Incentive Mechanism Considering Variety of User Cost in P2P Content Sharing

Kenichiro Sato, Ryo Hashimoto, Makoto Yoshino, Ryoichi Shinkuma, Tatsuro Takahashi
Communications and Computer Engineering, Graduate School of Informatics, Kyoto University
Yoshidahonnmachi, Sakyo-ku, Kyoto, 606-8501 Japan
Email: {sato, hashimoto, yoshino}@cube.kuee.kyoto-u.ac.jp, {shinkuma, ttakahashi}@i.kyoto-u.ac.jp

***Abstract*—Users in peer-to-peer (P2P) content sharing can share their content by contributing their own resources to one another. However, since there is no incentive for contributing contents or resources to others, users may attempt to obtain content without any contribution. To motivate users to contribute their resources to the service, incentive-rewarding mechanisms have been proposed. On the other hand, emerging wireless technologies, such as IEEE 802.11 wireless local area networks, beyond third generation (B3G) cellular networks and mobile WiMAX, provide high-speed Internet access for wireless users. Using these high-speed wireless access, wireless users can use P2P services and share their content with other wireless users and with fixed users. However, this diversification of access networks makes it difficult to appropriately assign rewards to each user according to their contributions. This is because the cost necessary for contribution is different in different access networks. In this paper, we propose a novel incentive-rewarding mechanism called EMOTIVER that can assign rewards to users appropriately. The proposed mechanism uses an external evaluator and interactive learning agents. We also investigate a way of appropriately controlling rewards based on the system service's quality and managing policy.**

I. Introduction

In P2P content sharing, users have to contribute to services by uploading their own content. However, such contributions are accompanied by offering their own resources such as bandwidths and storage. Therefore, users experience psychological dissatisfaction, namely the cost when they contribute resources. Because of this cost, most users attempt to obtain content without making any contributions. These kinds of users are called free riders. Incentive mechanisms that motivate users to contribute resources by rewarding them have been proposed [1]-[3] to solve this problem. In incentive mechanisms, the utility increased by rewards must outweigh the cost caused by contribution.

On the other hand, the recent developments in wireless broadband access and mobile terminals, such as IEEE 802.11 wireless local area networks, beyond third generation (B3G) cellular networks and WiMAX have led to the demand for P2P content sharing not only by fixed users but also by wireless users [4], as shown in Fig. 1, which is expected to be a killer service in the next generation mobile networks. However, since the resources of wireless terminals such as storage, bandwidths, and batteries are strictly limited, the cost of contribution by wireless users is much larger than that by fixed users [5]. Thus, as users of P2P content sharing become more diversified, the cost of contribution also becomes more diversified.

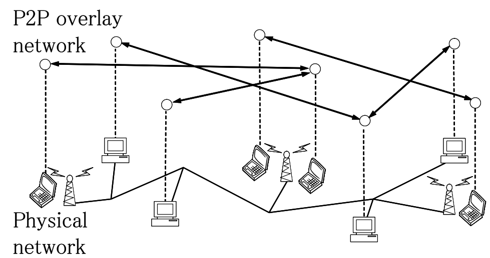


Fig. 1. P2P content sharing with wireless users.

In this situation, it is difficult to assign incentive rewards to users appropriately because the total reward that can be assigned to users is limited and the utility of each user increased by the reward must outweigh the cost caused by the contribution. However, conventional incentive mechanisms do not take into account whether the assigned reward compensates for the cost or not. In addition, they have assumed that all users have the same cost function even though, as previously explained, it varies in practical environments. In this paper, we propose a novel mechanism for incentive rewarding that can be applied to P2P content sharing by users who have varied costs and that uses an external evaluator and interactive learning agents. We also investigate a way of appropriately controlling rewards based on users' contributions and the system's service quality.

II. Conventional Incentive-rewarding Mechanisms

Several incentive mechanisms for P2P content sharing have already been proposed [1]-[3]. However, conventional mechanisms suffer from the following problems.

