Incentivizing Cooperation in P2P File Sharing

Indirect Interaction as an Incentive to Seed

Arman Noroozian, Mathijs de Weerdt, and Cees Witteveen

Delft University of Technology, The Netherlands
{a.noroozian,m.m.deweerdt,c.witteveen}@tudelft.nl

Abstract. The fundamental problem with P2P networks is that quality of service depends on altruistic resource sharing by participating peers. Many peers freeride on the generosity of others. Current solutions like sharing ratio enforcement and reputation systems are complex, exploitable, inaccurate or unfair at times. The need to design scalable mechanisms that incentivize cooperation is evident. We focus on BitTorrent as the most popular P2P file sharing application and introduce an extension which we refer to as the indirect interaction mechanism (IIM). With IIM BitTorrent peers are able to barter pieces of different files (indirect interaction). We provide novel game theoretical models of BitTorrent and the IIM mechanism and demonstrate through analysis and simulations that IIM improves BitTorrent performance. We conclude that IIM is a practical solution to the fundamental problem of incentivizing cooperation in P2P networks.

Keywords: Incentives for Cooperation, Peer to peer coordination, BitTorrent.

1 Introduction

It is widely agreed upon that BitTorrent's *Tit-for-Tat* (TFT) mechanism is successful in incentivizing selfish autonomous peers to contribute to others as long as they are still in need of their peers' contributions. However, this is insufficient to incentivize peers to contribute after they have downloaded the files they seek. Consequently many mechanisms have been developed for P2P systems that incentivize cooperation beyond this natural limit of *direct* TFT as implemented in BitTorrent.

In this body of work we introduce an Indirect Interaction Mechanism (*IIM*) that extends the BitTorrent TFT. It takes advantage of mining a large data set for intersting and behavioral patterns. The observed pattern is that peers tend to download multiple files simultaneously (*multi-swarming*) or are in possession of files that have been downloaded in the past. This means that contrary to the standard BitTorrent protocol in which pieces of a single file are bartered for pieces of the *same* file, pieces of *different* files can be bartered for each other in a coordinated manner that better matches the supply and demand of files. Our data demonstrates this is a simple task. We observe that a large number of peers are able to use IIM with little coordination required.

While other approaches focus on reward/punishment or reputation mechanisms for better cooperation [24,26,9,17], IIM incentivizes cooperation without the need for such mechanisms (e.g. sharing ratio enforcement). We demonstrate that discriminative indirect interaction is not only possible in but that it incentivizes cooperation in game theoretic terms and also improves or does not degrade the performance of the standard

BitTorrent protocol. IIM is suitable for online bartering applications such as P2P enabled *Video on Demand* or Online Music services. Furthermore, economic experiments also hint at the strategy being effective [18,21].

Dealing with the complexities of distributed open systems such as BitTorrent and designing effective cooperation and coordination among peers is an area which can greatly benefit from techniques for mining the large data sets available.

In Section 2 we first give background information on the BitTorrent protocol. In Section 3 we then build a game theoretical model of BitTorrent to formally investigate the incentives problem. In Section 4 we explain the concept of indirect interaction and present our indirect interaction mechanism as a solution to the problem along with its game theoretical model. Next we give our experimentation and validation in Section 5: We give our data mining results in Section 5.1 demonstrating easy coordination of peer interactions, and we validate our hypotheses in Section 5.2. In Section 6 we discuss experimental economics studies that indicate our approach to be effective, and then explore some of the notable related work on incentives and other work on file sharing (in Section 7). Finally we conclude and present some ideas for future work in Section 8.

2   Background

BitTorrent is a file sharing protocol designed by Bram Cohen [2]. In BitTorrent peers *barter* bandwidth. Each file is divided into pieces and subpieces. Peers interact in rounds during which they can download (sub)pieces from their neighboring peers.

All content that can be downloaded has an associated *.torrent* file in which information about the content, size, SHA-1 hashes of file pieces, etc. are stored. The *.torrent* file also contains the address of a central server referred to as a *tracker*. The tracker is the entity responsible for introducing peers to each other, and keeping information about the peers. A peer will contact the tracker to request a list of peers (typically random list of size 50) which it can contact to download content. It will periodically contact the tracker for more peers. Peers that are downloading the same content from the same tracker are organized into a logical *swarm*.

BitTorrent recognizes two types of peers. Ones that are downloading a file: *leechers*, and ones that posses the entire file: *seeders*. Leechers follow the aforementioned procedure to download. Seeders wait to be contacted to upload content. They get nothing in return for uploading. On the other hand leechers are obliged by the protocol to reciprocate other leechers that provide content to them in a tit-for-tat fashion. As explained below, a leecher's reciprocation is not necessarily proportional to the amount of contributions received (e.g. it can reciprocate with less bandwidth than it actually received).

The TFT mechanism is implemented in BitTorrent's *unchoking* algorithm and has been the main driver behind Bit Torrent's performance and popularity. Each peer's available upload bandwidth is split into upload slots. The number of slots vary depending on the available upload bandwidth (typically equals 5). The unchoking algorithm assigns upload slots to preferred neighboring peers. At the beginning of each bartering round a peer assesses the amount of bandwidth that it has received from its neighbors and gives away its upload slots to the highest ranking contributors. Seeders assign uploads slots to the fastest downloading peers in a round robin fashion. In order to explore the bandwidth capacity of other neighbors, peers assign one of their upload slots to a random

neighboring peer every few number of rounds. This is referred to as *optimistic unchoking*. Optimistic unchoking has a two fold purpose as it also allows newly arriving peers to receive a piece with which they can start the bartering process. This is referred to as *bootstrapping*.

