

Pricing for Maximizing Provider's Revenue in Multicast Content Delivery Services

Takehiro Kajita¹, Kyoko Yamori², Yoshiaki Tanaka³

¹Global Information and Telecommunication Institute, Waseda University, Tokyo, 169-0051 Japan

kajita000@ruri.waseda.jp

²Department of Business Administration, Asahi University, Mizuho-shi, 501-0296 Japan

kyamori@alice.asahi-u.ac.jp

³Research Institute for Science and Engineering, Waseda University, Tokyo, 162-0044 Japan

ytanaka@waseda.jp

Abstract. With the rapid growth of broadband computer networks, users are inclined to use delivery services that can handle large-size content. When offering download service of popular contents, it is more useful to use a multicast system and deliver content to all users who request the content than using unicast system. Using a multicast system, the content server can increase the number of users and decrease user's waiting time. As a result, increasing user's WTP (Willingness To Pay) and the provider's revenue can be expected. In this paper, we suggest two content delivery methods which can be employed in the multicast content download services. We also discuss the provider's revenue of each delivery method. Using user's behaviour model in which a user participates in the service and leaves it, we show the relation between price for the service and user's participation rate of the service through simulation. The relation between price and the provider's revenue is also demonstrated in the simulation. In addition, in each delivery method, we show an optimal price which maximizes the provider's revenue.

Please use the following format when citing this chapter:

Kajita, T., Yamori, K., Tanaka, Y., 2008, in IFIP International Federation for Information Processing, Volume 286; Towards Sustainable Society on Ubiquitous Networks, eds. Oya, M., Uda, R., Yasunobu, C., (Boston: Springer), pp. 173–183.

1 Introduction

With the rapid growth of broadband computer networks, users are inclined to use delivery services that can handle large-size content. In the case of unicasting, the content server has to send the same data to many receivers repeatedly. With multicasting, on the other hand, the content server has to send the data only once. At the split node of the routing tree, the node copies the data and then sends them. Thus, a multicast system can reduce the network traffic.

When offering download services for popular contents, a multicast system is more beneficial compared to a unicast system for content delivery. If a network bandwidth can be used effectively, we can increase the maximum number of users who can participate in the service. And because of the increase in bandwidth during delivery, user's waiting time decreases, which results in an increase in user's WTP (Willingness To Pay) for the service, thereby increasing the provider's revenue.

However, during multicast deliveries, the content server ignores available bandwidth of current users in the multicast tree, therefore, narrowband users, i.e. users with low available bandwidth, may experience packet loss. In content download services, packet loss can cause serious problems. When the delivery speed is greater than the available bandwidth of the user, packet loss occurs and the buffer will overflow at an intermediate router along the path.

In order to avoid packet loss, the content server has to set the delivery speed to the smallest available bandwidth of the user in the multicast group in a multicast download service. In other words, as long as a link with low available bandwidth exists, every user in the multicast tree will experience degrade in performance.

In a content download service, user's utility corresponds to user's feeling with regards to the service and is determined by waiting time. If the user's utility increases, user's WTP for the service will increase accordingly. When user's WTP is high, the service provider can set a high price for the service and thereby increasing revenue. Therefore, in a multicast content download service, it is necessary to increase user's utility and WTP by using some solutions for narrowband users for maximizing the provider's revenue.

As a solution for narrowband users, this paper proposes using multiple speed scenario, which divides narrowband users and broadband users into separate groups and performs separate multicast deliveries for each group.

This paper assumes two delivery methods for narrowband users, and discusses the provider's revenue in the multicast content download service.

2 Multicast Content Delivery Services

2.1 Service Overview

In this paper, we assume a multicast download-type content delivery service. An example is a service which sells software and delivers it with a multicast system on the network. Conventionally, a unicast system was the mainstream method for download-type content delivery services. However, on the start of the delivery of popular contents, it is expected that requests from many users converges for a short time at the server.

In the case of unicasting, the link bandwidth is shared by all registered users, and if many users send their request to the content server within a short time, the delivery

speed for all users will be too slow and network traffic congestion will occur. For this case, it is necessary to refuse the user's request in order to reduce the server load and network congestion. However, when a user's request is rejected by the service, the user's WTP will decrease accordingly [1]. Thus the provider will face a loss of revenue. When the provider's revenue is taken into consideration, the service provider should not refuse too many incoming requests.

