Over-contribution in discretionary databases

Mike Klaas
University of British Columbia
klaas@cs.ubc.ca

Abstract

Interest in the study of discretionary databases has grown as numerous large-scale public file-sharing systems have emerged as mainstream internet applications. In particular, much recent work has focused on the *free-rider* problem in peer-to-peer networks. We examine the related problem of *over-contribution* in the Usenet file-sharing network, which occurs when a surfeit of contribution causes collective harm. Using a game-theoretic model we show that, first, groups of users can self-regulate their network usage if the users are reputation-motivated, and second, this does not occur on a system-wide level. In the latter case, explicit incentive schemes are needed to motivate behavioural goals.

Introduction

Shared databases are an increasingly important means of facilitating dissemination and collection of information for groups in organizations and in the general public. A shared database is *discretionary* if users contribute to the database voluntarily. One issue that arises in real-world systems is the *over-contribution* problem, which occurs when a group of users is utilizing the database at a level beyond available resources. This issue is manifest strongly on Usenet, a distributed file-sharing system that predates peer-to-peer (P2P) networks by decades. To our knowledge, no prior study of the over-contribution problem nor of Usenet has been conducted.

Social dilemmas arise from discretionary databases when a user's personal interests are at odds with the collective good; see, eg. (Sweeney 1973; Kalman, Fulk, & Monge 2000). Perhaps the most infamous example of social dilemmas is the *free-rider* problem, which occurs when the database suffers from under-supply due to users not perceiving individual utility from contributing, choosing instead to free-ride. Several recent papers have examined this issue from a game-theoretic perspective, and have applied incentive models to resolve the dilemma (Golle *et al.* 2001; Buragohain, Agrawal, & Suri 2003). We share the view that game theory is the correct tool for incentive modelling in discretionary databases, and extend some of their models to tackle the contribution issue.

In our analysis we demonstrate that *reputation* in the Usenet community can influence a user's behaviour, and that a group of reputation-motivated users is self-regulating in the amount of content contributed. We also identify that reputation is solely a local effect in the network (i.e., one confined to small sets of newsgroups) and is thus insufficient to regulate behaviour on a global scale. Finally, we present some directions toward a solution to the latter problem.

Overview of Usenet

What started as a means of distributing textual discussion in the 1970's has become a significant means of sharing files, with over 1TB contributed to the system per day.[1] The file sharing side of Usenet is an interesting case study because it is a system that emerged without ever being designed for such use. Indeed, all infrastructure that currently exists to support file transfer arose in response to user activity rather than having inspired it.

When a file is contributed to the system, it is broken down into fixed-size parts that are called *articles*. All the articles corresponding to a file must be found by a user to retrieve the file. This is rendered difficult by the fact that the articles are only related by their subject line (which has some structure[2]) but this convention is only erratically adhered to. High volumes of content contributed to a group can significantly increase the time a user spends to locate and assemble the content he is interested in.

Usenet file transfer is similar to P2P file-sharing networks in that it is a distributed discretionary database for file transfer, but differs from them in several important ways:

- Free-riding is a major problem in P2P networks, but is not as important an issue on Usenet. Downloading does not put a critical load on the system, and heavy users pay commercial servers proportionately for their use.
- When shared, files are uploaded to the system even if they are not wanted by any consumers. In a peer-to-peer system, there is no intrinsic social cost to sharing a file that is never requested. Providing an incentive to share *desirable* content is hence crucial in the Usenet domain.
- Usenet exhibits a high level of community spirit, despite being semi-commercial. This creates a strong incentive to contribute to the system due to a greater identification

[1] As cited March 19, 2004 on `alt.binaries.news-server-comparison`.

[2] Eg. 'Julian Bream Ed - Vol11 [05/22] - yEnc "02 - Bream Ed CD11 - Paganini - Grand Sonata in A- Romanze.mp3" (12/20).'

with the system and palpable positive or negative feedback from other users.

In addition to the above, it is important to note that content contributed to Usenet is available only for a certain amount of time, and is then removed in a FIFO manner. The length of time that content is available then depends greatly on the disk space available on the server for file storage. This is another reason why large amounts of content generally detracts from a user's utility gained from the system.

