December 2003

# Free Riding, Altruism, and Cooperation on Peer-to-Peer File-Sharing Networks

Kedar Samant
*University of Illinois at Chicago*

# Free Riding, Altruism, and Cooperation on Peer-to-Peer File-Sharing Networks

**Kedar Samant**
College of Business Administration
University of Illinois at Chicago
Chicago, IL USA
**ksaman2@uic.edu**

## Abstract

*Researchers have acknowledged the existence of* free-riding *in peer-to-peer networks. Krishnan et al. (2002a) provide a plausible game theoretic explanation for the sustenance of cooperation in P2P networks and their ability to tolerate free-riding. Our paper investigates this issue further using a computational model. We find the Krishnan et al. (2002a) model to hold true only for a restricted set of assumptions. We argue that, in general, aggregation of individuals' utility is necessary to explain the ability of P2P systems to tolerate free-riders. From our experiments we observe that the stability of the network is sensitive to the underlying incentive structure of individual users. We suggest a detailed incentive structure for users participating on the P2P network and examine this incentive structure in light of existing data of P2P usage. The findings of this paper should be useful to researchers and practitioners for policy making, network design, regulating growth, and deploying novel business models on P2P file sharing networks.*

## Introduction

The recent growth in peer-to-peer systems has raised several interesting issues for researches and practitioners alike. P2P systems function with two important underlying notions: (1) users are able to search and obtain a desired good from the network and (2) users share their goods for other users of the network.

It has been argued that digital goods on P2P networks can be treated as public goods (non-excludability in supply and non-rivalry in demand) and there exists a social dilemma where individuals either contribute to the network and share goods or shrink and free ride at the expense of other users who share (Adar and Huberman 2000; Hardin 1968; Krishnan et al. 2002b). Traditional economic analysis predicts a tragedy of the commons, where free-riding leads to overall degradation in system performance or even collapse of the system (Adar and Huberman 2000; Hardin 1968). While researchers have conclusively observed a high level of free-riding on P2P networks (Adar and Huberman 2002; Saroiu et al. 2002), surprisingly enough these P2P systems have grown steadily without any degradation of performance or collapse (Dignan 2002).

Some recent work has attempted to explain this phenomenon by assuming altruism among users, whereby users share their goods in spite of the cost involved, and derive value from other users benefiting from the download of these goods (Golle et al. 2001). However, it is not reasonable to assume a significant level of altruism among users when (1) users are participating in an *ad hoc* network of anonymous users and (2) users desire to maximize their own utility through download and consumption of goods (Axelrod 1984; Bester 1998). Golle et al. (2001) acknowledge the existence of the free-riding problem and propose a micro-payment mechanism to encourage file sharing among users in P2P systems. Their model, however, does not explain the growth in current P2P systems that operate in the absence of any such payment mechanism. Krishnan et al. (2002a) provide a game theoretic analysis of individual users' incentive to participate—download and share—without ascribing altruism among users. Their model suggests a plausible explanation for the sustenance and growth of P2P systems in the absence of any kind of altruism or micro-payment mechanism. The model suggests cooperation and sharing based on the notion of belief that, by sharing, few users obtain a positive utility and they are strictly worse-off if they decide not to share. They argue that sharing is desirable for a certain number of users, as they believe that it would increase their utility to download. Using an agent-based experimental setup, we simulated this model over a range of values of the model variables. We found the model to hold true only for a restricted set of assumptions. We argue that aggregation of individuals' utility is necessary to explain the ability of P2P systems to tolerate free-riders. The results of our experiments indicate that the stability of P2P networks is sensitive to the underlying incentive structure of individual users. Changes in the individual's incentive structure can bring about a sudden collapse of the P2P file-sharing network. We carefully analyze the individual user's behavior and the underlying incentive structure and suggest the elements of a more detailed model that forms the subject of our ongoing research. We examine this model in light of the data on existing P2P file-sharing network. Our analysis should be useful to researchers and practitioners for policy making, network design, functional improvements, regulating growth, and deploying novel business models on P2P file sharing networks.

