A Methodology to Estimate the Unit Price of User Contribution in P2P Streaming System - A Case Study

Junghyun Kim[1], Sean Moon[2], Gunhee Kim[1], Heejung Kim[1], and Sungchoon Park[1]

[1]Broadcasting Technical Research Institute, Korean Broadcasting System, Seoul, Korea
[2]Peering Portal Inc., Seoul, Korea
E-mail: brilliant@kbs.co.kr

ABSTRACT

Peer-to-Peer content delivery technology recently begins to be used not only for file sharing applications such as eDonkey but also for every kind of content services. KBS, a broadcasting company in Korea, is aggressively driving to apply Peer-to-Peer technology to KBS's commercial internet video service but we found that there are two big huddles. First, end users may refuse to share their own resources for KBS's cost reduction using Peer-to-Peer content delivery technology. Second, it may cause that the number of free-riders increases and the efficiency of the overall system would fall. From commercial service provider's perspective, we have to avoid that end users have unfavorable impressions on the service and the usefulness of Peer-to-Peer technology decreases. In order to overcome these problems, we studied how to offer incentive to end users and how much incentive would be reasonable, and then applied the result to real service for verification.

Keywords: Peer-to-peer content delivery, user contribution, incentive, video on demand, streaming

I. INTRODUCTION

As Peer-to-Peer (P2P) technology is becoming widely used, problems on services using P2P technology due to end users' selfish behavior have been issued for a long time[1][2]. It is one of the major concerns that offering incentive to P2P users in order to overcome the problem, and many researchers studied on the subject[3][4]. However those researches are not seem to be suitable for commercial services such as KBS's internet video service because most of the researches are for problems of P2P file sharing system such as Gnutella and Napster. It is why we began our research.

In P2P file sharing systems, end users have a strong tendency to share their own resources because there are not distinct service providers who get advantages from their services using P2P and end users have concrete desire to download what they are looking for as fast as possible. On the contrary, there always should be service providers for commercial services who operate those services and achieve direct and/or indirect gains from huge number of end users. It is one of the main reasons that make end users hesitate to share their own resources with other end users. The difference between file sharing systems and commercial services from end users perspective requires more clear and definitive incentive system to draw end users' active participation for commercial services using P2P technology.

If we offer incentive to end users with currency or pseudo currency which can be used for pay service, we meet another problem: what criterion could be used to decide the amount of incentive for each end user. One of the main reasons to use P2P technology for commercial services is to offer high quality service to end users with relatively small cost, but the total account of incentive may exceed the reduced cost by P2P technology. Incentive system, from service provider's perspective, has to avoid this unprofitable situation. The amount of incentive should be big enough to attract end users, but the total amount of incentive should be small not to exceed reduced cost by P2P technology.

II. BACKGROUND

KBS, as a public broadcasting service company in Korea, is continuously requested to offer better service with higher bit rate video, and the number of end users of KBS's internet service keeps growing. These two facts are primary factors that increase network usage, in other words, expenses. Thus we had to consider P2P technology affirmatively to solve the two requirements without large extra cost. However the key issue in using P2P technology was to dodge end users' unfavorable response and efficiency falloff due to free-riders.

The simplest way that we invented was giving right to choose enabling or disabling P2P function to end users. End users who do not want to share their own resources may disable the P2P function, and they can utilize the service without sending data to other users. Meanwhile, end users who enabled P2P functions shares their resources get better service as a compensation for their resource, for example, higher speed or higher bit rate service without advertisement videos. However we did not believe that it was not a good solution because there still remained free-rider's problem among end users who enabled P2P functions. As a result, we had to find out another incentive system which offers intuitive and direct incentive proportional to the amount of user contribution for the whole system.

A. Requirements

Incentive given to end users must have sufficient value to attract them to share willingly their own resources to other end users. At the same time, the incentive system to be invented must satisfy the following four conditions.

First, it is to have fair criterion to measure the amount of user contribution in order to avoid unfair reward between end users. In order to meet this condition, incentive must be proportional to the amount of user contribution for the whole system.

Second, it is to be simple structure that uses only simple calculations to measure the amount of user contribution. If it uses complex calculations and it has complicated structure, then it is hard to apply easily and it is less cost effective.