Social dilemmas in Usenet

The classic social dilemma in discretionary databases is the free-rider problem. As discussed above, this problem occurs frequently in practice in real-world P2P networks (Adar & Huberman 2000). Usenet presents a theoretically similar but somewhat more dire case. Unlike some P2P systems, sharing is not default behaviour of the client, and requires tangible cost to perform. Also, whereas most P2P systems are non-commercial (which might inspire some level of altruistic "donations"), access to binary Usenet is dominated by commercial servers. Worst still, on these commercial servers, users must pay to consume and are given no credit for contributing!

This makes Usenet an intriguing case to analyze, as it does not appear to suffer from a contribution problem despite the disadvantages aforementioned. We believe the principal reason lies in Fulk's insight that users may *identify* with the collective and consequently derive some personal satisfaction in maximizing the *common good*. This effect manifests on Usenet at a much greater level than on P2P systems for several reasons. First, Usenet was originally a discussion medium, and files are intermixed with discussion in a newsgroup. This, combined with the very division of the system into newsgroups, contributes to foster a sense of community within a newsgroup which in turn contributes to the identification of a user. Additionally, users can and do provide explicit positive and negative feedback to contributors within the group.

Unfortunately, as noted in the previous section, a surfeit of contribution can potentially reduce the utility of a user. This is due to two main causes: the over-consumption of system resources (which can occur both on system-wide and local levels) and the time and complexity cost of finding and consuming content in large-volume groups.

Before analyzing this issue in detail, we present a novel game-theoretic model of discretionary databases. We believe it has sufficient flexibility to be used with a wide variety of motivational analysis problems, while not limiting the inclusion of specifics when they are important in the analysis of a particular system.

Problem statement

As in some of the work cited above, we built a game-theoretic model of the discretionary database, and used the concept of Nash equilibrium to analyze social dilemmas that occur in the database. Let $\mathcal{I} = \{a_1, a_2, \ldots, a_n\}$ be the set of agents that participate in the system. We model the actions of these agents over a fixed time interval as a Baysian game with actions, types, and rewards as defined below. Agents contribute and consume amounts of *content* divided into individual *items* (in Usenet, *items* are simply files). Deciding whether to measure contributions in terms of volume or number of items is critical to developing a realistic motivational model. We argue that for file-sharing systems, the balance lies at neither extreme.

Our game will represent one time slice of a repeated game. Thus when we speak of contribution and consumption amounts, they can be thought of as contribution and consumption *rates*, respectively.

Utility – general case

An agent's utility is composed of several factors divided into two types: utility of consumption actions (represented by the function u_i^{DN}) and utility of contributory actions (represented by the function u_i^{UP}). Total utility is the sum of these two factors $u_i = u_i^{UP} + u_i^{DN}$. We will assume risk-neutral agents throughout.

Consumption utility One of the principle motivations of this paper is to develop a model for consumption utility in a discretionary databases which better reflects reality. The general setting is as follows: Let $\mathcal{Q}$ be the set of content available for consumption. Then the valuation function $u_i^{DN}(C)$ gives the utility of agent $a_i \in \mathcal{I}$ for consuming some $C \subseteq \mathcal{Q}$.

A motivational model for consumption should consider the following factors:

- *Content Retrieved* The amount of data successfully retrieved from the system.
- *Variety* Agents are happier given a greater selection.
- *Heterogeneity* The agent is typically only interested in a subset of the available content, and has different levels of interest for different subsets of content.
- *Inherent Cost* Represents the cost of consuming data as a combined function of bandwidth and the agent's time.
- *Explicit Cost* There may be a explicit cost in terms of an internal currency or monetary dollars.

Current models typically only recognize a few of these factors. (Buragohain, Agrawal, & Suri 2003) use a model of utility linear in volume of content consumed ($u_i^{DN}(C) \propto size(C)$), augmented with a *benefit matrix* which allows differential valuation (*heterogeneity*). A benefit matrix $\mathbf{B} = \{b_{ij}\}$ is an n by n matrix with non-zero entry b_{ij} indicating a_i's interest in the content a_j provides. The expression for a_i's utility becomes $u_i^{DN}(C) = \sum_{j \neq i} b_{ij} c_j$, where c_j is a_j's contribution in C. We can drop the condition that $j \neq i$ by requiring that $\forall i$, $b_{ii} = 0$.

This model is useful in that it allows cost to be neglected (as it is rolled into the proportionality constant), and its linearity aids analysis. It does not, however, consider *variety*, and linearity is a crippling assumption: as we argue below, utility is inherently sub-linear in the size of the content.

