Fast Download but Eternal Seeding: The Reward and Punishment of Sharing Ratio Enforcement

Jia, AL; Rahman, R; Vinko, T; Pouwelse, JA; Epema, DHJ

DOI
10.1109/P2P.2011.6038746

Publication date
2011

Document Version
Accepted author manuscript

Published in
2011 IEEE International Conference on Peer-to-Peer Computing (P2P11)

Fast Download but Eternal Seeding: The Reward and Punishment of Sharing Ratio Enforcement

Adele L. Jia, Rameez Rahman, Tamás Vinkó, Johan A. Pouwelse and Dick H.J. Epema
Parallel and Distributed Systems Group
Department of Computer Science, Delft University of Technology, the Netherlands
Email: adele.lu.jia@gmail.com

***Abstract*—Many private BitTorrent communities employ Sharing Ratio Enforcement (SRE) schemes to incentivize users to contribute their upload resources. It has been demonstrated that communities that use SRE are greatly oversupplied, i.e., they have much higher seeder-to-leecher ratios than communities in which SRE is not employed. The first order effect of oversupply under SRE is a positive increase in the average downloading speed. However, users are forced to seed for extremely long times to maintain adequate sharing ratios to be able to start new downloads. In this paper, we propose a fluid model to study the effects of oversupply under SRE, which predicts the average downloading speed, the average seeding time, and the average upload capacity utilization for users in communities that employ SRE. We notice that the phenomenon of oversupply has two undesired negative effects: a) Peers are forced to seed for long times, even though their seeding efforts are often not very productive (in terms of low upload capacity utilization); and b) SRE discriminates against peers with low bandwidth capacities and forces them to seed for longer durations than peers with high capacities. To alleviate these problems, we propose four different strategies for SRE, which have been inspired by ideas in social sciences and economics. We evaluate these strategies through simulations. Our results indicate that these new strategies release users from needlessly long seeding durations, while also being fair towards peers with low capacities and maintaining high system-wide downloading speeds.**

I. Introduction

BitTorrent is a popular Peer-to-Peer (P2P) protocol for file distribution. A key of its success lies in its Tit-For-Tat (TFT) incentive policy, which works reasonably well in fostering cooperation among downloading peers (also known as *leechers*). However, TFT does not provide any incentive for peers to remain in the system after the download is complete, in order to *seed* the entire file to others. Therefore, peers are free to engage in “Hit and Run” behavior, the scenario under which a peer leaves immediately upon completing a download.

In recent years, there has been a proliferation of so-called *private* BitTorrent communities aimed at incentivizing seeding. These communities employ a private tracker based method that maintains centralized accounts and records the *sharing ratio* of each peer, i.e., the ratio between its total amount of upload and download. Community administrators specify some *threshold* above which all members are required to maintain their sharing ratios. This mechanism is known as *Sharing Ratio Enforcement* (SRE). Community members whose sharing ratios drop below the threshold are warned and then banned from downloading, or even expelled from the community. In this way, it is guaranteed that each peer provides a certain level of contribution to the community.

The main motivation for implementing SRE is to close the gap between bandwidth demand and supply as observed in public communities, where there is significantly more demand than supply [16]. Thus, the basic design goal of SRE is to achieve higher system-wide downloading speeds by increasing the bandwidth supply. Several measurement studies have shown that SRE is very effective in increasing supply [8], [15], [16], [21]. For instance, [16] reports seeder-to-leecher ratios that are at least 9 times higher in private communities than in public ones, while downloading speeds are measured to be 3-5 times higher. Although apparently the abundant supply of bandwidth leads to high downloading speed, in this paper we argue that oversupply also has some negative effects such as excessively long seeding times that are often unproductive. Previous work has dealt with analyzing the pros and cons of SRE schemes from a macroeconomic perspective [12], [19]. Our analysis considerably expands these works. The contributions of this paper are as follows:

1) We provide a fluid model of private BitTorrent communities and we consider two user behaviors, *lazy-seeding* (seed only the minimum amount as required by SRE) and *over-seeding* (seed more than required). We analyze the influence of these user behaviors on the performance of SRE. Besides the downloading speed, we consider additional performance metrics in our model, namely the average seeding time and the average upload capacity utilization. These new performance metrics are highly relevant to the user experience, but have not been considered in previous studies (Section III).
2) We show that while achieving high system-wide downloading speeds (the *reward* of SRE), SRE indirectly forces users to seed for extremely long times, during which their upload capacity utilizations are quite low (the *punishment* of SRE). Further, SRE discriminates against low-capacity peers and forces them to seed for longer durations than high-capacity peers (Section V).
3) We propose new strategies for SRE that alleviate the long and unproductive seeding, while still maintaining a reasonable system-wide downloading speed. Furthermore, we show that these strategies also reduce the

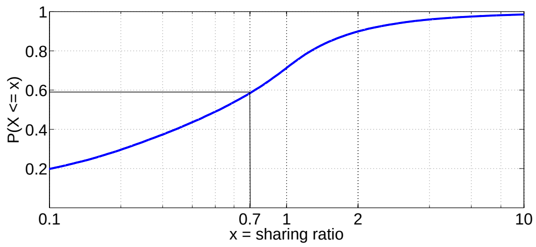


Fig. 1. Over-seeding behavior: the CDF of the sharing ratios of peers in BitSoup.org.

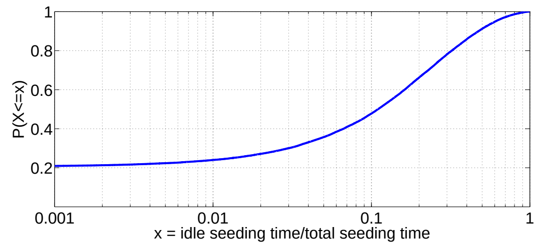


Fig. 2. Unproductive seeding: The CDF of the fraction of idle seeding time of peers with sharing ratios smaller than 1 in BitSoup.org.

discrimination against low-capacity peers. We evaluate our strategies using extensive simulations (Sections VI and VII).

II. Support from Real World Observations

In order to support our model and analysis of SRE, in this section we present real world observations of a private BitTorrent community, BitSoup.org. In this community, users are required to maintain sharing ratios greater than 0.7. The measurement trace [7] reports the upload and download amount, as well as the seeding time of each user. In total, information on nearly 87,000 users and 13,000 torrents was obtained during a period of two months. For more details we refer the reader to [7].

Previous measurement studies have already shown the effectiveness of SRE in boosting the cooperation of users, in terms of high seeder-to-leecher ratios and long seeding times leading to high downloading speeds [16], [15], [21], [8]. In addition to these results, we have two further interesting observations as follows.

