Enforcing Efficient Resource Provisioning in Peer-to-Peer File Sharing Systems

Panayotis Antoniadis
Athens University of Economics and Business
antoniad@aueb.gr

Costas Courcoubetis
Athens University of Economics and Business
courcou@aueb.gr

ABSTRACT

In this paper we focus on content availability as the main good provided by a p2p file sharing system and consider peer availability, the amount of time peers stay connected sharing their files, as their main contribution. We wish to study the effectiveness of incentive mechanisms which enforce contribution by somehow relating it with consumption. Towards this end, we propose a mechanism who wishes to regulate the time required for peers to stay on-line sharing their files by dictating a fixed upload throughput used by uploading peers and thus a certain average time for a download to finish. We formulate and analyze a suitable economic model focusing on peer availability in order to understand the role of this important system parameter and provide the means to efficiently tune it.

1. INTRODUCTION

A fundamental difference of p2p applications compared to traditional distributed systems is the fact that decisions of the individual peers are based on their own self-interest and this in principle leads to inefficient system operation. In particular, a rational peer would wish to participate in the system without sharing any resources following the so-called 'free riding' strategy [1]. A popular approach to incentivize contribution is the employment of system rules dictating the equation of consumption and contribution per peer either through direct exchange (bilateral [7] or multilateral [2]), or by using a virtual currency (e.g. tokens), or through some other means for accounting [9]. There are also less strict approaches defining a certain relation between the level of contribution and quality of service received, often modelled as a service blocking probability [10][6].

All these approaches do not consider the private preferences of peers in terms of utility and costs nor try to achieve a certain level of resource provision. Their focus is on fairness and cooperation. Actually, detailed modelling of the economic transactions carried out in a p2p file sharing system, required in order to follow a 'welfare economics' approach, is in general a very complex task. It is thus very important to make the necessary abstractions in order to construct meaningful and tractable economic models. In our modelling work [8][5] we have chosen to focus on the issue of content availability in p2p file sharing systems[1]. Content availability is a non-rivalrous resource since files are not consumed by downloading them, and hence it has the main property of what is called a 'public good'.

We seek resource provision rules that would maximize the overall system's efficiency. That is, the total utility derived due to the content availability achieved minus the corresponding cost. The most important challenge for maximizing efficiency in this context is to acquire the private information of the participants: their preference parameter expressing their utility. However, when exclusions are possible, as suggested by recent asymptotic results (see [8] and references therein), a fixed contribution scheme, where all peers contribute the same fee (computed using a simple optimization problem) when they choose to participate, is within $O(1/n)$ from the maximum social welfare that could be achieved using Mechanism Design (and the optimal game definition) as the number of peers in the system becomes large.

But the fully distributed and untrusted p2p environment makes even such a simple incentive scheme very difficult to enforce in practice since some sort of accounting is required, which notably faces very challenging attacks due to the ability of peers to easily change their identity in most p2p systems [11]. So, the design of 'memory-less' incentive mechanisms is of great interest in this context especially if they could be configured towards improving the system's economic efficiency. A standard mechanism of this kind is the direct exchange of resources (bilateral or multilateral) which relies on the identification of peers with mutual interest on each other's services offered. We believe that less strict mechanisms that ensure that peers just contribute to the common good, requiring from a peer to contribute her resources to *any* other peer and not necessarily to the peer from which they receive service, while consuming, constitute an interesting alternative for some p2p applications which allow their enforcement. Such mechanisms exploit the public good aspect of many p2p systems and provide the means to the system designer to tune the required contribution based on certain system parameters such as its size, the distribution of peers' types, and more, towards improving the overall economic efficiency achieved. Then the enforcement mechanism, run by the serving peer, should ensure the validity of the contribution, which is the most challenging problem in this context.

We focus on a proposed memory-less mechanism for content availability which ensures that peers contribute to the system during the interval of time they are consuming resources for themselves by dictating a fixed (not too high) throughput for uploading files and introduce a suitable eco-

[1][6] follows a similar modelling approach but focuses on the game theoretic aspects of the problem and not on the efficient provision of resources through regulation as we do.

nomic model; in this model, the cost is directly related to the time a peer is forced to stay on-line and hence contribute to others. We study in depth this model and show its attractive properties deriving analytical expressions for the efficient tuning of its most critical parameter: the average download time. Moreover, our model provides the means to compare the proposed mechanism to other system rules incentivizing contribution by imposing constraints on the consumption of resources (often the only realistic alternative in p2p systems), such as the popular scheme equating downloads and uploads performed by each peer.