Despite the success of the TFT mechanism [2,11], seeders are the neglected party in BitTorrent. They gain no utility by participating. As a result peers have no incentive to remain connected to a swarm after they have downloaded a file. A frequently observed behavior is that leechers leave a swarm as soon as they turn into seeders [15]. This phenomenon has lead to a drive towards various mechanisms for incentivizing cooperation such as reputation mechanisms and sharing ratio enforcement in private BitTorrent communities. Reputation systems suffer from problems such as whitewashing [5], communication overhead, sybil attacks [3], and have remained quite theoretical due to implementation complexity [15]. Private communities on the other hand suffer from exclusivity, imbalanced supply/demand and unfairness [7,15]. The need to design scalable incentive mechanisms that mitigate such problems is evident.

3   A Model of BitTorrent

In order to examine the incentives problem formally we have derived a novel game theoretical model of BitTorrent. Several key aspects of the protocol have directed us towards our choice of model the first of which is uncertainty. That is, a BitTorrent peer never knows whether it has uploaded with a high enough bandwidth in order to be reciprocated. This means that under standard game theoretical assumptions (rationality and utility maximization) BitTorrent peers would be maximizing their *expected* utility. A second aspect is interaction in rounds. At the end of each round peers will assess and readjust their strategies. Our model accordingly characterizes the utility of each player for every round. The third aspect is the protocol's design to prefer to interact with peers that have a higher upload bandwidth and rare pieces of the file. We capture these aspects of the BitTorrent protocol in our model which is based on a proposed framework by *Buragohain et al.* [1].

We characterize a peer's bandwidth contributions as a vector in which all neighbors are indexed and the values correspond to the bandwidth that was contributed to the peer with that index. The notation b_i denotes this characterization for a peer i. We have $b_i = (b_{i0}, \ldots, b_{i\|N_i\|})$ where N_i is the set of peer i's neighbors.

Peers can make certain contributions to each other which we characterize as the fraction of the file they possesses. Based on a peer i's pieces and another peer j's pieces the notation α_{ij} represents the fraction of the file that i can provide to peer j. If α_{ij} is relatively large there is a higher chance that peer i has a rare piece in which j is interested.

BitTorrent peers have a preference for rare pieces so we assume that each contribution has a certain value for a peer based on the rarity of the piece. We characterize this with a factor v_{ij} which denotes the value that peer i assigns to receiving a contribution from peer j, normalized such that a value of 1 represents the exact compensation for contributing the complete upload capacity.

However, the rarity of a piece is not the only factor influencing the choice of potential interaction partners. The amount of contribution of a peer in the previous round also

Table 1. Parameters of the Model

N_i	The set of peer i's neighbors.
b_i	The vector $(b_{i0}, \ldots, b_{i\|N_i\|})$, ($b_{ij}$ denotes the bandwidth peer i assigned to j).
v_{ij}	The value to peer i of a contribution from peer j to peer i.
α_{ij}	The fraction of the total pieces that peer i can provide to peer j.
$P(\alpha_{ij}, b_{ij})$	The probability of peer i receiving a contribution from j, having made a bandwidth contribution of b_{ij} while it has α_{ij} to provide to j.

determines the choice. We denote by $P(\alpha_{ij}, b_{ij})$ the probability of a peer i receiving a contribution from j to which it has made a bandwidth contribution of b_{ij} while it has an α_{ij} fraction of the pieces to provide. We refer to this probability as the *probability of reciprocation*. This probability $P()$ in fact can be seen to capture the history of interaction as a belief where all history is represented in the current state. We assume that the probability of reciprocation is a monotonically increasing function in both α_{ij} and b_{ij} and greater than zero (zero when a peer owns the entire file). The logic here is that contributing more bandwidth results in a higher chance of reciprocation and since the more pieces a peer can provide the more likely it is to have a rare piece, a higher α also results in a higher chance of reciprocation. Table 1 summarizes the parameters of our model.

A strategy of a peer i in our game theoretical model is equivalent to the vector b_i. Naturally the sum of peer i's bandwidth contributions to its neighbors cannot exceed its upload capacity. Using the notation b_{-i} to denote the strategy of peers other than i, we describe the expected utility of a peer i as follows:

$$u_i(b_i, b_{-i}) = -\|b_i\| + \sum_{j \in N_i} [P(\alpha_{ij}, b_{ij}) \times v_{ij} \times \|b_j\|] \tag{1}$$

Equation 1 states that the expected utility of i is equal to the sum over all neighbors j of the probability $P(\alpha_{ij}, b_{ij})$ that i receives a contribution from j multiplied by the value that i assigns to receiving the contribution (v_{ij}) and the bandwidth at which it receives the contribution, minus its own contribution $-\|b_i\|$. Note that we model i's expected share of j's bandwidth also within the probability $P()$.

According to this equation, a peer has to either increase its contribution, or the fraction of the pieces that it can provide to its neighbors in order to increase its utility. This is a well known fact about the BitTorrent protocol and how it operates. While the model is capable of explaining the basics of the protocol it is also capable of accounting for many observed phenomena of the protocol including, freeriding [5] (also see [13]), large view exploits [27], strategic piece revelation [12], minimum reciprocation winning uploads [22] and clustering of similar bandwidth peers [10]. For further details we refer the reader to [20].