1) They do not take into account whether the utility increased by the reward outweighs the cost caused by the contribution or not [1][3].
2) The cost function and utility function of users are defined uniquely and the logic is only proved based on the definition. Moreover, they assume that all users have the same cost function [1][2].
3) If all users act selfishly to maximize their own benefit, there is no guarantee that the quality of system service

This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the IEEE "GLOBECOM" 2008 proceedings.

can be optimized, which has not been discussed in [1]-[3].

4) Rewards can be given in any form as long as they are equal in values. However, the conventional mechanisms limit how rewards are given, resulting in decreased scalability and versatility [2][3].

III. Proposed Incentive-rewarding Mechanism

To solve these problems, we propose a novel incentive-rewarding mechanism called EMOTIVER that can stabilize the system at high-quality levels by assigning rewards to users appropriately even when the cost of contribution differs from user to user. To achieve this goal, our mechanism requires

- The rewards assigned to users to be appropriately controlled and
- User satisfaction to be monitored, i.e., whether the cost is compensated for by the assigned reward.

The following section explains the external evaluator and the interactive learning agents used in EMOTIVER.

A. External Evaluator

The system in incentive mechanisms has to maintain service quality such as content discovery probability and variety of content at high levels by appropriately assigning rewards to users. However, it is difficult to maintain high quality in a distributed manner because of the following problems.

- A great deal of information about users and the system is needed to maintain service quality. Users have to exchange and manage the information in a distributed manner.
- Users behave selfishly to maximize their own profits without taking the quality of system services into account.

To solve these problems, we used the external evaluator proposed by Bochi et al. [6]. The role of the external evaluator is to optimize the system by effectively using rewards. More concretely, it manages how much each user is contributing and it determines how much reward should be assigned to them. Although the external evaluator is centrally operated by a service provider, users directly exchange their content on a P2P basis, which is a similar structure to KaZaA's.

Rewards in the external evaluator can be determined according to the 'effort level' and 'results' of each user. The effort level indicates how much a user is willing to contribute, while the results indicate how much a user has actually contributed. In P2P, for example, the waiting time to receive requests from other users can be the effort level, while the amount of content uploaded to other users can be the result. The external evaluator manages the effort levels and the results for all users. Based on these, it distributes rewards to users at the end of period t as follows.

1) The external evaluator determines the evaluated value, $Eval_i(t)$, for user i by considering the set of effort level $E_i(t)$ and result $R_i(t)$. That is, $Eval_i(t)$ can be expressed as a function of variables $E_i(t)$ and $R_i(t)$:

$$Eval_i(t) = f(E_i(t), R_i(t)). \quad (1)$$

2) It distributes the total reward, $W_{total}(t)$, to users in proportion to their evaluated value.

We can control users' effort levels and the quality of system services by controlling the reward-distribution-function $Eval_i(t)$.

We do not limit the way of paying rewards to users not to degrade the scalability and the versatility; rewards can be paid in any form such as cash, virtual money, content discovery probability, content download speed and so on.

B. Interactive Learning Agents

Even if we only use the external evaluator described in the previous section, all users still have to search for the effort level that will maximize their expected profits, which is obtained by subtracting the contribution cost from the reward utility. However, searching for the optimal effort level incurs additional cost for users, and there is no guarantee that they can always find this.

To solve these problems, we introduce the interactive learning agent that Lee et al. have proposed [7]. That agent first suggests that the user choose a wireless service. It then receives satisfaction/dissatisfaction for the service fed back by the user, and finally finds the service that will maximize her/his satisfaction.

When we apply this agent to our incentive mechanism, as seen in Fig. 2, agent first chooses an effort level E and suggest it to user. It then receives evaluation for the effort level fed back by the user, and finally finds the effort level that will maximize her/his profit. Users can easily find out their optimal effort level by simple interactions with their learning agents, so we assume that users obey their agents and behave according to E.

C. System Flowchart

In this paper, we defined the effort level E as 'the waiting time for requests from other users during period t', which consists of three levels, 0, 0.5 and 1. A user with effort level 1 uploads content whenever she/he receives a request. If a user selects effort level 0.5, she/he waits for requests during half of period t. Moreover, we defined the result R as 'the number of uploaded files.' The flowchart for our incentive mechanism EMOTIVER is in Fig. 3.

