Master's Thesis

Title

Design principles for ISP networks with consideration to overlay routing behaviors

Supervisor

Professor Masayuki Murata

Author

Ryota Fukumoto

February 14th, 2007

Department of Information Networking

Graduate School of Information Science and Technology

Osaka University

Master's Thesis

Design principles for ISP networks with consideration to overlay routing behaviors

Ryota Fukumoto

Abstract

In recent years, network applications and services, such as P2P (Peer-to-Peer) file sharing, CDNs (Content Delivery Networks), and IP-VPNs (IP Virtual Private Networks), have been deployed over the Internet. These applications and services configure their own overlay networks. Each overlay network adopts its own routing algorithm to optimize users' criteria. On the other hand, the underlay network employs the IP routing. However, the IP routing does not necessarily take into account metrics for the overlay routing. That is why the overlay routing has been considered.

End users of overlay networks greedily choose better routes to optimize their own performance without considering the performance with regards to the network-wide criteria. Such selfish behaviors of end users cause unexpected increase of traffic volumes and traffic distributions on the underlay network. Since the underlay network cannot directly control the behaviors of the end users of overlay networks, it is important for network operators of underlay networks to appropriately design the networks by taking into account the selfish behavior of the overlay routing.

In this thesis, we first evaluate the impact of the overlay routing's selfish behavior on realistic router-level topologies from the perspective of maximum link load and average latency in the case that the fixed shortest path routing is adopted for the underlay networks. The results of our evaluation show that the overlay routing achieves near optimal latency but concentrates traffic on certain links, when the overlay routing is at equilibrium. In addition, our results show that the maximum link load and the average latency in realistic router-level topologies which have the power-law attribute are more affected by the overlay routing than those in POP level topologies.

We next consider the case that the routes in underlay networks are dynamically updated against the change of link load. In such cases, the routing in overlay networks and that in the underlay networks interfere with each other. We then evaluate the impact of the interaction between the overlay routing and the routing in the underlay networks in realistic router-level topologies. The results indicate that the interaction makes the maximum link load be over 1.0. We then investigate the impact of the interaction in cases that the number of nodes dynamically updating the routes in underlay networks is limited. We examine where we should allocate the nodes that dynamically update the routes. The results shows that the maximum link load is decreased to 0.65 from 1.0 when only the nodes classified as “provincial hubs” dynamically update the routes in the underlay networks against the change of link load.

Based on this observation, we finally investigate the network design of ISP networks with consideration to overlay routing behaviors. We then introduce overlay-optimal capacity design and compare the design to two conventional capacity designs in order to clear up which links we should enhance in ISP topologies. The results show that we need to enhance the capacities of some links more than that of the conventional design if we consider the overlay routing behaviors. It also indicates that there exist nodes that have important roles for reducing the impact of the overlay routing behavior.

Keywords

Overlay routing

ISP networks

Power-law

Network design

Router-level topologies

Contents

List of Figures

List of Tables

1 Introduction

In recent years, network applications and services, such as P2P (Peer-to-Peer) file sharing, CDNs (Content Delivery Networks), and IP-VPNs (IP Virtual Private Networks), have been deployed over the Internet. These applications and services configure their own overlay networks; nodes in the overlay networks are connected with logical links through one or more links in the underlay networks, which we assume IP networks. Each overlay network adopts its own routing algorithm to optimize users' criteria, such as latency and the other QoS (Quality of Service) metrics, which we call overlay routing. On the other hand, the underlay network employs the IP routing. That aims to minimize a hop distance of routes or minimizes the maximum link load for traffic engineering [1]. The IP routing does not necessarily take into account metrics for the overlay routing.

Overlay networks allow end users to control routes to other nodes in order to satisfy their QoS requirements. Therefore, end users of overlay networks greedily choose better routes to optimize their own performance without considering the performance with regards to the network-wide criteria. Such selfish behaviors of end users cause unexpected increase of traffic volumes and traffic distributions on the underlay network. Since the underlay network cannot directly control the behaviors of the end users of overlay networks, it is important for network operators of underlay networks to appropriately design the networks by taking into account the selfish behavior of the overlay routing.

Roughgarden et al. take a theoretical approach to evaluate the performance of a selfish routing, in which users select their routes selfishly [2]. They point out that the average latency is degraded seriously when each of end users ignores latencies of the other users and selfishly chooses more optimal routes for themselves. However, Qiu et al. show that the overlay routing may achieve near-optimal latency at equilibrium (the difference between the overlay routing and a latency optimal routing is usually close to 0 and is always within 30%), while the links in the underlay networks are significantly congested [3]. In Ref. [3], Qiu et al. use POP (Point of Presence) level networks as Internet-like environments. That is, they only show a possible application of overlay routing to networks which have at most 100 nodes. However, since the users of overlay networks are distributed all over the Internet, overlay networks are generally widespread over the large area of the Internet. More importantly, it is known that the degree distribution of nodes in the Internet exhibits the power-law attribute [4]. However, the POP level topologies discussed in [3] have too small number of nodes to reflect the power-law attribute. Therefore, it is important to show the performance of overlay routing on more Internet-like network topologies in terms of the number of nodes and the power-law attribute.

In this thesis, we first evaluate the impact of the overlay routing's selfish behavior on realistic router-level topologies from the perspective of maximum link load and average latency in the case that the fixed shortest path routing is adopted for the underlay networks. The results of our evaluation show that the overlay routing achieves near optimal latency but concentrates traffic on certain links, when the overlay routing is at equilibrium. In addition, our results show that the maximum link load and the average latency in realistic router-level topologies which have the power-law attribute are more affected by the overlay routing than those in POP level topologies. That is, much more traffic concentrates on certain links in the realistic router-level topology by the overlay routing. We also demonstrate that the overlay routing achieve higher maximum link load and latency, when the overlay routing is at non-equilibrium due to the oscillation of routes in overlay networks.

We next consider the case that the routes in underlay networks are dynamically updated against the change of link load. In Ref. [1], an IP routing that dynamically updates the routes against the change of link load is studied. In such cases, the routing in overlay networks and that in the underlay networks interfere with each other. In Ref. [5], Liu et al. evaluate the impact of an interaction between the overlay routing and the traffic engineering in the underlay networks. The authors reveal the possibility of the performance degradation of both of the overlay and underlay networks due to the interaction. However, they evaluate the influence of the interaction with only a small size topology that has 14 nodes. We again evaluate the impact of the interaction between the overlay routing and the routing in the underlay networks in realistic router-level topologies. The results indicate that the interaction makes the maximum link load be over 1.0. That is, the interaction between the overlay and underlay networks degrades the network performance. In order to relax the degradation of the network performance due to the interaction, we investigate the impact of the interaction in cases that the number of nodes dynamically updating the routes in underlay networks is limited. Due to this limitation we can expect that the magnitude of the interaction gets smaller and the convergence time of the overlay and underlay routing gets shorter. However, it is more important to show where we should allocate such nodes. We then examine where we should allocate the nodes that dynamically update the routes. The results shows that the maximum link load is decreased to 0.65 when only the nodes classified as “provincial hubs” dynamically update the routes in the underlay networks against the change of link load, which indicates that the overlay routing and the underlay routing can coexist.

Overlay routing has the huge influence on the network design because such routing is uncontrollable for network operators. If we understand overlay routing’s behaviors, we can use the insights for network design. Hence, based on this observation, we finally investigate the network design of ISP networks with consideration to overlay routing behaviors. We then introduce overlay-optimal capacity design and compare the design to two conventional capacity designs in order to clear up which links we should enhance in ISP topologies. We show that not only the nodes classified as “connector nodes” but also the nodes classified as “peripheral nodes” have important roles for reducing the impact of the overlay routing behaviors.

This thesis is organized as follows. Section 2 shows the network model that we use for our evaluation. In Section 3, we evaluate the performance of the overlay routing on ISP topologies. Next in Section 4, we evaluate the influence of the interaction of the overlay routing and the underlay routing in ISP networks. In Section 5, we evaluate capacity design on ISP topologies. Finally, Section 6 concludes our thesis and mentions the future work.

2 Network model

In this section, we show the network model that we use for our evaluations. We first show the underlay networks that we assume, IP networks. We next show the overlay networks that we assume.

2.1 Underlay networks

In the following sections, we show about underlay networks that we asumme, IP networks. We first show about the physical topologies. We next show about routings in the underlay networks, IP routing.

2.1.1 Physical topology

Recent measurement studies on Internet topology show that the connectivity of nodes exhibit a power-law attribute (e.g., see [4], [6]). That is, the probability $p(k)$, that a node is connected to k other nodes, follows $p(k) \sim k^{-\gamma}$. In recent years, considerable numbers of studies have investigated power-law networks whose degree distributions follow the power-law [7–11]. Here, the degree is defined as the number of out–going links at a node. The theoretical foundation for the power–law network is introduced in Ref. [12] where they also presents the BA (Barabashi-Albert) model in which nodes increase incrementally and links are placed based on the connectivity of topologies in order to form power–law networks. The resulting power-law networks have two main characteristics: (1) a small number of links are connected with numerous nodes, while a large number of links are connected with a few nodes, and (2) the number of hop-counts between nodes is small (*small–world* property).