Third, it is to use simple measuring and reward method for end users to understand easily. If end users cannot understand the logic, then they will not participate actively.

Fourth, it is to guarantee that the total amount of incentive for the whole end users does not exceed the amount of profit of service providers due to P2P technology. If not, there is no reason to use P2P technology for commercial services.

Our study does not cover the subject of the amount of incentive sufficient to attract end users because it is believed to be about psychology or economics.

B. Contribution Factors

According to the four conditions stated above, we decided that it is reasonable and logical to measure the amount of user contribution with user resources used for P2P transfer between end users because it is easy to make end users understand how the incentive system works and how to measure the amount of contribution of each end user for the whole system.

There are four user resources that may be used to measure user contribution: 1) CPU, 2) memory, 3) storage, 4) network. Among these four user resources, we thought that CPU, memory and storage are not suitable for measurement tool because there are large variations in user environments on those resources. The variations make it difficult to find out the relation between the amount of resources used by P2P technology and the amount of user contribution, and eventually, they can not measure user contribution fairly. For example, it is not easy to tell that the two users' contribution are same even when their resource usage have the same value because they have different CPUs, different size of memories and different size of hard disks. It means that we do not derive the amount of contribution from the amount of CPU, memory or storage usage by P2P technology. As a result, we had to use network usage as a criterion to measure the amount of user contribution to the system.

III. Network Resource as a Contribution Index

Network device also varies user by user, but the amount of transferred data to other end users by P2P technology reflects how much the user contributes to the system. It is because the only way for end users to contribute to the system is to transfer data to other end users. If a user has a low speed network environment, the user may transfer less data than higher speed network users during a given time period. Moreover, it is very simple and easy to understand. The more data an end user transfers to other end users, the more the end user contributes to the system. It means that we could calculate the amount of user contribution from the amount of transferred data. All that we have to do for incentive is collecting the amount of transferred data for each user.

A problem in applying this method to incentive system is how to decide the amount of incentive for the given amount of transferred data. In order to identify the relation between the amount of incentive and the amount of transferred data, we had to know the unit value of a given transferred data. In some countries, ISPs (Internet Service Providers) offer price schemes on the basis of the amount of transferred data for their customers, in other words, service providers. In this case, it is very easy for service providers to value transferred data. For example, 1GB data transfer would be 10 cents where the service provider uses ISP's internet service that costs 10 cents per GB. However, in some countries such as Korea, network cost is decided by used network bandwidth. For example, a service provider pays ISP proportional to the peak bandwidth used by the service ("peak used bandwidth") during a month. In this case, service providers have difficulties to calculate the amount of transferred data from the used bandwidth. As a result, it is very hard to identify user contribution to the system from the amount of transferred data and the unit price of user contribution. Therefore, we had to analyze the relation between peak used bandwidth and the amount of transferred data in order to presume the unit price of user contribution.

1) Network Usage Analysis

First of all, we assumed that there is relevance between the peak used bandwidth and the amount of transferred data even though there could be differences between services. In order to analyze the relation between the peak bandwidth and the transferred data, we collected log files in which used bandwidth and the amount of transferred data are recorded several commercial services. We have chosen six commercial services which consist of online game software delivery service, music on demand service and video on demand service. And then recorded used bandwidth and the amount of transferred data to log files for ten days. Table 1 shows the amount of data was transferred during a day for 1Mbps peak used bandwidth. As you see in Table 1, it is very hard to find differences between each service. It means that the

assumption that there is relevance between the peak used bandwidth and the transferred data is reasonable.

Table 1. Daily transferred data per 1Mbps bandwidth usage (GB)

Day	Service A	Service B	Service C	Service D	Service E	Service F
1	5.80	4.96	4.96	5.59	5.77	5.05
2	4.22	4.92	5.21	4.82	4.89	4.46
3	4.20	5.61	5.15	5.18	5.05	5.05
4	4.67	5.78	4.20	4.94	5.32	5.02
5	4.96	5.37	4.21	4.27	4.96	5.43
6	4.77	4.97	5.43	4.95	4.92	5.19
7	5.08	5.13	4.79	5.40	4.30	5.28
8	5.51	4.95	5.10	4.76	5.08	4.25
9	5.06	4.51	5.78	5.67	5.08	4.81
10	5.03	5.39	4.96	4.41	4.33	5.14

And then, we converted the amount of daily transferred data displayed in Table 1 into the amount of monthly transferred data because ISPs normally charge monthly and we need to know how much data transfer induced during a month by 1Mbps bandwidth usage. More precisely, we projected sixty "peak used bandwidth and the amount of daily transferred data" relation values to "peak used bandwidth and the amount of monthly transferred data" relation values. Fig. 1 shows the result.