## Agent-Based Model for P2P Networks

An individual's incentive to participate in the P2P network is based on the utility one derives from participation. We used an agent-based modeling(ABM) technique to simulate the P2P network. ABM is often used to model complex systems, consisting of heterogeneous agents with micro-level behaviors and interactions, where the overall system-level behavior emerges from the individual micro level interactions (Axtell 2000; Axtell et al. 2002; Epstein et al. 2002). Based on Krishnan et al. (2002a), we model the individual user's utility based on the value $v$ one derives by downloading an item and the cost $c$ one pays for sharing files on the network. The value is private to the user and depends on her preference and evaluation of that good. A user also carries a unique item that she can share if she decides to share. This cost is also private and will differ across all users. Sharing cost includes cost of hard-drive space, upload bandwidth consumption by other users, and fear of legal prosecution, among others.

For a user $i$ if $p$ is the probability that she finds the desired good on the network and is able to download it, then the utility is given by

$$U(i) = p * v_i - S_i c_i , \quad S_i \, \varepsilon \, [0,1])$$

In their model, Krishnan et al. (2002a) suggest the existence of a pure strategy equilibrium such that, out of $n$ original users, $k$ ($0 < k < n$) users decide to share their content, while the rest do not. This is based to two important underlying notions:

(1) The probability $p$ of finding a desired good is a concave function in $k$.

(2) There exist a certain number of users who have private value $v_i$ and private cost $c_i$ such that sharing is desirable for those users (Krishnan et al. 2002a).

Krishnan et al. (2002a) modeled the user interaction on the network as a one-step game where individual users make their decision to share or not to share based on the utility they obtain during the same time interval. In the ABM, we incorporate a *history* element whereby users consider not only their current payoff but also value their past experience on the network. Users, therefore, aggregate their utility over a period of time. This argument is reasonable as users typically participate on the P2P network for an extended a period of time and they decide to share as long as their *overall* utility from participation is positive. One must also note that participating users make the decision to download frequently, but the decision to share usually holds for extended periods.

## Experimental Setup

Based on individual's value for $c_i$ and $v_i$ there can be three categories of users: (1) $c_i > v_i$, (2) $c_i < v_i$, and (3) $c_i = v_i$. It is important to understand the dynamics and behavior of the model over a range of values for variables $v_i$, $c_i$, and aggregated utility history $H$. We modeled individual users as satisficing agents, i.e., they decide to share as long as their expected utility is positive. The P2P network model consisted of 10,000 users with a set percentage (50 percent) of users initialized to share. Variables $v_i$ and $c_i$ were varied over the range (0.01 to 1) and aggregation of utility over a period of time $H$ (1 and 2).

**Setup A:** $v_i$ = Uniform (0.0001 - 1); $c_i$ = 0.5, 0.1, 0.01; $H$= 1

In this setup, users do not aggregate their utility and decide to share based on the utility in the same time cycle. When $c_i = 0.5$, we have on average 50 percent of total users with $c_i > v_i$. When $c_i = 0.01$, 1 percent of total users have $c_i > v_i$ and the rest, 99 percent, have $v_i > c_i$. We observe that cooperation (sharing) is not sustained in the network and the network collapses (refer to Figure 1). Therefore, when users do not aggregate their utility, there is no sharing among users even when 99 percent of the users have $v_i > c_i$ (in other words, even a few users with cost of sharing greater than value of downloading can lead to network collapse).

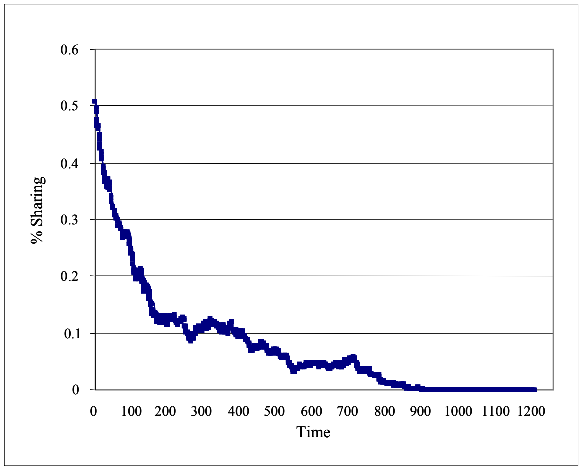


**Figure 1. Setup A with $v_i$ = Uniform (0.0001 - 1), $c_i$ = 0.01; $H$ = 1**
(Cooperation does not sustain in the network and the network collapses. Total percentage of shares decrease sharply over time and within t = 110 steps there are zero sharers from the initial 50 percent shares.)