On the other hand, in the case of multicasting, if many users send their request to the content server within a short time, the content server can collect their requests for one multicast flow and send the data set to all the registered users at once. Thus, a multicast system can reduce network traffic congestion and send the content to many users. Taking all this into consideration, it is more effective to use a multicast system to deliver popular contents. As a result, an increase in the provider's revenue can be expected. Therefore, in this paper, we assume a multicast download-type content delivery service.

In a multicast content download service, it is more reasonable to collect user requests over a period of time for later delivery over a multicast tree. This paper uses a service that performs multicast delivery to users within a given time interval, referred to as delivery interval.

In order to avoid packet loss, the content server sets the delivery speed to the smallest available bandwidth in the tree in a multicast download service. Therefore, user's waiting time increases greatly by the influence of narrowband users. In this paper, we suggest two delivery methods as solutions to the problem of narrowband users. And we compare the provider's revenue in each delivery method and examine the effectiveness.

The process of content delivery is shown in Fig. 1.

- (i) Delivery requests originate from users and are sent to the content server.
- (ii) The content server delivers content to all users using a multicast system.
- (iii) Users pay for the service to the service provider.

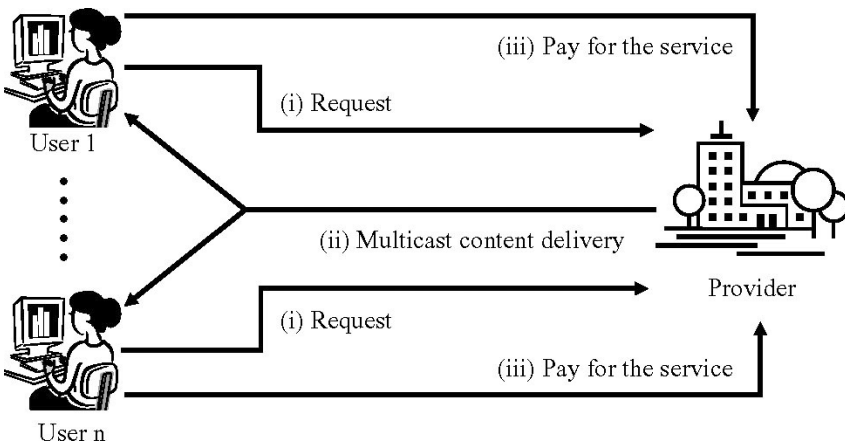


Fig. 1. The process of content delivery

In this paper, we assume a service which sells software, with content-based pricing scheme for the service pricing. Users who want to get content send requests to the content server. The content server collects their requests for one multicast flow and sends the data set to all the registered users at once. Users pay the content provider after receiving the contents. Because we assume that the pricing scheme for the service is content-based, it is necessary for the user to pay for the service whenever the user downloads contents. We assume pay per view.

2.2 Simple Scenario and Multiple Speed Scenario

We suggest two delivery methods as solutions to the problem of narrowband users which are exemplified in Fig. 2 and Fig. 3. The delivery method shown in Fig. 2 is called simple scenario in this paper. The simple scenario is a basic delivery method. In the simple scenario, the content server sets the delivery speed to the smallest available bandwidth in the multicast tree. If the simple scenario is adopted, the packet loss problem can be alleviated considerably. However, the simple scenario has a problem -- if only a single user in the whole tree experiences the bandwidth bottleneck, the decrease in delivery speed will also cause the drop in user's WTP. On the other hand, the delivery method shown in Fig. 3 is called multiple speed scenario in this paper. The multiple speed scenario divides narrowband users and broadband users into separate groups and performs separate multicast deliveries for each group as a solution for narrowband users. This method enables delivery to broadband users with a high delivery speed, while not blocking requests from narrowband users.

Because in those delivery methods, delivery speed is adjusted to the slowest user's available bandwidth in the multicast group, even in the situation where packet loss occurs, it is expected that the amount of packet loss is minimal. Therefore, we can resend the lost data by using reliable multicast technology [2].

