Master's Thesis

Title

Proposal and Analysis of Biologically-inspired Symbiotic P2P File-sharing Networks

Supervisor

Professor Masayuki Murata

Author

Tohru Morimoto

February 16th, 2009

Department of Information Networking

Graduate School of Information Science and Technology

Osaka University

Master's Thesis

Proposal and Analysis of Biologically-inspired Symbiotic P2P File-sharing Networks

Tohru Morimoto

Abstract

With emerging needs for application-oriented network services, various overlay networks have been widely deployed over physical IP networks. Since selfish behavior of overlay networks to satisfy demands of their applications and users often conflicts with each other, performance of the overall network system and quality of service offered to users easily deteriorate. To tackle the problem, our research group proposes the framework called *overlay network symbiosis* based on the biological symbiosis model where different bacteria coexist in the shared medium. In the overlay network symbiosis, overlay networks directly and/or indirectly interact with each other through the shared environment and accomplish cooperative or collaborative control.

In this thesis, to demonstrate an example of biologically-inspired symbiotic overlay networks, we propose a mechanism that enables different P2P file-sharing networks to cooperate and live together with mediation of a portal server. We introduce a portal server as the shared environment so that P2P file-sharing networks can cooperate with each other by exchanging search requests and shared files. In our proposed mechanism, the portal server provides users with transparent utilization of multiple P2P file-sharing networks by handling search requests and shared files in place of users. Users can search, get, and share files through the portal server without being aware of existence of P2P file-sharing networks. We model the above proposed mechanism based on the biological symbiosis model. Through numerical analysis based on the mathematical model, we show that the proposed mechanism improves the hit ratio of search requests in comparison to the scenario where P2P file-sharing networks are independent, especially in the case that number of search requests and shared files are too small for P2P file-sharing networks to endure.

Keywords

Biological Symbiosis Model

P2P File-sharing

Cooperative Network

Numerical Analysis

Contents

List of Figures

List of Tables

1 Introduction

With emerging needs for application-oriented network services, various overlay networks such as P2P (Peer-to-Peer) networks, Grid networks, and CDN (Content Delivery Network) have been widely deployed over physical IP networks. They are different in targeted application-oriented performance, network topology, and the amount and pattern of communication. For example, a Grid network requires transmission of vast amount of data with ideally small delay for high-performance distributed computing and data processing. On the contrary, a P2P file-sharing network wants to disseminate a small data, i.e. query message over the whole overlay network to find a desired file as fast as possible. Since each overlay network behaves in a selfish manner to satisfy demands of its applications and users, co-existence of multiple overlay networks often causes various problems [1, 2, 3, 4]. For example, when overlay networks share and compete for the same physical network resources such as link and router, chain of selfish control leads to performance degradation and even the instability of a system. Let's assume that an overlay network changes its topology to use less congested physical links to enhance throughput. Other overlay networks using those uncongested links experience performance degradation caused by increase of traffic.

Since the affected overlay networks are also selfish and greedy in improving their performance, they actively change their topology accordingly. It further triggers reaction of other overlay networks. Consequently, the influence extends to the whole network. As an another example, let us consider competition of P2P file-sharing networks for information resource. Each network attempts to attract more users and increase the number and kinds of shared files by a user-friendly interface, high hit ratio of search, fast file retrieval, and anonymity. Because of the diversity in usability, performance, and type of shared files, users may prefer one network to others and share their files on chosen P2P file-sharing networks. Consequently, the availability of files differs among networks as commonly observed in current P2P file-sharing services. Therefore, users need to participate in two or more P2P file-sharing networks to get their desired files or share their files with many other users. It implies there exist redundant and duplicated files in multiple P2P file-sharing networks leading to the waste of storage and network resources.

In order to improve the performance of the overall system, several cooperative mechanisms such as information exchange among overlay networks [5, 6, 7] and routing overlay [8, 9] have been proposed. In [6], the authors investigated a spectrum of cooperation among coexisting overlay networks. They described variety of cooperation such as sharing measurement information, sharing control information, cooperative query forwarding, and inter-overlay traffic engineering. As an example, they proposed an architecture called Synergy where overlay networks cooperated with each other in inter-overlay routing. In this architecture, nodes designated to inter-overlay routing are selected based on the number, location, and separation of nodes in each of the overlay networks. The synergy network consisting of those chosen nodes relays long-lived flows so that they traverse better paths than ones determined by the physical routing. It is shown that such inter-overlay routing improves performance in terms of latency, throughput, and loss. In [8], so-called RON (Resilient Overlay Networks) nodes in different routing domains compose a mesh overlay network and control path selection over RON. RON is proposed to accomplish faster failure recovery in routing. In general, BGP-4 running at the border routers between AS's takes a long time, on the order of several minutes, to converge to a new route after a link failure occurs over physical IP networks. To react to link failure faster, RON nodes monitor the quality of their virtual links established over physical IP networks frequently. They exchange the information of the observed quality of virtual links with each other for all RON nodes to maintain up-to-date information of the overlay network. When a link failure occurs in physical IP networks, RON nodes detect it immediately and select an alternative path over RON on the order of several seconds. P4P can also be considered as a framework of overlay network cooperation [10, 11]. In a P4P network, an information server called iTracker is prepared by a network service provider for the use of P2P networks. An iTracker offers up-to-date information about the status of physical network, e.g. topology, and it can be regarded as a mechanism for P2P networks to share the physical network information. By using the information, P2P networks can control their behavior in a cooperative manner.

Our research group considers the framework called *overlay network symbiosis* for cooperation among overlay networks that share and compete for network and information resources [12]. In the overlay network symbiosis, cooperation is based on the mathematical model of symbiotic living organisms in the ecosystem. In the ecosystem, symbiosis is often observed among living organisms of different species, groups, and individuals in the shared environment. Symbiosis emerges from direct and/or indirect interaction among organisms. In [13], the authors established the mathematical model of biological symbiosis where closely related bacterial strains lived together in a reactor by exchanging metabolites through their cell and the medium. Based on the biological symbiosis model, we can model and analyze symbiosis among overlay networks. We regard overlay networks as bacteria, direct interaction such as message exchanges and indirect interaction such as competition for shared resources as exchange of metabolites among cells, and the shared environment such as physical networks, inter-overlay network, and some mediation mechanism as a reactor. For example, in case of cooperative query forwarding, a reactor corresponds a mechanism which relays search and response messages, i.e. metabolites, among overlay networks. Chemical reactions of metabolites in bacterial strains corresponds to generation of a response message against a search message.

In this thesis, to demonstrate an example of biologically-inspired symbiotic overlay networks, we propose a mechanism that enables different P2P file-sharing networks to cooperate and live together with mediation of a portal server. We introduce a portal server as the shared environment so that P2P file-sharing networks can cooperate with each other in sharing files by exchanging search requests and shared files over a portal server. In our proposed mechanism, the portal server provides users with transparent utilization of multiple P2P file-sharing networks by handling search requests and shared files in place of users. Users can search, get, and share files through the portal server without being aware of existence of P2P file-sharing networks. We model the above proposed mechanism based on the biological symbiosis model by regarding a portal server as a reactor, P2P file-sharing networks as bacterial strains, and information resources as metabolites. Through numerical analysis based on the mathematical model, it is shown that the proposed mechanism improves the hit ratio of search requests in comparison to the scenario where P2P file-sharing networks are independent, especially in the case that number of search requests and shared files are too small for P2P file-sharing networks to endure.

The rest of the thesis is organized as follow. In Section 2, we briefly introduce the mathematical model of co-existence of bacterial strains and then explain the overlay network symbiosis. Next, we propose a mechanism and a model of biologically-inspired symbiotic P2P file-sharing networks in Section 3. Then, in Section 4, we show results of numerical analysis, where the effectiveness of symbiosis is evaluated by the hit ratio of search requests. Finally, we conclude the thesis and describe future work in Section 5.

2 Overlay Network Symbiosis

In this section, we introduce the mathematical model of co-existence of bacterial strains and the overlay network symbiosis proposed based on the biological symbiosis model.

2.1 Biological Symbiosis Model

In [13], the authors proposed a mathematical model of a mechanism that permitted two types of bacterial strains, i.e. E. coli, to live together by exchanging metabolites through shared media, i.e. culture in a reactor. Bacterial strains have a metabolic network of generating metabolite S_2 from other metabolite S_1. Metabolites diffuse in and out of a cell through membrane depending on the difference in metabolic concentrations (Fig. 1). Two strains are different in the speed of metabolite conversion in the metabolic network.

