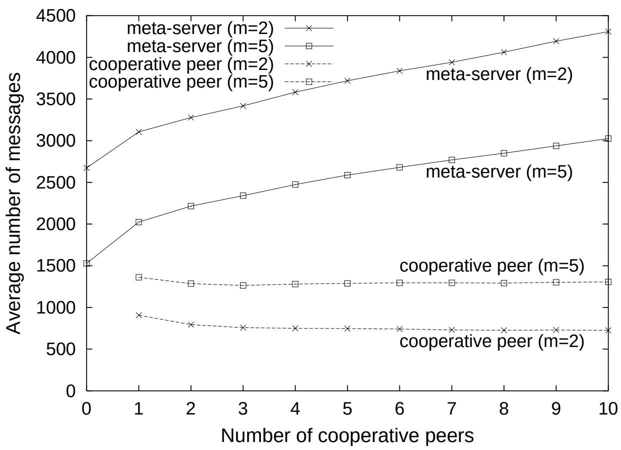

Figure 7. Influence of Number of Cooperative Peers on Load (1000:1000)

Figure 7 shows the average load on cooperative peers and meta-servers in network1 where the number of peers in each P2P network is 1000 and the number of cooperative peers is changed from 0 to 10. The number of cooperative peers 0 means two P2P networks are uncooperative. In the figure, m stands for the number of meta-servers in each P2P network. It is shown that the load on meta-servers almost linearly increases with the increase of the number of cooperative peers. In addition, we also find that the increase in the load on meta-servers by cooperation is almost the same regardless of the number of meta-servers. On the other hand, the load on cooperative peers does not change much. This is because that all cooperative peers relay the query messages independently of others. Consequently, the same query messages are injected into a host P2P network through multiple cooperative peers. This increases the load on meta-servers.

In Fig. 8, the number of meta-servers in each P2P networks changes from 1 to 10. In the figure, c stands for the number of cooperative peers. It is shown that the load on meta-servers decreases with the increase of the number of meta-servers and it is the same for both in cooperative and non-cooperative networks. On the other hand, the load on cooperative peers increases as the number of meta-servers increases. When the number of meta-servers increases, the number of peers per meta-server decreases. Thus, the amount of meta-information deposited in a meta-server decreases. Consequently, the probability that query message does not match any meta-information at the designated meta-server increases. Then, query messages are forwarded to neighboring meta-servers and the number of query messages that a cooperative peer receives increases. Therefore, from a view point of the load on cooperative peers, which is usually less powerful than meta-servers, the cooperation among P2P networks with a small number of meta-servers is desirable.

Figure 9 depicts variations of load with changes in the number of peers in network1 from 100 to 1000. The number of peers in network2 is fixed at 1000. It is obvious that the load on both cooperative peers and meta-servers increases as the number of peers of network1 increases. We also find that the difference in load on meta-servers becomes small with the decrease of difference in size among P2P networks.

To compare the increase of load on meta-servers in networks of different size, Fig. 10 shows the normalized load, derived as the ratio of the number of messages after cooperation to that before cooperation. It can be seen that the load on meta-servers in network1 increases more than that in network2, since more query messages are forwarded from network2 into network1. On the other hand, Fig. 6 indicates that the benefit of network2 in cooperation is smaller than that of network1. It means that the most

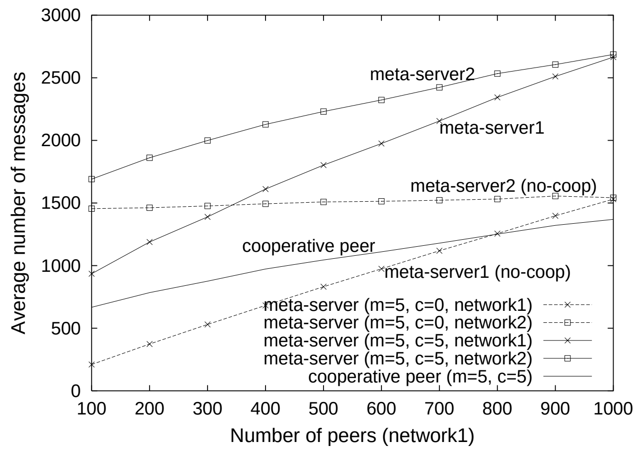

Figure 9. Influence of Size of P2P Network on Load (100∼1000:1000)