**Setup B:** $v_i$ = Uniform (0.5 - 0.6); $c_i$ = Uniform (0.3 – 0.4); $H$ = 1

In this setup, users do not keep track of their utility history and decide to share based on the utility in the same time cycle. Here, $v_i$ is strictly greater than $c_i$ for all users. We observe that cooperation does sustain (refer to Figure 2). There is sharing among users and the network is able to tolerate free riders.

Similar results were obtained from experiments for setup A and B when varying the level of initial sharers (10 percent, 30 percent, 70 percent, and 90 percent).

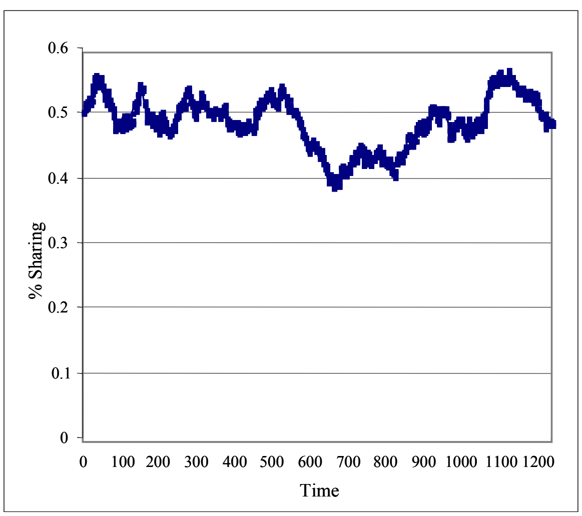


**Figure 2. Setup B with $v_i$ = Uniform (0.5 and 0.6); $c_i$ = Uniform (0.3, 0.4)**
(Cooperation sustains with total percentage of shares vary between 40 percent and 60 percent.)

**Setup C:** $v_i$ = Uniform (0.0001 - 1); $c_i$ = Uniform (0.0001 - 1); $H = 2$

In this setup, users aggregate their utility over two time cycles. Here $v_i$ and $c_i$ both vary uniformly over the same range of values, giving on average 50 percent of users having $v_i > c_i$. We observe that cooperation sustains (refer to Figure 3). There is sharing among users and the network is able to tolerate free riders.

We performed experiments for setup *C*, with varying levels of initial sharers (10 percent, 30 percent , 70 percent, and 90 percent), by varying the range of values for $v_i$ and $c_i$, and user's utility aggregated over period of time ($H > 2$). Similar results were obtained.

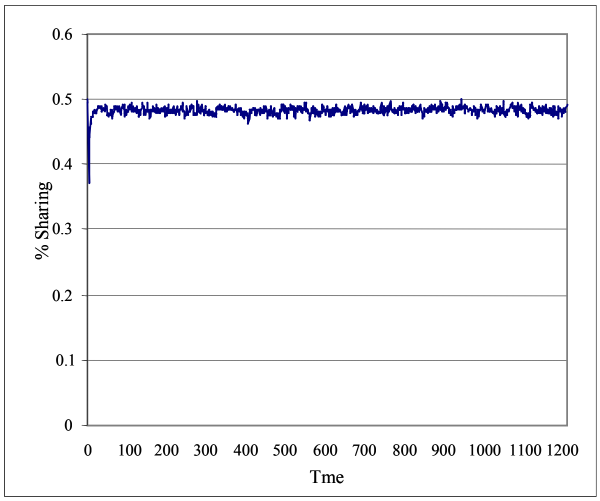


**Figure 3. Setup C with $v_i$ = Uniform (0.0001 - 1); $c_i$ = Uniform (0.0001 - 1); $H = 2$**
(Cooperation does sustain in the network. Total percentage of shares in the network stabilizes at around 48 percent.)

We observe that the model fails to stimulate cooperation when viewed as a one-shot game. A simple and intuitive scheme of adding a history element (i.e., aggregation of utilities over time) for the users is enough to sustain cooperation among users and make the P2P network tolerate free-riding.

From the experiments, we also observed that sharing in the P2P model is sensitive with respect to the values of model variables, calling for a detailed discussion on these variables. The next section suggests the elements of a more detailed model.

## A Detailed Incentive Structure

In the model above, the individual user's utility was formulated by a single equation in $v_i$ and $c_i$. Here we consider a more detailed formulation that analyzes an individual's incentive to download and share separately.