2. CONTRIBUTING WHILE CONSUMING

The main attributes of a peer's contribution towards content availability are: 1) the number of files shared and 2) her own availability (the time she stays on-line sharing and serving these files). We propose a memory-less incentive mechanism where the uploader of a file first checks whether a candidate *downloader* shares the amount of files required before providing a requested file. Moreover, she should ensure that these files are accessible to the rest of the group (e.g. by checking with the search mechanism for their availability). If the downloader refuses to actually upload one of his files to a potential *requestor*, the latter will inform the uploader to stop the transfer. The same would happen if the file sent by the downloader is not valid.

The second attribute of a peer's contribution is the time that she spends in the system sharing the required amount of files. Since we have assumed that peers can be forced to contribute only during the time they download files, this time is directly related to the upload throughput offered by the uploader. We thus propose to compute and dictate a certain value for the upload throughput used by all peers in the system which would be a good compromise between the benefit acquired from the increased peer availability and the corresponding cost (for having to wait more for one's downloads to finish).

Of course there are certain assumptions that should be made in order for this mechanism to be realistic. First of all, we consider uploading cost to be of limited importance, especially while downloading. Moreover, our focus on content availability (and especially on the 'long tail' of the content[2]) rather than bandwidth offered for uploads motivates us to also assume that: 1) congestion on the upload link of an individual peer is rare and 2) all files have a similar (low) rate of requests, either because they are unpopular or because they are distributed in proportion to their popularity. Finally, we assume that peers act rationally in their own self-interest and not maliciously.

In [4], we describe the main requirements of a p2p system implementing the proposed memory-less mechanism and propose some practical solutions to address the various additional incentive issues that arise in this context. The most serious weakness of applications of the proposed mechanism is likely to be attacks based on the provision of invalid content, particularly when the cost of provision of an individual piece of content is high and its probability of request is low. Newcomers may then often be able to get away with advertising invalid content for long enough to successfully complete their downloads. However, in other scenarios, where the cost of provision of new content is less high (e.g. when a user has anyway much content on her PC for her own use), then this could be a much less attractive attack. In other words, this mechanism would be best fitted for a user base like the one Direct Connect[3] used to have rather than this of BitTorrent-based applications. Interestingly, Direct Connect employed fixed contribution rules like those suggested by our public good model but enforced using central control.

Another important characteristic of our mechanism is that it relies on the existence of super peers. First, we rely on them to act as seeds for the content by providing a certain initial amount of files and by becoming the roots of the envisioned trees of downloaders which would further increase the availability of content in the system. Otherwise, our mechanism would require "cycles" of content requests to be formed resembling to a direct exchange mechanism (see [2]). Second, our super-peers are responsible for computing useful system information (such as the size of the system, the total number of files shared, etc.) and tuning important system parameters, such as the fixed upload throughput and the minimum number of files shared per peer. This is due to our effort to incorporate some sort of regulation towards improving the economic efficiency of the system as this is defined by our modelling work. Moreover, the existence of super-peers makes its implementation more realistic and helps avoid specific incentive issues that arise, discussed in [4]. Note that in practice there are often many peers with the suitable capabilities and, altruistic, incentives to play this role [12] and in any case their existence is necessary for even more fundamental system functionality such as service discovery. Thus, practical incentive mechanisms that do not make this assumption could unnecessarily restrict themselves to worse levels of efficiency than those that could be achieved otherwise —see Section 3.3.3.

3. ECONOMIC MODELLING

Suppose that peers $1, \ldots, n$ are to share the use of a public good: the expected number of distinct files made available in the system. The good can be provided at quantity Q for a rate of cost $c(Q)$. If N is the maximum number of distinct valid files, we use Q/N, the probability that a random request is satisfied, to express the content availability achieved in a system of size Q.