Temporal dynamics of metabolite concentrations in a cell of strain $i \in \{A, B\}$ are formulated as,

$$\frac{ds_1^{(i)}}{dt} = \frac{P}{V}(s_1^{(R)} - s_1^{(i)}) - (k_{1,2}^{(i)} + k_p)s_1^{(i)}, \tag{1}$$

$$\frac{ds_2^{(i)}}{dt} = \frac{P}{V}(s_2^{(R)} - s_2^{(i)}) + k_{1,2}^{(i)}s_1^{(i)} - k_p s_2^{(i)}, \tag{2}$$

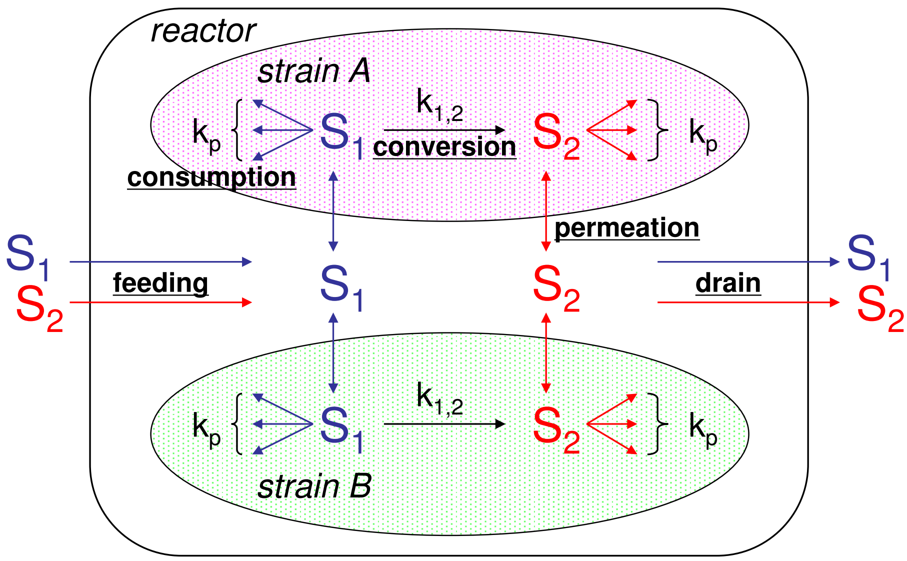


Figure 1: Symbiosis model of bacteria

where P stands for the permeation coefficient of cell membrane and V does for the average volume of a cell. $s_{\{1,2\}}^{(i)}$ and $s_{\{1,2\}}^{(R)}$ are metabolite concentrations in a cell of strain i and in the reactor, respectively. k_p is the metabolite consumption rate in a cell. $k_{1,2}^{(i)}$ is the metabolite conversion rate in a cell of strain i. The first term of Eqs. (1) and (2) stands for the permeation of metabolites through cell membrane depending on their concentrations in and out-of a cell. The second term of Eq. (1) means that the concentration of metabolite S_1 decreases for conversion and consumption. On the other hand, the concentration of metabolite S_2 increases for conversion from metabolite S_1 as expressed by the second term of Eq. (2) while being consumed within a cell by the third term.

Next, metabolite concentrations in the reactor evolve as,

$$\frac{ds_1^{(R)}}{dt} = D(s_1^{(0)} - s_1^{(R)}) + \sum_{i\in\{A,B\}} X^{(i)} P(s_1^{(i)} - s_1^{(R)}), \tag{3}$$

$$\frac{ds_2^{(R)}}{dt} = D(s_2^{(0)} - s_2^{(R)}) + \sum_{i\in\{A,B\}} X^{(i)} P(s_2^{(i)} - s_2^{(R)}), \tag{4}$$

where $X^{(i)}$ stands for the number of cells of strain i per volume in the reactor. The fresh medium containing metabolites of concentration $s_{\{1,2\}}^{(0)}$ is added to the reactor at the constant rate and the culture is drained at the same rate, i.e. chemostatic culture. The first term of Eqs. (3) and (4) corresponds to dilution of the culture where D is the dilution rate. Furthermore, the metabolite concentrations change for permeation with cells as the second term expresses.

Change in population of cells is formulated as,

$$\frac{dX^{(i)}}{dt} = \mu^{(i)} X^{(i)} - DX^{(i)}, \tag{5}$$

where the growth rate $\mu^{(i)}$ is defined as,

$$\mu^{(i)} = \alpha s_1^{(i)} s_2^{(i)}. \tag{6}$$

Eq. (6) implies that a cell with high metabolite concentrations grows fast. Here, $\alpha > 0$ is a constant.

Figures 2 and 3 show results of numerical analysis where $s_1^{(0)} = 10.0$, $s_2^{(0)} = 0.0$, $\frac{\alpha}{D} = 1.0$, $\frac{P}{D} = 1.0$, $\frac{k_p V}{P} = 1.0$, $\frac{k_{1,2}^{(A)} V}{P} = 5.0$, and $\frac{k_{1,2}^{(B)} V}{P} = 0.4$. In the figures, X axis corresponds to time in unit of D and Y axis shows the population of cells and the concentration of metabolite S_2, respectively. At first there is only bacterial strain A in the reactor. At time $10D$, bacterial strain B, which generates metabolite S_2 from metabolite S_1 ten-times slower than bacterial strain A, where

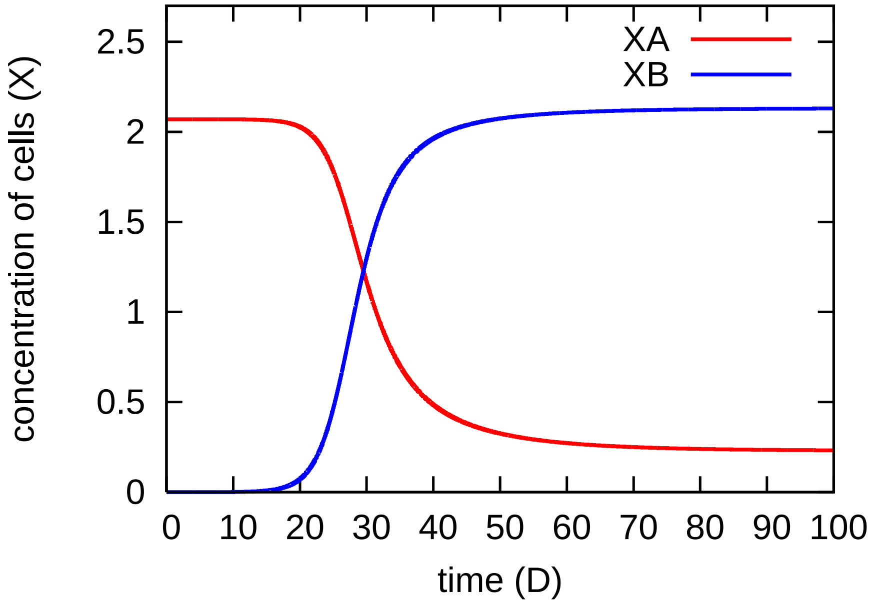


Figure 2: Population of bacterial strains

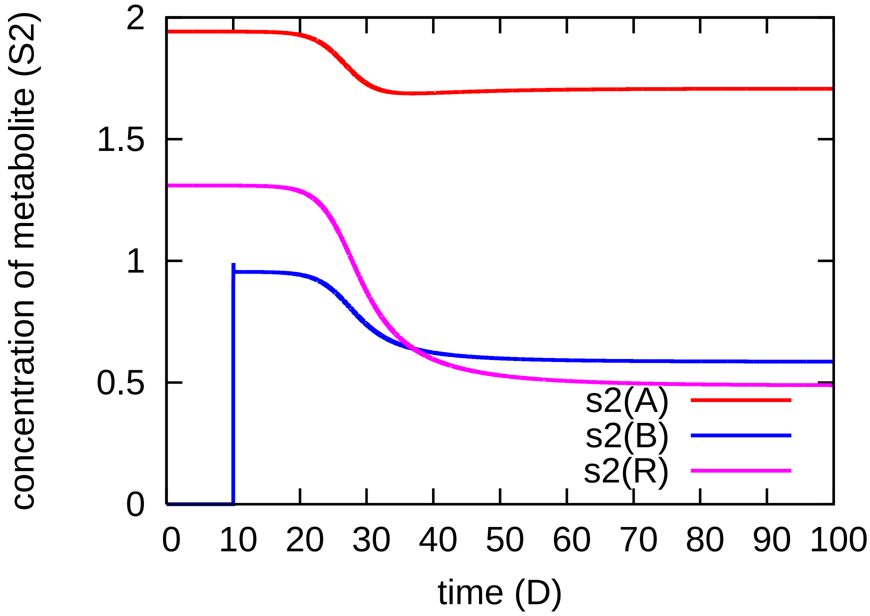


Figure 3: Concentration of metabolite S_2

$k_{1,2}^{(B)} < k_{1,2}^{(A)}$, is introduced into the reactor. As shown in Fig. 2, the population of bacterial strain A that consumes metabolite S_1 faster than bacterial strain B decreases after that after a while, the concentrations of bacterial strains in the reactor become constant at time $90D$ and both are larger than zero. That is, they live together. In Fig. 3, it can be seen that $s_2^{(R)} < s_2^{(B)} < s_2^{(A)}$ holds in the stable condition. It implies that metabolite S_2 permeates cell membrane of both bacterial strains A and B to the reactor. Depending on parameter setting, symbiotic conditions where both bacterial strains take metabolites from the reactor, i.e. $s_2^{(B)} < s_2^{(R)}$ and $s_2^{(A)} < s_2^{(R)}$, or one bacterial strain supplies metabolites to another bacterial strain, e.g. $s_2^{(B)} < s_2^{(R)} < s_2^{(A)}$, also appear.

2.2 Biologically-inspired Overlay Network Symbiosis

Our research group proposes the framework called *overlay network symbiosis* based on the biological symbiosis model [12]. In [14], we regarded a reactor as a system, bacterial strains as overlay networks that offered a service to users, metabolite S_1 as a group of users, metabolite S_2 as the shared resource, the metabolite conversion rate in a cell as the number of users served per unit time, i.e. the service rate or service capacity of network, and X as the size of a network. Based on the mathematical model, we investigated conditions that made competing overlay networks coexist. We showed that among available overlay networks more users preferentially received the service from a less loaded network, i.e. network with the lower metabolic concentration $s_1^{(i)}$. We also observed that network i with high metabolic concentration $s_2^{(i)}$ released the occupied resource for the use of other networks. More importantly, we revealed that there were conditions where a single overlay network could not survive alone but could live together by harmonious coexistence of other networks.

In the context of the overlay network symbiosis, we also proposed mechanisms for pure and hybrid P2P file-sharing networks to interact and cooperate with each other [15, 16]. For example, in [15], cooperation among pure P2P networks is accomplished by exchanging search and response messages through logical connections established among so-called cooperative peers. Through simulation experiments, it is shown that more provider peers are found within the proximity of a searching peer by cooperation of P2P file-sharing networks. However, to allow P2P file-sharing networks to exchange messages, we need to introduce a special program, called a cooperative program, to a peer. To have better and moderate cooperation, we propose a mechanism that P2P file-sharing networks cooperate with each other while users, peers, and P2P file-sharing networks are unaware of a driving force of cooperation in Section 3.