Equation 1 clearly demonstrates the seeding incentives problem. The utility of a seeding peer is always negative because it does not have demand for file pieces. This problem needs to be addressed in order to be able to sustain the operation of P2P systems and avoid the tragedy of the commons.

4   Indirect Interaction

Seeding incentives in BitTorrent can be created by allowing *indirect interactions*, where a peer *i* supplies pieces to peer *j* who can supply pieces of another file to some other peer *k*, who in its turn can supply *i*. Such a setting requires peers to be involved in up/downloading multiple files, i.e., they are *multi-swarming*.

Definition 1 *(**Multi-Swarming Peer**). A peer p is a **multi-swarming peer** if it is online in swarms $S = \{s_1, s_2, \ldots\}$ where $|S| \geq 2$ and $\exists s_i \in S$ in which p is a leecher.*

The extension of BitTorrent presented in this section enables such interactions. We therefore refer to this protocol as the *Indirect Interaction Mechanism* (*IIM*).

4.1   Indirect Interaction Mechanism

IIM is a mechanism that relies on the tracker and its information to assist multi-swarmers in finding suitable parties to interact with in the same way that a tracker introduces new sets of peers in the standard BitTorrent protocol. Additionally, however, the tracker uses a supply/demand graph of active swarms.

Definition 2 *(**Supply and Demand Graph**). A Supply and Demand graph is a directed graph $G = (V, E)$ with edge labels, where V is the set of swarms, and the edges in E are defined as follows.*

1. *(**Seeder-Leecher**) There is an edge (s_1, s_2) in E if there exists a peer p that is a leecher in s_1 and a seeder in s_2. The edge is labeled with the set of all such peers p.*
2. *(**Leecher-Leecher**) There is an edge (s_1, s_2) as well as an edge (s_2, s_1) in E if there exists a peer p that is a leecher in both s_1 and s_2. Again the edges are labeled with the set of all such peers p.*

Any peer *p* that is a seeder in both *s*1 and *s*2 (**Seeder-Seeder**) does not occur in any of the labels on edges. Figure 1 gives an example of a supply/demand graph. This graph, denoted by *G* above, is stored as an adjacency matrix at the tracker. The IIM then works as follows.

1. Multi-swarming peers seeking indirect interaction announce all swarms they are participating in to the tracker and do a request.
2. The tracker updates the Supply and Demand graph, and then uses breadth-first search to find a limited number of cycles of small length.
3. The tracker introduces the peers of the found cycles to the requesting peer.
4. The requesting peer contacts the peer succeeding it in the cycle with the information it received from the tracker and requests indirect interaction.
5. All subsequent peers in the cycle will do the same until the request message arrives back at the originating peer.
6. Once a cycle has been established peers can unchoke connections with the same TFT mechanism that they use for normal BitTorrent connections with the additional consideration that unchoking the first next peer along the cycle depends upon the amount of bandwidth received from the previous peer in the cycle.

Run time operations of the tracker scale linearly with the number of requesting peers (since the length of the cycles is bounded by a constant), and quadratically in memory requirements.

4.2   The IIM Model

For simplicity we only present the model of IIM for indirect interaction cycles of length 2. However, our model and analysis are generalizable to cycles of greater length.

Swarm	Peers	Peer Type
1	p_1	Leecher
	p_2	Leecher
	p_3	Seeder
	p_6	Seeder
2	p_1	Leecher
	p_2	Leecher
	p_4	Seeder
	p_5	Leecher
3	p_3	Leecher
	p_4	Leecher
	p_5	Leecher
	p_6	Seeder

Consider a peer i that is multi-swarming in two swarms. Let us assume that i has received a set of length 2 cycles C_i from the tracker and it decides to upload content to these cycles. We denote the contribution that i makes to each cycle by b_{ic} for $c \in C_i$. This creates a third term in the utility function as follows

Fig. 1. Example of a Supply and Demand graph

$$u_i(b_i, b_{-i}) = -\|b_i\| + \sum_{j \in \hat{N}_i} [P(\alpha_{ij}, b_{ij}) \times v_{ij} \times \|b_j\|] + \sum_{c \in C_i} [P'(\alpha'_{ic}, b_{ic}) \times v_{ic} \times \|b_c\|] \quad (2)$$

Notice that peer i is making the same *total* contribution as before, i.e., $\|b_i\|$. From the perspective of the normal peers nothing has changed about the strategy of peer i because they derive expected utility based on i's total contribution. The interpretation of the second term in the equation is exactly the same as before. Notice that this time however, the probability of reciprocation from cycles (3rd term) depends on the fraction of pieces of the second file (α'_{ic}) that peer i can provide to individual peers in C_i. That is, peer i can provide α' pieces of one file in return for α pieces of the file it itself is interested in. Equation 2 constitutes our model for the proposed IIM mechanism.

Consider now a peer i which has a strategic choice between contributing through seeding and indirect interaction through a cycle of length two, or directly within the swarm it is leeching in (*i.e.* IIM *vs.* standard BitTorrent). For cycles of length two, we can safely assume that the probability P' of receiving a contribution through a cycle is the same as the probability of direct reciprocity P. However, when peer i is seeding, $\alpha_{ic} = 1$, so assigning the same bandwidth to the cycle is expected to contribute more to the utility than direct interaction. Consequently, a former seeder can now gain positive utility from seeding content instead of a negative utility as was the case in the original BitTorrent model (Equation 1).