Peer i has a utility for the good of $\theta_i u(Q/N)$, where θ_i is a 'preference parameter' which is known only to peer i, but which is a random sample from a distribution on $[0, 1]$, with distribution function $H(\cdot)$ and density function $h(\cdot)$. The function $u(\cdot) \geq 0$ is assumed to be continuously differentiable, increasing and strictly concave in its argument (i.e. $u(x) = x^\beta$, where $\beta < 1$ is a positive constant). In order to build the public good Q, each peer has to share f_i files for a fraction t_i of time ($0 \leq t_i \leq 1$). Then, at some arbitrary time the total average number of not necessarily distinct files shared F, will be $F = \sum_{i=1}^{n}(f_i t_i)$. Due to duplication, the number of available distinct files Q will be in general a concave function of F, which when Q/N is not close to 1 and all files are equally popular, could be approximated by F, as shown in [5]. That is, in our range of parameters, $Q(F) \approx F$, and we can use Q instead of F, which is not a

[2]that large part of the set of content in which individual files are not popular, but which together constitute the majority of the total requests, and thus often generating larger value than the popular ones [3].

[3]See http://en.wikipedia.org/wiki/NeoModus_Direct_Connect.

crucial assumption for the qualitative results we obtain.

Since we do not account for uploading costs, no limitation is posed on the rate with which peers request and download files. Moreover, the number of files shared is considered a 'sunk' cost, incurred by a peer before entering the system and includes mainly the costs for acquiring (e.g. ripping a CD) and storing the content. Thus, in our analysis we have assumed that the system designer has determined beforehand a fixed and common number of files f required to be stored on each peer.

So, during system operation, the rate of cost peers have to contribute for building the public good Q is only due to the fraction of time t_i they have to stay on-line, sharing the fixed amount of f files. We assume that it is linear in t_i and the same for all peers. So, $c(Q) = \alpha \sum_{i=1}^{n} t_i$, where α converts time units to monetary units, and Q will be equal to $\sum_{i=1}^{n} t_i f$, and thus $c(Q) = \alpha Q/f$.

3.1 First-Best

Under complete information (when the payoff parameter θ_i of every peer i is known) and unlimited enforcement capabilities, the system designer should decide on the optimal amount of Q built and the fraction of time t_i that each peer i should stay on-line, solving

$$\underset{\{t_1,\ldots,t_n\},\,Q}{\text{maximize}} \sum_{i=1}^{n} \theta_i u(Q/N) - c(Q) \tag{1}$$
$$\text{s.t. } 0 \le t_i \le 1 \;\&\&\; \theta_i u(Q/N) \ge \alpha t_i,\ \forall i.$$

When the optimal size of the system Q^* is computed, it is then trivial to compute a set of feasible t_i^*s (i.e. such that $\theta_i u(Q^*/N) \ge \alpha t_i^*,\ \forall i$). This solution is called the 'first-best' provision since the corresponding efficiency level is the maximum possible.

3.2 Fixed contribution scheme

The fixed contribution scheme says that the system designer should choose $\bar{\theta}$ and Q according to the following

$$\underset{Q,\,\bar{\theta}}{\text{maximize}}\ \ nu(Q/N) \int_{\bar{\theta}}^{1} x dH - c(Q) \tag{2}$$
$$\text{s.t. } n(1 - H(\bar{\theta}))\bar{\theta} u(Q/N) = c(Q)$$

As demonstrated in [8], (2) above maximizes the expected social welfare over the choice of fixed fee policies and is within $O(1/n)$ of the second-best (the maximum possible under incomplete information). The optimal policy will correspond to the optimal values of the two variables Q and $\bar{\theta}$. Solving (2) with $c(Q) = \alpha Q/f$ and the additional constraint that $t^* \le 1$, we can compute Q^* and then the minimum contribution of each participating peer (with $\theta_i \ge \bar{\theta}$) would be $t^* = \bar{\theta} u(Q^*/N)/\alpha$.

Although the fixed contribution mechanism is in theory a very simple and attractive incentive scheme, it requires constant auditing of a peer's contribution and user memory (long-term tracking of peers' behaviour); both very difficult tasks in a realistic p2p system. And notice that the inability to incentivize peers to stay in the system longer than needed by their downloads to complete, could result to 'market failure', since in the limiting situation where the access lines of peers have infinite capacity, and hence the average download time is zero, a rational peer would set $t_i = 0$, without altering her own use of the system.