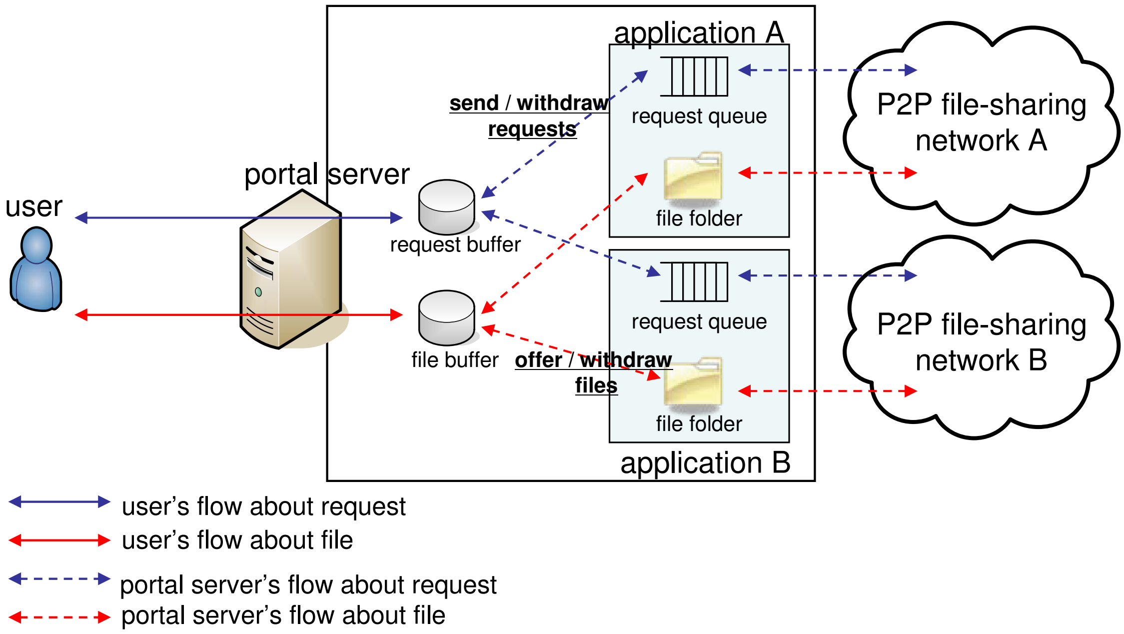


Figure 4: Symbiotic P2P file-sharing networks with a portal server

3 Biologically-inspired Symbiotic P2P File-sharing Networks

In this section, as an example of symbiosis of overlay networks based on our overlay network symbiosis, we propose a mechanism of symbiotic P2P file-sharing networks with a portal server and its mathematical model for analysis.

3.1 Symbiotic P2P File-sharing Networks with Portal Server

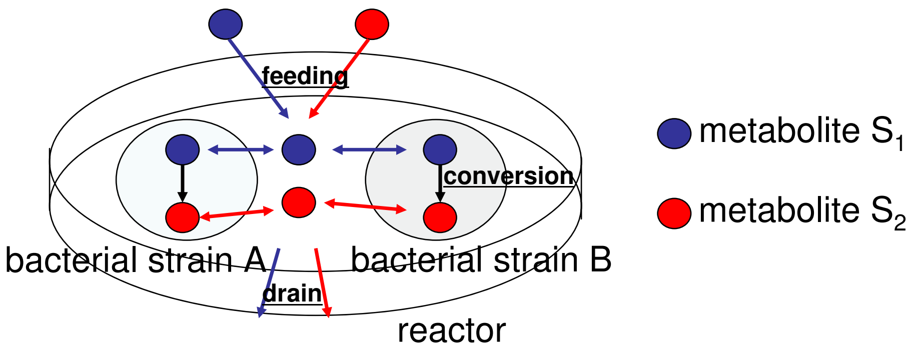
We assume that there are various P2P file-sharing networks. So that P2P file-sharing networks can cooperate with each other in sharing files by exchanging search requests and shared files, we introduce a portal server as the shared environment. Figure 4 illustrates the proposed architecture. A portal server belongs to multiple P2P file-sharing networks as a peer in order to send and cancel search requests and to offer and obtain shared files in place of users. Users can search, get, and share files through a portal server without being aware of existence of P2P file-sharing networks. Since a portal server belongs to a P2P file-sharing network as a normal peer, the network and other peers participating to the network are unaware of the existence of the cooperation mechanisms implemented on the portal server. That is, P2P file-sharing networks are made cooperative without noticing.

When a user registers information resources such as a search request and a file to share to a portal server, the portal server first deposits them in its corresponding buffers. Depending on condition of P2P file-sharing networks, it issues or withdraws a request in a request queue of a corresponding P2P file-sharing application and puts or withdraws a file in a shared file folder of a corresponding P2P file-sharing application. For example, when number of files shared in a P2P file-sharing network is small, a portal server supplies files from its file buffer to the network in order to foster sharing and exchanging files in the network. On the contrary, a portal server withdraws files from a loaded P2P file-sharing network and supplies them to other networks. When a request is served by a P2P file-sharing network and a portal server obtains a corresponding file from a peer participating in the network, it is deposited in the shared file folder or the file buffer while sending it to the requesting user.

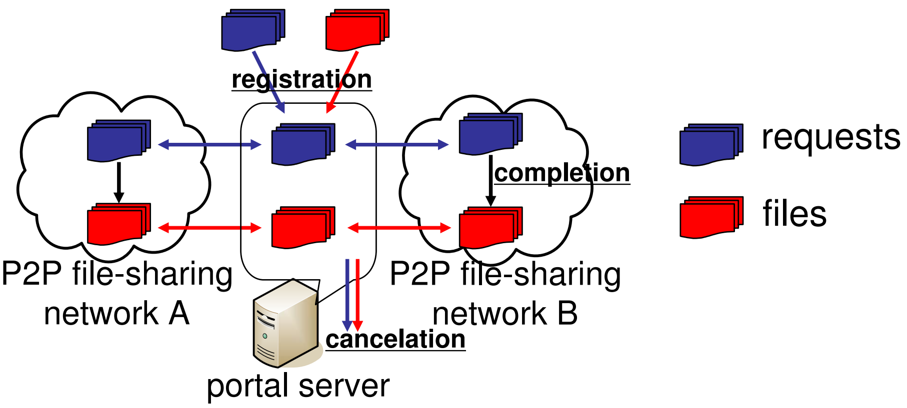
3.2 Biologically-inspired Model of Symbiotic P2P File-sharing Networks

We can model the above proposed mechanism based on the biological symbiosis model by regarding a portal server as a reactor, P2P file-sharing networks as bacterial strains, requests as metabolite S_1, and files as metabolite S_2 (Fig. 5). Registration of requests and files on a portal server by users corresponds to addition and drain of culture media to and from the reactor. A portal server adjusts the number of requests to be served by, and the number of files to be shared on P2P file-sharing networks depending on the condition of each network. When we regard requests and files as metabolites, this corresponds to exchange of metabolites between bacterial strains through the medium in a reactor.

However we cannot directly adopt the biological symbiosis model explained in Section 2.1 to model the symbiotic P2P file-sharing networks. In P2P file-sharing networks, there exist users participating to P2P file-sharing networks without mediation of a portal server. We call them direct users hereafter. Direct users are peers constituting P2P file-sharing networks. In contrast, users of a portal server are called portal users, which do not belong to any P2P file-sharing network under consideration. Direct users send requests and upload files directly to a P2P file-sharing network and obtain files directly from a P2P file-sharing network (Fig. 6). Such direct interaction of direct users with P2P file-sharing networks corresponds to direct injection and extraction of metabolites to and from bacterial strains. However, neither of dynamics of metabolite concentrations in a cell,



(a) Biological symbiosis model



(b) Model of symbiotic P2P file-sharing networks

Figure 5: Relation between biological symbiosis model and model of symbiotic P2P file-sharing networks

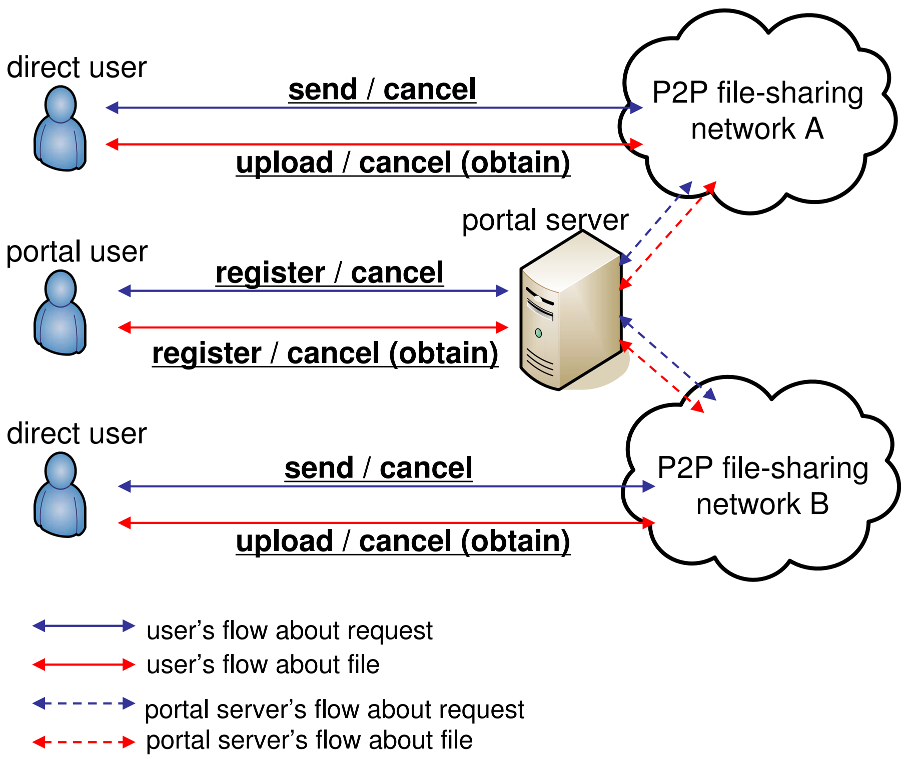


Figure 6: Behavior of direct user, portal user, and portal server

i.e. Eqs. (1) and (2) has such a term.