However, even if the degree distributions of some topologies are the same, more detailed characteristics are often quite different. A work by Li et al. [13] has enumerated various topologies with the same degree distributions, and has shown the relation between the characteristics and performances of these topologies, throughput and likelihood metrics. With the technology constraints imposed by routers, the degree of nodes limits the capacity of links that are connected to. Li et al. therefore point out that higher–degree nodes tend to be located at the edges of a network, and lower-degree nodes tend tot be located at the core of a network. That is, they point out that the structure of a topology generated by BA model is different from that of ISP topologies. As for the

Table 1: The number of nodes and links in each topology

	Sprint	AT&T	BA	ER
Nodes	467	523	467	467
Links	1280	1304	1280	1280

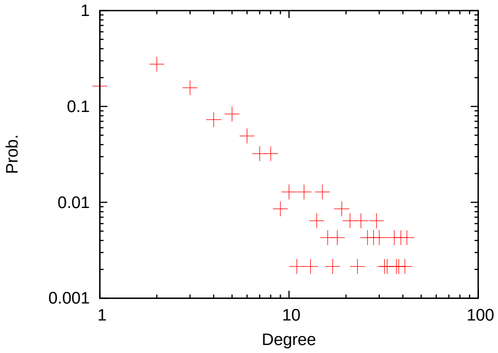
throughput and link load properties when applying routing, it is shown that a topology generated by BA model have different property from ISP topologies in [14].

In order to evaluate overlay routing behaviors in ISP topologies, we therefore use three ISP router-level topologies. The topologies are Sprint topology and AT&T topology measured in Ref. [15]. Sprint and AT&T are major ISP in United States. We call these topologies the Sprint topology, the AT&T topology, and the Verio topology respectively. The Sprint topology has 467 nodes and 1280 links. The AT&T topology has 523 nodes and 1304 links. The Figs. 1(a) and 1(b) represent the degree distribution of each topology. The figures show that these topologies certainly have power-law like distribution. That is, the distributions can be approximated by straight lines.

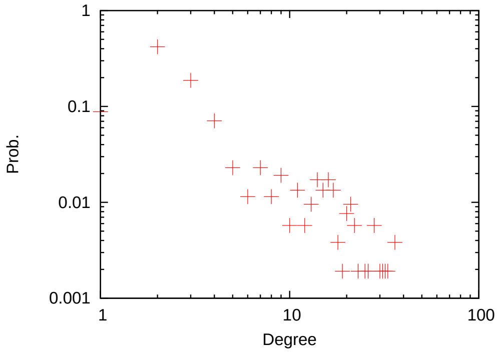
We also use BA topology generated by the BA (Barabashi-Albert) model [12], and ER topology generated by ER (Erdos-Renyi) model [16] for comparison. In ER model, links are randomly placed between nodes. We call these topologies the BA topology, and the ER topology respectively. The Figs. 1(c) and 1(d) show the degree distribution of these two topologies. In the figure, we can see that the BA topology has power-law degree distribution and the ER topology has Poisson degree distribution. These topologies are generated such that the numbers of nodes and links are the same as that for the Sprint topology. These all topologies consist of nodes and links shown in Table 1.

2.1.2 IP routing

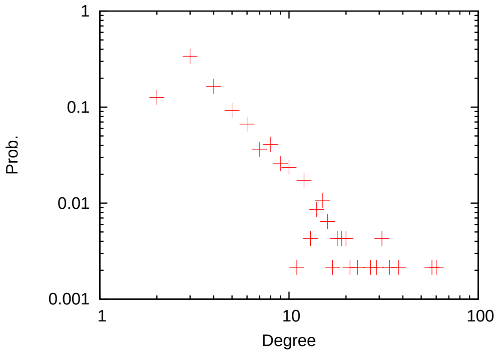
In underlay networks, traffic is transferred with IP routing. In our thesis, we consider a static routing and a dynamic routing. As static routing, we assume that traffic is transferred to minimize hop-counts between a source node and a destination node. We call this routing the minimum hop routing. As for the dynamic routing, we assume the routing that dynamically controls routes of the IP packets to detour congested links. We call this routing the dynamic IP routing. We call the nodes which dynamically controls routes “source routing nodes”. Each dynamic routing node



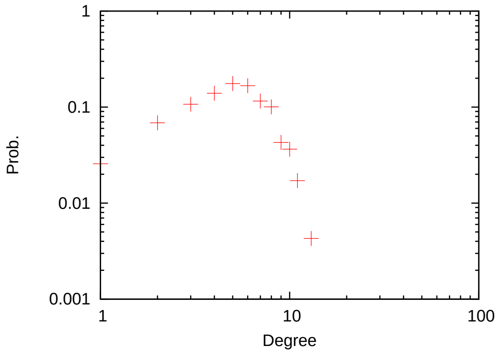
(a) The Sprint topology



(b) The AT&T topology



(c) The BA topology



(d) The ER topology

Figure 1: Degree distribution of each topologies

determines routes using Dijkstra's shortest path algorithm. In detail, the nodes assign the cost of link l, W_l as $W_l = 1/(C_l - f_l)$. Here, C_l is the capacity of link l, f_l is the amount of traffic which pass through link l. As f_l closes to C_l, W_l increases. Each dynamic routing node then can avoid congested links. Note that the nodes explicitly specify the routes of traffic generated on them (source routing). If a source routing node determines a route of the traffic, the node specifies all nodes that the traffic passes through.

2.2 Overlay networks

In the following sections, we show overlay network model that we assume. We first show about logical topologies. We next show about routings in the overlay networks, overlay routing.

2.2.1 Logical topology

Overlay network constructs logical topologies. Logical topologies consist of overlay nodes and logical links which connect between overlay nodes. Logical links consist of one or more physical links. Physical links which compose a logical link is depend on network level routing, IP routing. For instance, when we assume that the network level routing uses minimum hop routing, the logical link (a', d') consists of the physical link (a, e) and (e, d) as shown in Figure 2. If the network level routing is dynamic, the physical links which compose the logical link (a', d') changes as often as the routing changes routes.

In our evaluations, we assume that overlay nodes are fully connected as shown in Figure 2. It may be unrealistic that overlay nodes are fully connected. However, we can expect to evaluate the result got when overlay networks are constructed optimally.

2.2.2 Overlay routing

We assume that two types of traffic occur in networks. One is non-overlay traffic. Non-overlay traffic is stationary and transferred by IP routing. The other is overlay traffic. Overlay traffic is transferred by overlay nodes. Each overlay node transmits overlay traffic generated on each overlay node to another overlay node such that each overlay node maximizes its own metrics. In real networks, each overlay node measures the quality of logical links or paths. Since overlay nodes generally cannot know the states of physical network (e.g., physical links that compose a

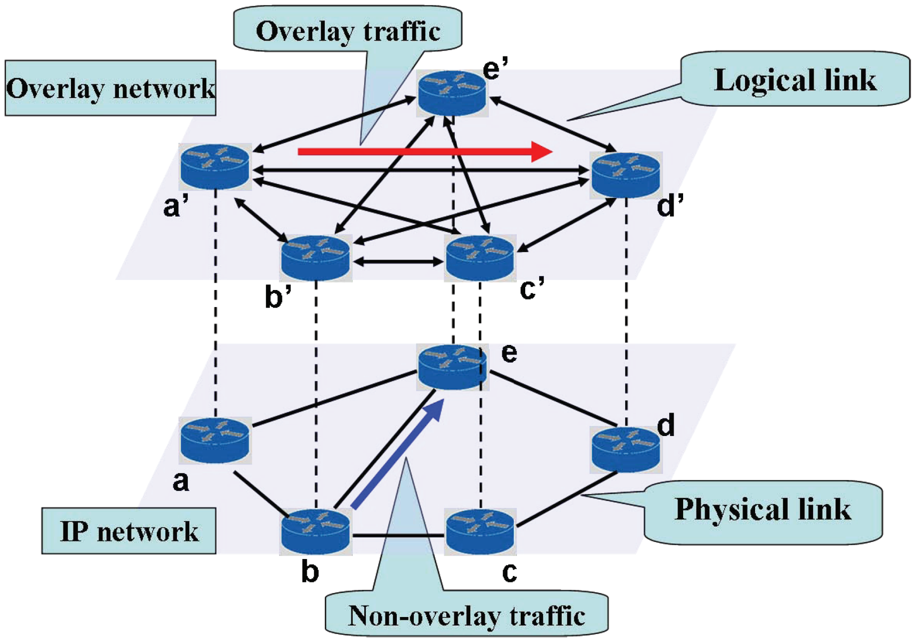


Figure 2: Network model

logical links), overlay nodes determine routes of overlay traffic by using only qualities of logical links or paths.

In our evaluations, we assume that each overlay node explicitly specifies logical path of traffic generated at the node (source routing). That is, if an overlay routing node determines a route of the overlay traffic, the overlay node specifies all overlay nodes that the traffic passes through.

In our evaluations, we consider two kinds of overlay routings. One is the overlay routing which aims to minimize the end-to-end delay. This type of overlay routing is assumed in [3]. The other is the overlay routing which is managed by the network operator in order to avoid congestions and to improve resilience (e.g., RON [17]). In our thesis, we call these overlay routings "User-oriented overlay routing" and "Network-oriented overlay routing". That is, we consider not only an overlay routing which aims to improve the user-wide criteria but also an overlay routing which aims to improve the network-wide criteria. Note that not only nodes that use the user-oriented overlay routing, but also nodes that use the network-oriented overlay routing does not consider with other overlay nodes. We then compare and evaluate these overlay routing.