4.3   IIM Properties

We now analyze the IIM model in a bit more detail. One consequence of the discussion above is that this also helps bootstrapping a peer. Consider a peer i that is partially or wholly in possession of one file and starting to acquire another file, while there are other peers j in possession of α and α' pieces of these files. At this stage of a download, $\alpha'_{ik} = 0 \; \forall k \in N_i$ (α is at its minimum) which means that the utility that peer i expects to derive from the second term of Equation 2 is at its lowest. Therefore i would have to wait to receive contributions as optimistic unchokes from its neighbors.
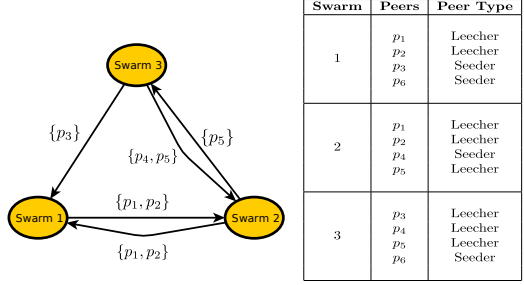
On the other hand, peer i possesses pieces of another file that it can use in indirect interaction with peer j. This means that i would no longer have to wait for the optimistic unchokes if it interacts indirectly with peer j. By symmetry the same argument holds for j. As a result, interacting indirectly would achieve a higher utility for both peers at bootstrapping phase because there is a higher chance that contributing in cycles will acquire reciprocation for both involved parties ($\alpha_{ij} > 0$ for peer i and $\alpha'_{ji} > 0$ for j). Hence we have the following hypothesis:

Hypothesis 1 *(**Faster Bootstrapping**). A multi-swarming peer will be able to bootstrap faster if it interacts indirectly in cycles.*

This time imagine a peer i which is leeching the file α and seeding the file α' (leecher-seeder) and another peer j which is doing the exact opposite (seeder-leecher). These two peers would gain additional utility by interacting indirectly in cycles because the fraction of the file that the pairs can provide to each other are maximal. Notice that if peer i already knows that peer j has chosen the option to interact indirectly with peer i, it would also decide to interact indirectly with peer j because it would derive a greater utility. The same argument is valid in reverse for peer j. As a result, peers i and j would at worst be indifferent to choosing either option.

Hypothesis 2 *(**Leecher-Seeder/Seeder-Leecher Pairs are Better Off Interacting Indirectly**). A multi-swarming peer i that is seeding file α' and leeching file α and a multi-swarming peer j that is doing the exact opposite will be better off to interact in cycles.*

Other combinations of the peers are less structured than the bootstrapping phase or the leecher-seeder/seeder-leecher pairs. These combinations are more difficult to derive analytically from the model. In such cases, the probability of reciprocation and the assigned value for each contribution from each peer have a more prominent role in determining the utility of this peer. However, by expectation interacting indirectly with a multi-swarming peer is similar to interacting with any other random peer in the set of neighbors. That is, we expect the fraction of pieces that the two multi-swarmers are able to barter and the values that they assign to each others' pieces be similar to those of a random peer in the set of their neighbors. Hence we can derive the following hypothesis regarding the other possible combinations of indirect interaction:

Hypothesis 3 *(**IIM Will not Reduce the Performance of the BitTorrent**). The standard BitTorrent protocol will not significantly outperform the indirect Interaction mechanism in terms of the time required to download an entire file.*

5   Experimentation and Validation

Our approach to validating claims consists of a measurement study and simulations of the indirect interaction mechanism. Our measurement study focuses on validating our assumption regarding the possibility of indirect interaction. It has helped us in designing IIM. Our simulations on the other hand focus on validating hypotheses that we derived from our game theoretic model of IIM. In short we demonstrate that IIM incentivizes potential multi-swarmers to choose indirect interaction over seeding in or disconnecting from swarms which could otherwise be used for indirect interaction.

5.1   Measurement Study

We have conducted a measurement study of a private BitTorrent tracker for the existence of potential supply/demand cycles. We have used the *FileList.org* community tracker data set previously studied by Roozenburg et al. [25]. Our choice of the data set has a two fold purpose. First, the possibility of indirect interaction is dependent on diversity and a large enough number of peers to locate nth parties. The FileList data is large enough to eliminate the possibility of not having enough diversity in supply and demand. Second, precise measurement of the possibility requires a data granularity at the level of individual peers. The FileList.org data contains information based on anonymized user names which allows us to uniquely identify users even if they reside behind a NAT device or Firewall or share the same IP address. The data set consists of a member base of some $110,000$ peers forming some 2972 swarms. The key idea in our measurements is the notion of a *multi-swarming* peer.