A. Existence of over-seeding behavior

Intuitively, the initial goal of users in a BitTorrent community is to download files. To achieve this, maintaining a sharing ratio close to the SRE threshold is sufficient. However, we observe that not all the users behave like this. As shown in Fig. 1, more than 40% of the users in BitSoup keep sharing ratios higher than 0.7[1] and more than 10% of the users keep them higher than 2. This phenomenon of peers seeding more than required and achieving sharing ratios that are (much) higher than the SRE threshold has also been observed in other communities (e.g., [15]).

From the above observation we abstract two user behaviors for our later analysis, *lazy-seeding* and *over-seeding*. Lazy-seeding peers seed the minimum amount required by SRE. They represent the users who are download-oriented, i.e., who only seed enough to maintain adequate sharing ratios to be able to start new downloads. On the other hand, over-seeding peers are deposit-oriented, and always maintain (much) higher sharing ratios than required. The behavior of such peers may be triggered by various motivations such as altruism, a desire to be part of the rich elite of the community, or a habit of storing sharing ratio for the future.

[1] It can be observed that a significant fraction of users have sharing ratios lower than the SRE threshold. This is because in BitSoup, as well as in other private communities, users are not immediately banned after their sharing ratios drop below the threshold: they are given a certain amount of time to increase their sharing ratios.

B. Unproductive seeding

It is clear that in order to achieve high sharing ratios, peers need to spend considerable amount of seeding time. In the case of over-seeding peers, long seeding times are to be expected. However, we observe that even many peers with considerably small sharing ratios suffer extremely long seeding times, and a significant part of their seeding time is spent idle without being able to upload anything to others. As a consequence, they have to wait for a long period until their sharing ratios are high enough to start new downloads.

Fig. 2 shows the CDF of the fraction of idle seeding time of peers with sharing ratios smaller than 1 in BitSoup. We can see that 10% of these peers spend at least half of their seeding time idle. It should be noted that Fig. 2 only shows the fraction of idle seeding time. It can be conjectured that the fraction of seeding time that is not completely idle yet still yields very low upload capacity utilization, would be much higher. We hypothesize that this problem of unproductive seeding is due to the oversupply under SRE.

III. A Simple Model for Private BitTorrent Communities

From the two real world observations described in Section II, in this section we propose a fluid model to analyze the effects of SRE as well as the influence of different user behaviors on the performance of SRE.

A. Model description

We follow a similar fluid modeling approach as in [18], [17], [14] and extend it by including SRE and considering multiple user behaviors derived from real world observations. The notation we use is shown in Table I.

We consider a private BitTorrent community in which SRE is adopted and we divide peers into N different classes according to their upload capacities. Let U_i represent the upload capacity of peers in class i; to simplify the presentation let $U_i = \gamma_i U_N$. Without loss of generality we assume that $\gamma_1 > \gamma_2 > ... > \gamma_{N-1} > \gamma_N = 1$. We further assume that the download capacity of peers is not a bottleneck.

TABLE I
NOTATION OF OUR BITTORRENT MODEL

Notation	Definition
F	the size of the file shared in a swarm.
α	the SRE threshold.
λ_i	the arrival rate of leechers in class i.
U_i	the upload capacity of a peer in class i.
$u_i(t)$	the average upload speed of a peer in class i at time t, u_i for the steady state.
$d_i(t)$	the downloading speed of a peer in class i at time t, d_i for the steady state.
$x_i(t)$	the number of leechers in class i at time t, x_i for the steady state.
$y_i(t)$	the number of lazy-seeding seeders in class i at time t, y_i for the steady state.
s_i	the number of over-seeding seeders in class i.
T_i	the average seeding time for lazy-seeding peers in class i.

As mentioned earlier we consider two user behaviors: lazy-seeding and over-seeding. To better understand the effect of SRE and the over-seeding behavior, we consider an idealized scenario of a swarm, in which there are s_i over-seeding peers with an infinite desired threshold for their sharing ratios. This implies that they stay in the swarm as seeders indefinitely. Lazy-seeding peers in class i join the swarm as leechers with an arrival rate equal to λ_i and sharing ratios equal to 0. After they finish their downloads, they calculate the sharing ratios they have achieved and, if necessary, they seed in this swarm until their sharing ratios reach the SRE threshold α. Then they leave the swarm. Throughout this paper we assume that $\alpha < 1$, which is the case for most private communities.

According to TFT, during the leeching process peers favor other peers who have recently reciprocated to them most. In this way, peers are roughly clustered according to their capacities, and peers with similar capacities have similar performance [17], [14]. Let $x_i(t)$ and $y_i(t)$ represent the number of leechers and lazy-seeding seeders in class i at a particular time t, and let T_i represent the average seeding time of lazy-seeding peers in class i, then the evolution of $x_i(t)$ and $y_i(t)$ can be described as:

$$\begin{aligned} \frac{dx_i(t)}{dt} &= \lambda_i - \frac{x_i(t)d_i(t)}{F}, \\ \frac{dy_i(t)}{dt} &= \frac{x_i(t)d_i(t)}{F} - \frac{y_i(t)}{T_i}, \qquad i = 1, 2, ..., N, \end{aligned} \tag{1}$$

where $d_i(t)$ represents the average downloading speed of peers in class i at time t. The term $x_i(t)d_i(t)/F$ specifies the rate at which leechers in class i turn into lazy-seeding seeders and $y_i(t)/T_i$ specifies the leaving rate of lazy-seeding seeders in class i. In a steady state, $dx_i(t)/dt = dy_i(t)/dt \equiv 0$. Letting x_i and y_i represent the number of leechers and lazy-seeding seeders in class i in a steady state, and d_i represent the average downloading speed of leechers in class i, from Eq. (1) we have:

$$\lambda_i = \frac{y_i}{T_i} = \frac{d_i x_i}{F}. \tag{2}$$

Depending on the peers arrival rates (λ_i) and the number of over-seeding peers (s_i), a steady state will be either one of the two cases described separately below.

B. Oversupplied

When there are a large number of seeders and a small number of leechers, and seeders cannot always fully utilize their upload capacities, we can characterize the swarm as *over-supplied*. Oversupplied is a typical phase for swarms in private BitTorrent communities. Given the abundance of seeders, it is realistic to assume that in an oversupplied swarm seeders perform most of the uploads. Further considering the *piece availability problem*[2], peers are more likely to download from seeders rather than other leechers. A previous measurement study [16] shows that in two private communities where the oversupplied situation exists, over 90% of the data comes from seeders. Accordingly in our model we assume that in an oversupplied steady state leechers do not contribute upload capacities. The condition for a swarm to be in an oversupplied steady state is:

$$\sum_i \lambda_i \Delta F < \sum_i (y_i + s_i)\gamma_i U_N \Delta, \tag{3}$$

which specifies that within a time interval Δ the total upload volume that *can* be provided by all seeders is larger than the total download volume required by the $\sum_i \lambda_i \Delta$ new leechers arriving in the same interval.