To take into account direct users, we extend the bacterial symbiosis model illustrated in Fig. 1 to a new model in Fig. 7. The difference is existence of arrows connecting inside of strains to outside of the reactor. When we define the metabolite concentrations added to the whole system as $s_{\{1,2\}}^{(U)}$ and the volume of reactor as V_R, $s_{\{1,2\}}^{(U)} V_R$ corresponds to the number of metabolites S_1 and S_2 in the fresh medium. Among them, $s_{\{1,2\}}^{(0)} V_R$ is added to the culture in the reactor and the remaining $V_R(s_{\{1,2\}}^{(U)} - s_{\{1,2\}}^{(0)})$ is directly added to bacterial cells. Here, $s_{\{1,2\}}^{(U)} - s_{\{1,2\}}^{(0)}$ means the metabolite concentration added to bacterial cells. Assuming that the fresh medium is evenly added to both strains, temporal dynamics of metabolite concentrations in a cell can be re-formulated as,

$$\frac{ds_1^{(i)}}{dt} = \frac{P}{V}(s_1^{(R)} - s_1^{(i)}) - (k_{1,2}^{(i)} + k_p)s_1^{(i)} + M_t\{\frac{1}{2}(s_1^{(U)} - s_1^{(0)}) - s_1^{(i)}\}, \quad (7)$$

$$\frac{ds_2^{(i)}}{dt} = \frac{P}{V}(s_2^{(R)} - s_2^{(i)}) + k_{1,2}^{(i)} s_1^{(i)} - k_p s_2^{(i)} + M_t\{\frac{1}{2}(s_2^{(U)} - s_2^{(0)}) - s_2^{(i)}\}, \quad (8)$$

where M_t stands for addition and drain rate of metabolites to and from bacterial strains per unit

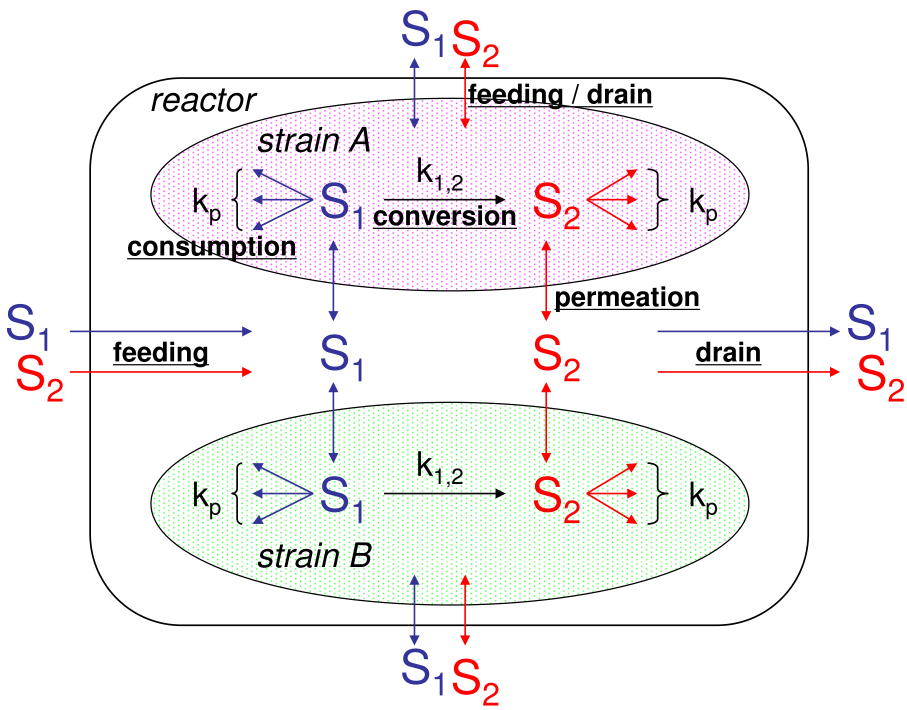


Figure 7: Extended bacterial symbiosis model

time. Other dynamics in the reactor still conform to the original model, i.e. temporal dynamics of metabolite concentrations in the reactor follow Eqs. (3) and (4) and change in population of cells follows Eq. (5).

We summarize parameter definition in the mathematical symbiosis model of P2P file-sharing networks in Table 1. In the table, assuming that volume of cell is identical and one, we regard concentration as number. Based on the definitions, we derive temporal dynamics of symbiotic P2P file-sharing networks as follows.

First, temporal change in the number $s_1^{(i)}$ of requests being served per peer in P2P file-sharing network i is given by the following differential equation.

$$\frac{ds_1^{(i)}}{dt} = P(s_1^{(R)} - s_1^{(i)}) - k_{1,2}^{(i)} s_1^{(i)} - k_p' s_1^{(i)} + \frac{1}{2} M_t (s_1^{(U)} - s_1^{(0)}). \tag{9}$$

where we denote $k_p + M_t$ as k_p'. $s_1^{(i)}$ is a quotient of the total number of requests divided by the number of participating peers in P2P file-sharing network i. $s_1^{(i)}$ changes in relation to the number $s_1^{(R)}$ of requests that a portal server holds (the first term in the right-hand side). The condition that $s_1^{(i)}$ is more than $s_1^{(R)}$ implies that more peers are searching or downloading files.

Table 1: Parameter definition

parameter	definition
$s_1^{(i)}$	the number of requests being served per peer in P2P file-sharing network i
$s_2^{(i)}$	the number of shared files per peer in P2P file-sharing network i
$s_1^{(R)}$	the number of requests that a portal server holds in buffer
$s_2^{(R)}$	the number of files that a portal server holds in buffer
$s_1^{(0)}$	the number of new requests that portal users register to a portal server per unit time
$s_2^{(0)}$	the number of new files that portal users register to a portal server per unit time
$s_1^{(U)}$	the total number of new requests that portal users and direct users issue per unit time
$s_2^{(U)}$	the total number of new files that portal users and direct users provide per unit time
$k_{1,2}^{(i)}$	rate of search completion in P2P file-sharing network i per unit time
k'_p	rate of disappearance of information resources from P2P file-sharing networks per unit time
P	rate of exchange of information resources between a P2P file-sharing network and a portal server per unit time
D	rate of registration and cancellation of information resources to and from a portal server by portal users per unit time
M_t	rate of uploading and downloading of information resources to and from P2P file-sharing networks by direct users per unit time
$\mu^{(i)}$	growth rate of P2P file-sharing network i
$X^{(i)}$	the number of participating peers in P2P file-sharing network i
α	growth coefficient ($\alpha > 0$)

Then, to reduce the load, a portal server withdraws requests from the P2P file-sharing network and consequently $s_1^{(i)}$ decreases. $s_1^{(i)}$ decreases when corresponding downloading finishes (second term) and decreases for cancellation (third term). $s_1^{(i)}$ increases when direct users issue requests on P2P file-sharing network i (fourth term).

Next, temporal change in the number $s_2^{(i)}$ of files shared per peer in P2P file-sharing network i can be given by the following differential equation.

$$\frac{ds_2^{(i)}}{dt} = P(s_2^{(R)} - s_2^{(i)}) + k_{1,2}^{(i)} s_1^{(i)} - k_p' s_2^{(i)} + \frac{1}{2} M_t (s_2^{(U)} - s_2^{(0)}). \quad (10)$$

$s_2^{(i)}$ is a quotient of the total number of shared files divided by the number of participating peers in P2P file-sharing network i. $s_2^{(i)}$ changes in relation to $s_2^{(R)}$ (first term). The condition that $s_2^{(i)}$ is more than $s_2^{(R)}$ implies that the P2P file-sharing network has a sufficient number of files. Then, a portal server stops offering files to the network to prevent excessive supply. $s_2^{(i)}$ increases when a portal server and direct users finish downloading files (second term) and decreases when a portal server and direct users stop sharing files (third term). $s_2^{(i)}$ increases when direct users upload files to share (fourth term).

Temporal change in the number $s_1^{(R)}$ of requests that a portal server holds in its buffer is given by the following differential equation.

$$\frac{ds_1^{(R)}}{dt} = D(s_1^{(0)} - s_1^{(R)}) + \sum_{i \in \{A,B\}} X^{(i)} P(s_1^{(i)} - s_1^{(R)}). \quad (11)$$

$s_1^{(R)}$ increases when portal users register requests and decreases for cancellation (first term). To search files efficiently, a portal server sends requests to a P2P file-sharing network with the small number of requests being served, i.e. $s_1^{(i)} < s_1^{(R)}$ or a P2P file-sharing network with the large number of participating peers (second term). On the other hand, a portal server withdraws requests from a P2P file-sharing network with the large number of requests being served, i.e. $s_1^{(i)} > s_1^{(R)}$.

Temporal change in the number $s_2^{(R)}$ of files that a portal server holds in its buffer is given by the following differential equation.

$$\frac{ds_2^{(R)}}{dt} = D(s_2^{(0)} - s_2^{(R)}) + \sum_{i \in \{A,B\}} X^{(i)} P(s_2^{(i)} - s_2^{(R)}). \quad (12)$$

$s_2^{(R)}$ increases when portal users register files and decreases for cancellation (first term). A portal server uploads or withdraws files in relation to $s_2^{(i)}$ and $X^{(i)}$, i.e. the number of participating peers (second term).