3.3 Fixed upload throughput mechanism

In our memory-less mechanism, we propose the control of the throughput b with which peers upload files to each other. We denote by d the average download time of one file, which is in general a decreasing function of b and we will thus use d as our control parameter in our analysis below and for simplicity assume that $d(b) = s/b$, where s is the average file size.[4]

Under this 'fixed upload throughput' mechanism, the contribution of peers will depend on their request rate multiplied by the probability that their requests are successful (and thus it will not be the same for everybody as in the fixed contribution scheme). We define $r(\theta)$ to be a function that maps a peer's type to its request rate. For our analysis we have chosen $r(\theta) = \theta^2$, assuming a convex relation between the type and the request rate of peers. Let x_i ($0 \le x_i \le \theta_i$) be the value to which a peer is willing to reduce her type, and hence her request rate, when facing an average download time d in a system of size Q. Then the fraction of time that peer i will be downloading in the system, say t_i, will equal the fraction of time that a $M/D/\infty$ queue has at least one customer present. Assuming that t_i is small, we have $t_i = r(x_i)(Q/N) \times d$, that is, the rate of successful requests, $r(x_i)(Q/N)$, times the average download time per successful request, d.

We describe now the iterative procedure that will converge to the limiting value of Q. Suppose that Q_0 files are made available by super-peers and Q_1 additional files are made available by the peers themselves while they are present and downloading in the system. Q denotes the total content in the system, i.e., $Q = Q_0 + Q_1$. Each peer i will choose an optimal x_i by solving the following local optimization problem.

$$\underset{x_i}{\text{maximize}} \Big\{ x_i u(Q/N) - \alpha r(x_i)(Q/N)d \Big\} \tag{3}$$
$$\text{such that } 0 \le x_i \le \theta_i \text{ and } r(x_i)(Q/N)d \le 1.$$

The cost is taken to be proportional to the fraction of time that peers are downloading (and hence to the fraction of time that they are making files available for upload). For $r(x_i) = x_i^2$, the solution will be where

$$x_i(Q) = \min\left\{\theta_i, \bar{\theta}(Q)\right\}\ ,\ \bar{\theta}(Q) = \frac{u(Q/N)}{2\alpha d(Q)/N}\,. \tag{4}$$

Note that $x_i(Q)$ is a decreasing function of Q. This fact remains true under the assumptions only that u is concave and g is convex. Assuming that while downloading, a peer makes f files available for uploading, her choice of x_i will

[4] In order for this to hold we would have to assume that once a download starts, then it will be completed (e.g. requiring from peers to wait for all uploads initiated before the end of their own download to complete as well). However, this is not straightforward to enforce in a realistic p2p system. But peers in practice make parallel requests, they don't disconnect immediately after their downloads complete, they keep searching for files while downloading, etc. What our model really requires is the ability to force peers to stay more time in the system per download and setting the upload throughput to a certain (low) value would have in general the desirable result. Thus, it is not unrealistic to make this assumption under which our model is valid.

have an effect on the number of distinct files, $Q_0 + Q_1$, that are available for others to upload. Suppose peer 1 measures the number of files in the system as $Q_0 + Q_1$ and decides to change her type from x_1 to $x_1(Q_0 + Q_1)$. Then Q_1 will change from the solution of $Q_1 = \sum_i r(x_i)((Q_0 + Q_1)/N)df$ to the solution of

$$Q_1' = \left[r(x_1(Q_0 + Q_1)) + \sum_{i \neq 1} r(x_i) \right] ((Q_0 + Q_1')/N)df.$$

And so repeating this iteratively for $k = 1, 2, \ldots$ we have $Q_1^{k+1} = \sum_{i=1}^{n} r(x_i(Q_0 + Q_1^k))((Q_0 + Q_1^k)/N)\, d\, f$. Our first observation is that this procedure will always lead to a fixed point $Q_1 = \frac{\sum_{i=1}^{n} r(x_i) Q_0\, d\, f}{N - \sum_{i=1}^{n} r(x_i)\, d\, f}$, or alternatively

$$\frac{\sum_{i=1}^{n} r(x_i)\, d\, f}{N} = \frac{Q - Q_0}{Q}\,. \tag{5}$$

But when $Q_0 = 0$ 5 becomes $\frac{\sum_{i=1}^{n} r(x_i)\, d\, f}{N} = 1$ and hence it must be that $\frac{\sum_{i=1}^{n} r(\theta_i)\, d\, f}{N} \geq 1$, since otherwise

$$Q_1^{k+1} \leq r(\theta_i)(Q_1^k/N)\, d\, f < \rho Q_1^k \leq \rho^k Q_1^{(0)}\,,$$

and so $Q_k \to 0$ as $k \to \infty$.