In such a steady state, once a new peer joins, seeders upload to it with their full upload capacities and its download will be finished quickly. After that, seeders will be idle and wait for the next upload opportunity. Hence, on average seeders cannot fully utilize their upload capacities, and because of the operation of TCP[3] their average upload speeds will be proportional to their upload capacities. Let u_i represent the average upload speed of a peer in class i, then we have $u_i = \gamma_i u_N$ and the total actual upload volume provided by all seeders should be equal to the total download volume required by all leechers, i.e.:

$$\sum_i \lambda_i \Delta F = \sum_i (y_i + s_i)\gamma_i u_N \Delta. \tag{4}$$

After the download is finished, a lazy-seeding peer in class i seeds for a period of length T_i until it achieves the SRE threshold, i.e., until $\alpha = T_i u_i / F$. Substituting $y_i = \lambda_i T_i$ from Eq. (2) into Eq. (4) we find

$$u_N = \frac{(1-\alpha)\sum_i \lambda_i F}{\sum_i s_i \gamma_i}. \tag{5}$$

Then,

$$T_i = \frac{\alpha F}{u_i} = \frac{\alpha F}{\gamma_i u_N} = \frac{\alpha \sum_i s_i \gamma_i}{(1-\alpha)\gamma_i \sum_i \lambda_i}. \tag{6}$$

Given Eqs. (3) and (6), we can rephrase the condition for a swam to be in an oversupplied steady state as:

[2]The piece availability problem specifies that two connected leechers may not perform actual download/upload, because they cannot find interesting pieces at each other.

[3]BitTorrent uses TCP as the transport layer protocol.

$$\frac{\sum_i s_i \gamma_i U_N}{\sum_i \lambda_i F} > 1 - \alpha. \quad (7)$$

In a private community where the SRE threshold equals α, when s_i, U_i, λ_i, and F fulfill the condition of Eq. (7), we say that the swarm is in an oversupplied steady state. With a relatively small peer arrival rate (λ_i) and a large number of over-seeding seeders (s_i), peers will experience very low upload capacity utilizations and extremely long seeding times. This situation gets even worse for peers with low capacities: the ratio between the seeding times of two peers is inversely proportional to the ratio of their upload capacities (γ_i). Moreover, over-seeding peers with higher upload capacities have a stronger influence (proportional to γ_i) on this situation: the seeding time incurred by one over-seeding seeder in class i ($i < N$) is equal to that incurred by γ_i ($\gamma_i > 1$) over-seeding seeders in class N.

C. *Undersupplied*

We recognize a swarm to be undersupplied if it is not oversupplied. We assume that in an undersupplied swarm both leechers and seeders can fully utilize their upload capacities, i.e., $u_i = U_i$. This assumption has been validated by previous studies [17], [14]. In this situation, within a time interval Δ the total upload volume that *can* be provided by all peers should be no larger than the total download volume required by all $\sum_i \lambda_i \Delta$ new peers, i.e.:

$$\sum_i \lambda_i \Delta F \geq \sum_i (x_i + y_i + s_i) \gamma_i U_N \Delta. \quad (8)$$

Peers contribute their upload capacities, hence gain sharing ratios both in the leeching and the seeding process. At the end, they achieve sharing ratios equal to the SRE threshold, i.e.:

$$\alpha = \frac{((F/d_i)U_i + T_i U_i)}{F} \quad \Rightarrow \quad T_i = \frac{\alpha F}{U_i} - \frac{F}{d_i}, \quad (9)$$

where $(F/d_i)U_i$ represents the upload volume provided by a leecher in class i in its leeching process.

With x_i leechers, y_i lazy-seeding seeders, and s_i over-seeding seeders in class i in a steady state, the average downloading speed of a peer in class i can be calculated by solving the system of equations proposed in our previous work [14]:

$$d_i = \frac{(\sum_j D_{ji} x_j + \sum_j S_{ji}(y_j + s_j))U_j}{x_i}, \quad (10)$$

where D_{ji} (S_{ji}) specifies the fraction of upload speed allocated from a leecher (seeder) in class j to leechers in class i in the BitTorrent protocol.

From Eqs. (2), (9), and (10), we can derive T_i, x_i, and y_i. Due to the space constraint we omit their derivation here. However, simply from Eq. (9) we already have $T_i = \alpha F/U_i - F/d_i < F/U_i$. This implies that in an undersupplied steady state, an upper bound for the seeding time of a peer is the ratio between the size of the shared file and its upload capacity, which is much better than in an oversupplied steady state (Eq. (6)).

Further applying Eqs. (9) and (10) to Eq. (8), we rephrase the condition for a swarm to be in an undersupplied steady state as follows:

$$\frac{\sum_i s_i \gamma_i U_N}{\sum_i \lambda_i F} \leq 1 - \alpha, \quad (11)$$

which is exactly the reverse of Eq. (7).

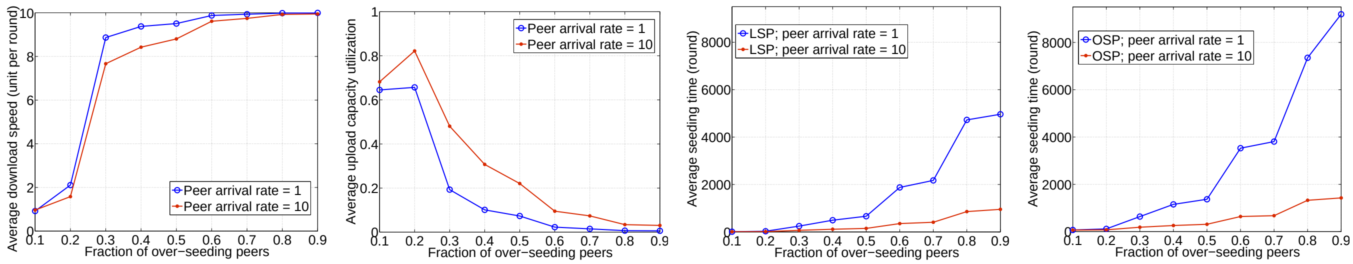
IV. Experimental setup

In our theoretical model we consider an idealized scenario of a private community where there is only one swarm, and we assume that the over-seeding peers have an infinite desired threshold for sharing ratios. Taking the insights obtained from our model, in the following sections we perform simulation-based analysis by considering more realistic scenarios. We simulate a private community that contains a number of different files (each associated to a different swarm). Each peer exhibits either one of two user behaviors: lazy-seeding or over-seeding. At any time a peer can only participate in one swarm, either as a leecher or a seeder. We do not consider parallel leechings or seedings, or a combination of these, since a peer who downloads and/or seeds n files simultaneously can be considered as being n different peers, each having $1/n$ of the original upload capacity. In our future work, we will consider the parallel leechings and seedings in which a peer's upload/download capacity is dynamically allocated among multiple swarms.