We propose the following model for a_i's consumption utility. We assume that content is only additive across some subsets of $\mathcal{Q}$ (we call these *content classes*. Formally, let $\mathcal{Q} = \bigcup_i^m Q_i$, such that the $\{Q_i\}$ form a partition of $\mathcal{Q}$

(i.e., they are pair-wise disjoint). Then $C \subseteq \mathcal{Q}$ is defined as $C = \bigcup_i^m C_i$ with $C_i \subseteq Q_i$. Let θ_i be a sub-linear function that maps volume into utility for agent a_i, and let w_{ik} be a weight which indicates a_i's interest in the content in class Q_k. Then a_i's utility for consuming C is defined as:

$$u_i^{DN}(C) = -cost_i^{DN}(C) + \sum_{k=1}^{m} w_{ik}\theta_i(size(C_k)) \quad (1)$$

Let us consider the model in light of the requirements outlined above. First, both *inherent cost* and *explicit cost* are modelled using the $cost_i^{DN}(\cdot)$ function, which includes bandwidth, time, and monetary cost. We consider *variety* in terms of the notion of *substitutability*. Two items are (partially) substitutable when they are in the same content class. Thus, we require that the marginal utility gain in a class goes to 0 as the total consumption in the class goes to infinity. As the number of classes increases, agents have a wider selection of non-substitutable content, and this is reflected in our model. Finally, we argue that the size of an item is inherently non-linear. For instance, it is doubtful that an agent will obtain the same utility from downloading one cd-image ($\approx 2^{29}$ bytes) than a thousand users downloading an image ($\approx 2^{18}$ bytes). This view is also supported by information-theoretic arguments for diminishing-return valuations for general resources (see (Lazar & Semret 1998)), and the known importance of *downward-sloping* valuations in combinatorial auction theory and economics. We formalize the sub-linearity of θ_i as follows.

Assumption 1. *For any $a_i \in \mathcal{I}$, the **class utility** function $\theta_i : [0,\infty) \times [0,\infty)$ satisfies:*

(i) $\theta_i \in \mathcal{C}^2(0,\infty)$

(ii) $\theta_i(0) = 0$

(iii) $\frac{d\theta_i(x)}{dx} > 0$ *(θ_i is non-decreasing)*

(iv) $\frac{d^2\theta_i(x)}{dx^2} < 0$ *(θ_i' is non-increasing)*

(v) $\lim_{x\to\infty} \frac{d\theta_i(x)}{dx} = 0$ *(marginal utility goes to 0)*

Two examples of functions that satisfy these properties are $\theta_i(x) = \sqrt{x}$ and $\theta_i(x) = \log(1+x)$.

Contribution utility We identify the following factors that should be addressed by a utility model for contributions:

- *Inherent Preference for Contributing* The personal satisfaction of agent for contributing to the system, or the agent's satisfaction gained for contributing as much personal content as possible to the system.
- *Inherent Cost* Time and bandwidth contribution cost.
- *Explicit Reward* In a micro-economic system, contributions may be explicitly rewarded.
- *Reputation* The agent may be motivated by positive or negative feedback from others.

We propose the following contribution model for a_i's utility for contributing a set of content C:

$$u_i^{UP} = -cost_i^{UP}(C) + gain_i^{UP}(C) + \sum_{j\neq i}^{n} v_j(a_i) \quad (2)$$

Here $cost_i^{UP}(C)$ is the cost to contribute C, $gain_i^{UP}(C)$ is the agglomerative inherent preference to contribute and explicit reward for contributing C, and $v_j(a_i)$ is the *feedback* factor, which is a measure of how much a_j valued a_i's contribution. Note that it is equivalent from the perspective of our model whether the inherent preference for contribution is due to an altruistic concern for other agents' utility, or to some other factor.[3]

Usenet utility model

To apply the framework to Usenet, we make the following modelling decisions. The cost to contribute or consume is assumed to be linear and governed by a factor γ_i unique to each agent. That is, $cost_i(x) = \gamma_i x$. The contribution gain function is also linear in size ($gain_i^{UP}(x) = \lambda_i x$). We also assume that all agents share a common class utility function $\theta_i = \theta_j = \theta$. A crucial decision is to decide how to partition content into classes for consumption utility. We assume that each agent contributes a set of partially-substitutable content, and that there is no substitutability among contributors. This simply involves assigning each agent's contribution to its own partition. We use the benefit matrix $\mathbf{B}$ for interest weighting.