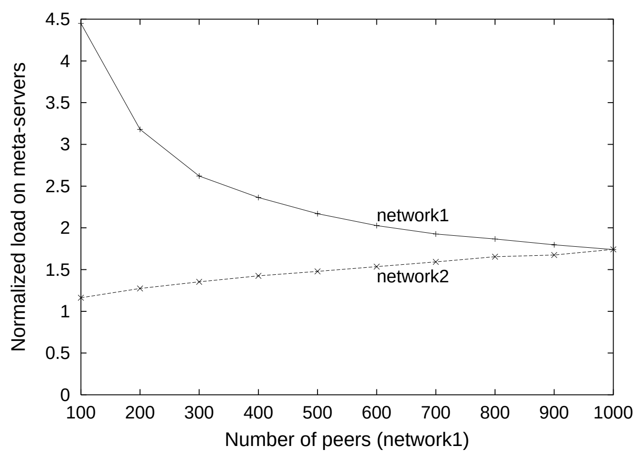

Figure 10. Normalized Load on Meta-servers (100∼1000:1000)

of query messages injected into network1 by network2 are redundant and meaningless. Therefore, we believe that caching meta-information at cooperative peers is effective in reducing the load on meta-servers.

5 Conclusions and Future Work

In this paper, we proposed two cooperative approaches by which two or more hybrid P2P file-sharing networks can efficiently cooperate with each other to improve their collective application-level QoS. Through simulation experiments, we have shown that our proposed cooperative mechanisms for the Shared-Peer-Based approach can improve the application-level QoS in terms of search efficiency, at the sacrifice of the increased load on meta-servers and cooperative peers. We also investigated the influence of network configurations such as the number of peers and meta-servers.

We are now considering an improved scheme to reduce the system load. As several ongoing researches, first we plan to investigate an efficient cache algorithm for cooperative peers, and consider details for deciding whether or not to cooperate. Furthermore we should evaluate the influence of cooperation among overlay networks on physical networks, and investigate cooperative mechanisms which takes into account characteristics of physical networks.

Acknowledgement

This research was partly supported by "The 21st Century Center of Excellence Program" of the Ministry of Education, Culture, Sports, Science and Technology, Japan.

References

[1] A. Nakao, L. Peterson, and A. Bavier, "A Routing Underlay for Overlay Networks," in *Proceedings of ACM SIGCOMM*, pp. 11–18, August 2003.

[2] M. Seshadri and R. H. Katz, "Dynamics of Simultaneous Overlay Network Routing," Tech. Rep. UCB//CSD-03-1291, EECS Department, University of California, Berkeley, November 2003.

[3] I. Stoica, D. Adkins, S. Zhuang, S. Shenker, and S. Surana, "Internet Indirection Infrastructure," in *Proceedings of ACM SIGCOMM*, pp. 78–88, August 2002.

[4] M. Kwon and S. Fahmy, "Toward Cooperative Inter-overlay Networking," in *Proceedings of IEEE ICNP*, November 2003. Poster.

[5] T. Yomo, W.-Z. Xu, and I. Urabe, "Mathematical Model Allowing the Coexistence of Closely Related Competitors at the Initial Stage of Evolution," *Researches on Population Ecology*, vol. 38, no. 2, pp. 239–247, 1996.

[6] N. Wakamiya and M. Murata, "Toward Overlay Network Symbiosis." to be presented at the Fifth International Peer-to-Peer conference (P2P2005), August-September 2005.

[7] R. Schollmeier, "A Definition of Peer-to-Peer Networking for the Classification of Peer-to-Peer Architectures and Applications," in *Proceedings of the First International Conference on Peer-to-Peer Computing (P2P2001)*, pp. 101–102, August 2001.

[8] "KaZaA Homepage." `http://www.kazaa.com`.

[9] J. Liang, R. Kumar, and K. W. Ross, "Understanding KaZaA." `http://cis.poly.edu/~ross/papers/UnderstandingKaZaA.pdf`.

[10] M. T. Schlosser and S. D. Kamvar, "Simulating A File-Sharing P2P Network." `http://www.stanford.edu/~sdkamvar/papers/simulator.pdf`.