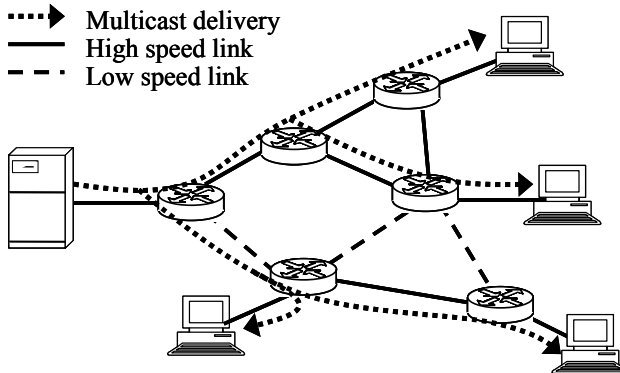


Fig. 2. Simple scenario

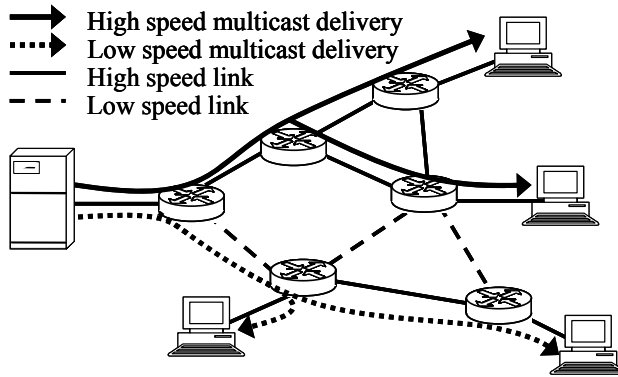


Fig. 3. Multiple speed scenario

3 User's Behaviour Model

3.1 User's Joining Model

In order to evaluate the services, we have to determine the user's behaviour model which in turn determines user participation. In this paper, we assume the user's behaviour model using user's WTP. WTP stands for the maximum amount of money that a user is willing to pay. WTP is often used for evaluation of services.

The following steps determine user participation: The user compares the price for the service and one's WTP. The result of this comparison determines whether the user joins in the service. If the user's WTP is higher than price for the service, the user joins in the service.

The WTP of user i is defined by Eq.(3.1).

$$U_i = \alpha \times v_i \tag{3.1}$$

where U_i is WTP of user i . v_i is the evaluation value of the service to waiting time of user i . α is the random parameter in the range of 0 to 1.

In this paper, v_i which is the evaluation value of the service to waiting time is given by:

$$v_i = 1 - 0.5 \times \frac{t}{x} \tag{3.2}$$

where t is the waiting time of user i . x is the mean waiting time in the simple scenario. Relation between user's waiting time and the evaluation value of the service is shown in Fig. 4. As shown in Fig. 4, v_i is 1 when the user's waiting time is 0. And v_i will decrease linearly with increase in waiting time.

We use the mean waiting time of the simple scenario as a standard in this paper. When the waiting time of user i is the same as the mean waiting time of the simple scenario, v_i becomes 0.5. Moreover, v_i has a variance within the range of -0.05 to 0.05.

Before user i joins in the service, user i does not know its own waiting time. In this case, user i uses the mean waiting time of users who joined in the service before user i tries to join in the service.

If the content has not been delivered to the user, the user will not know the mean time of the service. Therefore, in this case, WTP of users are decided randomly within the range of 0 to 1.

In this experiment, the evaluation value of the user is decided based on the mean waiting time of the simple scenario. It is necessary for this evaluation value to be more precisely determined via a questionnaire survey.

When the user's request arrivals, the user decide one's WTP using Eq.(3.1). However, because the waiting time cannot determine at this point, the user decides the evaluation value with the mean waiting time of the current service.

The user i compares the price p for the service and one's WTP. Where $U_i > p$, the user joins in the service. Where $U_i < p$, the user does not join in the service. Because we assume that the pricing scheme for the service is content-based, p is a price of one download per content. It is necessary for users to pay for the service whenever contents are downloaded. The condition of the decision and the probability of joining in the service are given by:

$$P_i = \text{Prob}(U_i > p) \tag{3.3}$$

P_i is the probability that user i joins in the service.

The user who participates in the service joins the multicast group of the requested content, and receives content with multicast content delivery.