By rolling the cost function into the sum, we can rewrite equation 1 as follows to obtain a_i's utility for consuming:

$$\hat{v}_{ij}^{DN} = b_{ij}\theta(c_j) - \gamma_i c_j \quad (3)$$

$$u_i^{DN} = \sum_{j\neq i}^{n} \hat{v}_{ij}^{DN} \quad (4)$$

By separating the equation in the way we obtain an expression for $\hat{v}_{ij}^{DN}$, which is the (dis-)utility contributed to a_i by a_j. We take this is the feedback factor $v_j(a_i)$. The analysis for other possible feedback assumptions (eg., giving feedback with some probability) is similar and does not change the resulting conclusions. We obtain the following expression for a_i's contribution utility:

$$u_i^{UP} = (\lambda_i - \gamma_i)c_i + \sum_{j\neq i}^{n} \hat{v}_{ji}^{DN} \quad (5)$$

Finally, users have a limit on the content they have available to contribute. We assume a_i's contribution amount c_i is bounded by some constant k_i.

A model of agent behaviour in Usenet

In the previous section we have developed an incentive model for Usenet file-sharing. In this section we will analyze user behaviour in the model in terms of Nash equilibria. A *Nash equilibrium* of a game is a set of agent strategies in which no single agent can gain by unilateral deviation (that is, every agent is playing a *best response* strategy). They are useful in predicting what stable strategies we might expect in play.

Although we have included a strong reputation term in our contribution utility, it should be noted that the level of community spirit varies widely from group to group. Thus, we will examine populations who are motivated by peer feedback, and those who are not.

[3] Selfishness, for instance, can be manifest in several forms. One user we interviewed confided that her principal motivation for sharing content on KaZaA was to "show off" her personal music collection.

Non reputation-motivated agents

The contribution utility for these agents does not include the feedback factor. Unsurprisingly, this results in an equilibrium with no incentive to contribute responsibly (see appendix for proofs of all propositions):

Proposition 2. *For agents not motivated by reputation, if* $\forall i, \gamma_i \neq \lambda_i$, *then*

$$\forall i, c_i = \begin{cases} k_i & \text{if } \lambda_i > \gamma_i \\ 0 & \text{otherwise} \end{cases}$$

is a unique Nash equilibrium.

We exclude the case where an agent's cost is perfectly balanced with their gain, as this leads to a degenerate solution.

This result indicates that agents who have an inherent preference to contribute will tend toward contributing maximally. Thus over-contribution can only be avoided if contributing agents have small contribution limits, or only a few agents contribute.

Reputation-motivated agents

Here we consider a population of agents whose contributions are motivated only by feedback ($gain_i^{UP}(\cdot) = 0$). Our analysis in this case is much more encouraging:

Proposition 3. *For reputation-motivated agents:*

- *There exist fixed* c_i^* *such that* $\forall i, c_i = \min\{c_i^*, k_i\}$ *is a unique Nash equilibrium.*
- *Given* θ *there exists a threshold* τ *such that if*

$$\sum_{j \neq i}^{n} b_{ij} \geq \tau \sum_{k}^{n} \gamma_k \tag{6}$$

then $c_i^* > 0$. *Otherwise,* $c_i^* = 0$.

This result suggests that agents will converge on a stationary strategy where contribution depends on the benefit others derived from the content and the cost they incur from its contribution. Depending on the form of θ, there exists a threshold (again based on benefit and cost) which determines whether an agent will contribute. Together, this suggests that feedback in a group of reputation-motivated agents is sufficient to regulate behaviour of individuals, as agents have incentive to optimize social welfare. Thus over-contribution will be avoided as the feedback received is reduced as system performance degrades.

Competition for system resources

In the previous section we showed how positive and negative feedback from other users can influence the contribution rate for reputation-motivated agents. This is a form of collective regulation of individual action. Unfortunately, the way Usenet is structured, this cannot solve the global resource allocation problem which we define below.

Let us return to the user-benefit matrix $\mathbf{B}$. Some insight can be gained by examining its structure. First, in a general file-sharing system, most of the entries will be 0 because of the wide variety of content available. Second, if there is interest between two agents in one direction, it is likely that their content is topically similar, so there will be a mutual interest expressed. Thus $\mathbf{B}$ exhibits symmetries (i.e., $b_{ij} = 0$, there is a high probability that $b_{ji} = 0$, and likewise for non-zero entries). To a lesser degree, matrix entries also tend to be transitive ($b_{ij} > 0 \wedge b_{jk} > 0 \Rightarrow b_{ik} > 0$).