Standard calculations based on the information of the distribution of peer types, (4), and (5), allow us to compute the value of d that would maximize social welfare based on the information of the distribution of peer types in all cases.[5]

3.3.1 Stability

We discuss now some interesting properties of our model and more specifically the role of the parameter d. We consider the case where $Q_0 = 0$ as before, which leads to simpler expressions that are easier to analyze but also focuses on the amount of content built by peers themselves. For the computation of the social welfare achieved under different choices of our control parameter d we will assume that there is some initial content for bootstrapping purposes, which is withdrawn after peers start downloading content from each other. Then the optimal value of d is $d^* = \frac{3.06\, N}{n f}$, for $u(Q) = Q^{1/2}$. Note that due to 5 the minimum possible value of d is $d_{min} = \frac{N}{(1/n)\sum_{i=1}^{n} r(\theta_i)\, f\, n}$. So, since $\frac{1}{n}\left[\sum_{i=1}^{n} r(\theta_i)\right] \approx \frac{3\, N}{n\, f}$ for large n, d_{min} is close to d^*.

Hence, using d^* as the optimal choice of d, if for a particular realization of the θ_is, $\frac{1}{n}\sum_{i=1}^{n} r(\theta_i) < 1/3.06$ (A), then the stability condition will not hold because in this case $d_{min}(\boldsymbol{\theta}) > d^*$. But since the mean of $r(\theta_i)$ is 1/3, the probability of this event (A) goes to zero very fast as a function of n (probably exponential fast, since (A) consists a large deviation). On the other hand, if $\frac{1}{n}\sum_{i=1}^{n} r(\theta_i) > 1/3.06$, there is a value $d^*(\boldsymbol{\theta})$ for which the social welfare would be greater than this achieved for $d = d^*$. But note that both $d_{min}(\boldsymbol{\theta})$ and $d^*(\boldsymbol{\theta})$ are quantities that need full information to compute. So, $d^* = \frac{3.06\, N}{n f}$ is an approximation of $d^*(\boldsymbol{\theta})$ for large n that uses the information of the distribution of the θ_is.

We depict in Figure 1 above how social welfare is affected by our choice of d by plotting its value at equilibrium for different values of d above d_{min}, for a realization of $\boldsymbol{\theta}$ where

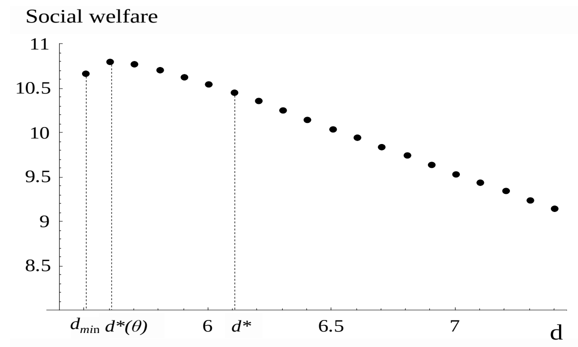

Figure 1: Social welfare as a function of d ($n = 100$, $\beta = 0.5$, $\alpha = 0.3$, $N = 10^4$, $f = 50$)

$\frac{1}{n}\sum_{i=1}^{n} r(\theta_i) > 1/3.06$. Notice that the complete information optimal value of $d^*(\boldsymbol{\theta})$ is again close to d_{min}. Moreover, as explained above, in this case where n is small, d^* is not the optimal selection for d. However, we should stress that for large n, d^* is a very good approximation of $d^*(\boldsymbol{\theta})$.

3.3.2 Evaluation

Having characterized the efficiency achieved by the first-best and the fixed contribution mechanism a natural question is how the efficiency of the fixed upload throughput mechanism is compared with them. But note that there are two differences of the model analyzed above with the fixed contribution scheme (and first-best).