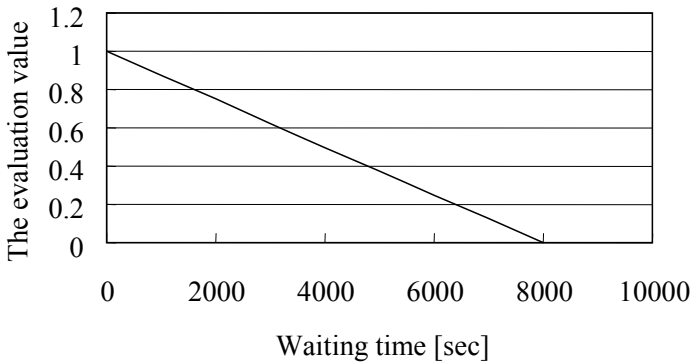


Fig. 4. The evaluation value versus waiting time

3.2 User's Leaving Model

Users pay the price p for the service after the content delivery. And then, users recalculate the evaluation value v_i' by using the user's waiting time of this content delivery using Eq.(3.2). New WTP U_i' of user i is decided by using this v_i' .

Where $U_i^* < p$, the user might not be satisfied in spite of having paid for the price for the service. Subsequently, the user leaves service in time of one delivery interval and will not send any more requests to the service during that time.

In this paper, the arrival of user requests follows Poisson distribution. Arrival rate is set to $\lambda = 0.5$ [1/sec]. Basically, the number of users on the network is set to 5000. Arrival interval of a request from one user is set to 10000[sec]. The value of λ is calculated as

$$\lambda = \frac{5000}{10000}$$

During the time when some users leave the service, arrival rate λ' is calculated as

$$\lambda' = \frac{(5000 - d)}{10000}$$

where d is the number of leaving users.

4 Evaluation Model

4.1 Network Model

This paper estimates the provider's revenue by performing a simulation. To generate network topology for the simulation, a random network of $N = 50$ nodes was created using the method proposed by Waxman et al [3].

Since Dijkstra shortest path search algorithm is used to define a path, there is a tendency for uneven utilization of network links due to concentration of traffic in a small portion of the network while the rest of the network is underutilized [4-6].

In our simulation, H-links with small background traffic, i.e. high available bandwidth, constitute the majority of all links in the network. L-links with big background traffic, i.e. low available bandwidth exist locally. Narrowband users are located on paths with at least one L-link, and in the case of broadband users, all links along the path are H-links.

In this paper, it is assumed that position of L-links changes with time. According to traffic measurement results in [7], the shortest interval between large changes in network traffic is about 1 hour. Therefore, in the simulated network the interval between changes in the link state is randomly selected each time from the range between 1800 and 5400 seconds. Links can become L-links as the results of the state change with the probability of β . Duly, the probability of change into an H-link is $1 - \beta$. The probability β , therefore, directly controls the ratio of narrowband users within the network. For the sake of simplicity, locations for bottleneck links are selected randomly. This paper does not consider the self-similarity of traffic.

In this paper, we set the bandwidth of H-link to 100Mbps and the bandwidth of L-link to 10Mbps. The probability of β which decides the ratio of L-link is set to 2%. Even if the probability β is low, there are cases where L-links can not be avoided. Therefore, in this paper, we do not consider methods which ignore L-links when multicast trees are constructed.

4.2 System Model

In this paper, we assume that ten kinds of content exist in the content server and users send requests to the server to get content. The size of the content in total constitutes

$C=450$ Mbytes. The content is not rated by popularity, and each user is assumed to request all contents without partiality.

In this paper, arrival rate is set to $\lambda=0.5$ [1/sec]. λ is used for all 10 contents. Therefore, the arrival rate λ_i of request for content i is easily calculated as $\lambda_i=\lambda/10$. Additionally, arrival rate λ is the arrival rate for the whole network and is constant throughout the whole simulation.

The server uses a fixed delivery interval. The server begins the content delivery, and starts a new delivery interval for the next delivery at the completion of each delivery interval. In the multiple speed scenario, because the server performs low-speed delivery for narrowband users and high-speed delivery for broadband users simultaneously, the server uses a fixed delivery interval for a high-speed delivery (H-Interval) and a fixed delivery interval for a low-speed delivery (L-interval) respectively.

When a user makes a request to the server, it decides which group this user belongs to according to the available bandwidth of the user, and adds this user to an appropriate multicast group. Narrowband users who are located on paths with at least one L-link are added to low-speed delivery group. And in the case of broadband users, since all links along the path are H-links, they are added to the high-speed delivery group.