When these structural properties hold, we can permute the indices of agents to obtain a block-diagonal form of $\mathbf{B}$ (with the caveat that $b_{ii} = 0$):

$$\mathbf{B} = \begin{pmatrix} \blacksquare & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \blacksquare & \cdots & \mathbf{0} \\ \vdots & & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \blacksquare \end{pmatrix}$$

where boldface **0**'s denote block of zeros. This corresponds to a partitioning of the agents into groups of mutual topical interest. This effect is most apparent on Usenet, where the partitions correspond to newsgroups (or sets thereof).

This is a problem in Usenet because agents are only motivated by feedback from other agents in the same group. However, the system resources are shared by all groups. Since there is little to no feedback or collective identification between an agent and a group to which he doesn't belong, there is no incentive for groups to act non-selfishly in acquiring system resources.

Modelling the competition

We assume that each group of agents can through feedback achieve a desired collective action. Thus we consider *entire groups* as individual players in this game, and model bandwidth competition as a Baysian game.

Let κ be the amount of a system performance resource, such as the total bandwidth per time slice that can be propagated without loss. A player's utility is proportional to the expected amount downloaded.[4] Let $c_i \in [0, k_i]$ be the amount uploaded per time slice for player i, and denote by s_{-i} the strategy of all other players . If $\sum_i c_i \leq \kappa$, then an agent's utility is $u_i \propto c_i$. Otherwise, the contribution exceeds the system's capacity. Let $f_i(c_1, c_2, \ldots, c_n, \kappa)$ be a function that calculates the expected fraction of player i's content that isn't dropped. Then the expected utility for player i is:

$$\mathbb{E}(u_i) \propto \begin{cases} c_i & \text{if } \sum_j c_j \leq \kappa \\ c_i \cdot f_i(c_1, c_2, \ldots, c_n, \kappa) & \text{otherwise} \end{cases} \tag{7}$$

In Usenet, the most common way to choose f_i is to drop content with probability proportional to its size,[5] so

$$\forall i, f_i = f(c_1, c_2, \ldots, c_n, \kappa) = \frac{\kappa}{\sum_j c_j}$$

Proposition 4. *In the* n*-player bandwidth competition game,* $S = \{k_1, k_2, \ldots, k_n\}$ *is the unique Bayes-Nash equilibrium.*

This is a problem for servers: users in a group have incentive to post amounts of content beyond what they value to gain a larger slice of the global resource pie. This increases load unnecessarily and causes other performance problems.

[4] This holds in our model as long as the number of classes is proportional to c_i, i.e., variety increases with size.

[5] Implemented by dropping articles with uniform probability.

Contribution Valuation

In the previous sections we have developed a general model for utility in discretionary databases and applied it to Usenet. In this section we will discuss how these ideas could be used as components when designing new systems in general and to work toward solving the global bandwidth competition dilemma in particular.

It is immediately apparent that our utility model cannot be *directly* incorporated into a micro-economic or differential service scheme. This is due the difficulty present in eliciting agent utility parameters (the class weights being an example). Instead, we present a notion of contribution valuation that takes these factors into account indirectly by measuring the popularity, availability, and size of an item.

Value of an item

The expected value to the system of an item contributed has many potential uses. It can be used to reward the contributor, to determine contributor statistics, or to tune the performance of the system by guiding resource allocation.

There are three key values relating to a item f that will be useful. They are the item's size f_{size}, the item's availability f_{avail}, and the consumption count f_{con}. In Usenet, f_{con} is the number of times the file is downloaded and f_{avail} the number of copies present on the system. It is not uncommon for files to be re-contributed as they roll off the server; the same file is occasionally uploaded to several different groups simultaneously. We are only concerned with some finite history, so f_{con} and f_{avail} are integrals of downloads and supply over a finite-history window, respectively.

We define the value of a contributed item as follows:

$$v(f) = \int \frac{f_{con}}{f_{avail}} \cdot \theta(f_{size}) dt \quad (8)$$

This incentivizes contribution in many desirable ways. First, popular files have higher value, but only up to a certain supply level. This carries forward from our earlier model where contribution increases in the sum of the interest in the content and decreases in the social cost associated with it. Second, the first user to post a popular file will receive the highest reward, as there won't be competing copies of the item on the server. Finally, although we don't try to determine content classes explicitly, files in the same class will receive fewer downloads due to consumers' interest in variety. Thus this measure captures substitutability implicitly.