User-oriented overlay routing The user-oriented overlay routing assumes that end users control routes considering their own metrics. With the user-oriented overlay routing, each overlay node (end user) transmits traffic generated on itself. The overlay nodes specify all overlay nodes that exist on the way to the destination node in order to optimize their own metrics. The metric is the average end-to-end delay of the traffic. They change their routes selfishly and dynamically. Therefore overlay nodes keep changing their routes until they can find more effective routes. Here, we define equilibrium as a state where no overlay node can improve their metrics by changing the amount of traffic which it sends along different logical paths. That is, overlay nodes keep changing their routes until the network state becomes equilibrium.

Network-oriented overlay routing We introduce the network-oriented overlay routing. This overlay routing assumes that routers control routes considering their own metrics. But the metrics are network costs. We define the network cost as the following function.

$$\sum_l 1/(C_l - f_l)$$

Here, C_l is a capacity of link l, f_l is the amount of flow on link l. The overlay nodes change their routes without cooperating with the other overlay nodes. Therefore the overlay nodes keep changing their routes while they can find more effective routes.

Network-oriented optimal overlay routing As well as above-mentioned two overlays routing, we introduce the network-oriented optimal overlay routing for comparison. With the network-oriented overlay routing, each overlay node selfishly controls routes of traffic considering the network cost. However, with the network-oriented optimal overlay routing, overlay nodes cooperate with each other. The network-oriented optimal overlay routing therefore can optimize the network cost. However we consider this overlay routing for comparison with the network-oriented overlay routing.

3 Evaluation of overlay routings in power-law networks

We apply the above-mentioned overlay routings to power-law networks. As mentioned above, overlay routing changes routes dynamically. So we evaluate the overlay routings by computing their performance at equilibrium similarly in Ref. [3]. We use the flow deviation method (Frank-Wolfe algorithm) [18] to compute their performance at equilibrium. We then evaluate maximum link load and average latencies at equilibrium. We also evaluate overlay routings when overlay nodes change routes periodically, overlay nodes change their routes at the same time, in order to compare with the overlay routing at equilibrium.

3.1 Simulation model

We use the Sprint topology and the AT&T topology as ISP topologies. We also use BA topology and ER topology for comparison. In our evaluation, we assume that each node-pair generates the same amount of traffic at a unit time and all nodes are overlay nodes. We also assume that all links have 10 Gbps capacity and that the propagation delay of all links is 1 msec. Because we focus on the link load property, we set the processing delay on nodes 0 msec. Therefore the latency is the total of propagation delays and transmission delays on links. As for network level routing, we consider the minimum hop routing. Logical links of overlay networks are decided by minimum hop routing.

Computing equilibrium with the flow deviation method

As mentioned above, we use the flow deviation method to compute the equilibrium of overlay routings. The flow deviation method incrementally changes the flow assignment along feasible and descent directions. Given objective function T, the method set w_e as a partial derivative with respect to F_e, where F_e is the amount of traffic that traverses link l. Then, the new flow assignment is solved by using the shortest path algorithm in terms of w_e. By incrementally changing from the old to the new flow assignment, optimal flow assignment is determined. We compute the equilibrium of overlay routings with this method. In this thesis, we set objective function T to

$$T_u = \sum_e \int l_e(x_e)dx \quad (User - oriented)$$

$$T_n = \sum_e \int (1/(C_e - x_e))dx \quad (Network - oriented)$$

$$T_o = \sum_e 1/(C_e - x_e) \quad (Network - orientedoptimal)$$

where C_e is the capacity of link e, x_e is amount of traffic on link e, and l_e is latency on link e.

3.2 Evaluation metrics

In this section, we evaluate the following two metrics. One is maximum link load. In this thesis, we define the link load of link i as F_i/C_i. We then define the maximum link load as the maximal value of link loads in the network. We also use the average latency in the network as evaluation metrics.

3.3 Evaluation and discussions

In following sections, we evaluate the overlay routings on ISP topologies using above mentioned metrics. We first focus on the maximum link load on the topologies. We next investigate the average latency of overlay traffic. We then evaluate link loads in more detail.

3.3.1 Maximum link load

We show the maximum link load for representative topologies, as we vary the overlay traffic factor in Figure 3. The vertical axis represents maximum link load in the topologies. The horizontal axis represents overlay traffic factor. Overlay traffic factor represents the amount of traffic generated on the overlay network. When the overlay traffic factor is 1.0, the amount of overlay traffic equals to the amount of traffic generated on the physical network, non overlay traffic. In our evaluations, we set the amount of non overlay traffic to the amount which maximum link load achieves 0.5 with the network level routing.

We can make the following observations. At first, 3(c), we can confirm that the network-oriented overlay routing achieves much lower maximum link load than the user-oriented overlay routing in Figs. 3(a), 3(b). We can also confirm that the network-oriented overlay routing can achieve maximum link load equivalent to that of the network-oriented optimal overlay routing. The similar observations are shown on the Sprint POP level topology (Ref. 3(e)).

On the other hand, we can confirm that the overlay routings at equilibrium achieve almost the same maximum link loads in the ER topology (Ref. 3(d)). This is due to the difference between the structure of the ER topology and that of the other topologies. As indicated in [14], the ER topology decentralizes link loads with the minimum hop routing while the other topologies concentrate link load to some links. That is, almost all links have a certain level of load. Therefore the ER topology hardly has links for the overlay routings to use for diverting traffic.

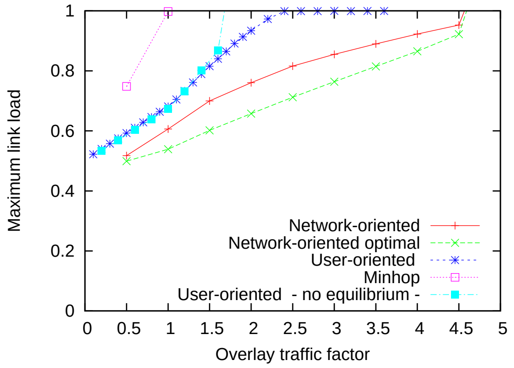
3.3.2 Average latency

We next show the average latency for representative topologies, as we vary the overlay traffic factor in Figure 4. The vertical axis represents the average latency of all node pair. The horizontal axis represents overlay traffic factor. We can confirm that the average latencies are about the same with any overlay routings. We therefore can say that the network-oriented overlay routing, which aims to decentralize the link loads, achieves much lower maximum link load than the user-oriented overlay routing, which aims to lower the average latency, while keeping the average latency approximately equivalent to that of the user-oriented overlay routing.

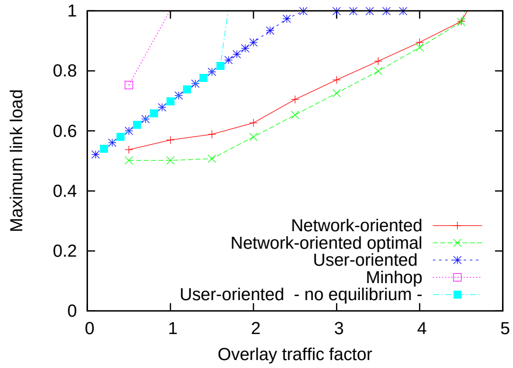
These results are got by computing the equilibrium. However, these overlay routings don't always converge to the equilibrium in the Internet. So we compare the performance on the equilibrium with the performance got when the end users periodically change the routes. The lines captioned "user-oriented - non equilibrium – " in Figs. 3 and 4 describe the results when end users periodically change their routes considering their own latency. In Fig. 3, we can confirm that the maximum link load of "user-oriented – non equilibrium – " overlay routing achieve 1.0 when the amount of overlay traffic is smaller than that of the user-oriented overlay routing. We can suppose that these overlay routings in real networks, work worse than that in these result, ex. due to oscillations.

3.3.3 Maximum link load versus average link load

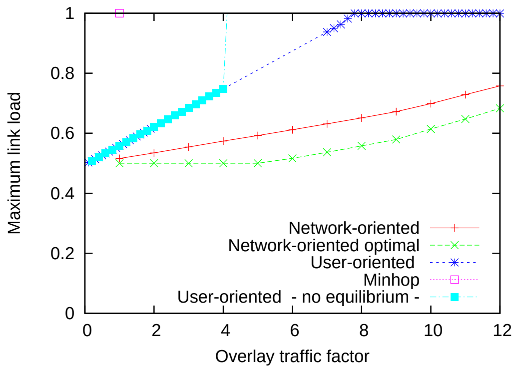
In Fig. 5, we show the maximum link load for representative topologies, as the average link load increases by varying the overlay traffic factor in Figure 4. The vertical axis represents average link load in the topology. The horizontal axis represents the maximum link load. The overlay routing is the user-oriented overlay routing. This figure show that ISP topologies, the Sprint, the AT&T, and the Sprint (POP level) topologies, increase the maximum link loads more rapidly than the BA and the ER topologies as the average link loads increase. That is, ISP topologies increase the maximum link load as the amount of overlay traffic increases. When we focus on the Sprint