The results of this measurement are depicted in Figure 2. Due to the complexity of finding all possible cycles of interaction we have limited our measurements to cycles of length no greater than 3 where a cycle of the form $p_a \rightarrow p_b \rightarrow p_a$ between peers p_a and p_b is considered of length 2 and a cycle of the form $p_a \rightarrow p_b \rightarrow p_c \rightarrow p_a$ is of length 3. Figure 2 reveals a surprising characteristic of the supply/demand structure within BitTorrent communities. As depicted there is a high probability that multi-swarming peers can interact in cycles of length 2 and 3. This means that for example there is a high probability that within the community a peer p_a that is leeching in swarm s_1 and seeding in swarm s_2 (leecher-seeder) can be matched with a second peer p_b that is doing the exact opposite by seeding in swarm s_1 and leeching in swarm s_2 (seeder-leecher).
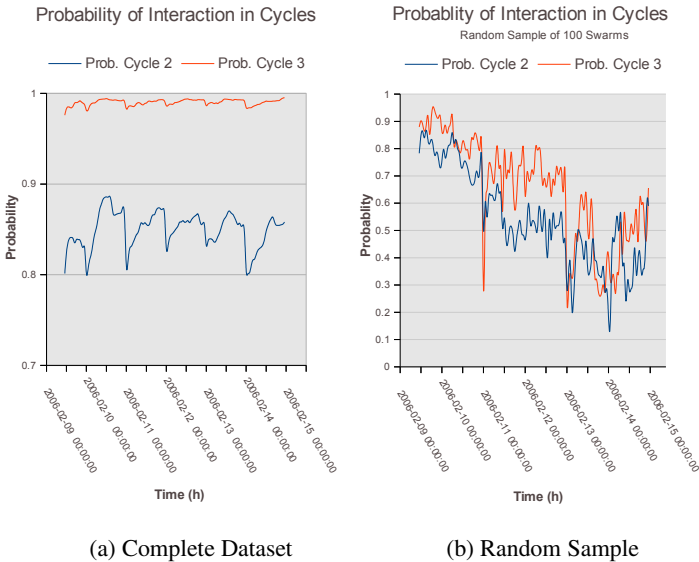

(a) Complete Dataset   (b) Random Sample

Fig. 2. Probability of multi-swarming peers forming cycles of length 2 to 3

Given that the *FileList.org* community implemented weak sharing ratio enforcement it is possible that our measurements overestimate the potential for indirect interaction. In order to address this issue we have conducted a series of measurements by randomly choosing 100 swarms from the data set and repeating our measurements with the select few. The random nature and small sample size provide us with a fair estimation of how probable indirect interaction is without sharing ratio enforcement. A sample result is depicted in Figure 2. Similar results have been consistently observed. This demonstrates that sharing ratio enforcement does not affect our overall probability measurements. Potential cyclic interactions seem to be an intrinsic property of file sharing behavior. Other measurement studies [6] also suggest observation of peers downloading multiple files (up to 80%) at a certain moment which lead to multi-swarming. However such studies use IP addresses as peer identifiers which result in overestimation. Our data set contains much lower numbers of multi-swarmers (average of 30%).

Also note that to make an estimate of the effectiveness of IIM, we are not interested in how often peers currently are multi-swarming in public swarms, but in their potential to multi-swarm in case an incentive for seeding is given. The *FileList.org* community provides such an incentive by enforcing a (weak) sharing ratio, and is thus quite representative of expected results with IIM.

5.2   Simulations

In order to verify Hypotheses 1, 2 and 3 we have simulated IIM with a modified version of *TriblerSim* [17] Our modifications allow us to simulate a simple setting of IIM which involves two swarms and two multi-swarming peers (one seeder-leecher and one leecher-seeder) which can have a cyclic interaction. Figure 3 demonstrates the types of scenarios that our tool is capable of simulating.

The upload bandwidth of peers is assumed to be the bottleneck; therefore all peers are assumed to have an unlimited download bandwidth. Communication between peers is assumed to be instantaneous (i.e. no network delay). A typical simulation starts with leechers having no pieces of the file. All peers start by announcing their presence to the tracker in a random order at time 0.

Demonstrated results have been obtained with two identical swarms of 1 seeder and 25 leechers. Seeders upload a file for ever while leechers disconnect as soon as they have finished downloading. Additionally there are 2 multi-swarmers that that are present in both swarms (making the total number of peers in each swarm 28). Each multi-swarming peer leeches one file and is in possession of an entire second file such that an indirect interaction cycle of length two is possible. Multi-swarming peers withhold the contents

Fig. 3. Possible simulation scenarios of the Indirect Interaction Mechanism

of their second file from other leechers. They also disconnect as soon as they have finished downloading the file. All peers have the same upload speed (homogeneous environment). The results have been obtained with 50 simulation runs. These simulations are mirrored with the exact same number of runs in which the multi-swarmers do not interact indirectly (act as standard leechers).
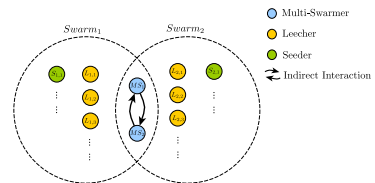
Hypothesis 1 and 2. Figure 4 plots the bootstrap time of one multi-swarming peer using IIM against one using standard BitTorrent. We have measured the bootstrap time of a peer as the time required to download 10% of the pieces. Figure 4 demonstrates that a multi-swarming peer bootstraps faster by expectation (Hypothesis 1). Hence our hypothesis can be confirmed as our paired t-test comparison of means demonstrates: $\alpha = 0.01, t(df = 62.3769) = -11.1429, P = 0.0000$. We attribute this observed effect to the fact that a multi-swarming peers does not have to wait for optimistic unchokes to receive an initial piece and can directly barter a piece of a file that it possess for the first piece of the file that it seeks.

It turns out that our expectations regarding download time are also met. Similar to Figure 4 multi-swarmers manage to stay ahead of standard leechers. Hypothesis 2 is confirmed with a paired t-test comparison of means $\alpha = 0.01,\ t(df = 50.3766) = -32.5107, P = 0.0000$.

Hypothesis 3. In order to test our hypotheses more rigorously we have also simulated IIM in limited other settings. We only briefly report on some of our observations in such settings due to space limitations and refer the reader to [20] for more details.