First, the fact that the latter doesn't consider the possibility for peers to decide on their request rate and most importantly on their type, which was originally assumed fixed. However, under the fixed contribution scheme peers wouldn't have the incentive to reduce their type (and request rate) since the amount of time they will stay on-line (their contribution) is predefined and enforced by some external enforcement mechanism. So, if we assume that this relation of type and request rate holds in both cases and that peers have the option to decide on their type in the case of the fixed contribution scheme as well, participating peers would always choose $x_i = \theta_i$, since choosing a x_i such that $0 \leq x_i < \theta_i$, wouldn't reduce their cost; only their utility. Obviously, such an option wouldn't change the decision of excluded peers (not to participate) as well. The same holds for the case of first-best provision as well.

Notably, the second difference plays a more important role. More specifically, in the 'fixed upload throughput' mechanism we have assumed that peers request content with a rate that depends on their true type. As we discuss in the following the existence of this relation and the shape of the corresponding function $r(.)$ have a great impact on the efficiency achieved by our mechanism. Moreover, in this new game setup created we don't know which is the optimal mechanism under incomplete information (this is a very interesting topic for future work in this context) and now the fixed contribution scheme is to be treated as just an alternative incentive mechanism —and not the one that asymptotically achieves the second-best efficiency. Actually, as it turns out the fixed contribution scheme is no longer, always, the best choice. We demonstrate this fact through numerical experiments.

Figure 2 depicts the average efficiency achieved by the

[5]Due to space limitations we have omitted the details of our stability analysis which can be found at the extended version of this paper, available at http://nes.aueb.gr/p2p.html.

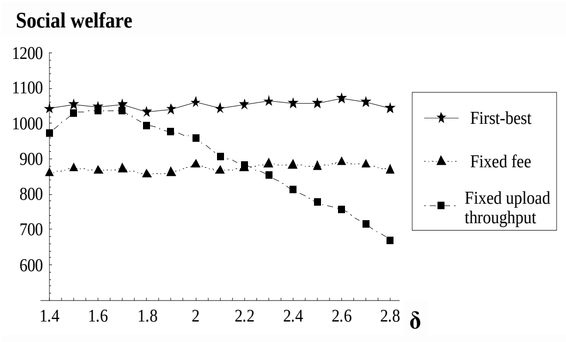

Figure 2: Evaluation of the 'fixed upload throughput' mechanism ($n = 10^4$, $\beta = 0.5$, $\alpha = 0.3$, $N = 10^6$, $f = 50$)

first-best provision, the fixed contribution scheme, and the fixed upload throughput mechanism as a function of a variable δ expressing the shape of the function $r(.)$ (i.e. $r(\theta) = \theta^{\delta}$) [6]. As one can see, the proposed mechanism achieves better efficiency than the fixed contribution scheme for $\delta < 2.3$. Notably, for a small range of values of δ (for $1.5 \leq \delta \leq 1.7$) the efficiency of the fixed upload throughput mechanism gets very close to first-best. The reason is that because of our assumption that the request rate of peers depends on their type, choosing d appropriately gets them to reveal important information concerning their actual preference parameters (by choosing the corresponding request rate assumed for the computation of d), which would be otherwise not available. Note that for large n, d^* converges to the optimal $d^*(\boldsymbol{\theta})$ and thus by using $d = d^*$ is like having complete information of peers' types. Additionally, our mechanism allows the system designer to impose a different contribution on peers depending on their actual type, and avoid exclusions without loosing the ability to acquire significant contributions from high-value peers.

However, the performance of our mechanism highly depends on the shape of this function $r(.)$, which we should stress that it is an external parameter not controlled by the system designer. And as is shown in Figure 2 there are cases where it is worse than this of the fixed contribution scheme. So, it is an interesting direction for future work the assessment of the proposed mechanism under different assumptions on the properties of the peers' request rate.

3.3.3 Upload/download ratio

In our mechanism we have defined the contribution of a peer to be the time required to stay on-line sharing a fixed amount of files independently from incoming requests. It would be however interesting to assess how it would perform in terms of actual service provisions performed (i.e. uploads) on average.