### *Download*

The value that a user derives from downloading a digital good from the network is private. However, downloading also inherits a certain cost $C_d$ for the user, i.e., like search cost, risk of getting bad quality goods. This cost is also private to the user but is influenced by the network itself (design, traffic, etc.), and the actions of all users. There also exists a fear of prosecution $F_d$ for downloading illegal goods. This cost, though private, can be influenced by environmental factors, e.g., users fear increases if they hear about a legal action against a P2P user by the RIAA (Recording Industry Association of America). The overall value one gets from searching, downloading, and consuming a digital good is a sum total of the benefits and cost associated with it. At every unit of time on the network, users decide to download as long as the total value derived from downloading is positive.

Utility $U_d$ derived by user *i* from download is

$$U_d(i) = p * v(i) - C_d(i) - F_d(i)$$

Where *p* is the probability that a user *i* finds a desired item, $C_d(i)$ is the cost of downloading the desired item , and $F_d(i)$ is the fear of legal prosecution for user *i*.

### Share

Users who share digital goods bear a cost for sharing and making the goods available for other users. The decision to share or not is denoted by *S*. This sharing cost $C_s$ includes the cost of allocating hard drive space, making upload bandwidth available for other users, and keeping the machine and the application up and running. This cost is private to the user and is not influenced by the network or any external factors. There also exists a fear of legal prosecution, $F_s$ , for sharing illegal goods. $F_s$ is private to the user, but can be influenced by environmental factors as explained above. However, sharing is also desirable. As explained by Krishnan et al. (2002a), at an overall network level, the probability of obtaining a desired digital good by an individual user increases as more users share. With this belief, which is modeled as a transform B(.) of the utility of downloading $U_d$, users assign a positive value to sharing, as it would increase their probability of finding a desired good when they download. This increases their download utility and users share as long as their overall benefit from sharing is greater than zero.

Utility $U_s$ derived by user *i* from sharing is

$$U_s(i) = S(i) ( B_i(U_d(i)) - C_s(i) - F_s(i)) \quad (S(i) \varepsilon (0,1))$$

Where $B_i(U_d(i))$ = Belief of user *i* that sharing increases downloading utility, $C_s(i)$ is the cost of sharing and $F_s(i)$ is the fear of legal action for user *i*.

The overall utility to participate is modeled as $U(i) = U_d(i) + U_s(i)$. If $U(i) < 0$, a user will not participate, i.e., will neither download not share.

## Examining the Proposed Incentive Scheme in Light of Existing Data

A measurement study of one such P2P file sharing network (Saroui et al. 2002) suggests that the users on the P2P network can be classified into three broad categories: (1) users with high downloads and high sharing, (2) users with high download and low or no sharing, and (3) users with very low or no downloads and high sharing.

Existence of type-1 users can be explained by (1) low cost and low fear of participation ($C_d$, $C_s$, $F_d$, $F_s$ ~0), and (2) high value of belief that sharing increases download utility ($B_i (U_d(i))>>0$).

Existence of type-2 users can be explained by (1) low cost and low fear of download ($C_d$, $F_d$ ~0), (2) high cost and fear of sharing ($C_s$ , $F_s$ >>0), and (3) low value of belief that sharing increases downloading utility ($B_i (U_d(i)) \sim 0$).

However, existence of type-3 users is difficult to explain based on the proposed utility structure. These users have $U_s(i) > 0$ as they share. But they have $B_i(U_d(i)) \sim 0$ as they do not intend to download. Even if we assume the extreme situation where the cost and fear of sharing for these users is near zero ($C_s \sim 0$, $F_s \sim 0$), there is no reasonable explanation for these users to share (i.e., $U_s(i) > 0$). Such users can be explained by assuming an element of altruism. Such users may include P2P network owners who have a vested interest in the success of the system, and rebels who support P2P networks as a matter of principle, not utility. We can thus have

$$U_s(i) = B_i (U_d(i)) - C_s(i) - F_s(i) + A(i)$$

where A(i) represents the degree of altruism in user *i*, $A(i) \sim 0$ for most of the users and $A(i) > 0$ for a very small proportion of users.

This model can be used to examine conditions for the sustenance and growth of P2P networks in the presence of free-riding. For example, one may argue that P2Pnetworks can sustain, grow, and at the same time tolerate free-riders, as long as the incentive structure satisfies the following :

(1) Moderate proportions of users have belief $B_i$ ($U_d$(i)) > 0
(2) Moderate proportions of users have cost of sharing $C_s \sim 0$
(3) Large proportions of users have low fear of download and share and low cost for download, $F_d$, $F_s$, $C_d \sim 0$
(4) Very small proportions of users have altruism A(i) > 0

It will be useful to examine the conditions essential for sustained cooperation. Variables $C_s$ and A(i) are private and are not affected by the environment. However, variables $B_i(U_d(i))$, $F_s$, $F_d$, and $C_d$,, although private, are influenced by the environment. Individuals $B_i(U_d(i))$ and $C_d$ can be altered by other P2P network users and P2P application. $B_i(U_d(i))$ will decrease significantly if users fail to obtain their desired product from the network for extended periods of time.

