Incentive P2P Networks: A Protocol to Encourage Information Sharing and Contribution[*]

Richard T.B. Ma, C.M. Lee, John C.S. Lui, David. K.Y. Yau

1. INTRODUCTION

The rapid growth of decentralized and unstructured peer-to-peer (P2P) networks [3, 7] points to a new efficient paradigm for information exchange on the Internet. A P2P network may exhibit a power-law topology [5] such that it can propagate queries quickly and, if implemented efficiently [7], it can locate objects in $\log(N)$ time, where N is the number of nodes in the network. However, there are remaining problems with the P2P paradigm which complicates its deployment. *Free-riding* and *tragedy of the commons* are two major problems. As shown in [1], nearly 70% of Gnutella users do not share any file with the P2P community and nearly 50% of all search responses come from the top 1% of content sharing nodes. Therefore, nodes that share resources are always congested and the tragedy of the commons [2] occurs. Another problem is that many users misreport their connection types so as to discourage others from going to their sites for file download. Worse yet, there is no service differentiation between users who do not share any information and users who contribute significantly to the P2P community. The objective of this paper is to design a mechanism that provides *incentives* for users to share information and offers preferential service to users who contribute to the P2P community. In particular, we address the following questions:

1. How can we utilize file transfer resources more efficiently?
2. How can we fairly serve different nodes which may have different connection types and contributions to the P2P community?
3. How can we avoid the free-riding problem?

In this paper, we explore a scheduling policy for file transfer service such that we can provide *incentive* to the P2P nodes. The assigned transfer capacity for each request is a function of the requesting node's connection type, its utility function, and its relative contribution compared with the other requesting nodes. Our objective is to make efficient use of P2P network resources and to provide both fairness and incentive to all nodes in the P2P community.

2. INCENTIVE P2P NETWORKS

The current P2P protocols provide no incentive and service differentiation. Therefore, the free-riding problem [1] can occur wherein each node only behaves like a client most of the time [6]. The free-riding problem also leads to the *tragedy of the commons* problem wherein most file requests are directed towards a limited number of nodes who are willing to share information and provide services. The design of an incentive mechanism for P2P networks is imperative. In determining the proper transfer bandwidth allocation for requesting nodes, the decision should be based on the requesting node's type, utility function, and contribution. We first introduce the following notations for our incentive P2P network.

$\mathcal{N} =$ A set representing all nodes in the P2P system with $|\mathcal{N}| = N$.

$\Lambda =$ $[\lambda_{i,j}]_{N*N}$ wherein $\lambda_{i,j}$ represents the average file transfer request rate from node i to node j.

$\Psi =$ $(u_1, u_2, \ldots, u_N)$ wherein u_i represents the *maximal* upload bandwidth (in Mbps) of node i, where $i \in \mathcal{N}$.

$\mathcal{D} =$ $(d_1, d_2, \ldots, d_N)$ wherein d_i represents the *maximal* download bandwidth (in Mbps) of node i, where $i \in \mathcal{N}$.

$\mathcal{R}_i =$ A set representing all nodes which may request file download from node i, i.e., any node j wherein $\lambda_{j,i} > 0$.

$\Theta =$ $(\theta_1, \theta_2, \ldots, \theta_N)$: It is a vector which represents the *connection type* of all nodes in the P2P network. In particular, $\theta_i \in \Omega$ is the connection type of node i and it is a function of the declared upload bandwidth u_i and download bandwidth d_i. The set Ω represents all possible connection types in our incentive P2P network.

$C_i(t) =$ represents the *cumulative contribution* of node i at time t wherein $C_i(t) \in \{\mathbb{R}^+ \bigcup 0\}$.

$x_i(t) =$ represents the bandwidth allocated to node i when i requests a file transfer. The bandwidth assignment is based on our incentive mechanism.

$U_i(\theta_i, x_i) =$ a non-negative function which represents the utility of node i when it declares its connection type to be θ_i and receives a file transfer service rate of x_i. The utility function takes on a concave, bounded, and normalized form. The utility function is:

$$U_i(\theta_i, x_i) = U_i(d_i, x_i) = \begin{cases} \log(\frac{x_i}{d_i} + 1) & \text{if } x_i \leq d_i \\ \log 2 & \text{if } x_i > d_i. \end{cases}$$

---

[*] Richard T.B.Ma, C.M. Lee and John C.S. Lui are from the Computer Science & Engineering Department, the Chinese University of Hong Kong. David K.Y. Yau is from the Computer Science Department, Purdue University.

Each node in the system, say i, will have a cumulative contribution indicator $C_i(t)$ at time t. The indicator $C_i(t)$ will increase if node i provides service to the community (e.g., by transferring files to other requesting nodes), or it will decrease if node i requests some service from the community (e.g., requesting file transfers from other nodes). Let us first state some *important properties* of our mechanism.