The simulation starts with 100 initial peers. Each of them joins a random swarm and is assigned a random number between 0 and 2 as its sharing ratio, in a way that the average sharing ratio for all peers is equal to 1. A peer may start a new leeching or a new seeding process if its randomly assigned sharing ratio is above or below the threshold, which, for lazy-seeding peers is the SRE threshold, and for over-seeding peers is their desired threshold. In this way we have created a steady state for the system.

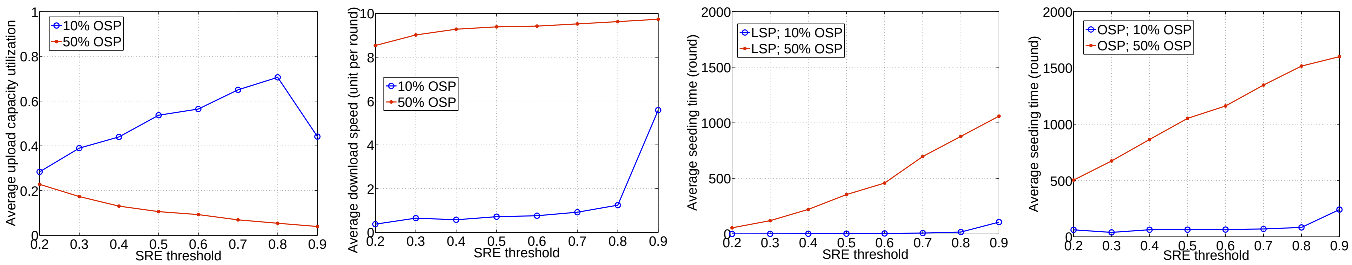
The simulation model is based on *rounds*. Each round represents a unit of time in which each peer is activated and may perform some activities, such as initiating new leeching or seeding sessions, or uploading and/or downloading data from other peers. In each round, leechers arrive according to a pre-assigned arrival rate and they join a random swarm. A new peer can start its first download freely, after which it maintains a sharing ratio above the threshold. Each peer attempts to download all files, in random order. We run the simulation for 2000 rounds and we keep a record of those peers who finish downloading all the files by the end of the simulation. All the results represent the average of 5 runs.

We consider three performance metrics, the average downloading speed, the average upload capacity utilization, and the average seeding time. The upload capacity utilization is calculated as the ratio between the upload speed and the upload capacity of a peer. It reflects system effectiveness in utilizing the upload capacities of peers. The average seeding time is



(a) Downloading speed (b) Upload capacity utilization (c) Seeding time of LSP (d) Seeding time of OSP

Fig. 3. System performance under different fractions of lazy-seeding peers (LSP) and over-seeding peers (OSP), and different peer arrival rates.



(a) Upload capacity utilization (b) Downloading speed (c) Seeding time of LSP (d) Seeding time of OSP

Fig. 4. Influence of the SRE threshold under different fractions of lazy-seeding peers (LSP) and over-seeding peers (OSP).

calculated separately for peers in different categories, like lazy-seeding and over-seeding peers, and high-capacity and low-capacity peers.

We always consider a bandwidth-homogeneous BitTorrent system unless otherwise indicated. The parameter settings[4] used in our simulation are introduced in Table II.

TABLE II
SIMULATION PARAMETER SETTINGS

SRE threshold	0.7	over-seeding threshold	2
file size	10 units	no. of files (swarms)	5
piece size	1 unit	no. of initial peers	100
upload capacity	1 unit per round	simulation rounds	2000

V. THE REWARD AND PUNISHMENT OF SRE

In this section we show the performance improvement and deterioration under SRE. Based on simulations we examine the influence of several parameters and we conclude the main reasons for the reward and punishment of SRE.

A. The imbalance of bandwidth supply and demand

SRE was designed to close the gap between bandwidth demand and supply as observed in public communities. However, under SRE, the presence of over-seeding peers completely reverses the situation and in private communities, swarms tend to be extremely oversupplied [16], [15], [21], [8].

In our first experiment we vary the fraction of over-seeding peers, thus generating different levels of oversupply. As shown in Fig. 3(a), with the fraction of over-seeding peers increasing from 0.1 to 0.9, the average downloading speed is increased nearly 10 times. However, the disadvantage of oversupply is more crucial: the average upload capacity utilization is significantly deteriorated and the seeding time is increased dramatically. With 50% over-seeding peers, on average each peer can only utilize less than 20% of its upload capacity (Fig. 3(b)). With this low upload capacity utilization, all peers have to stay for extremely long times (compared to their downloading times) to achieve the sharing ratio required by SRE (Figs. 3(c) and 3(d)). In our experiment with 50% over-seeding peers, the seeding time of a lazy-seeding peer is nearly 200 times more than its downloading time, and for over-seeding peers, it even increases to over 400 times.

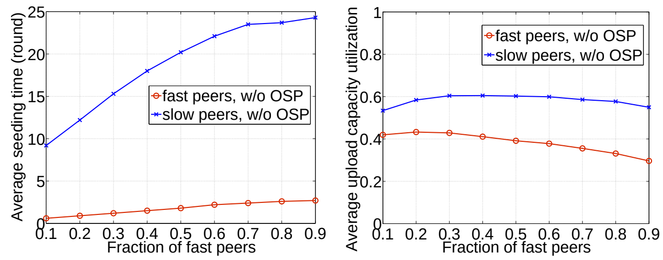
On the other hand, with a smaller peer arrival rate (which means a smaller demand) the imbalance is even worse and so is the performance. As shown in Fig. 3, when the peer arrival rate decreases from 10 to 1 peer per round, with the same fraction of over-seeding peers, the average upload capacity utilization is decreased 2-3 times and the average seeding time is increased 2-5 times.

Our simulation results show that, under SRE, the existence of over-seeding peers makes the swarms oversupplied. As a consequence, with a relatively large fraction of over-seeding peers and a small peer arrival rate, peers have to seed for extremely long times, though their seeding is not very productive. This is consistent with the our model results presented earlier (Eq. (6)).