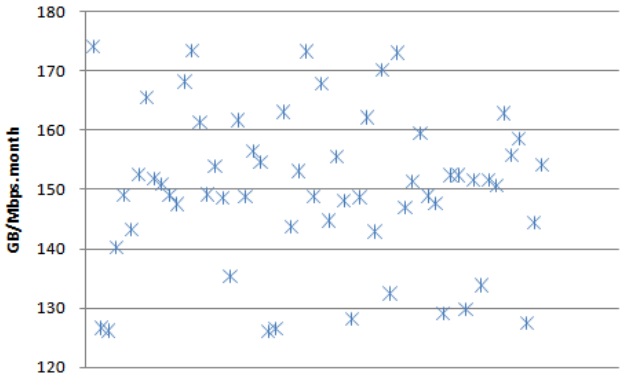

Fig. 1. Monthly peak bandwidth and the amount of transferred data

According to the central limit theorem, the sum of a sufficiently large number of identically distributed independent random variables each with finite mean and variance will be approximately normally distributed[5]. It means that the sufficiently large number of "peak used bandwidth and the amount of monthly transferred data" relation values also would be normally distributed. Using this assumption and the sixty "peak used bandwidth and the amount of monthly transferred data" values, we found that 1Mbps peak used bandwidth generates 150GB data transfer per a month in average and the standard deviation of the distribution is 12.9GB.

It is trivial that all the services do not generate exactly 150GB data transfer per 1Mbps used bandwidth each month. The exact amount of transferred data of a service may differ from that of other's, and it may vary as the time to measure the relation varies. That is why we assume that the relation is normally distributed because we had known the probability that the "peak used bandwidth and the amount of monthly transferred data" relation is right. For example, if someone assumes that 1Mbps used bandwidth will generate 150GB data transfer per a month, the probability that the amount of transferred data is less than or equal to 150GB is 50% as the following equation, a cumulative distribution function (CDF) of the normal distribution, as in (1).

(1)

When it happens that the amount of transferred data exceeds the assumed "peak used bandwidth and the amount of monthly transferred data" relation, the service provider must pay for incentive larger than the profit from P2P technology. And the above equation says that the probability of loss is 50% when a service provider assumes that 1Mbps used bandwidth generates 150GB data transfer.

There are two ways to avoid the loss caused by wrong prediction on "peak used bandwidth and the amount of monthly transferred data" relation. First, apply "1Mbps=150GB per month" relation, and then adjust rate for incentive in order to avoid loss. Second, adjust "peak used bandwidth and the amount of monthly transferred data" relation not to have loss even when incentive rate is 100%. We chose the latter one.

2) Unit price of user contribution

Everyone who knows statistics guesses easily that there should be possibility that an expected value from random events of probability distribution is wrong. For example, when we assume that "1Mbps=150GB per month" is right where the relation is normally distributed which average is 150GB, the probability that the assumption is not wrong is only 3.09% as shown in (2).

(2)

In our case, wrong prediction on the "peak used bandwidth and the amount of monthly transferred data" incurs extra expenses for incentive because the relation is used to decide unit price for incentive. As a result, we need to choose conservatively the "peak used bandwidth and the amount of monthly transferred data" relation value. It is because that

the total incentive will be less than the gain by P2P technology when the actual transferred data by P2P is less than the expected. It means that the "peak used bandwidth and the amount of monthly transferred data" relation value needs to be big sufficiently.

From the complementary CDF of the normal distribution, it is only 6.05% probability that the actual transferred data is greater than expectation where 170GB per month is used as "peak used bandwidth and the amount of monthly transferred data" relation value. The probability is expressed in (3).