(1) Conservation of the cumulative contribution and social utilities: The aggregate contribution of all nodes at any time $t > 0$ equals the aggregate cumulative utility of all nodes from time 0 to t. Formally,

$$\sum_{i=1}^{N} C_i(t) \;=\; \sum_{i=1}^{N} \int_0^t U_i(\theta_i, x_i(\tau))d\tau \qquad \forall t > 0. \quad (1)$$

The implication of this property is that the contribution by any node in the P2P network via the file transfer service is translated to utilities in the P2P community.

(2) Maximize social fairness during resource allocation: Given a node k and all requesting clients in $\mathcal{R}_k$, if $C_i(t) = C_j(t)$ for all $i, j \in \mathcal{R}_k$, our mechanism will distribute the upload resources of node k in such a way that the aggregate utility of all requesting nodes is maximized. Formally, we have

$$x(t) = \arg\max\left(\sum_{i\in\mathcal{R}_k} U_i(\theta_i, x_i(t))\right) \;\; s.t. \;\; \sum_{i\in\mathcal{R}_k} x_i(t) \le u_k.$$

The implication of this property is that our incentive P2P system maximizes the *social welfare* of the community.

(3) Incentive resource distribution: We provide incentive to rational users by assigning different transfer bandwidths to different requesting nodes. Given a node k and all requesting clients in $\mathcal{R}_k$, we have two cases:

- *No Congestion:* If the aggregate download bandwidth at time t of all requesting nodes is less than or equal to the upload bandwidth of node k, then all nodes in $\mathcal{R}_k$ will receive a transfer bandwidth equal to their respective download bandwidth and they will have equal utility. Formally, if $\sum_{i\in\mathcal{R}_k} d_i \le u_k$, then

  $$\begin{aligned} x_i(t) &= d_i \\ U_i(\theta_i, x_i(t)) &= U_j(\theta_j, x_j(t)) \qquad \forall i, j \in \mathcal{R}_k. \end{aligned}$$

  The implication of this property is that whenever node k has sufficient resources, all requesting nodes are equally happy.

- *Congestion:* When there is congestion for node k (i.e., $\sum_{i\in\mathcal{R}_k} d_i > u_k$), resource distribution will be a function of the contribution and download bandwidth of all the requesting nodes. Formally, for any two requesting nodes $i, j \in \mathcal{R}_k$, if the ratio of contribution to download bandwidth of i is greater than or equal to that of j, node k will distribute the transfer bandwidth resource such that the utility of node i is greater than or equal to that of node j. I.e.,

  $$\frac{C_i(t)}{d_i} \ge \frac{C_j(t)}{d_j} \quad \rightarrow \quad U_i(\theta_i, x_i(t)) \ge U_j(\theta_j, x_j(t)).$$

  The implication of this property is that our incentive system will provide a higher utility for those nodes who have a higher contribution per unit data request.

3. MECHANISM TO PROVIDE INCENTIVE

In this section, we briefly describe our mechanism. For the formal proof of the claim properties and the implementation algorithm, please refer to [4]. We assume that all the nodes report their true connection type θ and their contribution whenever they request file transfer. Because of limited space, we will not address a mechanism that provides the truth revealing property. Rather, we concentrate on the mechanism for providing incentive in the P2P system.

Our mechanism aims to distribute the resource u_k by considering the social welfare of the community and, at the same time, the contributions of all requesting nodes. I.e., it tries to provide a higher utility for a requesting client having a higher contribution. It is not difficult to show that, when one tries to maximize the social welfare, it implies that for any two requesting nodes $i, j \in \mathcal{R}_k$, one needs to enforce $(y_i(t) + d_i)/(y_j(t) + d_j) = 1$. To provide incentive, we distribute the resource u_k such that

$$\frac{x_i(t) + d_i}{x_j(t) + d_j} \;=\; \frac{C_i(t)}{C_j(t)} \qquad \forall i, j \in \mathcal{R}_k. \qquad (2)$$

It is easy to observe that when $C_i(t) = C_j(t)$, this is equivalent to maximizing the social welfare. To determine the resource allocation policy, we have

$$\begin{aligned} \sum_{l\in\mathcal{R}_k} \frac{x_l(t) + d_l}{x_i(t) + d_i} &= \sum_{l\in\mathcal{R}_k} \frac{C_l(t)}{C_i(t)} \\ \Rightarrow \frac{x_i(t) + d_i}{\sum_{l\in\mathcal{R}_k}(d_l + x_l(t))} &= \frac{C_i(t)}{\sum_{l\in\mathcal{R}_k} C_l(t)} \quad \forall\ i \in \mathcal{R}_k. \end{aligned}$$