B. The influence of the SRE threshold

In this subsection we analyze the influence of varying the SRE threshold. Fig. 4 shows that, which is consistent with our

[4]We choose 0.7 as the default value of the SRE threshold, as this value is used in many private communities [1], [2]. And we choose 2 as the default value of the desired threshold of over-seeding peers. We conjecture that for different values the tendency of this problem would be the same.



(a) Seeding time (b) Upload capacity utilization

Fig. 5. SRE's discrimination: without over-seeding peers (OSP).

intuition, when the SRE threshold is increased from 0.2 to 0.9, the upload capacity utilization is decreased while the average downloading speed and the average seeding time are increased. Further, the reward and punishment of SRE are limited when the fraction of over-seeding peers is small. This is consistent with our previous results (Eq. (6) and Fig. 3).

Surprisingly, in Fig. 4(a) we see that when there are 10% over-seeding peers, the upload capacity utilization is increased when the SRE threshold increases from 0.2 to 0.8, and then drops when it further increases to 0.9. We believe this is due to, what we term as, the *seeder's dilemma*: with either a very small or a very large number of seeders, peers cannot well-utilize their upload capacities. The former case is due to the piece availability problem. When there are not enough seeders, leechers have to exchange data with each other, which is not always possible since they only hold a part of the entire file. The latter case is due to the insufficient download demand. Without enough demand, even though seeders have the will, they cannot find enough leechers to upload to.

C. SRE's discrimination against peers with limited capacities

Until now we considered only bandwidth-homogeneous systems. However, would SRE have the same effects even in a bandwidth-heterogeneous system, and would these effects be the same on peers with different capacities? We answer these questions by extending our previous experiments to bandwidth-heterogeneous systems. More specifically, we simulate a system with two classes of peers, namely slow and fast peers. All the other settings are the same as in previous experiments, except that the upload capacity of slow peers is 1 unit per round and for fast peers it is 4 units per round.

1) Discrimination exists even without over-seeding peers: We first consider a private BitTorrent system without over-seeding peers, and we change the fraction of fast peers from 0.1 to 0.9. The simulation results are shown in Fig. 5.

We see that fast peers barely need to do any seeding work, but their existence increases the seeding times of slow peers (Fig. 5(a)). This result is consistent with our previous work [13] where we show that high-capacity peers manage to upload considerably more during the leeching process, and thus need to seed for shorter times. Meanwhile, consistent with our earlier model result, Fig. 5(b) shows that the upload capacity utilizations of both fast and slow peers do not change much with the fraction of fast peers. However, in general slow peers have better upload capacity utilizations. We believe this is due to the fact that slow peers stay as seeders longer than fast peers. Normally seeders can achieve better upload capacity utilizations, since they are not influenced by the piece availability problem.

While fast and slow peers both put all their effort in participating in the community, slow peers need to seed longer. We term this as SRE's *discrimination* against low-capacity peers. Next we show that when there are over-seeding peers, this discrimination is even more severe.

2) Discrimination is more severe with over-seeding peers: Similar to the results of previous bandwidth-homogeneous experiments, Fig. 6 shows that the negative aspects of SRE are more severe with a higher fraction of over-seeding peers. Interestingly, with the existence of 30% over-seeding peers, fast and slow peers now achieve similar upload capacity utilizations (Figs. 6(a) and 6(b)). We believe this is due to the fact that both fast and slow peers need to seed. Meanwhile, as shown in Figs. 6(c) and 6(d), when the fraction of over-seeding peers is increased from 0 to 30%, slow peers need to seed 200 to 500 rounds more than fast peers, while originally they only needed to seed 20 rounds more. In general, slow peers need to seed 4 times as long as fast peers, which is the same as the ratio between the upload capacity of a fast and a slow peer. This result is exactly consistent with our model prediction (Eq. (6)).

Clearly, the long seeding time, the low upload capacity utilization, and the discrimination against low-capacity peers severely deteriorate the user experience in private communities. In the following sections, we propose several strategies to alleviate these problems.

VI. Description of Proposed Strategies

Inspired by ideas in social sciences and economics, in this section we propose four strategies aimed at alleviating the negative effects of SRE, which require only a minor revision of the original SRE strategy.

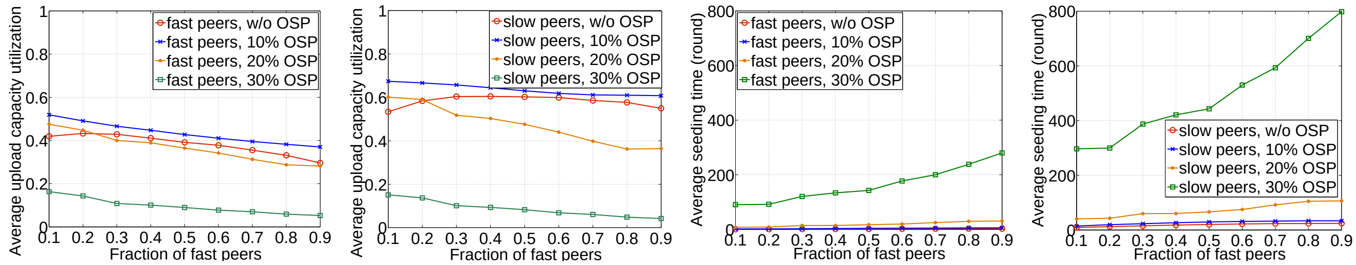
A. Negative taxation

The idea of *negative taxation* is that people earning below a certain amount receive supplemental pay from the government [3]. As we have already shown, due to the keen competition in uploading, peers may seed for long times, but still achieve very low sharing ratios. We take inspiration from the concept of negative taxation and devise a new strategy in which the upload amount of a peer is calculated as its its actual upload amount multiplied by coefficient $\mathcal{T}$ defined as:

$$\mathcal{T} = \max\{\min\{1/SR,\ \theta\},\ 1\},$$

where SR represents the sharing ratio of a peer and $\theta > 1$ represents the *maximum negative taxation degree*.

It is easy to see that a) when $SR \geq 1$, $\mathcal{T} = 1$, b) when $1/\theta \leq SR < 1$, $\mathcal{T} = 1/SR > 1$, and c) when $SR \leq 1/\theta$, $\mathcal{T} = \theta > 1$. By using this new strategy, to gain the same sharing ratio, poor peers ($SR < 1$) seed less and rich peers ($SR \geq 1$) seed the same amount as when using the original SRE.