There is an interesting observation to be made towards this end. First notice that the average number of uploads a peer offers during a unity of time being connected, equals to $U = \sum_{j=1}^{n} r(x_j)(f/N)$ (the total request rate multiplied with the probability that one of her f files is requested). So, the total expected number of uploads a peer i will offer under our scheme will be equal to $r(x_i)(Q/N)\, d\, U$, i.e. the fraction of time she stays on-line (due to sucessfull content requests) multiplied by the average number of uploads offered per unity of time. But at equilibrium, when $Q_0 = 0$, from (5), $U\, d = 1$. Hence, in this case peers will actually upload as many files as downloaded on average ($r(x_i)(Q/N)$), independently from the value of d. Similarly, when $Q_0 > 0$, from (5) we have that the upload/download ratio at equilibrium will be $\frac{Q-Q_0}{Q}$.

But more interestingly, having estimated the average number of uploads per download peers offer in our system, we can now compare our scheme with the very popular mechanism enforcing the equation of downloads with uploads using virtual currencies or other means (we will refer to this scheme as the '1-1 scheme' since it dictates a 1-1 upload/download ratio for each peer). In order to do this we will assume that the number of files all peers share under this scheme is again f[7]. We also assume that the cost of a peer participating in a system with a rule dictating a 1-1 uploads/downloads ratio will be again the time she will be required to stay on-line sharing her files. Notice that this time now depends directly on the request rate of the rest of the peers, since a peer should stay on-line as long as it is needed for her to acquire the necessary 'credit' so as to satisfy her own demand. This is the most fundamental differentiation between the two mechanisms: in our mechanism we impose a certain fixed contribution cost on each download in order to maximize a certain objective (i.e. economic efficiency) while in a virtual market of uploads the only task of the system designer is to impose a certain constraint on peers' behaviour independent from their preference parameters or the size of the system.

Returning to the notation of our model, there is again a certain amount of time d, let it be denoted by d_{1-1}, that a peer will have to stay on-line sharing her files for each download, which now depends on the actual total request rate in the system. That is, $d_{1-1} = \frac{N}{\sum_{i=1}^{n} r(x_i)\, f}$. To see why, notice that d_{1-1} is the time required by a peer to stay on-line on average to offer one upload. And this is exactly the time that should be 'contributed' by each peer for each download. (We have also assumed for simplicity that without any restriction on the upload throughput the average download time is zero).

It is not easy to compute analytically the corresponding equilibrium. However, we expect that the efficiency

[6] To compute the efficiency achieved by the fixed upload throughput mechanism we again assumed a certain number of files initially available (but withdrawn after the first step of our simulation) and used for the local maximization of each peer $x_i = \min\left\{\theta_i, \left(\frac{u(Q/N)}{\delta\alpha(Q/N)d}\right)^{\frac{1}{\delta-1}}\right\}$. To avoid complicated calculations, for each δ and realization of $\boldsymbol{\theta}$, we used $d = d^*(\boldsymbol{\theta})$ (computed numerically), which as already explained, for large n is very close to the value d^* we would have computed for the specific δ. That is, we have used a close approximation of the efficiency that would have been achieved under incomplete information. Notably, as it turns out from our experiments, $d^*(\theta)$ is always close to d_{min} but their distance increases as δ grows.

[7] In reality this decision is part of peers' strategy in such a system and there is actually a 'congestion game' [13] to be played amongst peers to that respect: each peer would try to share the files that generate the more incoming requests in order to reduce the time they will have to stay on-line for acquiring the necessary credit to satisfy their demand. This would lead to some files becoming 'congested' and this would lead to different file sharing strategies, and so on.

achieved under the 1-1 scheme will be in general less than this achieved by our mechanism. The reason is that in our case we compute the optimal value of d so as to maximize efficiency, while in the case of the 1-1 rule there is no reason to expect that in equilibrium the value of d_{1-1} will be the globally optimal. The numerical experiments conducted indicate that when an equilibrium is reached, its efficiency is on the average 40% less than this of the fixed upload throughput scheme. Additionally, as also indicated by our numerical experiments, the existence of equilibrium is not guaranteed under the 1-1 scheme but depends in general on the initial conditions of the experiment (i.e. the amount of content assumed that is initially available for bootstrapping purposes).