(3)

Using the complementary CDF of the normal distribution, we changed the amount of transferred data from 160GB to 190GB by incrementing 10GB. The result was 22.91%, 6.05%, 1.00%, 0.10%, respectively. As the amount of transferred data increases, the probability of loss decreases. However, at the same time, the amount of incentive for end users will decrease, too. In our study, we set the amount of transferred data as 170GB in order to reduce the probability of loss and to maximize attraction for end users.

It means that we would not have loss with 93.95% probability when "1Mbps used bandwidth = 170GB per month" is used. Hence, we calculated the unit price of transferred data by P2P content delivery, in other words, the unit price of user contribution, as in (4).

$$1\text{Mbps} = 15{,}000 \text{ Won per month} \quad (4)$$

$$1\text{Mbps} = 170\text{GB per month}$$

$$170\text{GB per month} = 15{,}000 \text{ Won per month}$$

$$1\text{GB} \cong 88 \text{ Won}$$

IV. Verification using Real Service

KBS had applied P2P content delivery to a trial service for 2008 Beijing Olympic Games prior to applying to the commercial service. The trial service had two purposes as an advance preparation for commercial service using P2P content delivery. First, we had to verify the performance of P2P content delivery in the service quality and cost saving aspects. Second, we needed to have confidence on the business model with incentive system that the probability of loss is very low.

During the eighteen days for the trial service, the total number of users who played video clips was 136,182 end users, the total number of played video clips was 457,041 times, and the total time to play video clips was 3,515,695 minutes. We thought that it was not a small service even though it was trial service for a limited period, and we believed that it was large enough to be an evaluation platform.

The trial service was a video on demand service with 2Mbps bit rate, and we could reduce network bandwidth usage from 2.5Gbps to 1Gbps by P2P technology as shown in Fig. 2. Since the network price in Korea is 15 Million Won per 1Gbps, KBS could save 22.5 Million Won during the service period, as expressed in (5).

$$1500\text{Mbps} \times 15{,}000 \text{ Won} = 22.5 \text{ Million Won per month} \quad (5)$$

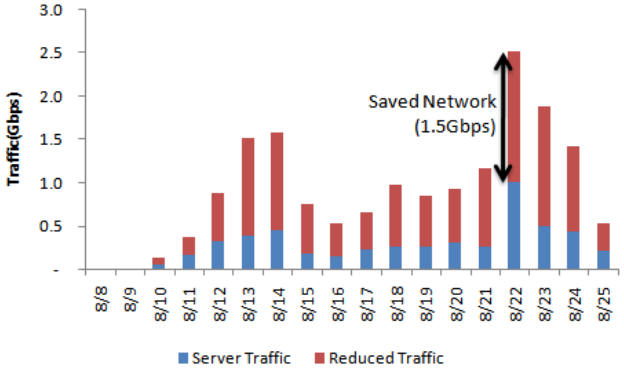

Fig. 2. Reduced traffic using P2P content delivery

We collected every information during the eighteen days that is helpful to analyze the service: the server traffic that end users receives data from server, the number of concurrent users, the number of video clips, the amount of transferred data from server and the amount of transferred data between end users. Before analyzing these data, we excluded data for the first eight days because only small number of users visited the service and the number of video clips was small. For the first eight days, the number of video clips was less than 300, and we thought that it might distort the result due to users' requests concentrated on the relatively small number of video clips. Table 2 shows the amount of transferred data by P2P for the remaining ten days.

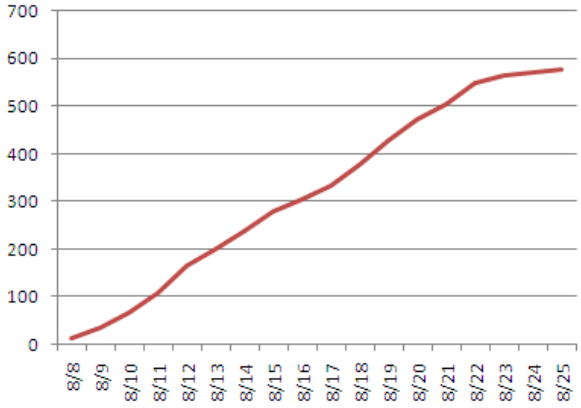

Fig. 3. The number of video clips in the trial service

Table 2. Amount of transferred data by P2P

Day	Transferred Data by P2P (GB)
1	4,246
2	5,988
3	7,206
4	7,021
5	6,950
6	7,536
7	15,959
8	12,083
9	12,868
10	4,381