(a) Upload capacity utilization of fast peers (b) Upload capacity utilization of slow peers (c) Seeding time of fast peers (d) Seeding time of slow peers

Fig. 6. SRE's discrimination under different fractions of over-seeding peers (OSP).

The maximum negative taxation degree controls the maximum negative taxation a peer can get, which alleviates the threat of free-riding.

B. Welfare for the rich

The term *welfare for the rich* is used to describe the bestowal of grants and tax-breaks to the wealthy [4]. Taking inspiration from this concept, we devise another strategy to alleviate the long seeding time, i.e., accelerating the seeding process of an over-seeding peer by giving welfare to it. The upload amount of a peer is calculated as its actual upload amount multiplied by coefficient $\mathcal{W}$ defined as:

$$\mathcal{W} = \max\{\min\{SR,\ \varphi\},\ 1\},$$

where $\varphi > 1$ represents the *maximum welfare degree*.

By using this strategy, to gain the same sharing ratio, poor peers ($SR < 1$) seed the same amount and rich peers ($SR \geq 1$) seed less than when using the original SRE. The maximum welfare degree controls the maximum welfare a peer can get, to prevent the over-seeding seeders from achieving their desired sharing ratios too quickly.

In our simulation we choose $\theta = \varphi = 2$. We conjecture that for different values of θ and φ the tendency of performance of the strategies would be the same.

C. Remuneration according to effort

In participatory economics, the *maxim of remuneration according to effort* has been introduced [5]. Under this scheme, people are paid according to the effort they put in rather than the amount of contribution. Taking inspiration from this concept, we propose the third strategy which takes into account the effort of users in terms of their seeding times. Previous studies have shown that the effort-based incentive policy applied in the leeching process improves the system-wide performance [20], [14]. We expect the same improvement when this effort-based methodology is applied in a private community.

More specifically, by applying SRE with *counting seeding time*, a peer can start a new download when either it has achieved the SRE threshold or it has seeded for a sufficiently long time. In this way, peers that were stuck in long seeding process in oversupplied swarms may leave and perform further downloads. The new demand generated by these peers helps to balance the bandwidth demand and supply in the system.

Clearly, the definition of "a sufficiently long period" is quite vague. Community administrators may choose various values, like 4 hours, 10 hours, or one day. In our simulations, we simply assume that it equals the size of the shared file divided by the upload capacity of a peer. It should be noted that since over-seeding peers are deposit-oriented, they still start new downloads only when they have achieved their desired sharing ratios.

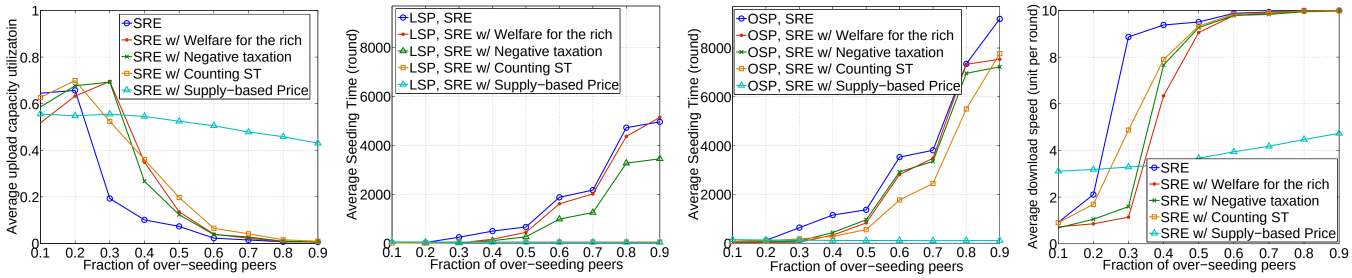
D. Supply-based price

According to the law of supply and demand, if the demand remains constant and the supply increases, the price of an item decreases and vice versa. We take inspiration from this insight to devise our fourth strategy, i.e., SRE with *supply-based price*. The basic idea is that the price a downloader needs to pay for downloading one unit of data should be inversely correlated with the supply in the swarm, i.e., the higher the seeder-to-leecher ratio, the less a downloader should pay and vice versa. In this way, in an oversupplied swarm, a leecher pays less and potentially achieves a higher sharing ratio by the end of its leeching process. Hence it is less likely for it to have an insufficient sharing ratio and thus stay as a seeder, which indirectly solves the oversupply problem in this swarm. On the other hand, in an undersupplied swarm, a leecher pays more and potentially achieves a smaller sharing ratio, which makes it stay as a seeder with a higher possibility than using the original SRE. In this way, the undersupply problem is also alleviated indirectly.

Simplified from our model result, i.e., Eqs. (7) and (11), we use the *seeder-to-leecher ratio* (SLR) as a metric to decide whether a swarm is oversupplied or undersupplied. Community administrators can set different SLR values as the threshold, but we simply assume that when $SLR \geq 1$ the swarm is oversupplied and when $SLR < 1$ the swarm is undersupplied. The download amount of a peer is calculated as its actual download amount multiplied by coefficient $\mathcal{P}$ defined as:

$$\mathcal{P} = \max\{1/SLR,\ \phi\},$$

where ϕ represents the *lowest price* for downloading one unit



(a) Upload capacity utilization (b) Seeding time of LSP (c) Seeding time of OSP (d) Downloading speed

Fig. 7. Strategy performance in alleviating the eternal seeding problem: under different fractions of lazy-seeding peers (LSP) and over-seeding peers (OSP).

of data, which is used to alleviate the threat of free-riders. In our simulation we choose $\phi = 0.1$.

VII. Strategy Evaluation

In this section we evaluate the performance of the new strategies proposed in Section VI. The experimental setup is the same as in Section V and results are shown in Figs. 7 and 8.

A. *Higher upload capacity utilization and shorter seeding time*

From Fig. 7 we see that by using any of the new strategies, peers achieve higher upload capacity utilizations, as well as smaller seeding times. As shown in Fig. 7(a), when there are 40% over-seeding peers the upload capacity utilization is increased 2-3 times compared to using the original SRE. While all other strategies have decreasing upload capacity utilizations with an increasing fraction of over-seeding peers, SRE with supply-based price performs stably. Given any fraction of over-seeding peers, on average peers can utilize at least 40% of their total upload capacities while for the original SRE it drops to less than 1% when there are 90% over-seeding peers.

With the improved upload capacity utilization, the average seeding time is reduced significantly. As shown in Figs. 7(b) and 7(c), when there are 60% over-seeding peers, SRE with welfare for the rich reduces at least 10% of the original seeding time for both lazy-seeding and over-seeding peers. SRE with negative taxation deals with lazy-seeding peers directly, hence it achieves an even better performance in reducing the seeding time of lazy-seeding peers, which is a 50% improvement compared to that achieved by SRE with welfare for the rich.