Another weakness of the 1-1 scheme in the context of our model is that it doesn't take advantage of the number of files Q_0 possibly contributed by super-peers as does our mechanism, under which the more is Q_0 the less is the uploads/downloads ratio performed by each peer. In particular, according to (5), this ratio will be on average $\frac{Q-Q_0}{Q}$ at equilibrium. But even more importantly, the 1-1 scheme would always fail to reach an equilibrium when $Q_0 > 0$ in the context of our model. And the reason is exactly the fact that it doesn't allow for a different ratio of uploads/downloads than 1-1 as the stability condition (5), for $Q_0 > 0$, requires. Hence, in the case of the 1-1 scheme, the upload/download ratio should be ideally adapted according to Q_0 in order for the system to be stable, but this is not practical in general.

Finally, there is an additional qualitative difference between the two approaches. The 1-1 scheme requires an ex ante contribution from peers in order to acquire service. This requirement in our scheme is somehow ex post since everyone is free to consume with no restrictions and contribution comes as a consequence from consumption, which is more friendly in terms of simplicity and elasticity, and it wouldn't harm significantly the 'community spirit' inherent in many p2p applications, which actually seems to play a very important (often decisive) role for their success.

4. CONCLUSION

Mechanisms enforcing contribution while consuming constitute a very interesting class of incentive mechanisms for p2p systems since they do not require sophisticated accounting functionality to be implemented and they allow the system designer to configure the contribution of each peer towards maximizing the economic efficiency of the system avoiding at the same time the synchronization problems that arise in the case of the direct exchange of resources. We have proposed a mechanism belonging to this class for increasing content availability in p2p file sharing systems and formulated a corresponding economic model which provides useful insights for the efficient tuning of its most critical parameter: the fixed upload throughput.

But such an approach could be also followed in any p2p application which allows this enforcement of resource provisioning towards the common good while consuming. For example, one could consider the design of a 'contribute while consuming' mechanism in the case of packet forwarding in ad-hoc networks, requiring from peers to piggyback a certain number of additional packets in order for their packet to be forwarded. It is part of our on-going work to explore under which assumptions and specific applications such a mechanism would be meaningful in this context.

5. REFERENCES

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

[2] K. G. Anagnostakis and M. B. Greenwald. Exchange-based Incentive Mechanisms for Peer-to-Peer File Sharing. In *Proceedings of ICDCS04: 24th IEEE International Conference on Distributed Computing*, 2004.

[3] C. Anderson. The Long Tail. Wired Magazine, http://www.wired.com/wired/archive/12.10/tail.html, 2004.

[4] P. Antoniadis, C. Courcoubetis, and B. Strulo. Incentives for Content Availability in Memory-less Peer-to-Peer File Sharing Systems. *ACM SIGecom Exchanges*, 5(4):11–20, July 2005.

[5] P. Antoniadis, C. Courcoubetis, and R. Weber. An Asymptotically Optimal Scheme for P2P File Sharing. 2nd Workshop on Economics of Peer-to-Peer Systems, Harvard University, 2004.

[6] C. Buragohain, D. Agrawal, and S. Suri. A Game Theoretic Framework for Incentives in P2P Systems. In *Proceedings of the Third IEEE International Conference on Peer-to-Peer Computing (P2P 2003), Linkoping, Sweden*, 2003.

[7] B. Cohen. Incentives Build Robustness in BitTorrent. In *Workshop on Economics of Peer-to-Peer Systems, Berkeley, CA*, 2003.

[8] C. Courcoubetis and R. R. Weber. Incentives for large p2p systems. accepted for publication in IEEE Journal on Selected Areas in Telecommunications, available at `http://nes.aueb.gr/p2p.html`, 2005.

[9] M. Feldman, K. Lai, I. Stoica, and J. Chuang. Robust Incentive Techniques for Peer-to-Peer Networks. ACM E-Commerce Conference (EC'04), 2004.

[10] M. Feldman, C. Papadimitriou, J. Chuang, and I. Stoica. Free-Riding and Whitewashing in Peer-to-Peer Systems. In *Proceedings of 3rd Annual Workshop on Economics and Information Security (WEIS04)*, 2004.

[11] E. Friedman and P. Resnick. The Social Cost of Cheap Psedonyms. *Economics and Management Strategy*, 10(2):173–199, 2001.

[12] J. Liang, R. Kumar, and K. W. Ross. Understanding kazaa. submitted, 2004.

[13] R. W. Rosenthal. A class of games possessing pure-strategy nash equilibria. *InternationalJournal of Game Theory*, 2:65–67, 1973.