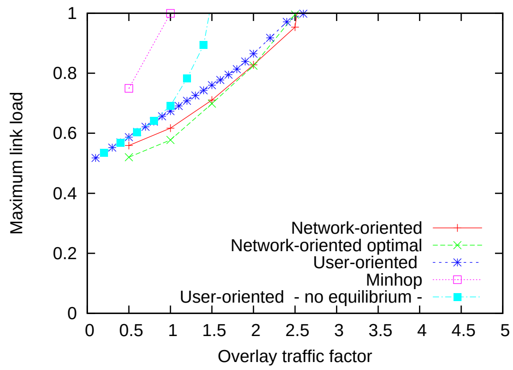
(a) The Sprint topology



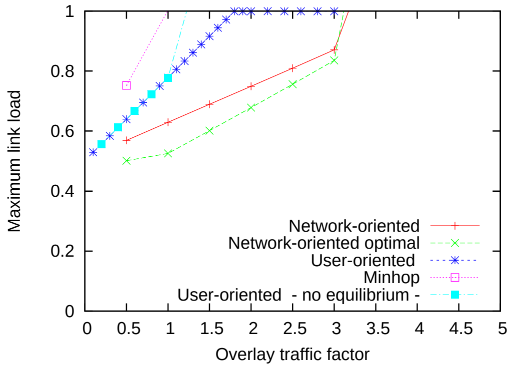
(b) The AT&T topology



(c) The BA topology

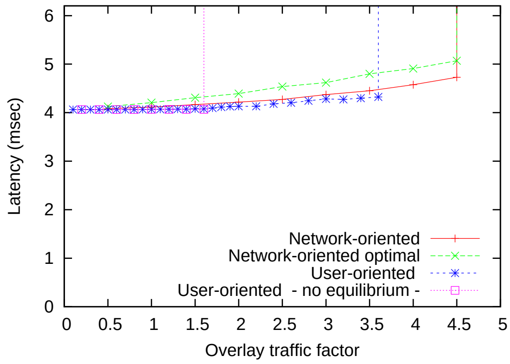


(d) The ER topology

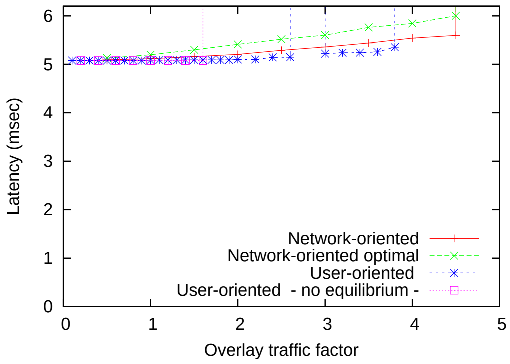


(e) The Sprint topology（POP level）

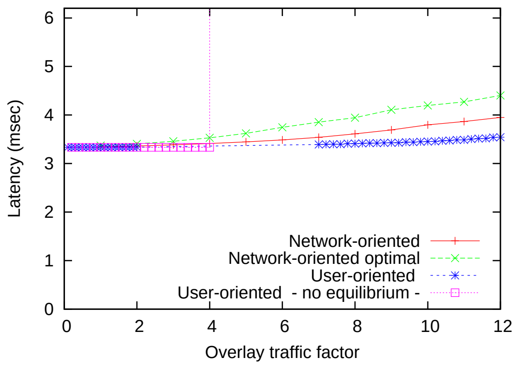
Figure 3: Maximum link load versus overlay traffic factor



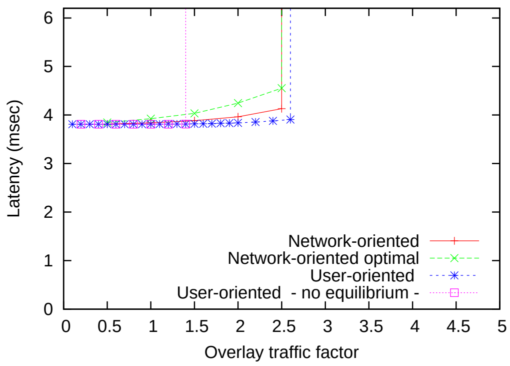
(a) The Sprint topology



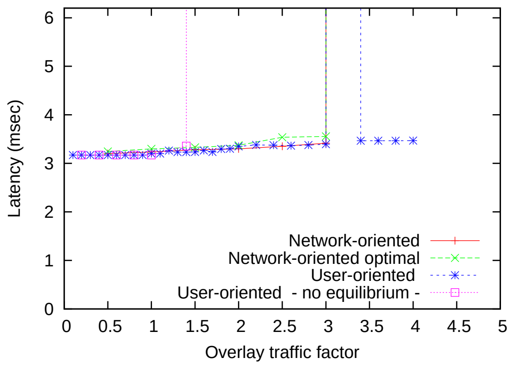
(b) The AT&T topology



(c) The BA topology



(d) The ER topology



(e) The Sprint topology (POP level)

Figure 4: Average latency versus overlay traffic factor

topology, we can confirm that the router level topology increase the maximum link load more rapidly than the POP level topology. In study [3], Lili Qiu et al. evaluate the overlay routing in POP level topologies. In addition, we show that the larger topologies, router-level topology, are affected by overlay routings from the perspective of maximum link load.

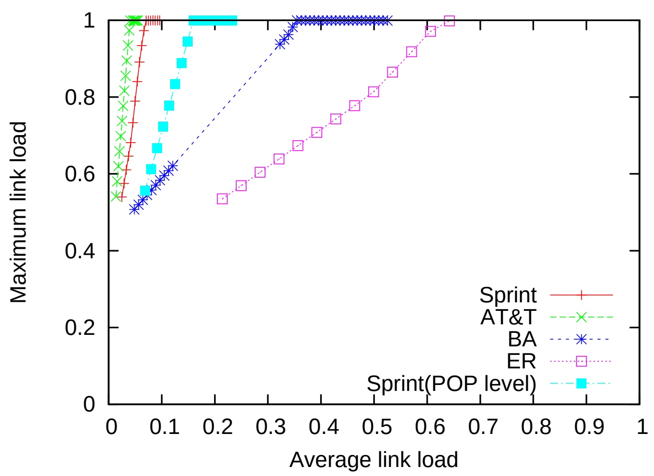


Figure 5: Maximum link load versus average link load

3.3.4 Distribution of link load

We next focus the distribution of link load. In the following, we evaluate link load of minimum hop routing and the overlay routings described in Section 2 to ISP topologies. In our evaluations when the minimum hop routing is applied, we set the amount of non-overlay traffic such that the maximum link load becomes 1.0. We then generate the same amount of traffic as the traffic when applying the overlay routings.

In the following figures, we also show the result of minimum hop routing, denoted as "Min-hop", that of user-oriented overlay routing, denoted as "User-oriented", that of network-oriented overlay routing, denoted as "Network-oriented", and that of network-oriented overlay routing, denoted as "Network-oriented optimal".

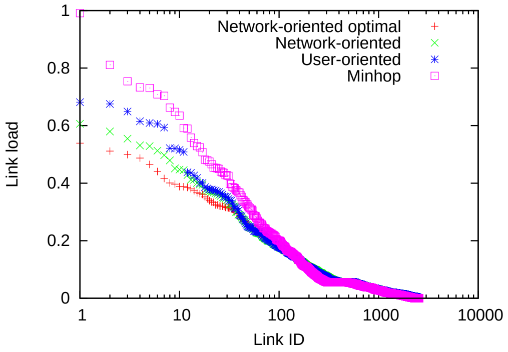
Minimum hop routing Fig. 6 shows the distribution of link load. In the graph, the vertical axis represents the link load. The horizontal axis represents the rank of each link. The graph shows that in ISP topologies, the Sprint topology and the AT&T topology, the differential of link load between the most congested link and the other link is larger than in BA topology. In the graph, the largest value of link load is 1.0, but the second largest value of link load is about 0.8 in the Sprint topology. Also in AT&T topology, that of link load is about 0.7. However, we can see more than five links that have load of over 0.8 in the BA topology. This shows that the traffic tends to concentrate to certain links in ISP topology rather than in the BA topology.

Overlay routings In Fig. 6, we can see that the maximum link load of user-oriented overlay routing is higher than the others. In addition, the figure shows that the load of many links is widely reduced in the BA topology but that of few links is reduced by the difference of the link load between minimum hop routing and the overlay routings.

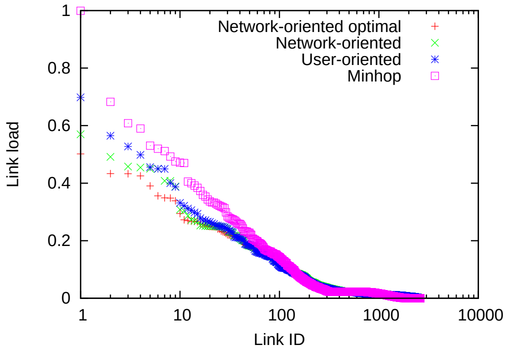
Next, Fig. 7 shows that how to change the link load between the minimum hop routing and the overlay routings. The vertical axis represents the link load of minimum hop routing, and the horizontal axis represents link load change with overlay routing. In other word, the horizontal axis means the difference of link load between minimum hop routing and overlay routing. Bigger value on the horizontal axis means that the link load increases with overlay routing. The figure shows that when the link load is high with minimum hop routing, it is reduced by overlay routing. Especially, the link load with overlay routing is reduced in proportion to the link load with minimum hop routing and we can approximate it by the downside line.