Temporal change in the number $X^{(i)}$ of participating peers in P2P file-sharing network i is given by the following differential equation.

$$\frac{dX^{(i)}}{dt} = \mu^{(i)} X^{(i)} - D X^{(i)}. \tag{13}$$

$X^{(i)}$ increases when a new user participates in and decreases for leave of peers. The growth rate $\mu^{(i)}$ is defined as a product of the number $s_1^{(i)}$ of requests and the number $s_2^{(i)}$ of shared files in P2P file-sharing network i as,

$$\mu^{(i)} = \alpha s_1^{(i)} s_2^{(i)}, \tag{14}$$

where $\alpha > 0$ is a constant.

4 Numerical Analysis

In this section, we evaluate biologically-inspired symbiotic P2P file-sharing networks through numerical analysis based on the mathematical model.

4.1 Analysis Setting

We set P, k_p, and α at 1.0, and D at 0.01 [14]. Since the rate of registration and withdrawal of information resources to and from a portal server by portal users per unit time conforms to the rate of uploading and downloading of information resources to and from P2P file-sharing networks by direct users per unit time, we set M_t at 0.01. The total number $s_1^{(U)}$ of new requests per unit time is set at 15.0 and the total number $s_2^{(U)}$ of new files per unit time is set at 5.0. In the case of Gnutella, the average number of new search requests per minute is about 12000 in 2006 in [17]. On the other hand, to the best of our knowledge, there is no observation on the number of new files offered to a P2P network per unit time in literatures. Therefore, in this thesis, assuming that a user is eager to get files while hesitating in providing his files to other users [18], we empirically set the ratio of the number of new requests to the number of new files as 1/3.

Among new requests and files, those registered at the portal server are $s_1^{(0)} = 7.5$ and $s_2^{(0)} = 2.5$ assuming that the half of new requests and files are from portal users. We will change these ratio in the numerical analysis in section 4.2. The ratio U_R of the requests to be registered at the portal server and the ratio U_F of the files to be registered at the portal server by portal users are formulated as,

$$U_R = \frac{s_1^{(0)}}{s_1^{(U)}}, \tag{15}$$

$$U_F = \frac{s_2^{(0)}}{s_2^{(U)}}. \tag{16}$$

Following the above parameter setting, the ratio U_R and U_F are 0.5 respectively.

We assume that there are two P2P file-sharing networks A and B whose service rate are $k_{1,2}^{(A)} > k_{1,2}^{(B)}$. First, we change the service rate $k_{1,2}^{(A)}$ from 0.1 to 4.0 and the service rate $k_{1,2}^{(B)}$ from 0.01 to 0.6 to investigate whether two kinds of P2P file-sharing networks with the different service rate can live together. Next, we change the ratio U_R from 0.01 to 1.0 and the ratio U_F from 0.01 to 1.0 by changing $s_1^{(0)}$ and $s_2^{(0)}$ accordingly to investigate whether two kinds of P2P file-sharing networks, e.g. $k_{1,2}^{(A)} = 1.0$ and $k_{1,2}^{(B)} = 0.1$, or $k_{1,2}^{(A)} = 2.0$ and $k_{1,2}^{(B)} = 0.1$, can live together under different condition of utilization of a portal server. We consider that a P2P file-sharing network is alive when the number of participating peers is larger than threshold H, which is empirically set at 0.00002. We should note here that absolute values of parameters are not realistic values. However we can analyze system behavior from their relative relationship as stated above.

We use the hit ratio as a performance measure. The hit ratio is the ratio of requests that can find a desired file in P2P file-sharing networks to the total number of requests. It is formulated as,

$$Hit\ ratio = \frac{\sum_{i\in\{A,B\}} X^{(i)} s_1^{(i)} k_{1,2}^{(i)}}{D s_1^{(0)} + M_t(s_1^{(U)} - s_1^{(0)})}. \tag{17}$$

In numerical analysis, we consider two scenarios. Scenario 1 is the case where there are both of direct and portal users. That is, two P2P file-sharing networks are mediated by a portal server to operate cooperatively. Scenario 2 is the case where the portal server doesn't exist and there are only direct users. That is, P2P file-sharing networks are independent from each other and compete for shared files. To analyze scenario 2, we set parameters as $P = 0.0$, $s_1^{(0)} = 0.0$, $s_2^{(0)} = 0.0$, $s_1^{(R)} = 0.0$, and $s_2^{(R)} = 0.0$.

4.2 Numerical Results

In Fig. 8, each point indicates a scenario which leads to the higher hit ratio with combinations of the service rate $k_{1,2}^{(A)}$ and $k_{1,2}^{(B)}$. The region with points indicates conditions where P2P file-sharing networks in scenario 1 live together. The truth is that the hit ratio of scenario 2 is always 0 in the figure. That is, independent P2P file-sharing networks cannot survive when $k_{1,2}^{(A)}$ and $k_{1,2}^{(B)}$ are within the analyzed region, because the rates $s_1^{(U)}$ and $s_2^{(U)}$ are too small to cultivate two P2P file-sharing networks. It further means that cooperation through a portal server enables P2P file-sharing networks to live and offer service to users in the nonviable condition, while the total amounts of new search requests and shared files are the same between scenario 1 and scenario 2. Another important finding is that the larger difference in the service rates leads to the higher chance of cooperation and survival. The reason can be explained as follows. P2P file-sharing network B with the small service rate has the excessive number of requests and the insufficient number of shared files. The other P2P file-sharing network A with the large service rate has the insufficient number of requests and the sufficient number of shared files. When the difference in the service rate is large among the P2P file-sharing networks, they can supplement each other by supplying insufficient information resources with mediation of a portal server. On the other hand,

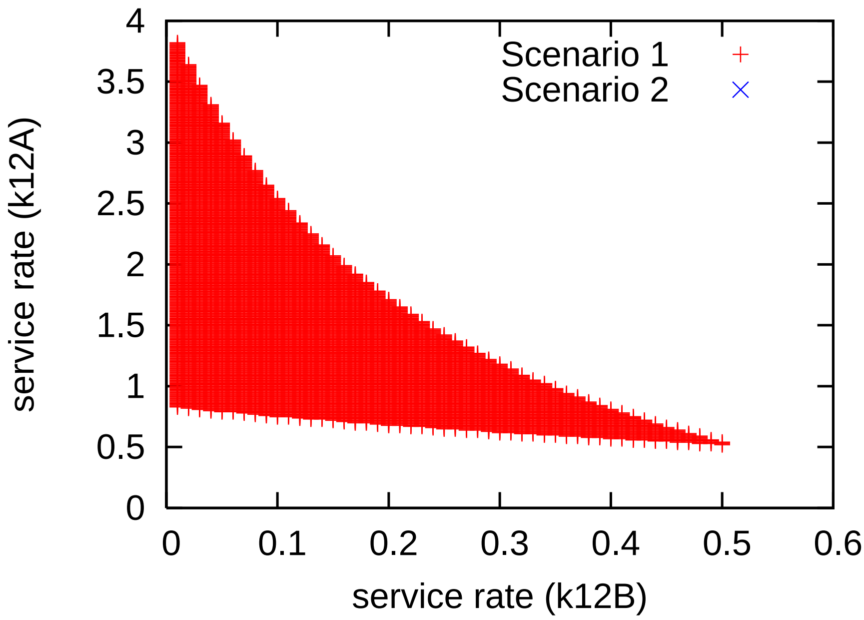


Figure 8: Scenarios leading to higher hit ratio

when the difference in the service rate is small among them, both P2P file-sharing networks lack both of search requests and shared files and cannot help each other. What is even worse, they cannot maintain themselves.

Figure 9 shows transition of the hit ratio when $k_{1,2}^{(A)} = 1.0$ and $k_{1,2}^{(B)} = 0.1$. As shown in Fig. 9, the hit ratio decreases and becomes zero, when two networks are independent in scenario 2. The reason can be explained as follows. Since the total number $s_1^{(U)}$ of new requests to P2P file-sharing networks is small, the number $s_1^{(i)}$ of requests and the number $s_2^{(i)}$ of shared files do not increase enough. It implies that networks are not effectively used or activated enough to grow. Users leave from P2P file-sharing networks and the networks eventually die. On the other hand, the hit ratio increases and becomes constant at time $50D$ when a portal server is introduced in scenario 1. This is because the portal server efficiently utilizes the small number of requests and files by using P2P file-sharing networks cooperatively.

To analyze behavior of the portal server, we compare the number of requests and files that the portal server holds to that shared in P2P file-sharing networks in Fig. 10. Case 1 indicates the condition where $s_1^{(R)} > s_1^{(A)}$, $s_1^{(R)} > s_1^{(B)}$, $s_2^{(R)} > s_2^{(A)}$, and $s_2^{(R)} > s_2^{(B)}$ hold. That is, a portal server supplies both P2P file-sharing networks with both of search requests and shared files