1) Interactive agent ag_i of user us_i chooses an effort level, $E_i(t)$, for period t from effort level table T_i, which shows all selectable effort levels, and proposes it to us_i.
2) us_i behaves according to $E_i(t)$, and produces a result, $R_i(t)$.
3) ag_i obtains $E_i(t)$ and $R_i(t)$ and reports them to the external evaluator.
4) The external evaluator determines the degree of reward for us_i, $W_i(t)$, according to the reward-distribution function.
5) us_i feeds back her/his satisfaction level $SAT_i(t)$ to ag_i based on $W_i(t)$ and cost $C_i(t)$.
6) ag_i inputs the set of $E_i(t)$ and $SAT_i(t)$ to the learning algorithm and learns us_i's preferences.

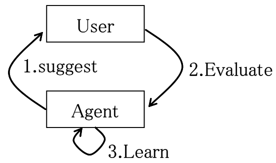


Fig. 2. Interactive learning agent.

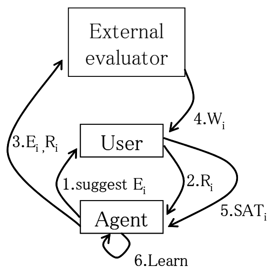


Fig. 3. Proposed incentive mechanism EMOTIVER.

7) The flow from steps 1 to 6 is repeated until the optimal effort level has been found.

The satisfaction level, $SAT_i(t)$, of us_i is given as,

$$SAT_i(t) = W_i(t) - C_i(t) \tag{2}$$

where $W_i(t)$ is the reward that us_i obtained and $C_i(t)$ is the cost that she/he incurred. $W_i(t)$ and $C_i(t)$ can be converted into willingness to pay (WTP) and willingness to accept compensation (WTA) [8]; they can be treated within the dimension of money.

IV. Proposed Reward-distribution Function

It is of utmost importance to design the reward-distribution function appropriately in our incentive mechanism. In this paper, we try to motivate users to contribute by adopting a simple evaluation function based on result or effort level, which is evaluated through computer simulations.

A. Reward Distribution Based on Result

This method distributes rewards in proportion to R^α. The reward assigned to user i, $W_i(t)$, is represented as:

$$W_i(t) = W_{total} \times \frac{\{R_i(t)\}^\alpha}{\sum_j \{R_j(t)\}^\alpha} \tag{3}$$

where α (> 0) is a control parameter. When α = 1, reward is assigned to a user in proportion to the number of files she/he uploaded, resulting in the compensation for the cost caused by the upload. By increasing α to more than 1, we can give priority to users producing better results, which becomes a stronger incentive for choosing higher effort levels. If we set α to less than 1, the difference in rewards between users with different results is reduced, which results in increasing the number of contributing users.

B. Reward Distribution Based on Effort Level

Although the reward distribution described in the previous section can probably compensate for the cost caused by upload, 'the indirect cost' can not be compensated for. For wireless users, because of the limitation of battery, connecting to the service and waiting for requests from other users cause additional cost even if they do not finally upload any file. To consider this cost, called indirect cost, in this section, we propose another way of distributing rewards, which gives rewards to users in proportion to E^α. The reward assigned to user i, $W_i(t)$, is represented as:

$$W_i(t) = W_{total} \times \frac{\{E_i(t)\}^\alpha}{\sum_j \{E_j(t)\}^\alpha}. \tag{4}$$

Since, using $\alpha = 1$, rewards are assigned in proportion to the effort level (=the waiting time to receive requests from other users) of each user, it is expected that the cost caused by the waiting time is compensated for by the given reward. When $\alpha > 1$, users can obtain large rewards by choosing the high effort level even if they did not give so much result, resulting in the increase of the effort level of every user. On the other hand, since, when $\alpha < 1$, the higher effort level is given less evaluation value, the total number of contributing users increase though the individual effort level may decrease.

V. Simulation Descriptions

The following sections discuss our evaluation of our reward distribution by computer simulation. Basically, it is impossible to compare our method with the conventional methods mentioned in Sect. II, because, as explained in the section, they do not have a mechanism that monitor satisfactions of users and distribute rewards according to the subjective user factors. However, when α is fixedly set to 1 in the result-based reward-distribution function, it can be considered as a conventional reward-distribution method [9].