When focusing on the links that have low load with minimum hop routing, the link loads increases in all topologies. Although the load of links that have low load with minimum hop routing in Sprint and AT&T increase about 0.2, the increase in the BA topology is not so league, about 0.05. This shows that the overlay routings can distribute the traffic which aggregate to the link with high load in minimum hop routing, and reduce the link load in the BA topology.

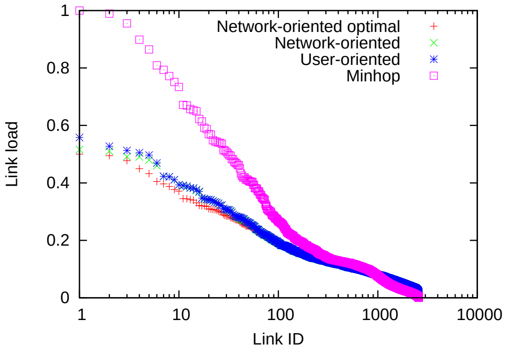
In ISP topology, it is similar to the case of the BA topology that Overlay routing can distribute the traffic through high load link. However, since it distributes the traffic through certain links, link load becomes higher. Compared to the BA topology, in other word, the number of links of which load changes significantly is high when the overlay routing is applied in ISP topology. In such topologies, the traffic tends to concentrate more than network administrator expects.



(a) The Sprint topology

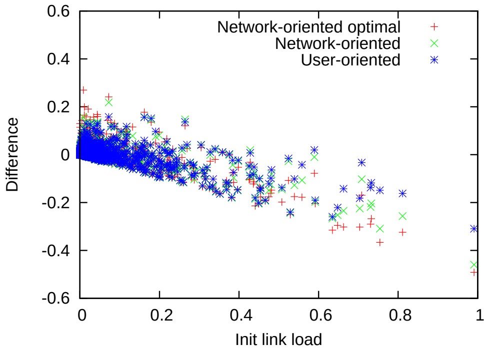


(b) The AT&T topology

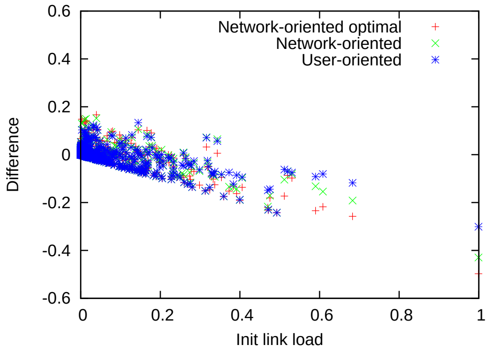


(c) The BA topology

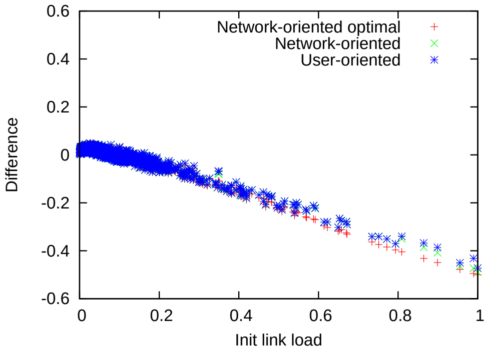
Figure 6: Distribution of link load



(a) The Sprint topology



(b) The AT&T topology



(c) The BA topology

Figure 7: Corelation between the link load with minimum hop routing and with overlay routing

4 Interaction between overlay routing and underlay routing

4.1 Interaction between two routing algorithms

As well as the overlay routing, the IP routing that dynamically controls the routes of the IP packets to detour the congested links have been considered by conventional researches. However, it is possible that the routing of the overlay routing and routing on the IP network interfere with each other, which deteriorate the performance of IP networks. We call this state the interaction. The interaction occurs when the overlay routing changes routes and degrade the performance of IP routing's routes. We then evaluate how much overlay routing and dynamic IP routing affects each other.

4.2 Simulation model

In this section, we show the simulation models for evaluating the interaction between the overlay and the IP routing. We use the Sprint topology and the AT&T topology as ISP topologies. We also use BA topology for comparison. In our evaluation, we assume that each node-pair generates the same amount of traffic at a unit time and all nodes are overlay nodes. We also assume that all links have 10 Gbps capacity and that the propagation delay of all links is 1 msec. Because we focus on the link load property, we set the processing delay on nodes 0 msec. Therefore the latency is the total of propagation delays and transmission delays on links. As for the network level routing, we assume the dynamic IP routing as shown in Sec. 2.1.2. As for the overlay routing, we assume the user-oriented overlay routing. In our simulation, the dynamic IP routing and the user-oriented overlay routing changes their routes of packets periodically. We assume that after the overlay routing continuously changes routes four times, the dynamic IP routing changes routes. We also assume that all nodes are overlay nodes in networks.

4.3 Evaluations and discussions

At first, we evaluate the mutual interference between overlay and IP routings. Next, we show the structural characteristic of nodes using the classification method in Ref. [19].

4.3.1 Evaluation results

Fig. 8 shows the maximum link load in the topologies with the overlay and the dynamic IP routings. The horizontal axis represents the number of source routing nodes. As the number of the source routing node increase, the amount of traffic which is dynamically controlled in IP network becomes large. In Fig. 8, source routing nodes are located by descending order of the number of the degree of nodes. In simulation experiments, the maximum link load always changes because overlay and source routing change the routing path freely and dynamically. So, the figures show the maximum, minimum, and average of link load during the simulation.

In the figure, we can see that the minimum link load is lower than the link load with no source routing node in all topologies. However, with the increase of source routing nodes, the maximum and average link load become higher. This indicates that the mutual interference caused by overlay and IP routing often makes the maximum load higher even though overlay routing may reduce the maximum link load with dynamic routing in IP network.

However, we can see that the maximum link load is reduced only in the case of a few source routing nodes on AT&T topology. To clear up the reason of the result, we show the difference of the role of nodes with high degree among Sprint, BA, and AT&T topology in next subsection.

4.3.2 Classification of node function

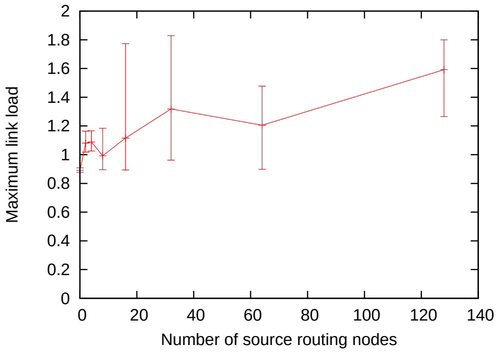
Roger Guimera et al. have proposed the classification method of node functions in Ref. [19]. The method divides a network to some modules and defines the within modularity, Z, and the participation coefficient, P, for each node. The P and Z of node i are defined as follows.

$$Z_i = \frac{k_i - < k_{s_i} >}{\sigma_{s_i}} \tag{1}$$

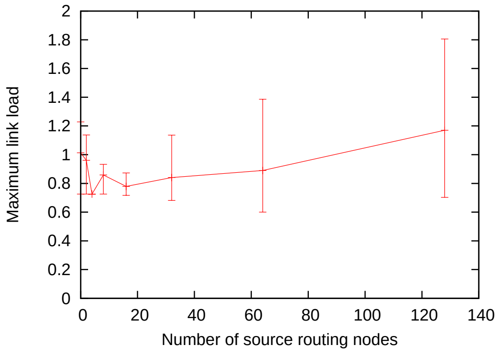
Here, k_i indicates the degree of node i. The σ_{s_i} indicates the variation of the degree in module s to which node i belongs. The $< k_{s_i} >$ means the average of the degree in the module s. Thus, Z_i becomes high when the number of degree of node i is high for the average degree in the module to which node i belong.

The definition of participation coefficient P for node i is as follows.

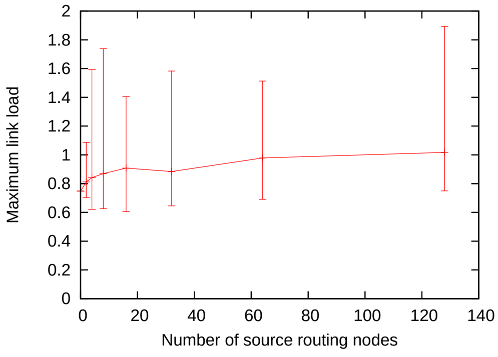
$$P_i = 1 - \sum_{s=1}^{N_m} (\frac{k_{is}}{k_i})^2 \tag{2}$$



(a) The Sprint topology



(b) The AT&T topology



(c) The BA topology

Figure 8: The maximum link load : source routing nodes are located by descending order of the degree