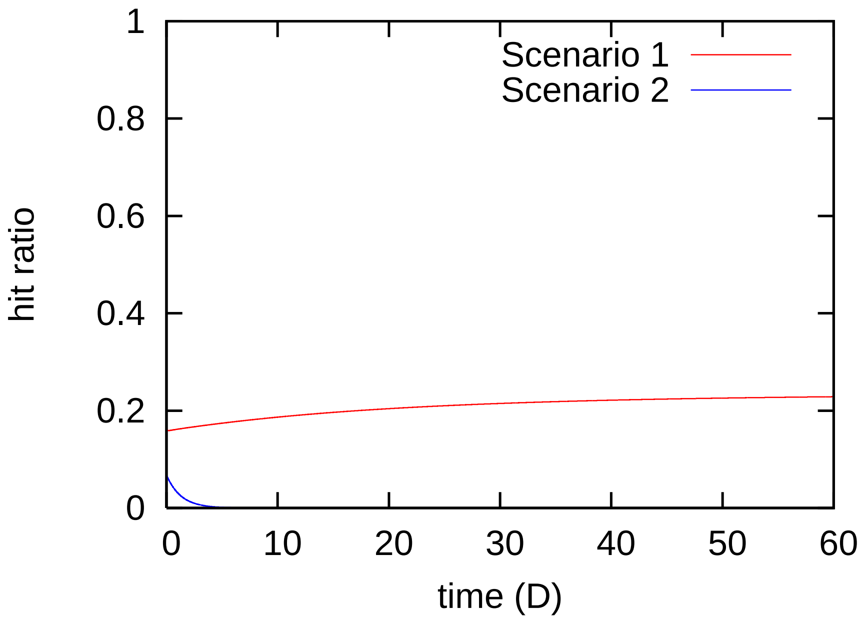


Figure 9: Transition of hit ratio

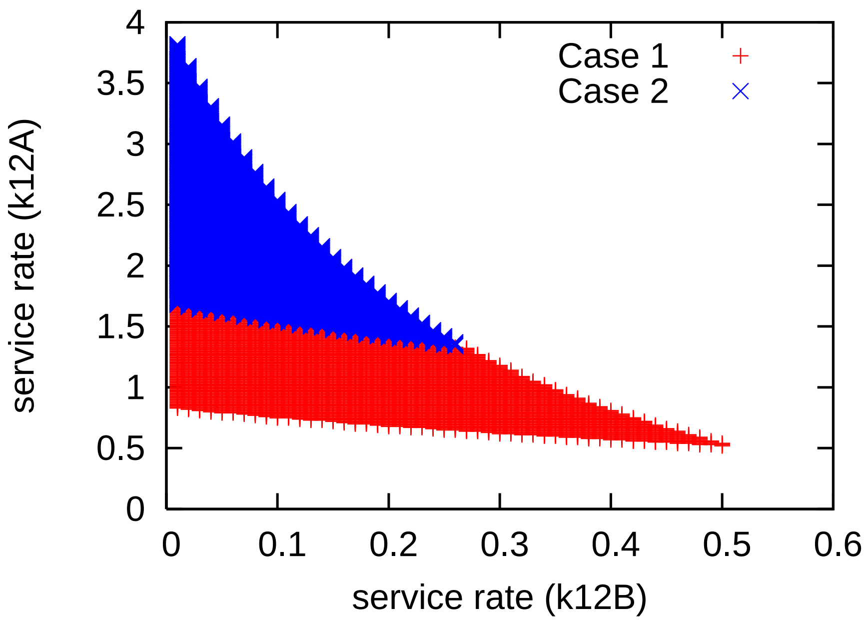
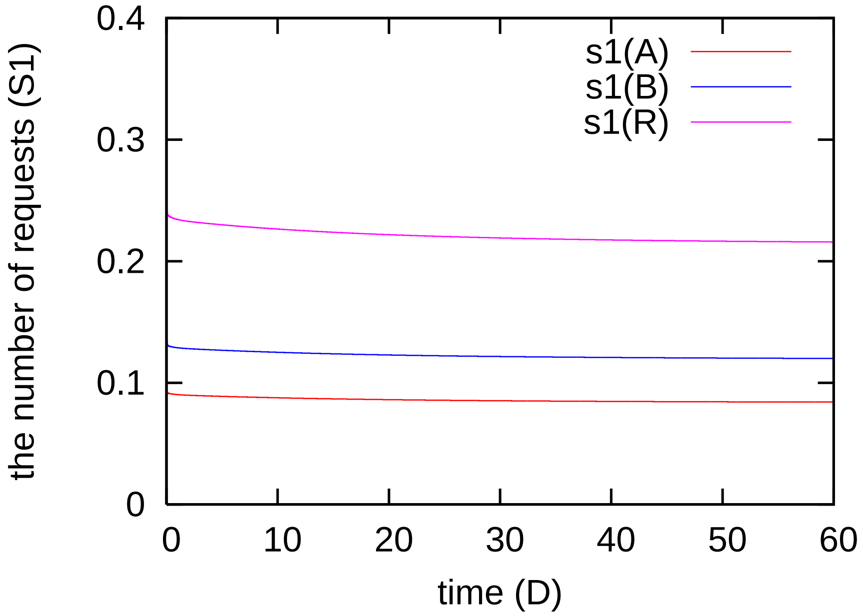


Figure 10: Case of the number of requests and files that the portal server holds to that shared in P2P file-sharing networks

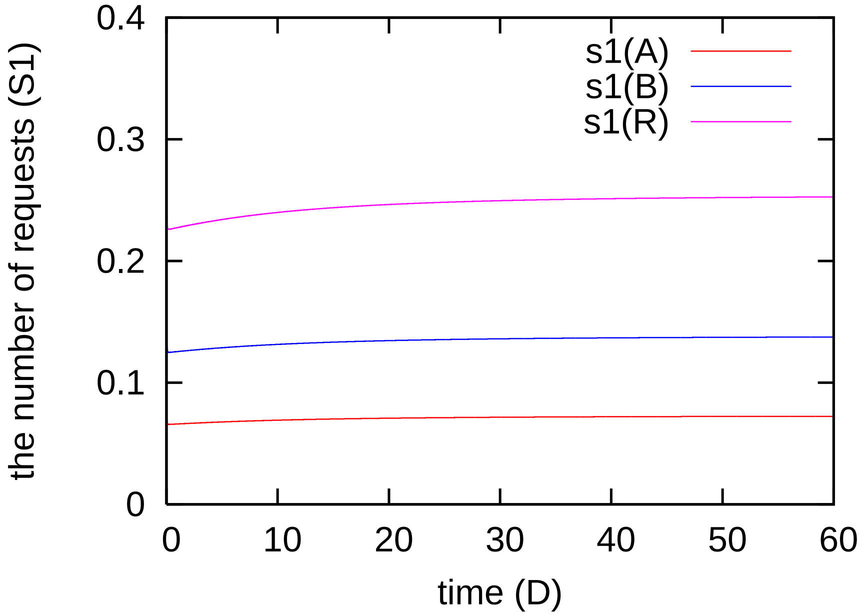
from its buffers. Case 2 indicates the condition where $s_1^{(R)} > s_1^{(A)}$, $s_1^{(R)} > s_1^{(B)}$, $s_2^{(R)} < s_2^{(A)}$, and $s_2^{(R)} > s_2^{(B)}$ hold. That is, although a portal server provides both P2P file-sharing networks with search requests, the portal server fosters effective file sharing by providing more files to P2P file-sharing network B with the small service rate with not only files registered by portal users but also files obtained from the other P2P file-sharing network A with the large service rate. Since the number $s_1^{(0)}$, i.e. 7.5, of new requests registered to a portal server is larger than the number $\frac{1}{2}(s_1^{(U)} - s_1^{(0)})$, i.e. 3.75, of new requests issued to P2P file-sharing network A regardless of the service rate within the analyzed region in Fig. 10, the number $s_1^{(R)}$ of requests buffered at the portal server is larger than the number $s_1^{(A)}$ of requests in P2P file-sharing network A. Therefore, following Eq. (12), the portal server always supplies both P2P file-sharing networks with search requests in the region of Fig. 10 regardless of the service rate. When the service rate $k_{1,2}^{(A)}$ is small, the increase of the number $s_2^{(A)}$ of shared files becomes also small by the second term in Eq. (10). Since the number $s_2^{(0)}$, i.e. 2.5, of new files registered to a portal server is larger than the total of the number $\frac{1}{2}(s_2^{(U)} - s_2^{(0)})$, i.e. 1.25, of new files to P2P file-sharing network A and the number $k_{1,2}^{(A)} s_1^{(A)}$ of files completed in P2P file-sharing network A per unit time, the number $s_2^{(R)}$ of files buffered at the portal server is larger than the number $s_2^{(A)}$ of files shared in P2P file-sharing network A. Because of smaller service rate $k_{1,2}^{(B)}$, the condition is the same for P2P file-sharing network B. Therefore, following Eq. (12), the portal server supplies both P2P file-sharing networks with shared files and the lower region in Fig. 10 belongs to case 1. On the other hand, when the service rate $k_{1,2}^{(A)}$ is large, the increase of the number $s_2^{(A)}$ of shared files becomes large by the second term in Eq. (10). Since the number $s_2^{(0)}$ of new files registered to a portal server is smaller than the total of the number $\frac{1}{2}(s_2^{(U)} - s_2^{(0)})$ of new files and the number $k_{1,2}^{(A)} s_1^{(A)}$ of files completed in P2P file-sharing network A per unit time, the number $s_2^{(R)}$ of files buffered at the portal server is smaller than the number $s_2^{(A)}$ of files shared in P2P file-sharing network A. Therefore, the upper region in Fig. 10 belongs to case 2 where the portal server obtains shared files from P2P file-sharing network A.

To see and compare transitions of the number of search requests and the number of shared files in these two cases, Figs. 11 and 12 are shown for the cases with $k_{1,2}^{(A)} = 1.0$ and $k_{1,2}^{(B)} = 0.1$, and $k_{1,2}^{(A)} = 2.0$ and $k_{1,2}^{(B)} = 0.1$, respectively. The difference in the number $s_1^{(A)}$ of requests in P2P file-sharing network A and the number $s_1^{(B)}$ of requests in P2P file-sharing network B in Fig. 11(a) is smaller than the difference in the number $s_1^{(A)}$ of requests in P2P file-sharing network A and the number $s_1^{(B)}$ of requests in P2P file-sharing network B in Fig. 11(b). Since the service rate of P2P file-sharing network A in Fig. 11(b) is higher than that in Fig. 11(a), the number $s_1^{(A)}$ of requests in P2P file-sharing network A in Fig. 11(b) becomes smaller than that in 11(a) with the same number $\frac{1}{2}(s_1^{(U)} - s_1^{(0)})$ of new requests injected. Similarly, as a result of the higher service rate $k_{1,2}^{(A)}$, P2P file-sharing network A in Fig. 12(b) increases the number $s_2^{(A)}$ of shared files more than that in Fig. 12(a). Consequently, the difference in the numbers $s_2^{(A)}$ and $s_2^{(B)}$ of shared files in Fig. 12(b) is larger than that in Fig. 12(a).