Global resource allocation

In this section, we examine some of the issues faced when tackling global resource allocation in Usenet. There are two main resources that need to be considered: bandwidth and retention, the latter being a measure of hard drive capacity to store uploaded files. Before discussing its allocation, we need a model of user utility for retention. Clearly, the difference in value of raising retention from zero to two days is significantly higher than going from ten to twelve. A sub-linear model will again be appropriate; the sharply-diminishing return of the setting suggests that a $\log$ function might be the best choice.

The problem setting is as follows. Given a limited quantity of bandwidth, limited hard disk space, and some set of contributions c_i from n groups, how can we allocate both resources in a way that maximizes our objective (whether it be social welfare, profit, incentive compatibility, etc.)? Micro-economic incentives would be difficult to apply to entire groups, and the general impracticality makes them inapplicable. Differential service schemes are much more promising. Current attempts at regulating retention focus on the latter. The most common approach is to classify groups by volume and reward small-volume group-types with longer, manually-chosen retention. This works significantly better than nothing, but has the problem of manipulation (concealing the true size of the group through its name) and the types are so broad that a very similar analysis to the above shows that resource competition still occurs within these types.

Very rarely, more sophisticated schemes are used that automatically determine retention on a per-group basis based on its volume. These methods also have serious manipulation issues–posting a single huge file in a low-volume group, for instance, can have a disastrous effect on its retention. We believe the failure of these methods is primarily due to lack of information: volume alone is an unreliable indicator. We believe that the file value measure presented in the previous section can be used to allocate resources in a more intelligent manner (in particular, to decrease manipulation possibilities).

A detailed exploration is a topic for future research, but we suggest the following simple scheme which is attractive. Assign retention based on individual files based on marginal file value ($v(f)/f_{size}$). This system rewards popular, small files, while being difficult to manipulate. The most apparent disadvantage is that implementing file-specific retention would be costly on the architecture of current servers.

Conclusion

Over-contribution, like free-riding, is a social dilemma that afflicts file-sharing discretionary databases such as Usenet. By constructing a game-theoretic incentive model of the system, we have shown that local over-contribution effects can be mitigated by group self-regulation. This depends on a sufficient community spirit being present. Previous work has identified the same potential solution to the effects of free-riding. An interesting consequence of this result is that it may be more effective for current systems to foster communication among users rather than impose regulations if the goal is to blunt the effects of undesirable dilemmas.

We have also demonstrated that this effect is insufficient for a system's population to self-govern on a system-wide level, due to the partitioning of the population into disjoint communities. This is the same problem that occurs for agents who are not reputation-motivated. In both cases incentivizing desired behaviour requires more explicit mechanisms; we suggest that a differential service scheme based on item value could achieve these ends.

Two immediate directions of future study are, first, empirical verification of the theoretical results presented, and second, examination of the effect of different levels of non-reputation-motivated agents to the local equilibrium.

Acknowledgements

I am indebted to Kevin Leyton-Brown for his invaluable assistance and advice in preparing this document.

References

Adar, E., and Huberman, B. 2000. Free riding on gnutella. *First Monday* 5(10).

Buragohain, C.; Agrawal, D.; and Suri, S. 2003. A game theoretic framework for incentives in P2P systems. In *Proc. of the Third IEEE International Conference on Peer-to-Peer Computing*.

Golle, P.; Leyton-Brown, K.; Mironov, I.; and Lillibridge, M. 2001. Incentives for sharing in peer-to-peer networks. In *Proc. of The Second workshop on Electronic Commerce (WELCOM'01)*.

Kalman, M. E.; Fulk, J.; and Monge, P. 2000. A motivational model for resolving social dilemmas in discretionary databases.

Kounavis, M. E. 1997. Auctions for resource sharing.

Lazar, A. A., and Semret, N. 1998. Design, analysis and simulation of the progressive second price auction for network bandwidth sharing. *ISDG* 8th.

Sweeney, J. 1973. An experimental investigation of the free-rider problem. *Social Science Research* 2.

Appendix: proofs of propositions

Proposition 2. *For agents not motivated by reputation, if* $\forall i, \gamma_i \neq \lambda_i$*, then*

$$\forall i, c_i = \begin{cases} k_i & \text{if } \lambda_i > \gamma_i \\ 0 & \text{otherwise} \end{cases}$$

is a unique Nash equilibrium.