A. Model

As mentioned in Sect. I, users' cost functions vary in practical environments. As the first step of our study, to simplify the simulation, we assumed that there were two types of users called Type 1 and Type 2. Type-1 users feel they have incurred cost c per upload, while Type-2 users feel they have incurred 2c. We assumed that there were 100 users for each type. Since, in our simulation model, users in each type are symmetrical, the results would not change even if the number of users was 10000. Then, while keeping the above ratio constant (= 2), we increased c to assess the influence of cost.

When a user requests content, the request is equally assigned to users who are waiting for requests from others. This assumption is reasonable because more users have more popular content; in the long term, requests are uniformly distributed to users. When a user uploads content, 1 is added to her/his result. Moreover, the user who obtained the content pays 1 as WTP, which is added to the total reward, W_{total}; total reward W_{total} increases in proportion to uploaded content. As mentioned in Sect. III-C, we can treat reward and cost within the same dimension and represent cost c as the relative value of WTP = 1. For example, c = 1 means that the user will accept the cost if 1 is paid to her/him.

We adopted this simple and versatile model because we want to apply our model to various actual P2P content sharing services in the future.

B. Learning Algorithm

Figure 4 outlines the timescale flow for the incentive rewarding and learning we did in the simulation. Every user fixes her/his effort level during each period. One event in Fig. 4 is defined as follows.

1) Each user determines whether she/he is waiting for requests or not, according to her/his effort level.
2) A user requests content.
3) An upload user is chosen randomly from users who are waiting for requests.

In this simulation, one period consists of 1000 events. At every end point of the periods, the external evaluator assigns a reward to all users and all agents learn based on $SAT_i(t)$, which is fed back from the users. $SAT_i(t)$ in Eq. (2) was used as the reinforcement signal for learning in the simulation. To update the Q value of each effort level in reinforcement learning [10], we used a 'profit sharing updating Q value based on the accumulated value'. That is, when the Q value of effort level $E_i(t)$ is represented as Q_{E_i}, Q_{E_i} is updated by

$$Q_{E_i} \leftarrow Q_{E_i} + SAT_i(t). \tag{5}$$

We used the ϵ-greedy method [10] to select the effort levels, which is proposed to users at every starting point of the periods. In, the ϵ-greedy method, the effort level having the maximum of Q value is chosen with probability $1 - \epsilon$, while one of the other effort levels is randomly chosen with probability ϵ. However, if ϵ is constant, agents randomly select effort level with the constant probability even after the optimum effort level is determined for its user. To avoid this problem, we updated ϵ by using $\epsilon = 0.3 - \frac{t}{60} \times 0.3$; since ϵ is reduced period by period, agents can finally select only one effort level, which has the maximum Q value.

Although this learning algorithm may not always be the best in practical use, the design of learning algorithms is not essential in this paper and has been left for future work.

We investigated how many users chose each effort level after learning for 60 periods in the simulation.

C. Considering Indirect Cost

Users might not only feel the cost when uploading but also when waiting for requests from other users, i.e., in waiting for requests, they have to maintain a connection to the P2P network, which causes additional cost especially in wireless users. Since this cost is not always in proportion to their results, we called it indirect cost, c_{in}. users waiting for requests incur indirect cost c_{in} uniformly for each event; when c_{in} is 0.001 and one period consists of 1000 events, the total cost incurred by users choosing effort level 1.0 is equal to 1.0. In the simulation, we set c_{in} to 0 for non indirect-cost case (Case I) and 0.008 for the large indirect-cost case (Case II). We can consider wireless users as users incurring large indirect cost because they can not permanently connect to the service because of the battery capacity.

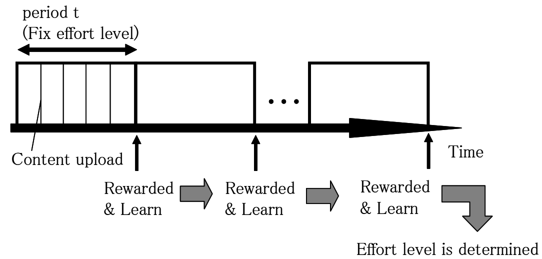


Fig. 4. Learning cycle.