Next, in Fig. 13, we change combinations of the ratio U_R and U_F to see how the ratio of search requests and shared files registered to a portal server to new requests and files influences the symbiotic behavior. Again, with parameter setting in Fig. 13, no P2P file-sharing network can survive in scenario 2, since the number of new requests and files are too small for independent P2P file-sharing networks to be individually activated enough. Depending on the value of service rate $k_{1,2}^{(A)}$, the region where P2P file-sharing networks cooperatively live is different between Fig. 13(a) and Fig. 13(b). Now consider the case in Fig. 13(a). When the ratio U_R of search requests to be registered at the portal server by portal users is large, the number $s_1^{(R)}$ of requests buffered at the portal server becomes large for the first term in Eq. (11). The excess requests at the portal server are offered to P2P file-sharing networks for the first term in Eq. (9). The numbers $s_1^{(A)}$ and $s_1^{(B)}$ of requests in P2P file-sharing networks become large by the first term in Eq. (9). However, the numbers $s_2^{(A)}$ and $s_2^{(B)}$ of shared files do not grow enough for the small service rates $k_{1,2}^{(A)}$ and $k_{1,2}^{(B)}$ to keep them alive unless they are supplied with shared files by direct users or the portal server. However, as shown in Fig. 8, P2P file-sharing networks cannot survive only with shared files supplied by direct users. So that the portal server can offer the enough amount of shared files to P2P file-sharing networks, U_F should be large enough and a larger U_F leads to the wider range of U_R leading to cooperation. On the other hand, when the service rate $k_{1,2}^{(A)}$ is large as in Fig. 13(b), P2P file-sharing network A can keep alive for the sufficient number $s_2^{(A)}$ of shared files regardless of the ratio U_R of requests. However, P2P file-sharing network B cannot survive. When the ratio U_F is small, e.g. 0.01, the number $X^{(A)}$ of participating peers at the stable condition is larger than the number $X^{(B)}$ regardless of the ratio U_R of requests. As modeled in Eq. (12), a portal server obtains more files from P2P file-sharing network A with the large service rate and the large population and provides them to P2P file-sharing network B which lacks shared files to live. As a result, both P2P file-sharing networks can keep alive regardless of the ratio U_R with small U_F. On

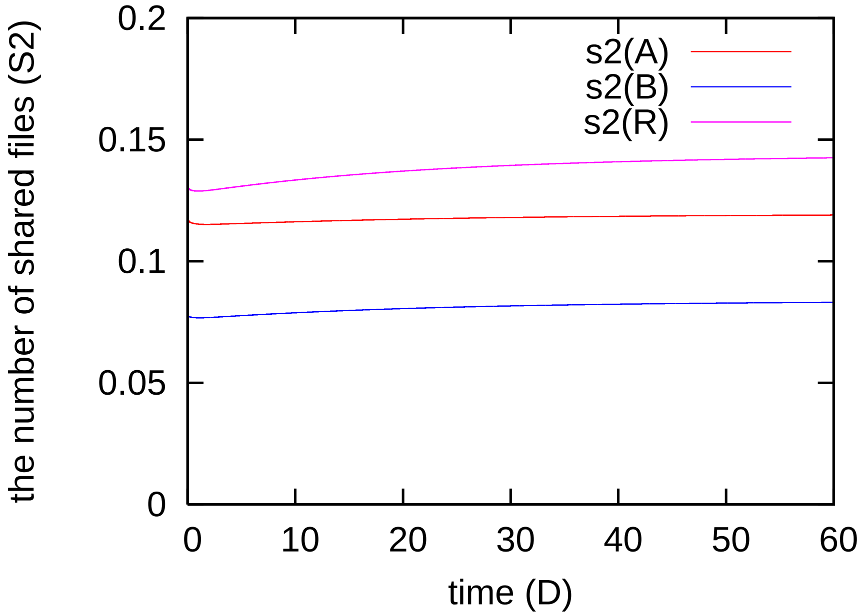


(a) $k_{1,2}^{(A)} = 1.0, k_{1,2}^{(B)} = 0.1$

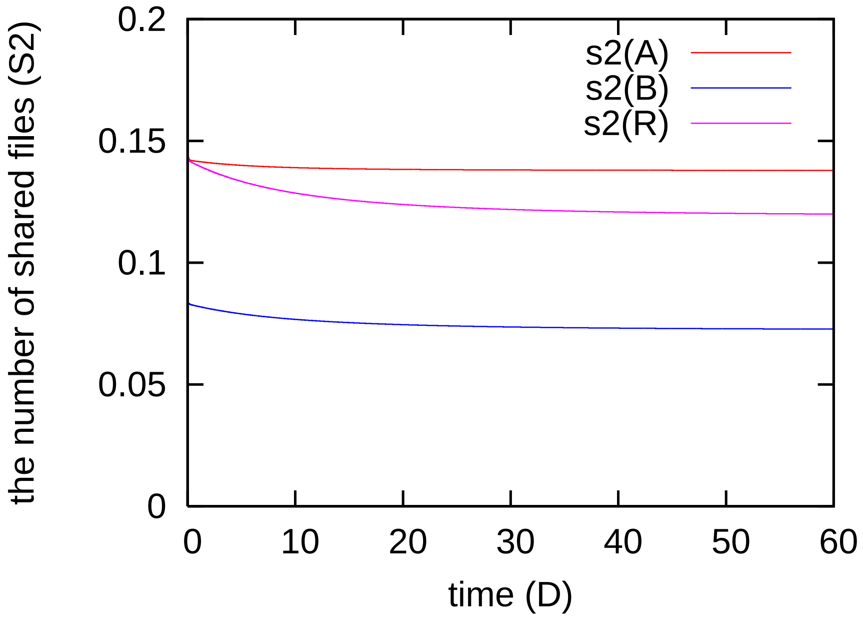


(b) $k_{1,2}^{(A)} = 2.0, k_{1,2}^{(B)} = 0.1$

Figure 11: Transition of the number of requests



(a) $k_{1,2}^{(A)} = 1.0, k_{1,2}^{(B)} = 0.1$

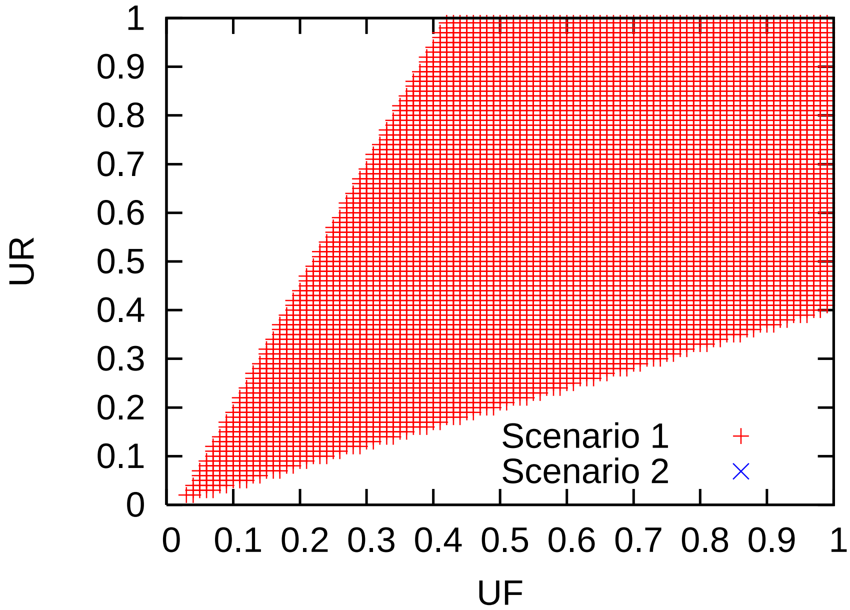


(b) $k_{1,2}^{(A)} = 2.0, k_{1,2}^{(B)} = 0.1$

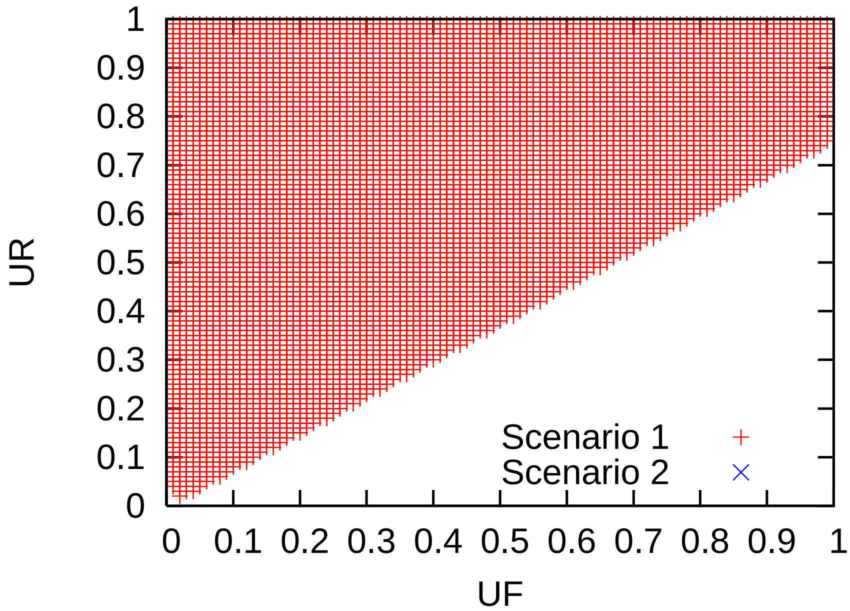
Figure 12: Transition of the number of shared files

the other hand, when the ratio U_F becomes large, e.g. 0.5, the number $X^{(A)}$ is smaller than the number $X^{(B)}$. Although the number $s_2^{(A)}$ of shared files in P2P file-sharing network A is large, a portal server cannot obtain shared files from P2P file-sharing network A enough to make P2P file-sharing network B survive for the small population $X^{(A)}$. It means that, to allow both P2P file-sharing networks to live, a portal server needs portal users to supply more files. This is a reason that the region with points exists only in the upper part with large U_F in Fig. 13(b).