Fig. 4. Comparison of standard BitTorrent and IIM bootstrapping for the first multi-swarming peer

We have conducted a series of simulations with three classes of peers with upload bandwidths of $2048, 1024$ and $512 kbps$. In each simulation we have used two multi-swarmers as before, varied their class and compared their download times when using IIM and standard BitTorrent. Our observations indicate that when the classes of the multi-swarming peers are not too far apart or belong to the faster classes, the same advantages as before exist for using IIM even though one multi-swarmer is interacting with a slower multi-swarmer. This leads us to a conjecture that IIM improves download speed by expectation when the indirectly interacting parties have similar upload bandwidths (see Hypothesis 3).

An interesting class of results have been observed when the multi-swarming peers have the maximum difference in upload bandwidths. An example of such a simulation is demonstrated in Figure 5a. The observation here is that standard BitTorrent outperforms IIM in terms of download speed for the faster multi-swarmer but vice versa for the slower multi-swarmer. However, the performance loss for the faster peer is very small (rejected null hypothesis in t-test comparison of means). This indicates that even though download speed has been sacrificed, access to rare pieces has been able to compensate for much of the lost download speed. This indicates the possiblity of smart unchoking algorithms that can trade off between download speed and piece rarity in BitTorrent. Another series of our simulations in a homogeneous setting focuses on the effect of having more seeders in the swarm. The expectation here is that more seeders reduce the attractiveness of multi-swarming peers because of reduced piece rarity. Figure 6
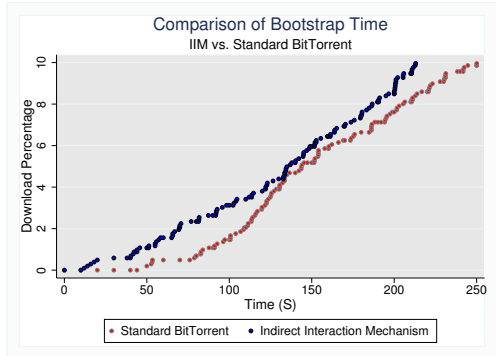
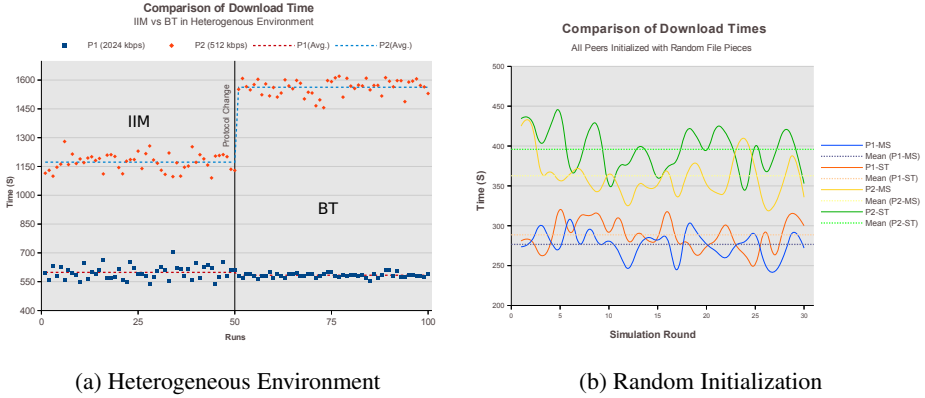(a) Heterogeneous Environment   (b) Random Initialization

Fig. 5. Comparison of standard BitTorrent and IIM for multi-swarming peers. 2 seeders with 2048 *kbps* upload; Other peers have 512, 1024 and 2048 *kbps* upload speeds. For Random Initialization $P1$ and $P2$ have same bandwidth.

demonstrates the ratio of one of our multi-swarming peer's download time with IIM to standard BitTorrent. While IIM's efficiency drops with more seeders, our hypothesis that IIM will not be outperformed holds. Finally Another series of simulation focuses on the scenario where the multi-swarming peers begin interaction at some later stage during their download process. This better matches real world scenarios. Figure 5b demonstrates the download time of two such peers. Starting interaction at a later stage gives the peers less time to capitalize on their access to rare pieces. As before our hypothesis holds. Here, P_1 and P_2 are the two multi-swarming peers. We compare their average download time with IIM and standard BitTorrent.

6   Discussion

BitTorrent (and generally P2P systems) have been studied game theoretically in numerous occasions [28,4].

    One of the more popular models for P2P systems is the prisoners' dilemma (PD) game which captures a social dilemma between cooperation or defection among two players. In PD games the rational strategy is a non cooperative choice even though not the optimal choice. One of the differences between this and previous studies is that IIM allows us to study incentives from the per-

Fig. 6. Efficiency of IIM with increasing ratio of $\frac{seeders}{leechers}$. (MS = multi-swarming, ST= standard BitTorrent).

spective of a larger group than 2 players in the context of a more general game referred to as the Public Good game (PG).

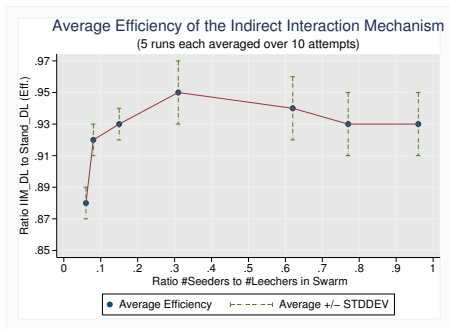PG can be considered a more general form of PD. Its most common form consists of players which have been bestowed some amount of private good (tokens). The public good is represented by a public pot in which tokens are deposited. All players have to decide simultaneously on the amount of contribution that they would make to the pot. Once the contributions have been made, the pot is multiplied by some factor and divided equally among all players as a payoff for their investment.