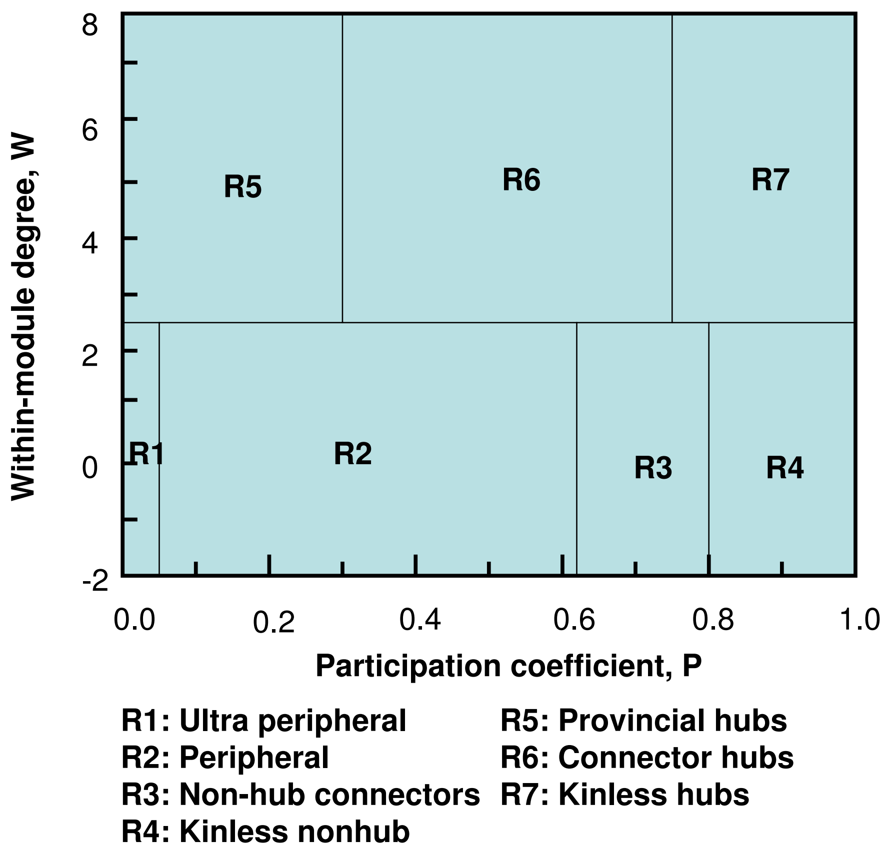
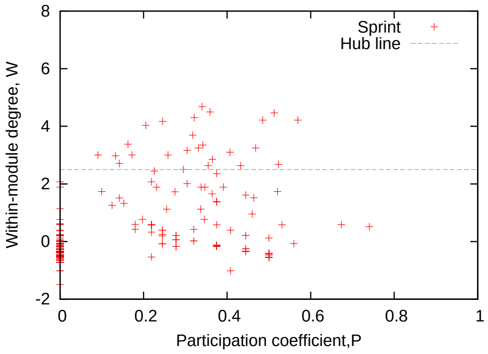
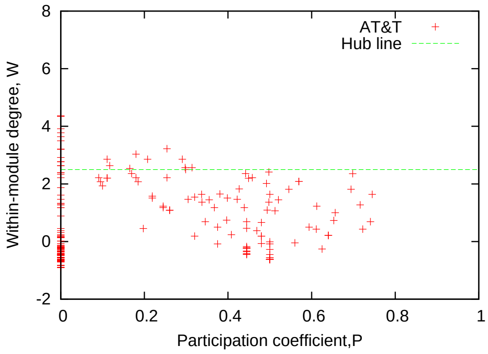


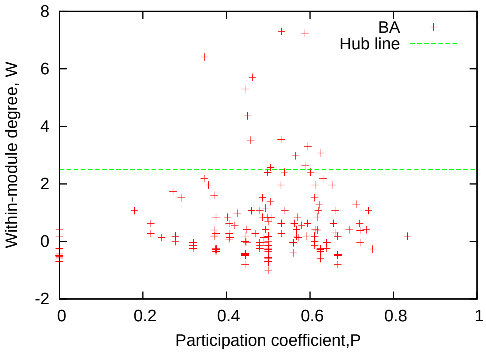
Figure 9: Classification of node function with participation coefficient and within module degree



(a) The Sprint topology



(b) The AT&T topology



(c) The BA topology

Figure 10: Classification of node function in each topology

Here, k_{is} means the number of links which connect to node i and its destination nodes belong to the same module as that of node i. That is, when the destination nodes of all links connected to node i belong to the same module as that of node i, P_i becomes 0. When the destination nodes of all links connected to node i belong to the different module, P_i becomes 1. Roger Guimera et al. [19] shows that the role of nodes is categorized by the value of Z_i and P_i, as shown in Fig. 9.

Fig. 10 shows the result of application of classification method, illustrated in Fig. 9, to Sprint, AT&T, and BA topology using the method [20]. In Fig. 10, the vertical axis indicates the participation coefficient. The horizontal axis represents within module degree. In this figure, the source routing nodes with high degree in the Sprint and the BA topologies are categorized as "Connector hubs". "Connector hubs" have high degree and transfer large amount of traffic between modules. Therefore, when source routing node is located on "Connector hubs", the source routing node reroute large amount of traffic. As the result, overlay and source routing interfere with each other.

On the other hand, Fig. 10(b) shows that there are not so many "Connectors hub" in AT&T topology but many "Provincial hubs". In hierarchical network topologies, "Provincial hubs" play a role that it aggregates traffic from lower layer and lays off to higher layer. So "Provincial hubs" have smaller chance to transfer the traffic between other modules than "Connector hubs". Therefore, by locating source routing nodes on "Provincial hubs", the mutual interference between overlay routing and source routing in IP network becomes small. "Provincial hubs" can reroute the traffic from lower layer to low load links.

Following that discussions, we next evaluated the interaction when we locates the source routing nodes on "Provincial hubs" in AT&T topology under the same condition as described in section 4.3.1. We could confirm that the maximum link load becomes 0.65 and the difference among maximum and minimum link load is reduced although the maximum link load becomes 1.0 with no source routing node.

5 Design principles with consideration to overlay routing

It is more important to appropriately design the network with the understanding of overlay behaviors. In this section, we evaluate the capacity design on ISP networks. We assume the three capacity designs. One is the capacity design that optimizes capacity settings with consideration to overlay routing. We call this capacity design "Overlay optimal design". The overlay optimal design minimize the network capacity cost and keep all link load below constant value under a given traffic demand. Here we define the network capacity cost as the sum of the link capacities in the topology. The capacity setting with this network design assumes the optimal capacity setting with consideration to overlay routing. Another capacity design is the capacity design that is given by adding link capacity which have high load. We call this capacity design "Conventional design". We assume that this network design is almost equal to the realistic one. The other is the capacity design that does not consider the overlay traffic. In detail, if a traffic demand is given, the capacity design allocates link capacities assuming that the traffic is all non-overlay traffic. We call this capacity design "Fixed design".

We evaluate these three capacity designs on ISP topologies. We first evaluate the network performance in ISP topologies where capacities are determined by theses three design. We next evaluate the capacity settings given by these three designs.

5.1 Approaches for capacity design in ISP networks

We explain the approaches for the two capacity design. We obtain the capacity settings of "Overlay optimal design" and "Conventional design" by using computer simulations. With the computer simulations, we repeat the capacity design and overlay routing until the capacity design do not changes capacity setting as shown in Fig. 11. The difference of this approach between the overlay optimal design and the conventional design is how to set the new capacity. The overlay optimal design reset link capacities so that all link have constant load. So the overlay optimal design also degrades link capacities which have small load in order to minimize the capacity cost. On the other hand, the conventional design reset link capacities so that all link have less than constant load. So the conventional design does not degrade the link capacities. In our evaluations, we assignee the constant values of the overlay optimal design and the conventional design to 0.5.