SRE with counting seeding time further relieves lazy-seeding peers from the long seeding process in a more effective manner. As shown in Fig. 7(b), they only need to seed for a negligible time compared to when using the original SRE, or either of the above two new strategies. Interestingly, by applying SRE with counting seeding time, the seeding time of over-seeding peers is also decreased (Fig. 7(c)), even though they still desire the high sharing ratios as when using the original SRE. We believe this is due to the fact that with lazy-seeding peers finishing their seedings sooner, the upload competition is reduced and over-seeding peers can achieve their desired threshold more quickly. Meanwhile, when the lazy-seeding peers are released from the seeding process, they join other swarms as new leechers, which indirectly alleviates the oversupply in those swarms.

Finally, the best performance in reducing the seeding time for all peers is achieved by SRE with supply-based price. The seeding time of both lazy-seeding and over-seeding peers is reduced by three orders of magnitude. In our view the main reason for the success of SRE with supply-based price is that it adaptively adjusts the supply and demand in a swarm. When the swarm is oversupplied, the price for downloading one unit of data is lower and peers can finish downloads at less expense, which directly reduces their consequent seeding amount and hence avoids adding more seeders in this oversupplied swarm. In this way, the imbalance of bandwidth supply and demand is mitigated, and the strategy gives a way to escape out of the seeder's dilemma as described in Section V-B. A similar argument can also be applied to an undersupplied swarm.

B. *Tradeoff: slightly decreased downloading speed*

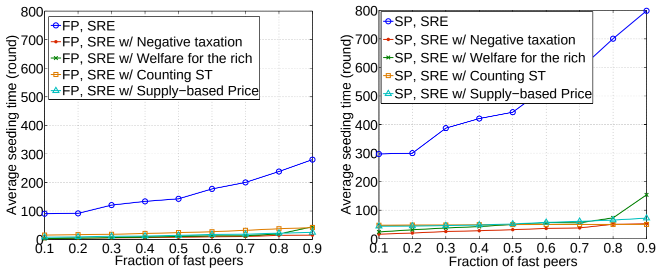
By adopting any of the new strategies, while the seeding time is dramatically reduced, as a trade-off, the average downloading speed is decreased (Fig. 7(d)), hence the downloading time is increased. However, given that in our simulations we consider files with size equal to 10 units, the increase of the downloading time (tens of rounds) is negligible compared to the decrease of the seeding time (hundreds or even thousands of rounds).

C. *Reduced discrimination*

To examine the effectiveness of the proposed strategies in alleviating SRE's discrimination against peers with limited capacities, we repeat our experiments by further considering a bandwidth-heterogeneous system with two classes of peers, fast and slow. From Fig. 8 we see that *all* the proposed strategies effectively alleviate SRE's discrimination against low-capacity peers. With 30% over-seeding peers, originally slow peers need to seed 200-500 rounds more than fast peers do. By applying any of the new strategies, this difference is reduced to within tens of rounds.

VIII. Discussion: The Change of User Behavior?

It could be argued that the new strategies proposed in this paper may trigger a change in user behavior. Specifically, users from the two classes that we defined, lazy-seeding and over-seeding, might be incentivized to switch their classes under the



(a) Seeding time of fast peers (FP) (b) Seeding time of slow peers (SP)

Fig. 8. Strategy performance in alleviating discrimination: with 30% over-seeding peers (OSP).

new strategies. It could be conjectured that as a consequence, the system performance might be adversely affected.

However, we note that there is only a very small fraction of strategic users in BitTorrent communities [6]. So the likelihood of peers changing behavior is quite small. Especially regarding over-seeding peers, we argue that in general the possibility of such peers changing to lazy-seeding peers is quite low. This is because we conjecture that the behavior of over-seeding peers is motivated by either one or a combination of the following three reasons: a) Over-seeding peers always want to be relatively more well-off in terms of sharing ratio as compared to average users so that they can be among the “rich elite” of the community and gain some potential benifits[5]; b) They are altruists who want to help the community as much as they possibly can; and c) They are hoarders who desire to conserve sharing ratio for “rainy days” i.e., those time periods when they feel they might engage in heavy downloading activity and might as a result be expelled from the community due to low sharing ratios.

Nevertheless, in this section we would like to analyze what happens if users do change their behavior. We consider each proposed strategy in turn and discuss the possible effects of the change in user behavior.

A. SRE with negative taxation and welfare for the rich

Under SRE with negative taxation, peers with lower sharing ratios gain sharing ratio more easily. Hence, lazy-seeding peers have no incentive to change their behaviors, while over-seeding peers may change to lazy-seeding peers. Similarly, when SRE with welfare for the rich is applied, strategic lazy-seeding peers could become over-seeding peers, while over-seeding peers will not change their behaviors.

The worst case scenario of applying either of these two new strategies is that all peers exhibit the same behavior, i.e., either all peers become lazy-seeding or all become over-seeding. In the former case every member of the community would still be forced to maintain a minimum sharing ratio required by SRE, i.e., every peer would continue to provide a certain level of contribution. This outcome would still be better than the situation in a public community where every peer has the option to leave immediately after downloading. In the latter case, i.e., when all peers are over-seeding, it is less likely for them to complain, since an over-seeding behavior, i.e., a desire for higher sharing ratios, automatically implies a long seeding time.

[5] Such as priority in downloading popular files, the possibility to send invitations to others, etc [1], [2].

B. SRE with counting seeding time

Under SRE with counting seeding time, users may set their upload speeds to zero and pretend to be seeding. However, according to the TFT policy in BitTorrent, in an undersupplied swarm the upload speed of a peer directly influences its downloading speed. If a user sets its upload speed to zero, it would hardly be able to download at reasonable speeds in undersupplied swarms, which are normally swarms providing new and popular content.

Further, private community administrators could set another SRE threshold, which is smaller than the original one, and stipulate that users who cannot achieve the original SRE threshold must seed for a predefined time, as well as achieve the smaller SRE threshold. In this case, the potential threat of free-riders is alleviated.

C. SRE with supply-based price

By adopting SRE with swarm-based price, the only advantage that users could gain is that they may opt to download files which have a lower price. However, this is unlikely to happen, since the choice of file to download mainly depends on user’s interest in the content, rather than the price. Even if some users might be tempted to download a file simply based on its price, this would have little influence on the performance of SRE with swarm-based price, because this strategy is self-organizing and will adjust the balance of supply and demand automatically.