In the multiple speed scenario, at the completion of H-Interval, the server begins to deliver ten different contents at the request of broadband users. Similarly for narrowband users, ten different contents are delivered upon request at the completion of L-Interval.

Assuming best effort network delivery, in order to deliver 10 distinct contents separately, the speed for each delivery should be 1/10 of link capacity or available bandwidth at the time.

In the multiple speed scenario, when a low-speed delivery uses too much bandwidth, there will be a negative effect on high-speed delivery. In order to avoid this negative influence on high-speed delivery, the server sets the upper limit speed of the low-speed delivery. In this paper, the speed limit is set to 1Mbps. And the L-Interval for low-speed delivery is set at double the value necessary to deliver content within the speed limit.

In the following examination, only the high-speed delivery interval is used as a parameter in order to find the optimal delivery interval which would maximize the provider's revenue. An optimal delivery interval has been already discussed by [8]; therefore, this paper assumes an optimal delivery interval being set without the need to evaluate the argument for an optimal delivery interval.

5 Numerical Results

5.1 Relation between Price for the Service and User's Participation Rate

Relation between price for the service and user's participation rate of the service is shown in Fig. 5. User's participation rate is calculated using the number of users who actually participated in the service and the maximum number of users participating in the service theoretically.

As per the results in Fig. 5, user's participation rate in each delivery method decreases with the growth in the price for the service. However, the user's participation

rate of the multiple speed scenario almost exceeds it in comparison with the user's participation rate of the simple scenario.

In the multiple speed scenario, if the user can join in the high-speed delivery group, the user can get content with a short waiting time. As a result, the user's evaluation value of the service becomes higher, and the probability that the user leaves the service becomes lower. Moreover, the evaluation value during user participation becomes higher because the mean waiting time of the multiple speed scenario is shorter than mean waiting time of the simple speed scenario. As a result, an improvement of the participation rate is seen compared with the simple scenario.

In this paper, the value which multiplies the evaluation value of the service and the random value of 0 to 1 is set to user's WTP. Therefore, even if there is a big difference in the evaluation value of the simple scenario and the multiple speed scenario, the difference of WTP becomes small. When it becomes clear that the value of parameter α is close to 1 through a questionnaire survey, it is expected

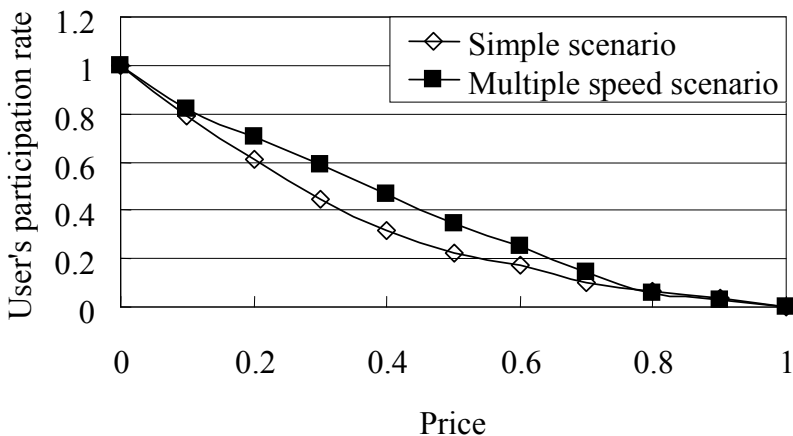


Fig. 5. Relation between price for the service and participation rate of the service

that the difference of the participation rate of the simple scenario and the multiple speed scenario becomes bigger.

5.2 Relation between Price for the Service and the Provider's Revenue

Relation between price for the service and the provider's revenue in one day is shown in Fig. 6. As per the results in Fig. 6, the optimal price which maximizes the provider's revenue is 0.4 in the multiple speed scenario. On the other hand, the optimal price which maximizes the provider's revenue is 0.3 in the simple scenario.

Because user's participation rate rises in the multiple speed scenario compared to the simple scenario, the price in the multiple speed scenario can be set higher. As a result, high provider's revenue can be obtained compared with the simple scenario as shown in Fig. 6.

In addition, when the server performs the delivery of popular content using unicast system, it is necessary to block user's request to reduce network congestion. However, when a user's request is rejected by the service, the user's WTP will decrease.