D. Results

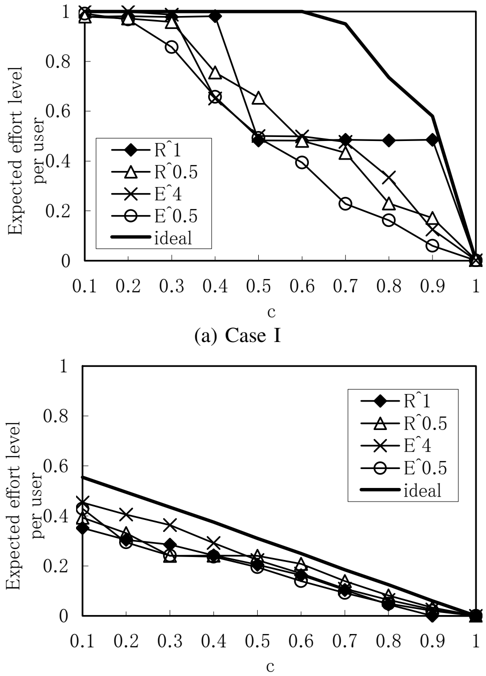
Figures 5 and 6 show the results of expected effort level and ratio of contributing users in Cases I and II. We compared R^α and E^α with $\alpha = 0.5, 1, 2, 4$ and here plotted the results of R^1, $R^{0.5}$, E^4 and $E^{0.5}$, which provided good performances.

The high expected effort level per user leads to the long existing time of users in the service, meaning that the increase in the probability of discovering a desired content. On the other hand, the ratio of contributing users means the ratio of users choosing $E > 0$. Therefore, when this value is high, we can expect the high variety of contents and the effect of balancing upload-cost. Here we ignore the lines of 'ideal'.

As we can see in Figs. 5 (a), (b), to achieve high expected effort level in Cases I, where there is no indirect cost, we should use R^1, which compensates for the cost caused by upload. However, in Case II, since the indirect cost was large, the highest effort level was given by using E^4, which can compensate for the indirect cost.

As seen in Figs. 6 (a), (b), to maximally increase the ratio of contributing users, $R^{0.5}$ should be used in every cost-case. This is because, as mentioned in Sect. IV-A, this function distributes rewards to many users by reducing the amount of reward for users who provided better results. Although we predicted that $E^{0.5}$ would achieve the same performance as $R^{0.5}$, as seen in these figures, the performance was always inferior to $R^{0.5}$. This is because $E^{0.5}$ fixedly gives reward to a user based on the effort level she/he is choosing, independently of the number of her/his uploaded files. Therefore, this function can not appropriately compensate for the cost caused by upload.

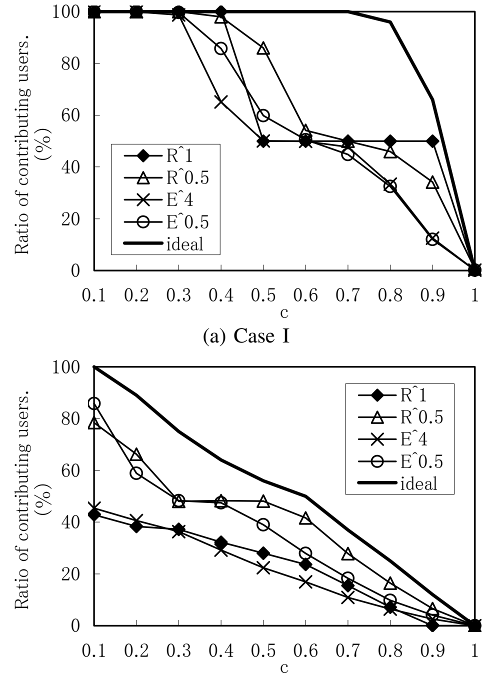
Finally, we compare the four reward-distribution functions with 'ideal,' which indicates the maximum possible values in each figure and was found from all the possible combinations of effort levels users choose. To obtain these values, we need to directly know the absolute values of cost experienced by users and solve an optimization problem for optimally assigning rewards to each user. In Figs. 5 (a) and 6 (a), we can find the obvious difference between ideal and our designed reward-distribution functions around c = 0.7. This is because, for example, though Type-1 and Type-2 users feel different costs for upload, our functions give the same amount of reward as long as they provided the same results or chose the same effort level. In this case, the reward is more than Type-1 users need to compensate for their incurred costs, while the reward is unsatisfactory to compensate for the costs incurred by Type-2 users. However, to optimize our reward-distribution