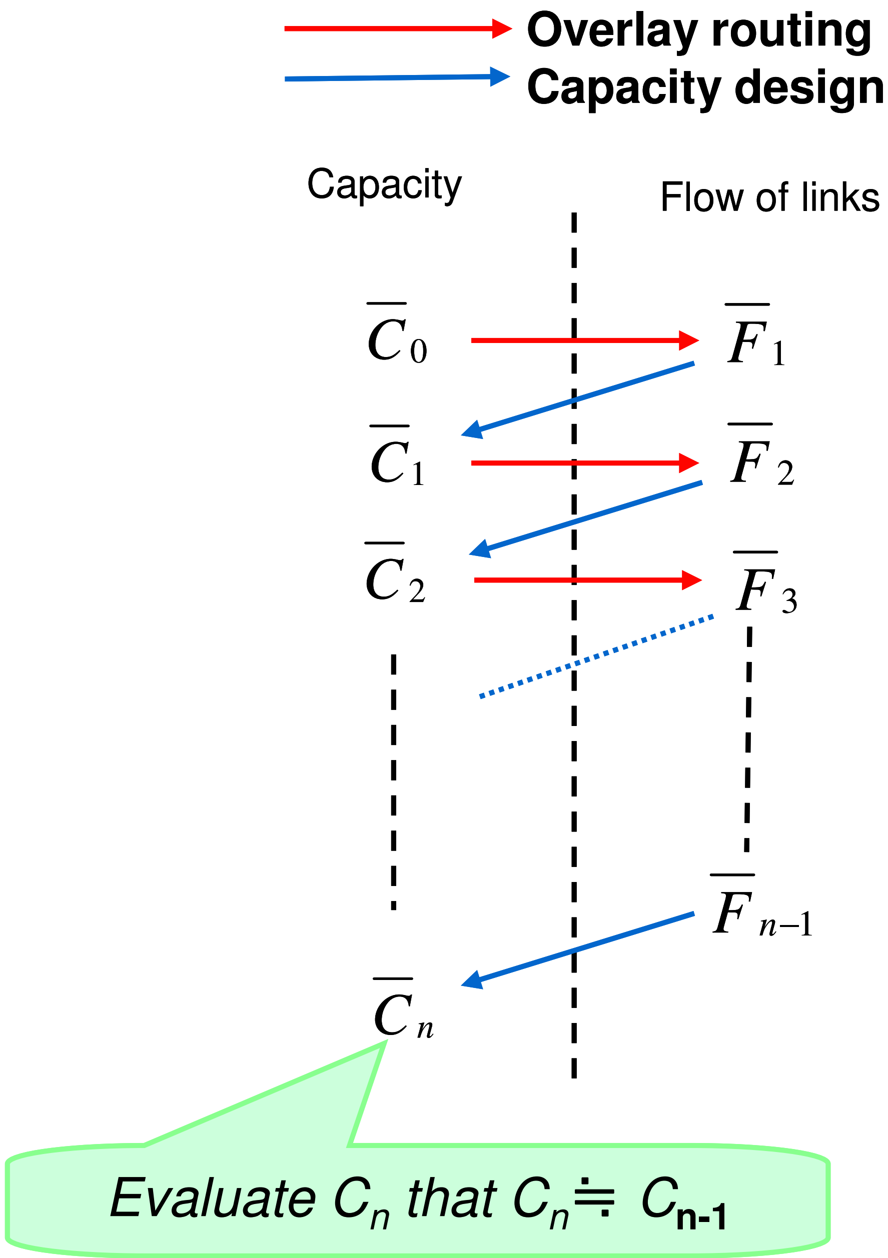


Figure 11: Approache for capacity design

5.2 Evaluations and discussions

In following section, we evaluate these three capacity designs on ISP topologies. We first evaluate the network performance in ISP topologies where capacities are determined by theses three design. We next evaluate the capacity settings given by these three designs.

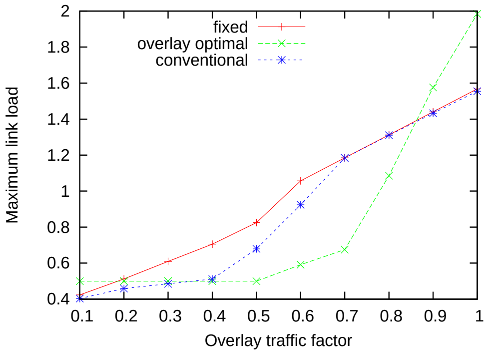
5.2.1 Network performance

We evaluate the network performance in ISP topologies where capacities are determined by theses three design. We evaluate the maximum link load and the average latency of overlay traffic and non-overlay traffic. In our evaluation, we set the network capacity cost the same with all capacity designs.

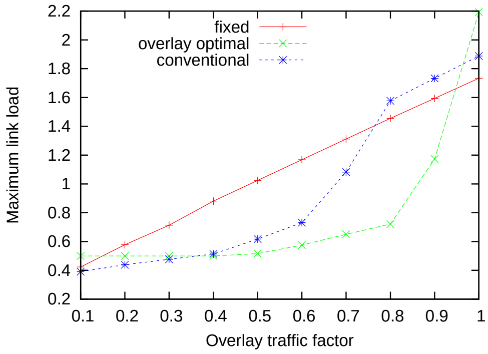
Fig. 12 shows the maximum link load. The vertical axis represents the maximum link load in the network. The horizontal axis represents the overlay traffic factor. The overlay traffic factor is same index in Section 3. The graph shows that the all topologies achieve the lowest maximum link load when the link capacities are designed by the overlay optimal design. That is, the overlay optimal design can design the robust network to overlay traffic.

Fig. 13 shows the average latency of overlay traffic in each topology. The vertical axis represents the average latency of overlay traffic. The horizontal axis represents the overlay traffic factor. The graph shows that the all topologies achieve the smallest average latency of overlay traffic when the link capacities are designed by the overlay optimal design. The difference of the latency is not so large when the amount of overlay traffic is small. However, as the overlay traffic increases, the topologies that designed by the overlay optimal design achieves low latency.

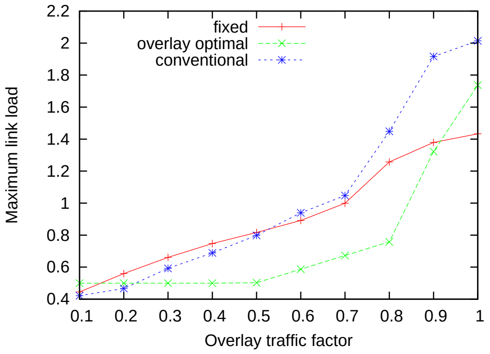
On the other hand, as for the average latency of non-overlay traffic, we can see the countertrend. Fig. 14 shows the average latency of non-overlay traffic in each topology. The vertical axis represents the average latency of non-overlay traffic. The horizontal axis represents the overlay traffic factor. The graph shows that the all topologies achieve the smallest average latency of overlay traffic when the link capacities are designed by the fixed design. However, as the overlay traffic increases, the topologies that designed by the overlay optimal design becomes to achieve lower latency than the topologies designed by the other two design.



(a) The Sprint topology

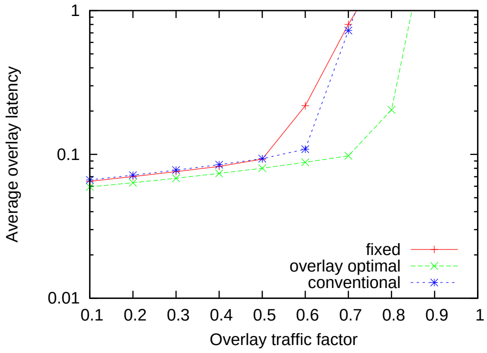


(b) The AT&T topology

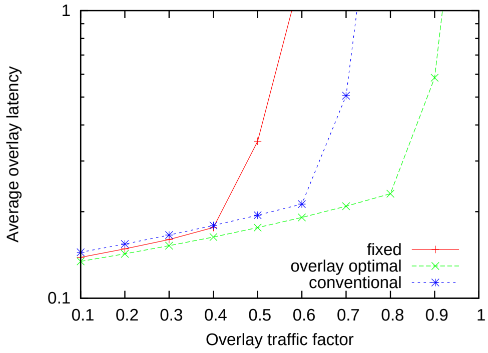


(c) The BA topology

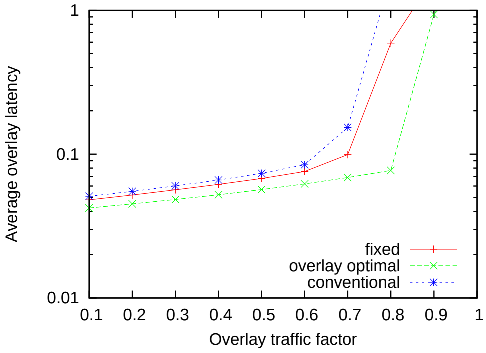
Figure 12: Maximum link load in each topology



(a) The Sprint topology

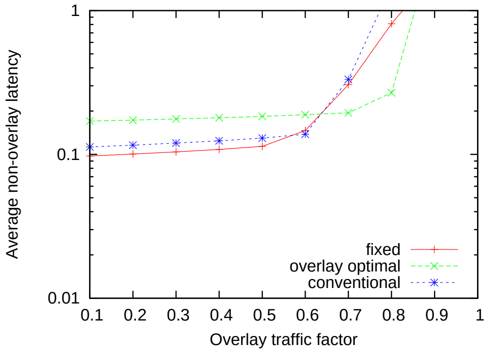


(b) The AT&T topology

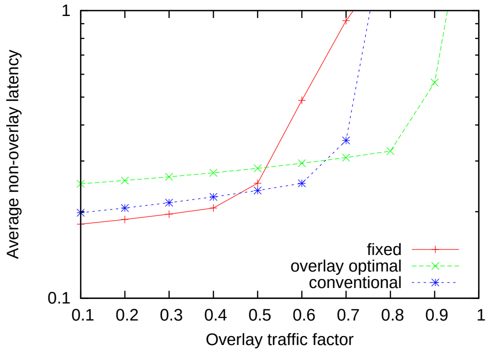


(c) The BA topology

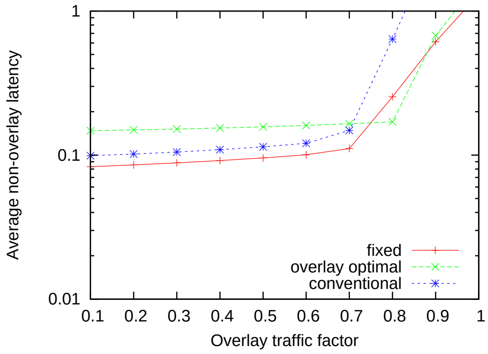
Figure 13: Average latency of overlay traffic in each topology



(a) The Sprint topology



(b) The AT&T topology



(c) The BA topology

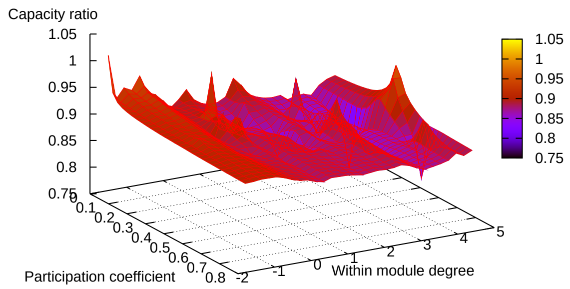
Figure 14: Average latency of non ovelay traffic in each topology

5.2.2 Characteristic of link capacity

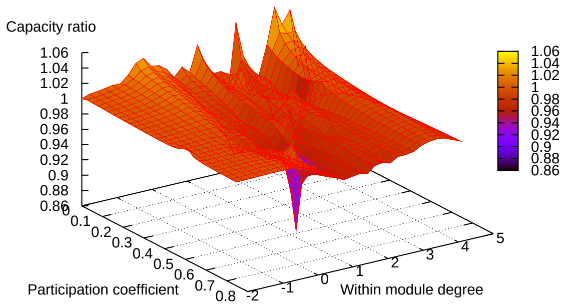
We next evaluate the capacity settings given by these three designs. Fig. 15 shows the ratio of the capacity designed by the overlay optimal design to the capacity designed by the conventional design. The x axis represents the participation coefficient of source node of each link. The y axis represents the within module degree of source node of each link. And the z axis represents the ratio of the capacity designed by the overlay optimal design to the capacity designed by the conventional design. If a links have the larger ratio, the link should be enhanced when considering overlay behaviors. That is, we should enhance the links more aggressively. If a links have the smaller ratio, the capacity of the link should be decreased.