Note the similarity between the utility in PG and the BT and IIM model. IIM can roughly be modeled as a repeated PG in which the initial tokens correspond to the peers's bandwidths. This allows us to derive a claim on user behavior from social experiments conducted with PG among a group of students by *Milinksi et al*. They show that people react positively towards *potential* reciprocation [18] (see also [21]). This experiment engages players in altering games of PG and indirect reciprocity. After some rounds of play groups are divided into two series in which one series knows that no more indirect reciprocity games follow. As a result the first group's contributions drop to zero while the second group player's continue to contribute. This suggests that as long as players have some expectation that indirect reciprocity will follow they still contribute for a series of rounds. Therefore we expect that with IIM a small probability of indirect interaction to be sufficient for people to seeding more often.

7   Related Work

Many incentive mechanisms appear in the literature among which, reputation systems, sharing ratio enforcement and private communities are the most notable.

Reputation mechanisms such as EigenTrust [9] are based on the concept of each peer having a global trust value that reflects the past experiences of other peers in the network with that peer. The main idea is to generalize local trust values to a global scale by weighting against the reputation scores of the assigning peers. Other approaches are based on only local trust values. There are practical issues with the storage and complexity of computing the reputation.

A second notable reputation mechanism is BarterCast [17]. BarterCast is a fully distributed sharing ratio enforcement protocol that uses an overlay network for peers to communicate their sharing ratio. An epidemic protocol is used for peer discovery and exchange of random lists of peers. A reputation metric and a Max-Flow algorithm are applied to the local view of each peer to gain an aggregate subjective view of the reputation scores.

Most reputation mechanisms are vulnerable to attacks ranging from simple lying about scores to more sophisticated attacks. Reputation systems generally cannot prevent such attacks [19]. Also they have remained quite theoretical and for example both EigenTrust and BarterCast face practical problems of communication overhead, collusion of malicious peers and sybil attacks [3]. A main advantage of IIM over reputation mechanisms is that IIM creates potential benefits almost in real-time, while reputation mechanisms only provide some promising future. We refer the reader to [14,18,4,5] for fundamental research on reputation systems.

Perhaps the most commonly used incentivizing mechanism is *sharing ratio enforcement* [17] in *private communities* [16]. This approach resorts to creating quantifiable

metrics as a measure of cooperation. The cooperation metrics are used to create a standard of membership. Low standards result in disqualification. Despite all the advantages, private communities can lead to a phenomenon referred to as *BitCrunch* [7]. Informally, it refers to high bandwidth peers uploading most of the content while leaving little demand for the services of lower bandwidth peers. As a result it is becomes very hard for lower bandwidth peers to sustain high enough ratios to maintain memebrship or experience the benefits. This also leads to unfairness as these peers can experience long waiting times even though they are cooperative by definition and willing to share content. Another problem with private BitTorrent communities is their exclusivity.

Finally, other notable parts of the literature are analytical fluid models and game theoretical models of BitTorrent like systems. Fluid models [23,28] study the effect of churn and ratios of peer types on the properties of the systems and demonstrate that for certain scenarios Nash equilibria may exist that make uploading content more beneficial than freeriding. Other game theoretic approaches demonstrate strategies of better utilizing upload capacity [12,8].

8   Conclusions

We have outlined the Indirect Interaction Mechanism (IIM) as a mechanism for incentivizing seeding. The idea is that a peer can barter pieces of the file it has for pieces of the file it seeks. In addition, we show how this idea can even be extended to a cycle of contributions. IIM avoids most of the pitfalls of other incentivizing mechanisms: it is fair, it imposes reasonable overhead, and it can be used in public BitTorrent trackers. IIM creates close to real-time benefits for cooperating with other peers. The real-time nature of IIM makes it robust to malicious behavior.

The positive utility for seeders allows the use of the public goods game as a model for studying a group of peers. Social experiments with the game suggest that users react positively to indirect interaction given that some expectation of its possibility exist. Therefore, users are expected to choose IIM over standard Bittorrent.

Despite IIM being a centralized mechanism, it scales linearly with the number of multi-swarming peers [20]. The theoretically attractive solution would be to eliminate the centralized component in the system. There are two important aspects to decentralizing IIM: 1) Storage of Information regarding peer activities in swarms. 2) Finding potential indirect interaction partners.

A promising approach is the exploitation of gossip protocols such as ones utilized in BarterCast. Given a gossip protocol a multi-swarmer can individually discover potential partners in the gossiped information. We can speculate that there is a high chance that this type of information can be found because potential partners are both active within shared swarms and highly likely to receive information regarding each others' activities with low latency.

Distributed Hash Tables (DHTS) present another solution for the storage of information regarding peer activities in swarms. Peers would have to manually search the DHT for suitable indirect interaction partners.