And because the link bandwidth is shared by all registered users in the case of unicasting, user's waiting time is longer than waiting time of the simple scenario and the multiple speed scenario. Therefore, the evaluation value of the service is lower than the simple scenario. Along with it, the provider's revenue is lower than the provider's revenue of the simple scenario.

As a result, using a multicast system for content download services effectively increases the provider's revenue. And we show that if the server uses the multiple speed scenario, the provider's revenue is maximized.

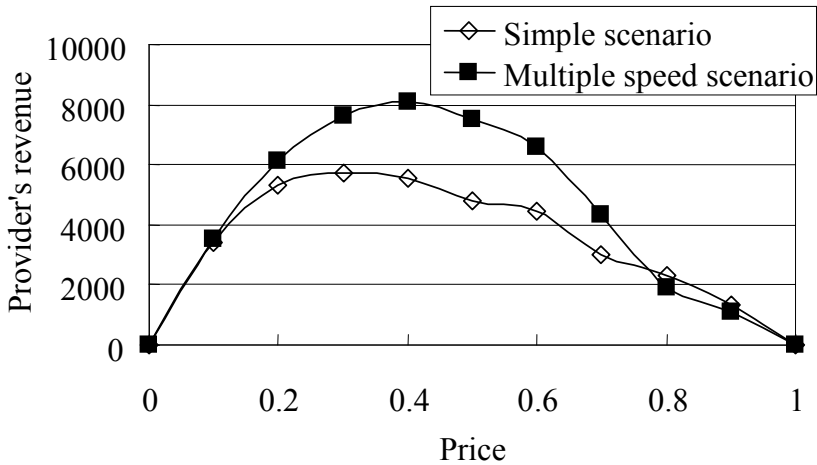


Fig. 6. Relation between price for the service and the provider's revenue

6 Conclusion

In this paper, we assumed the content download service with multicasting. And we assumed two multicast content delivery methods for narrowband users. This paper discussed the relation between price for the service and user's participation rate and provider's revenue in each delivery method.

As a result, this paper showed that if the server uses the multiple speed scenario, user's participation rate can be improved compared with the simple scenario. And the provider's revenue can be maximized by using multiple speed scenario. We also showed the optimal price for the service is 0.4.

In this paper, we used random parameter α to decide the evaluation value of the service. In future work, we plan to do a questionnaire survey to decide parameter α . Based on this, our result will be more reliable.

Reference

- 1 T. Kajita, K. Yamori, Y. Tanaka, "Relation between call blocking probability and willingness to pay in content download services," 2006 IEICE Technical Report, NS2006-37, pp.5-8, June 2006, in Japanese.

- 2 K. Robertson, K. Miller, M. White, and A. Tweedly, "StarBurst multicast file transfer protocol (MFTP) specification," Internet Draft, IETF, 1998.
- 3 B.M. Waxman, "Routing of multipoint connections," *IEEE J. Sel. Areas Commun.*, vol.6, no.9, pp 1617-1622, Dec. 1988.
- 4 K. Fujimoto, S. Ata, and M. Murata, "Statistical analysis of packet delays in the internet and its application to playout control for streaming applications," *IEICE Trans. Commun.*, vol.E84-B, no.6, pp.1504-1512, June 2001.
- 5 T. Ogura, J. Suzuki, A. Chugo, M. Katoh, and T. Aoyama, "Stability evaluation of a dynamic traffic engineering method in a large-scale network," *IEICE Trans. Commun.*, Special Issue on Internet Technology III, vol. E86-B, no.2, pp.518-525, Feb. 2003.
- 6 S. Shioda, T. Yagi, and K. Mase, "A new approach to the bottleneck bandwidth measurement for an end-to-end network path," *IEEE International Conference on Communications (ICC 2005)*, Vol.1, pp.59-64, Seoul, Korea, May, 2005.
- 7 R. Kawahara, K. Ishibashi, T. Hirano, H. Saito, H. Ohara, D. Satoh, S. Asano, and J. Matsukata, "Traffic measurement and analysis in an ATM-based internet backbone," *Comput. Commun.*, vol.24, Issues 15-16, pp.1508-1524, Oct. 2001.
- 8 T. Kajita, E. Takahashi, K. Yamori, and Y. Tanaka, "Optimal delivery interval and call admission control of multicast content delivery based on users' utility," *International Network Optimization Conference (INOC 2005)*, Session TB4, pp.B2.434-B2.441, Lisbon, Portugal, March 2005.