Fig. 15(a) shows the result on the Sprint topology. The graph shows that there are some higher areas. One of the areas has about 0.4 participation coefficient and 4.5 within module degree. This area corresponds to the connector hubs. We can see that we should enhance the links that are connected to the connector hubs. If we design the capacities considering only the loads of links, the network may hardly accommodate overlay traffic.

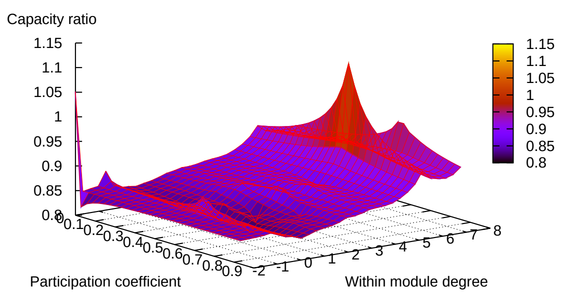
Another area where the value of the graph is high has about 0–0.2 participation coefficient and -2–0 within module degree. This area corresponds to the ultra-peripheral or peripheral. This result may be against our intuition. However we think that this result is affected by the ISP topologies. ISP router-level topologies have higher cluster coefficient and ISP topologies are locally clustered as shown in Ref. [21]. That is, there are some alternate paths which have almost equal latency in the same module. Therefore the oscillation of the overlay routes often occurs and the maximum link load increases. As well as the Sprint topology, we can see that the AT&T topology has the similar area where the ratio is high in Fig. 15(b). On the other hand, we can also see that the BA topology, which is not ISP topology, does not have such area as shown in Fig. 15(c).



(a) The Sprint topology



(b) The AT&T topology



(c) The BA topology

Figure 15: Capacity ratio of the overlay optimal design to the conventional design

6 Conclusion

In this thesis, we focused on overlay routing. We first evaluated the impact of the overlay routing's selfish behavior on realistic router-level topologies from the perspective of maximum link load and average latency in the case that the fixed shortest path routing is adopted for the underlay networks. The results of our evaluation shows that the overlay routing achieves near optimal latency but concentrates traffic on certain links at equilibrium. In addition, our results show that the maximum link load and the average latency in realistic router-level topologies which have the power-law attribute are more affected by the overlay routing than those in POP level topologies. That is, much more traffic concentrates on certain links in the realistic router-level topology by the overlay routing. We also demonstrate that the overlay routing achieve higher maximum link load and latency, when the overlay routing is at non-equilibrium due to the oscillation of routes in overlay networks.

We next consider the case that the routes in underlay networks are dynamically updated against the change of link load. We evaluated the impact of the interaction between the overlay routing and the routing in the underlay networks in realistic router-level topologies. The results show that the interaction between the overlay routing and underlay routing degrades the network performance. However the results also have shown that the overlay routing and the underlay routing can coexist when only the nodes classified as "provincial hubs" dynamically update the routes in the underlay networks against the change of link load.

We finally investigate the network design of ISP networks with consideration to overlay routing behaviors. We show that not only the nodes classified as "connector nodes" but also the nodes classified as "peripheral nodes" have important roles for reducing the impact of the overlay routing behaviors on ISP networks.

I would like to appreciate to my supervisor, Professor Masayuki Murata of Osaka University, for his pointed feedbacks, valuable comments, and encouragement.

All of my work would not be achieved without the support of Research Associate Shin'ichi Arakawa. I sincerely appreciate for him to improve my work and to support preparations of this thesis.

I am most grateful to Professors Koso Murakami, Makoto Imase, Teruo Higashino, and Hirotaka Nakano of Osaka University, for their appropriate guidance and invaluable firsthand advice.

I am particularly thankful for the continued input of Associate Professor Naoki Wakamiya, Associate Professor Go Hasegawa, and Research Associate Masahiro Sasabe.

I owe a lot to my friends, colleagues and the Advanced Network Architecture Research Group of Osaka University. Our conversations and work together have greatly influenced this thesis.

Last, but not least, I thank my parents for their invaluable support and constant encouragement during my undergraduate studies.

References

[1] B. Fortz and M. Thorup, "Internet traffic engineering by optimizing OSPF weights," in *Proceedings of IEEE INFOCOM 2000*, pp. 519–528, Mar. 2000.

[2] T. Roughgarden and E. Tardos, "How bad is selfish routing?," *Journal of the ACM (JACM)*, vol. 49, no. 2, pp. 236–259, Mar. 2002.

[3] L. Qiu, Y. R. Yang, Y. Zhang, and S. Shenker, "On selfish routing in Internet-like environments," in *Proceedings of ACM SIGCOMM '03*, pp. 151–162, Aug. 2003.

[4] M. Faloutsos, P. Faloutsos, and C. Faloutsos, "On power-law relationships of the Internet topology," in *Proceedings of ACM SIGCOMM '99*, pp. 251–262, Aug. 1999.

[5] Y. Liu, H. Zhang, W. Gong, and D. Towsley, "On the interaction between overlay routing and traffic engineering," in *Proceedings of IEEE INFOCOM '05*, Mar. 2005.

[6] B. Zhang, R. Liu, D. Massey, and L. Zhang, "Collecting the Internet AS–level topology," *ACM SIGCOMM Computer Communication Review*, vol. 35, no. 1, pp. 53–61, Jan. 2005.

[7] A. Akella, S. Chawla, A. Kannan, and S. Seshan, "Scaling properties of the Internet graph," in *Proceedings of the Twenty-second Annual Symposium on Principles of Distributed Computing*, pp. 337–346, July 2003.

[8] W. Willinger, R. Govindan, S. Jamin, V. Paxson, and S. Shenker, "Scaling phenomena in the Internet: Critically examining criticality," in *Proceedings of The National Academy of Sciences of The United States of America*, pp. 2573–2580, Feb. 2002.

[9] C. Gkantsidis, M. Mihail, and A. Saberi, "Conductance and congestion in power law graphs," in *Proceedings of ACM SIGMETRICS*, pp. 148–159, June 2003.

[10] K. L. Goh, B. Kahng, and D. Kim, "Universal behavior of load distribution in scale–free networks," *Physical Review Letters*, vol. 87, no. 27, Dec. 2001.

[11] R. Cohen, S. Havlin, and D. Avraham, *Handbook of Graphs and Networks – From the Genome to the Internet*, ch. 4. WILEY-VCH GmbH & Co., 2003. Structural Properties of scale–free networks.

[12] A. Barabasi and R. Albert, "Emergence of scaling in random networks," *Science*, vol. 286, pp. 509–512, Oct. 1999.

[13] L. Li, D. Alderson, W. Willinger, and J. Doyle, "A first–principles approach to understanding the Internet's router–level topology," *ACM SIGCOMM Computer Communication Review*, vol. 34, no. 4, pp. 3–14, Oct. 2004.

[14] R. Fukumoto, S. Arakawa, and M. Murata, "On routing controls in ISP topologies: A structural perspective," in *Proceedings of Chinacom*, Oct. 2006.

[15] N. Sprint, R. Mahajan, D. Wetherall, and T. Anderson, "Measuring ISP topologies with rocketfuel," *IEEE/ACM Transactions on Networking*, vol. 12, no. 1, pp. 2–16, Feb. 2004.

[16] P. Erdös and A. Rényi, "On the evolution of random graphs," *Publications of the Mathematical Institute of the Hungarian Academy of Sciences*, vol. 5, pp. 17–61, 1960.

[17] D. Andersen, H. Balakrishnan, F. Kaashoek, and R. Morris, "Resilient overlay networks," *ACM SIGOPS Operating System Review*, vol. 35, no. 5, pp. 131–145, Oct. 2001.

[18] L. Fratta, M. Gerla, and L. Kleinrock, “The flow deviation method: An approach to store-and-forward communication network design,” *Networks*, vol. 3, pp. 97–133, 1973.

[19] R. Guimera and L. A. N. Amaral, “Functional cartography of complex metabolic networks,” *Nature*, vol. 433, p. 895, 2005.

[20] M. E. J. Newman, “Modularity and community structure in networks,” *PROC.NATL.ACAD.SCI.USA*, vol. 103, no. 888, p. 8577, 2006.

[21] R. Fukumoto, S. Arakawa, T. Takine, and M. Murata, “Analyzing and modeling router–level Internet topology,” in *Proceedings of The International Conference on Information Networking (ICOIN)*, Jan. 2007.