Table 2 shows that the amount of transferred data by P2P for the ten days is about 84TB, and it means that it would be 252TB if we extend the service period to one month. In other words, total incentive would be compensation for the 252TB transferred data for one month. As calculated above, total incentive is 22.2 Million Won for 252 TB transferred data because 1GB is worth 88 Won. It is very close to 22.5 Million Won that is saved by P2P content delivery.

$$252{,}715\text{GB} \times 88 \cong 22.2 \text{ Million Won per month} \quad (6)$$

Fig. 4 shows comparison between saved network cost by P2P technology and the amount of total incentive to end users. This comparison means: first, it can be deduced that out analysis stated above is reasonable from the fact that saved network cost and total amount of incentive by applying network usage analysis and unit price of user contribution study stated above is similar, and second, return on investment for P2P technology can be adjusted by adjusting the percentage of incentive.

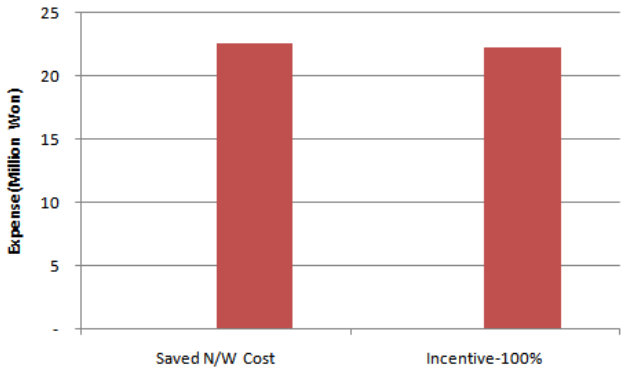

Fig. 4. Expense comparison

V. CONCLUSION

In this paper, we proposed an incentive system, that is believed to be the most simple and fair standard for user contribution, based on the amount of transferred data by P2P content delivery. We analyzed the relation between the peak used bandwidth and the amount of transferred data in order to calculate unit price of user contribution, and applied the result to a trial service for 2008 Beijing Olympic Games in order to compare the cost for incentive and the profit by P2P content delivery.

According to our analysis, 1Mbps network bandwidth used by a service corresponds to 150GB data transfer per a month in average with standard deviation of 12.9GB, and we proposed an incentive system with the assumption that the "peak used bandwidth and the amount of transferred data" relation is a normal distribution.

We assumed that 1Mbps network bandwidth usage generates 170GB data transfer from the "peak used bandwidth and the amount of transferred data" distribution in order to avoid loss caused by a probable error. With the assumption, the actual amount of transferred data will be less than or equal to 170GB with 93.95% probability, and 1GB data transfer is equivalent to about 88 Won values. We applied the unit price of user contribution to the incentive system of the trial service. As a result, we identified that the saved network cost by P2P content delivery was about 22.5 Million Won and the total incentive for end users would be 22.2 Million Won. This result reveals that the methodology used to identify the unit price of user contribution is reasonable and the incentive system is useful for commercial services.

REFERENCES

[1] E. Adar, and B. Huberman, Free Riding on Gnutella, First Monday, Vol. 5, No. 10, http://www.firstmonday.dk/issues/issue5_10/adar/

[2] S. Saroiu, PK Gummadi, and SD Gribble, A Measurement study of peer-to-peer file sharing systems, Proceedings of Multimedia Computing and Networking, 2002

[3] P Golle, K Leyton-Brown, I Mironov, and M. Lillibridge, Incentives for sharing in peer-to-peer networks, Electronic Commerce: Proceedings of the 3'rd ACM conference, 2001

[4] M. Feldman, K. Lai, I. Stoica, and J. Chuang, Robust incentive techniques for peer-to-peer networks, Proceedings of the 5'th ACM conference on Electronic commerce, 2004

[5] Rice, John, Mathematical Statistics and Data Analysis (Second ed.), 1995