When the resource of node k is less than the demand (i.e., $\sum_{i\in\mathcal{R}_k} d_i > u_k$), we have $\sum_{i\in\mathcal{R}_k} x_i(t) = u_k$. The incentive resource distribution is therefore:

$$x_i(t) = \frac{C_i(t)}{\sum_{l\in\mathcal{R}_k} C_l(t)}\left(u_k + \sum_{l\in\mathcal{R}_k} d_l\right) - d_i \quad \forall\ i \in \mathcal{R}_k.$$

The equation above provides a guideline of distributing transfer bandwidth among requesting nodes. Again, please refer [4] on how to update the contribution variable $C_i(t)$ for a requesting node i.

4. EXPERIMENTAL RESULTS

We use Contribution Dependent Progressive Filling (CDPF) [4] algorithm, which tries to satisfy equation (2) ,to distribute bandwidth resources.

Here, we present simulation results showing that our mechanism can provide higher aggregate utility than other scheduling disciplines like FCFS and processor-sharing. We compare the efficiency of our incentive mechanism with that of the FCFS and processor-sharing disciplines. The average file transfer request rate matrix, Λ, is randomly generated in 10,000 experiments. There are fifty servents and they can make requests to each other. There are five different connection types and each servent has an equal probability of being any of the connection type. The file request rate and the file service rate are Poisson. Under the FCFS discipline, there are at most five servents receiving service at the same time. Any further requests are queued and served in FCFS order. Under the processor sharing discipline, each requesting servent gets an equal share of the available bandwidth from the provider servent. The distribution for the incentive mechanism is as described above. The probability density function for the aggregate utility under these three resource distribution algorithms are illustrated in Figure 1. The x-axis is the value of the aggregate utility and the y-axis is the frequency achieving the value of aggregate utility. The proposed incentive mechanism *always* gives a higher aggregate utility than the other algorithms.

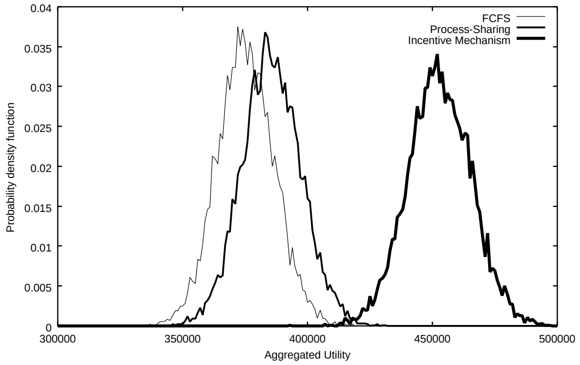

Figure 1: Probability density function for aggregate utility under FCFS, process-sharing and incentive mechanism.

5. CONCLUSION

We have presented an incentive mechanism for P2P networks. Our mechanism distributes resources among servents based on each servent's utility function, connection type, and contribution. Our mechanism achieves higher aggregate utility and fairness for a P2P network. Under our mechanism, the contribution value of a servent will be increased if it provides service to the P2P community. A servent who has a larger contribution value will receive a higher utility when it competes with other servents for file download services. Therefore, servents in the community have incentive to share information, thereby resolving the free-riding problem. Furthermore, our mechanism may decrease the contribution values of servents who access a congested resource. Therefore, it also provides incentive for servents to access information from non-congested servents and resolves the tragedy of the commons problem.

6. REFERENCES

[1] E. Adar and B. Huberman. Free riding on Gnutella. *Technical report, Xerox PARC, 10 Aug. 2000. FirstMonday*.

[2] G. Hardin. The tragedy of the commons. *Science*, 162:1243–1248, 1968.

[3] Q. Lv, P. Cao, E. Cohen, K. Li, and S. Shenker. Search and replication in unstructured peer-to-peer networks. *Proceedings of the 16th annual ACM International Conference on supercomputing (ICS) 2002; Proc. 2002 ACM SIGMETRICS conference (2-page poster). 2002.*

[4] T. B. MA, S. LEE, J. LUI, and D. YAU. Incentive for information sharing and service contribution in peer-to-peer networks: Theory and implementation. *technical report, CUHK, CSE Department, May 2003*.

[5] M. Ripeanu. Peer-to-peer architecture case study: Gnutella network. *Computer Science Dept., University of Chicago*, 2001.

[6] J. Shneidman and D. Parkes. Rationality and self-interest in peer to peer networks. *Int. Workshop on Peer-to-Peer Systems (IPTPS)*, 2003.

[7] I. Stoica, R. Morris, D. Karger, M. F. Kaashoek, and H. Balakrishnan. Chord: A scalable peer-to-peer lookup service for internet applications. *ACM SIGCOMM*, 2001.