IX. Related Work

Most existing studies on BitTorrent incentive policies focus on TFT and its variations [10], [11], [17], [9], [20], [14]. To date, only few works have analyzed private communities. Zhang *et al.* [21] investigate hundreds of private trackers and depict a broad and clear picture of the private community landscape. Chen *et al.* [8] compare system behaviors among 13 private trackers and 2 public trackers, and they show their differences regarding user viscosity, single torrent evolution, user behaviors, and content distribution. Liu *et al.* [15] also perform measurement studies and further develop a model to show that SRE indeed provides effective incentives, but is vulnerable to collusion. Andrade *et al.* [7] focus on the dynamics of resource demand and supply, and they show that users typically try to increase their contribution levels by seeding for longer and not by providing more bandwidth to the system. However, our paper shows that providing limited bandwidth is not the will of users, but it is a consequence of the oversupply in private communities. While these studies all focus on demonstrating the high seeding level achieved by private communities, Authors in [12] study it from a macroeconomic point of view and show that such communities can suffer from problems leading to serious inefficiencies. Authors in [19] further study these problems in private communities, and they provide a novel mechanism that proactively prevents the system from seizing. Following their ideas, our paper analyzes both the reward and punishment of adopting SRE based on measurement, theoretical model, and extensive simulations. Further we propose new strategies to alleviate SRE's punishment, which are evaluated to be very effective through simulations.

X. Conclusion

In this work, with the support of real-world observations we have provided an analytical model that captures the essence of SRE adopted by private communities. Our model predicts the average downloading speed, the average seeding time, and the average upload capacity utilization for users in communities that employ SRE. We extend the analysis of SRE using extensive simulations to demonstrate its reward and punishment. Given the existence of over-seeding user behavior, our simulation results show that by adopting SRE, swarms tend to be extremely oversupplied. Under this oversupply, peers achieve high downloading speeds but with significant tradeoffs, which include low upload capacity utilizations and extremely long seeding times. Under certain scenarios, a peer may seed over 1000 times as long as its downloading time, while on average only utilizing 1% of its upload capacity. Further, SRE discriminates against peers with low capacities and forces them to seed for even longer durations. To alleviate these problems, we propose four strategies and the simulation results show that they are all very effective. Particularly, SRE with supply-based price, while maintaining a system-wide high downloading speed, achieves very stable high upload capacity utilization and reduces seeding durations by three orders of magnitude as compared to the original SRE.

We leave the considerations of swarms with different popularities and parallel leechings and/or seedings in which a peer's upload/download capacity is dynamically allocated among multiple swarms, for our future work.

Acknowledgements

We would like to thank Prof. Dah Ming Chiu from Chinese University of Hong Kong for his valuable advices and comments. This work was supported by the Future and Emerging Technologies programme FP7-COSI-ICT of the European Commission through project QLectives (grant no.: 231200) and the EU FP7 project P2PNEXT (grant no.: 216217).

References

[1] http://www.bitsoup.org/.
[2] http://hdchina.org/.
[3] http://en.wikipedia.org/wiki/Negative_income_tax.
[4] http://en.wikipedia.org/wiki/Welfare_for_the_rich.
[5] http://en.wikipedia.org/wiki/Participatory_economics.
[6] K. Anagnostakis, F. Harmantzis, S. Ioannidis, and M. Zghaibeh. On the impact of practical p2p incentive mechanisms on user behavior. *NET Institute Working Paper*, 2006.
[7] N. Andrade, E. Santos-Neto, F. Brasileiro, and M. Ripeanu. Resource demand and supply in bittorrent content-sharing communities. *Computer Networks*, 53, 2008.
[8] X. Chen and X. Chu. Measurements, analysis and modeling of private trackers. In *Proceeding of IEEE P2P*, 2010.
[9] A. L. H. Chow, L. Golubchik, and V. Nisra. Bittorrent: An extensible heterogeneous model. In *Proceedings of IEEE INFOCOM*, 2009.
[10] B. Fan, D. M. Chiu, and J. C. Lui. The delicate tradeoffs in bittorrent-like file sharing protocol design. In *Proceeding of IEEE ICNP*, 2006.
[11] L. Guo, S. Chen, Z. Xiao, E. Tan, X. Ding, and X. Zhang. Measurements, analysis, and modeling of bittorrent-like systems. In *Proceedings of the 5th ACM SIGCOMM Conference on Internet Measurement*, 2005.
[12] D. Hales, R. Rahman, B. Zhang, M. Meulpolder, and J.A. Pouwelse. Bittorrent or bitcrunch: Evidence of a credit squeeze in bittorrent? In *Proceeding of Wetice*, 2009.
[13] A.L. Jia, L. D'Acunto, M. Meulpolder, and J.A. Pouwelse. Modeling and analysis of sharing ratio enforcement in private bittorrent networks. In *Proceeding of IEEE ICC*, 2011.
[14] A.L. Jia, L. D'Acunto, M. Meulpolder, J.A. Pouwelse, and D.H.J. Epema. Bittorrents dilemma: Enhancing reciprocity or reducing inequity. In *Proceeding of IEEE CCNC*, 2011.
[15] Z. Liu, P. Dhungel, D. Wu, C. Zhang, and K.W. Ross. Understanding and improving incentives in private p2p communities. In *Proceeding of ICDCS*, 2010.
[16] M. Meulpolder, L. D'Acunto, M. Capota, M. Wojciechowski, J.A. Pouwelse, D.H.J. Epema, and H.J. Sips. Public and private bittorrent communities: A measurement study. In *Proceeding of IPTPS*, 2010.
[17] M. Meulpolder, J.A. Pouwelse, D.H.J. Epema, and H.J. Sips. Modeling and Analysis of Bandwidth-Inhomogeneous Swarms in BitTorrent. In *Proceeding of IEEE P2P*, 2009.
[18] D. Qiu and R. Srikant. Modeling and performance analysis of bittorrent-like peer-to-peer networks. In *Proceeding of SIGCOMM*, 2004.
[19] R. Rahman, D. Hales, T. Vinko, J.A. Pouwelse, and D.H.J. Sips. No more crash or crunch: Sustainable credit dynamics in a p2p community. In *Proceeding of HPCS*, 2010.
[20] R. Rahman, M. Meulpolder, D. Hales, J.A. Pouwelse, D.H.J. Epema, and H.J. Sips. Improving efficiency and fairness in p2p systems with effort-based incentives. In *Proceeding of IEEE ICC*, 2010.
[21] C. Zhang, P. Dhungel, Z. Liu Di Wu, and K.W. Ross. Bittorrent darknets. In *Proceeding of IEEE INFOCOM*, 2010.