To analyze behavior of the portal server, we compare the number of requests and files that the portal server holds to that shared in P2P file-sharing networks in Fig. 14. As shown in the figure, all points with small service rate $k_{1,2}^{(A)}$ belong to case 1 and all points with large service rate $k_{1,2}^{(A)}$ belong to case 2. That is, the behavior of the portal server is not influenced by the ratio of new search requests and shared files registered to the portal server but by service rate of P2P file-sharing networks.

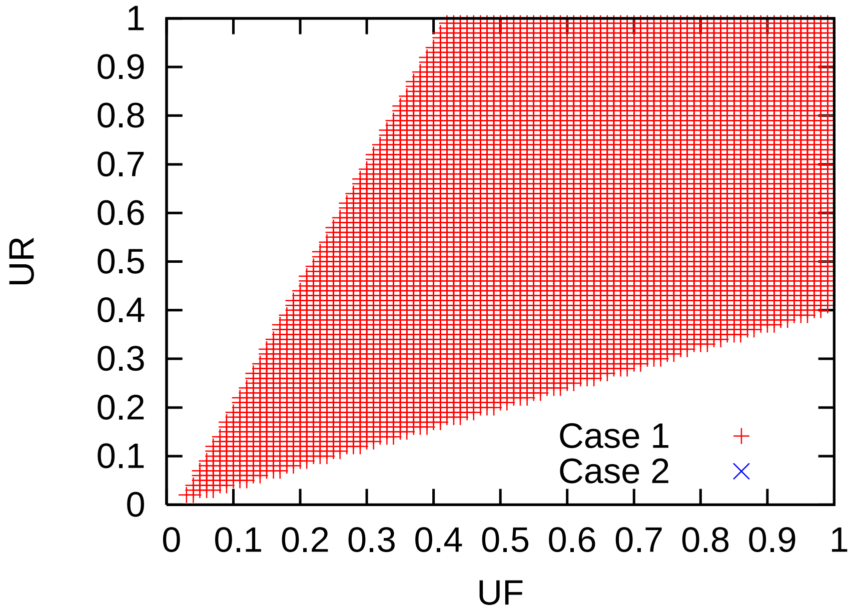


(a) $k_{1,2}^{(A)} = 1.0, k_{1,2}^{(B)} = 0.1$

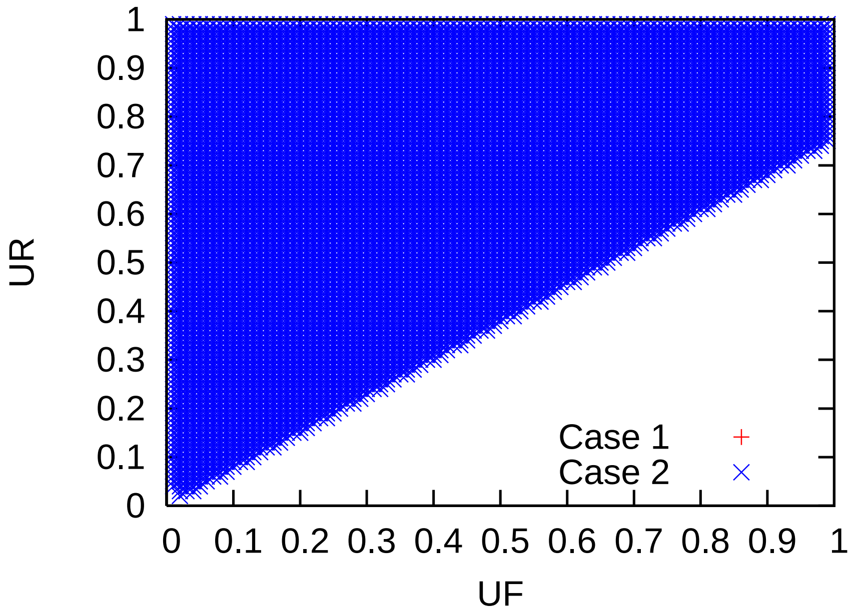


(b) $k_{1,2}^{(A)} = 2.0, k_{1,2}^{(B)} = 0.1$

Figure 13: Scenarios leading to higher hit ratio



(a) $k_{1,2}^{(A)} = 1.0, k_{1,2}^{(B)} = 0.1$



(b) $k_{1,2}^{(A)} = 2.0, k_{1,2}^{(B)} = 0.1$

Figure 14: Case of the number of requests and files that the portal server holds to that shared in P2P file-sharing networks

5 Conclusion

In this thesis, to demonstrate an example of biologically-inspired symbiotic overlay networks, we proposed a mechanism that enabled different P2P file-sharing networks to cooperate and live together with mediation of a portal server. We modeled the proposed mechanism based on the biological symbiosis model. Through numerical analysis based on the mathematical model, it was shown that the proposed mechanism improved the hit ratio of P2P file-sharing networks in comparison to the scenario where P2P file-sharing networks were independent. Especially, in the case that number of search requests and shared files were too small for P2P file-sharing networks to endure, a portal server fostered effective file sharing by providing more files to a P2P file-sharing network with the small service rate.

As future research work, we need to perform realistic simulation experiments taking into account network topology and other physical influence to investigate detailed behavior of P2P file-sharing networks mediated by a portal server.

Acknowledgments

I would like to express my deepest gratitude to my supervisor, Professor Masayuki Murata of Osaka University, for his expensive help and continuous support through my studies of this thesis.

I would like to express my appreciation to Associate Professor Naoki Wakamiya of Osaka University who has always given me appropriate guidance and valuable advice. All of my works would not have been possible without his supports.

I am very grateful to Professors Koso Murakami, Makoto Imase, Teruo Higashino, and Hirotaka Nakano of Osaka University, for their appropriate guidance. I am also indebted to Associate Professor Go Hasegawa, Specially Appointed Associate Professor Kenji Leibnitz, Assistant Professors Shin'ichi Arakawa, Masahiro Sasabe, and Yuichi Ohshita of Osaka University, who gave me helpful comments and feedback.

Finally, I want to give thanks to my friends and colleagues in the Department of Information Networking, Graduate School of Information Science and Technology of Osaka University.

References

[1] M. Seshadri and R. H. Katz, "Dynamics of simultaneous overlay network routing," Technical Report UCB//CSD-03-1291, Electrical Engineering and Computer Science (EECS), University of California Berkeley (UCB), Nov. 2003.

[2] L. Qiu, Y. R. Yang, Y. Zhang, and S. Shenker, "On selfish routing in Internet-like environments," in *Proceedings of ACM SIGCOMM Conference 2003*, pp. 151–162, Aug. 2003.

[3] R. Keralapura, N. Taft, C.-N. Chuah, and G. lannaccone, "Can ISPs take the heat from overlay networks?," in *Proceedings of ACM HotNets-III*, Nov. 2004.

[4] M. Ammar and S. Seetharaman, "Routing in multiple layers: Challenges and opportunities," in *Proceedings of Workshop on Internet Routing Evolution and Design (WIRED)*, Oct. 2006.

[5] M. Kwon and S. Fahmy, "Toward cooperative inter-overlay networking," in *Proceedings of the 11th IEEE International Conference on Network Protocols (ICNP)*, poster paper, Nov. 2003.

[6] M. Kwon and S. Fahmy, "Synergy: An overlay internetworking architecture and its implementation," Technical Report, Purdue University, Oct. 2005.

[7] I. Stoica, D. Adkins, S. Zhuang, S. Shenker, and S. Surana, "Internet indirection infrastructure," in *Proceedings of ACM SIGCOMM Conference 2002*, vol. 32, pp. 73–88, Aug. 2002.

[8] D. Andersen, H. Balakrishnan, F. Kaashoek, and R. Morris, "Resilient overlay networks," in *Proceedings of 18th ACM Symposium on Operating Systems Principles*, vol. 35, pp. 131–145, Oct. 2001.

[9] A. Nakao, L. Peterson, and A. Bavier, "A routing underlay for overlay networks," in *Proceedings of ACM SIGCOMM Conference 2003*, pp. 11–18, Aug. 2003.

[10] H. Xie, A. Krishnamurthy, A. Silberschatz, and Y. R. Yang, "P4P: Explicit communications for cooperative control between P2P and network providers." available online at `http://www.dcia.info/documents/P4P_Overview.pdf`, 2007.

[11] H. Xie, Y. R. Yang, A. Krishnamurthy, Y. Liu, and A. Silberschatz, "P4P: Provider portal for applications," in *Proceedings of ACM SIGCOMM Conference 2008*, pp. 351–362, Aug. 2008.

[12] N. Wakamiya and M. Murata, "Toward overlay network symbiosis," in *Proceedings of the 5th International Conference on Peer-to-Peer Computing (P2P2005)*, pp. 154–155, Aug. 2005.

[13] 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, pp. 239–247, Dec. 1996.

[14] N. Wakamiya and M. Murata, "Bio-inspired analysis of symbiotic networks," in *Proceedings of 20th International Teletraffic Congress (ITC-20)*, June 2007.

[15] J. Konishi, N. Wakamiya, and M. Murata, "Design and evaluation of a cooperative mechanism for pure P2P file-sharing networks," *IEICE Transactions on Communications*, vol. E89-B, pp. 2319–2326, Sept. 2006.

[16] H. Fu, N. Wakamiya, and M. Murata, "A cooperative mechanism for hybrid P2P file-sharing networks to enhance application-level QoS," *IEICE Transactions on Communications*, vol. E89-B, pp. 2327–2335, Sept. 2006.

[17] W. Acosta and S. Chandra, "Trace driven analysis of the long term evolution of gnutella peer-to-peer traffic," in *Proceedings of Passive and Active Network Measurement*, pp. 42–51, June 2007.

[18] M. Feldman, C. Papadimitriou, J. Chuang and I. Stoica, "Free-riding and whitewashing in peer-to-peer systems," *IEEE Journal on Selected Areas in Communications*, pp. 1010–1019, May 2006.