Proof. Since an agent's contributory utility is simply a linear term of cost and inherent preference for contributing, it is clear that a_i will be harmed by deviating from this strategy. If the agent values contributing more than his marginal cost ($\lambda_i > \gamma_i$), then u_i^{UP} strictly increases in c_i, thus a_i should play k_i. Otherwise, u_i^{UP} is strictly negative for $c_i > 0$, so a_i should not contribute. This is a dominant strategy, thus a unique equilibrium. □

Note that if we allow $\gamma_i = \lambda_i$, then all strategies give the same payoff, thus are all weak Nash equilibria. We do not believe that this situation would arise in practice.

Proposition 3. *For reputation-motivated agents:*

- *There exist fixed* c_i^* *such that* $\forall i, c_i = \min\{c_i^*, k_i\}$ *is a unique Nash equilibrium.*
- *Given* θ *there exists a threshold* τ *such that if*

$$\sum_{j \neq i}^{n} b_{ij} \geq \tau \sum_{k}^{n} \gamma_k \tag{9}$$

then $c_i^* > 0$*. Otherwise,* $c_i^* = 0$*.*

Proof. First note that an agent's consumption utility is only affected by other agents' contribution actions, which are fixed. Thus a strategy in equilibrium will depend solely on the contribution amount c_i. If u_i^{UP} has a global unique maximum at some c_i^*, it is a best response strategy for a_i.

To search for extrema, we combine equations 3 (feedback) and 5 (definition of u_i^{UP}), differentiate, and set to zero and obtain:

$$\theta'(c_i) = \frac{\sum_k^n \gamma_k}{\sum_{j \neq i}^n b_{ji}} \tag{10}$$

This has at most one solution as $\theta'(x)$ is injective (being monotonically decreasing). Since $\theta'(x)$ is continuous on $(0, \infty)$ and goes to zero as x goes to infinity (assumption 1), $\theta'(x)$ is onto $(0, a)$ for some a (by the IVT). Thus (10) has a unique solution if

$$\frac{\sum_k^n \gamma_k}{\sum_{j \neq i}^n b_{ji}} < a \tag{11}$$

$$\sum_k^n \gamma_k < a \sum_{j \neq i}^n b_{ji} \tag{12}$$

$$\tau \sum_k^n \gamma_k \leq \sum_{j \neq i}^n b_{ji} \tag{13}$$

where τ is defined by

$$\tau = \inf\left\{ \theta'\left(\frac{1}{k}\right) \middle| k = 1, 2, \ldots \right\} \tag{14}$$

When condition 14 holds, equation 10 has a solution c_i^* on the domain of θ' (i.e., $(0, \infty)$). If $k_i < c_i^*$, choosing $c_i = k_i$ maximizes u_i^{UP}, as u_i^{UP} increases on $[0, c_i^*)$. Thus $c_i = min\{c_i^*, k_i\}$ maximizes u_i^{UP} on $[0, k_i]$.

When condition 14 does *not* hold, equation 10 has no solution on $(0, \infty)$. But u_i^{UP} is continuous on $[0, \infty)$ and tends to $-\infty$ as c_i tends to infinity, thus takes a maximum at $c_i = 0$.

This holds for all a_i, hence there exists a unique Nash equilibrium $c_i = min\{c_i^*, k_i\}$ subject to condition 9. □

Proposition 4. *In the* n*-player bandwidth competition game,* $S = \{k_1, k_2, \ldots, k_n\}$ *is the unique Bayes-Nash equilibrium.*

Proof. Consider player i's strategy, keeping s_{-i} fixed. For $c_i \leq k_i$, we have

$$\mathbb{E}(u_i) \propto \frac{c_i \cdot \kappa}{c_i + \sum_{j \neq i}(c_j)} \propto \frac{c_i}{c_i + \sum_{j \neq i}(c_j)}$$

But the partial derivative wrt c_i is

$$\frac{\partial \mathbb{E}(u_i)}{\partial c_i} \propto \frac{\sum_{j \neq i} c_j}{\left(c_i + \sum_{j \neq i} c_j\right)^2} > 0$$

so $\mathbb{E}(u_i)$ increases monotonically in c_i and has a global maximum at $c_i = k_i$ (since $c_i \in [0, k_i]$). Since this is a dominant strategy equilibrium, it is unique. □