Existence of poor quality media or incorrect content on the P2P network can increase $C_d$. Increase in $C_d$ beyond a threshold value may lead to destabilization of the entire network. Record label companies are known to adopt this policy by posting bad quality music or corrupted media on the P2P network. It must be noted that individual users assign their value to $B_i(U_d(i))$ and $C_d$ by interacting with the entire network over a period of time. Thus, as the size of the P2P network increases it becomes extremely difficult for a individual user or a group of peers to significantly change the value of $B_i(U_d(i))$ and $C_d$. P2P networks can, therefore, exhibit a high level of tolerance against individuals attempt to destabilize the network by posting bad quality content.

Such is not the case with $F_d$ and $F_s$. These variables can be altered by environmental factors. For example, by airing legal actions against a few individual users, RIAA can increase the fear of participation. This can significantly decrease individual user's utility to share and download and can cause the entire network to collapse.

## Conclusion

Using Krishnan et al. (2002a), we develop an asgent-based model of P2P file sharing networks. We introduce the notion of history and find that, in the general case, aggregation of utilities is essential to sustain cooperation and tolerate free-riding on the network. This is achieved without the need for altruism on the part of the users. The model is sensitive to the underlying incentive structure of individual users.

To investigate this further, we propose a detailed incentive structure for individual users on P2P networks. Such an incentive structure can lead to a P2P network that is robust with respect to free-riding, and poor quality content. However, it may be susceptible to destabilization caused by factors like fear of prosecution. A model based on this incentive structure forms the basis of our ongoing study.

### *References*

Adar, E., and Huberman, B. A. "Free-Riding on Gnutella," *First Monday* (5:10), 2000.
Axelrod, R. *The Evolution of Cooperation*, Basic Books, New York, 1984.
Axtell, R. "Non-Cooperative Dynamics of Multi-Agent Teams," in *Proceedings of the International Conference on Autonomous Agents and Multi-Agent Systems*, Bologna, Italy, July 15-19, 2002.
Axtell, R. "Why Agents? On the Varied Motivations for Agent Computing in the Social Sciences," Working Paper No. 17, Center on Social and Economic Dynamics, Brookings Institution, November 2000.
Bester, H., and Güth, W. "Is Altruism Evolutionary Stable," *Journal of Economic Behavior and Organization* (32:2), February 1998, pp. 193-209.
Dignan, L. "Study: Kazaa, Morpheus Rave On," CNN, August 14, 2002 (available online at **http://news.com.com/2100-1023-949724.html**).
Epstein, J. D., Cummings, A. T., Chakravarty, S., Singa, R. M., and Burke, D. S. "Toward a Containment Strategy for Smallpox Bioterror: An Individual-Based Computational Approach," Working Paper No. 31, Center on Social and Economic Dynamics, Brookings Institution, December 2002.

Golle, P., Leyton-Brown, K., and Mironov, I. "Incentive for Sharing in Peer-to-Peer Networks," in *Proceedings of the 2001 ACM Conference on Electronic Commerce*, ACM Press, New York, 2001.

Hardin, G. "The Tragedy of the Commons," *Science* (168), 1968, pp. 1243-1248.

Krishnan, R., Smith, M. D., Tang, Z., and Telang, R. "The Virtual Commons: Why Free-Riding Can Be Tolerated in File Sharing Networks," in *Proceedings of the Twenty-Third International Conference on Information Systems*, L. Applegate, R. Galliers, and J. I. DeGross (eds.), Barcelona, December 2002a, pp. 807-812.

Krishnan, R., Smith, M. D., and Telang, R. "The Economics of Peer-to-Peer Networks," Working Paper, Carnegie Mellon University, Pittsburgh, PA, 2002b.

Saroui, S., Gummadi, P., and Gribble, S. "A Measurement Study of Peer-to-Peer File Sharing Systems," in *Proceedings of the Conference on Multimedia Computing and Networking*, San Jose, CA, 2002.