(a) Case I

(b) Case II

Fig. 5. Expected effort level per user.



(a) Case I

(b) Case II

Fig. 6. Ratio of contributing users.

function according to users who differently feel cost or to assign rewards more appropriately to them by using another complicated algorithm, we have to directly know the absolute values of costs users experienced, which is extremely difficult in practical environments. Therefore, what we can do is to adaptively use reward distribution functions according to the condition of cost and the management policy; for example, when the ratio of wireless users in a service is high, to increase effort levels chosen by users, we should use E^4 because indirect cost is dominant; if the goal is to decrease free riders (=not contributing users), we should use $R^{0.5}$.

VI. Conclusion

We proposed a novel mechanism for incentives that uses an external evaluator and interactive learning agents for P2P content sharing over heterogeneous access networks, which can distribute rewards appropriately according to the subjective user factors. Then, we designed the ways of distributing rewards based on result R (uploaded files) or effort level E (waiting time), which was evaluated through computer simulations. From the results, we concluded that, to increase average effort level of users, R^α ($\alpha \sim 1$) is basically most effective, while R^α ($\alpha \sim 0.5$) can minimize the ratio of free riders. However, when indirect cost is dominant, we should assign rewards based on E^α ($\alpha \sim 4$). Furthermore, we discuss the comparison of our designed reward-distribution functions with the ideal case. We found that, to improve the optimality, we have to differentiate users who provided the same results or chose the same effort level, which is included into future work.

Acknowledgment

This work is supported in part by KMRC R&D Grant for Mobile Wireless from Kinki Mobile Radio Center Foundation and the Research Grant from the Okawa Foundation for Information and Telecommunications (no. 07-07).

References

[1] S.M.Lui, et al., "Participation Incentive Mechanisms in Peer-to-Peer Subscription System" HICSS, pp. 3925-3931, Jan. 2002.

[2] T.L.Hong, et al., "Auction Incentive Mechanism in P2P " MUE '07, pp. 941-945, Apr. 2007.

[3] W.Tao, et al., "A Novel Incentive Mechanism for P2P Systems," Sixth International Conference on Parallel and Distributed Computing Applications and Technologies, pp.801-803, Dec. 2005.

[4] T. Hobfeld, et al., "Mapping of file-sharing onto mobile environments: Feasibility and performance of eDonkey with GPRS," WCNC, no. 1, pp. 2453-2458, New Oeleans, LA, USA, March 2005.

[5] Masato Yamada, et al., "Incentive Service Differentiation for P2P Content Sharing by Wireless Users," IEICE Trans., Commun., vol.E90-B, no.12, pp. 3561-3571, Dec.2007.

[6] Y.Bochi, et al., "A Direct-Indirect Reward Sharing Model in Multiagent Reinforcement Learning," AAMAS, 2003.

[7] G.Lee, et al., "Learning user preferences for wireless services provisioning," Autonomous Agents and Multiagent Systems Proceedings of the Third International Joint Conference on, pp.480-487, 2004.

[8] I.J. Bateman, et al., "Valuing Environmental Preferences: Theory and Practice of the Contingent Valuation Method in the US, EU, and Developing Countries," Oxford University Press, 1999.

[9] A.Habib, et al., "Incentive Mechanism for Peer-to-Peer Media Streaming," Technical report, School of Information Management and Systems, University of California, Berkeley. Dec. 2003.

[10] Ah-Hwee Tan, et al., "Self-Organizing Cognitive Agents and Reinforcement Learning in Multi-Agent Environment," IEEE/WIC/ACM, pp.351-357, 2005.