References

1. Buragohain, C., Agrawal, D., Suri, S.: A game theoretic framework for incentives in P2P systems. In: IEEE Conf. on Peer-to-Peer Computing (2003)
2. Cohen, B.: Incentives build robustness in BitTorrent. Technical report (2003), bittorrent.org
3. Douceur, J.R.: The Sybil Attack. In: Druschel, P., Kaashoek, M.F., Rowstron, A. (eds.) IPTPS 2002. LNCS, vol. 2429, pp. 251–260. Springer, Heidelberg (2002)
4. Feldman, M., Lai, K., Stoica, I., Chuang, J.: Robust incentive techniques for peer-to-peer networks. In: Proc. of the 5th ACM Conf. on Electronic commerce, pp. 102–111 (2004)
5. Feldman, M., Papadimitriou, C., Chuang, J., Stoica, I.: Free-riding and whitewashing in peer-to-peer systems. In: Proc. of the ACM SIGCOMM Workshop on Practice and Theory of Incentives in Networked Systems, pp. 228–236 (2004)
6. Guo, L., Chen, S., Xiao, Z., Tan, E., Ding, X., Zhang, X.: Measurements, analysis, and modeling of bittorrent-like systems. In: Proc. of the 5th ACM SIGCOMM Conf. on Internet Measurement (2005)
7. Hales, D., Rahman, R., Zhang, B., Meulpolder, M., Pouwelse, J.: Bittorrent or bitcrunch: Evidence of a credit squeeze in bittorrent? In: 18th IEEE Int. Workshop on Enabling Technologies (2009)
8. Jun, S., Ahamad, M.: Incentives in bittorrent induce free riding. In: P2PECON 2005: Proc. of the 2005 ACM SIGCOMM Workshop on Economics of Peer-to-Peer Systems, pp. 116–121 (2005)
9. Kamvar, S.D., Schlosser, M.T., Garcia-Molina, H.: The eigentrust algorithm for reputation management in P2P networks. In: Proc. of the 12th Int. Conf. on World Wide Web, pp. 640–651 (2003)
10. Legout, A., Liogkas, N., Kohler, E., Zhang, L.: Clustering and sharing incentives in bittorrent systems. In: Proc. of the 2007 ACM Int. Conf. on Measurement and Modeling of Computer Systems, pp. 301–312 (2007)
11. Legout, A., Urvoy-Keller, G., Michiardi, P.: Rarest first and choke algorithms are enough. In: Proc. of the 6th ACM SIGCOMM Conf. on Internet measurement, pp. 203–216 (2006)
12. Levin, D., LaCurts, K., Spring, N., Bhattacharjee, B.: Bittorrent is an auction: analyzing and improving bittorrent's incentives. In: Proc. of the ACM SIGCOMM 2008 Conf. on Data Communication, pp. 243–254 (2008)
13. Locher, T., Moor, P., Schmid, S., Wattenhofer, R.: Free riding in bittorrent is cheap. In: SIGCOMM 2006 (2006)
14. Marti, S., Garcia-Molina, H.: Taxonomy of trust: categorizing P2P reputation systems. Computer Networks 50(4), 472–484 (2006)
15. Meulpolder, M.: Managing supply and demand of bandwidth in peer-to-peer communities. PhD thesis, Delft University of Technology (March 2011)
16. Meulpolder, M., D'Acunto, L., Capota, M., Wojciechowski, M., Pouwelse, J., Epema, D., Sips, H.: Public and private bittorrent communities: A measurement study. In: IPTPS 2010 (2010)
17. Meulpolder, M., Pouwelse, J., Epema, D., Sips, H.: Bartercast: Fully distributed sharing-ratio enforcement in bittorrent. Technical report, Delft University of Technology; Parallel and Distributed Systems Report Series (2008)
18. Milinski, M., Semmann, D., Krambeck, H.-J.: Reputation helps solve the 'tragedy of the commons'. Nature 415(6870), 424–426 (2002)
19. Nisan, N., Roughgarden, T., Tardos, E., Vazirani, V.V.: Manipulation-Resistant Reputation Systems. In: Algorithmic Game Theory, ch. 27. Cambridge University Press (2007)

20. Noroozian, A.: Incentivizing seeding in bittorrent - indirect interaction as an incentive to seed. Master's thesis, Delft University of Technology (2010)
21. Nowak, M.A., Sigmund, K.: Evolution of indirect reciprocity. Nature 437(7063), 1291–1298 (2005)
22. Piatek, M., Isdal, T., Anderson, T., Krishnamurthy, A., Venkataramani, A.: Do incentives build robustness in bittorrent. In: NSDI 2007 (2007)
23. Qiu, D., Srikant, R.: Modeling and performance analysis of bittorrent-like peer-to-peer networks. In: Proc. of the 2004 Conf. on Applications, Technologies, Architectures, and Protocols for Computer Communications, pp. 367–378. ACM, New York (2004)
24. Rockenbach, B., Milinski, M.: The efficient interaction of indirect reciprocity and costly punishment. Nature 444(7120), 718–723 (2006)
25. Roozenburg, J.: Secure decentralized swarm discovery in Tribler. Master's thesis, Delft University of Technology (2006)
26. Sigmund, K., Hauert, C., Nowak, M.A.: Reward and punishment. Proceedings of the National Academy of Sciences of the United States of America 98(19), 10757–10762 (2001)
27. Sirivianos, M., Park, J.H., Chen, R., Yang, X.: Free-riding in bittorrent networks with the large view exploit. In: 6th Int. Workshop on Peer-to-Peer Systems (2007)
28. Yu, J., Li, M., Wu, J.: Modeling analysis and improvement for free-riding on bittorrent-like file sharing systems. In: Int. Conf. on Parallel Processing Workshops (2007)