Gifting Technologies: A BitTorrent Case Study[♦]

Matei Ripeanu[1], Miranda Mowbray, Nazareno Andrade[2], Aliandro Lima[2]
Enterprise Systems and Software Laboratory
HP Laboratories Bristol
HPL-2007-26
February 19, 2007*

peer-to-peer,
gifting,
communities

This paper is concerned with gifting: giving not motivated by a direct, immediate, or obvious benefit. We analyze a popular technology used for gifting: the BitTorrent file sharing system. We determine features associated with high levels of gifting and suggest changes to the protocol and to the design of associated BitTorrent Web sites to promote it. We then extend our conclusions and suggestions to gifting technologies in general.

* Internal Accession Date Only
[♦]First Monday, December 2006
[1]University of British Columbia, Canada
[2]Federal University of Campina Grande, Brazil
Approved for External Publication

Gifting Technologies: A BitTorrent Case Study[1]

Matei Ripeanu, Miranda Mowbray, Nazareno Andrade, Aliandro Lima

Abstract

This paper is concerned with gifting: giving not motivated by a direct, immediate, or obvious benefit. We analyze a popular technology used for gifting: the BitTorrent file-sharing system. We determine features associated with high levels of gifting and suggest changes to the protocol and to the design of associated BitTorrent Web sites to promote it. We then extend our conclusions and suggestions to gifting technologies in general.

1. Introduction

With the advent of peer-to-peer technologies collaborative resource usage has reached amazing popularity. Millions of members of peer-to-peer file-sharing communities share the content of digital files (Oram, 2001). Similarly, millions of members of peer-to-peer computing projects share the computing power of their desktops (Chien, 2004).

In such communities resource owners can allow other community members to use their resources. Often, this transfer of usage rights does not bring a direct benefit to the contributor. This model of collaborative resource usage has, however, limited efficiency: over time some such communities selectively attract members who *freeride,* that is, they consume resources without contributing. For example, (Hughes et al., 2005) report that up to 85 percent of members of the Gnutella file sharing community freeride.

Two complementary approaches can be used to encourage resource contribution and reduce freeriding levels. First, in traditional economic models based on reciprocation members are rewarded in kind for making resources available to their community. For example they gain the ability to exploit resources which they would otherwise be unable to access. Thus members' desire to access community resources is harnessed to motivate them to contribute their own resources.

Second, designs can cater to motivations of members other than that of receiving direct rewards for sharing their resources. In a First Monday paper, (McGee and Skågeby, 2005) show members of a file-sharing community contributing resources without direct, immediate or obvious beneft to themselves. They call this behavior *gifting*, and distinguish it from *trading*, that is, contribution motivated by the likelihood of immediate reciprocation.

While in this paper we analyze some trading-based solutions, our main aim is to identify factors associated with increased gifting in resource sharing communities. To do this, we study content-sharing communities powered by the BitTorrent protocol.

In BitTorrent communities, in addition to the digital content, the bandwidth enabling access to content is a second shared resource. New content is not directly transferred between the peer that introduces it to a community and each peer interested in downloading it. Instead, as we explain in Section 33, multiple peers contribute to content distribution resulting in lower network load at the content originator.

This collaborative mechanism has made BitTorrent a popular solution for distributing files from a single source to a large number of recipients. For instance, RedHat Inc. use it successfully to distribute RedHat, one of the most popular Linux distributions (Izal et al., 2004).

Members of a BitTorrent community can contribute by making new content available; by contributing bandwidth while they download a file; or by contributing bandwidth after they have obtained the whole file. We investigate how these types of contributions vary between different BitTorrent communities and between different members of the same community. We find significant variation in the level of contribution in both cases:

- The wide variation in the level of contributions by members of the same community support the hypothesis that individual members' motivations to contribute are diverse, complex, and contingent. Nevertheless, in comparison to file-sharing communities using other protocols (such as Gnutella or eDonkey), freeriding is rare in all the communities we study. Since the applications and demographics in these communities are similar, the low level of freeriding suggests the design of the BitTorrent protocol is successful in attracting higher overall user contribution levels.
- We link the variation in contribution levels across communities to different, community-specific motivations for gifting, in some cases successfully encouraged by mechanisms deployed by each community.

While some of our BitTorrent data has been previously published (Andrade et al., 2005), our analysis of the implications of this data for the support and promotion of gifting is new.

This paper is organized as follows. Section 2 surveys motivations for contributing, and their interaction. Section 3 identifies technological and social factors associated with high contribution levels in BitTorrent: technical features that make contributing easier or more effective, economic incentives for contributing, default contributing behavior by the software, legality of the digital content shared, and a culture of sharing within the community. Section 3.5 presents BitTorrent-specific suggestions to increase contribution levels and Section 4 presents the implications of our findings for the design of gifting technologies in general. In particular, it focuses on the "economics of effort": the somewhat obvious, but powerful, idea that people are more likely to gift if it is easy for them to do so, or if it costs them some effort *not* to gift.

2. Motivations for contributing, and the crowding-out effect

In many communities, and file-sharing communities are no exception, most members value the goods obtained during participation higher than their participation costs. Thus, when these members contribute to the community, their contribution can be explained as trading, rather than gifting.

For gifters however, the perceived reward is often tied to the impact of their contribution, e.g., the number of other participants that can benefit from it. To support high levels of gifted contributions a system designer has to understand the characteristics and consumption patterns of shared system resources as well as gifters' motivations.

(Frey and Jegen, 2001) divide motivations for giving into *extrinsic* and *intrinsic*. Extrinsic motivations are external to the giver, such as financial reward, reciprocated gifts, fame, social capital, the avoidance of punishment, or recognition by friends. Intrinsic motivations are driven by the giver's own psychology, ideology and spirituality - such as reinforcement of the giver's self-valuation as a good and generous person, or promotion of a cause the giver supports.

A cause promoted by a gift might be specific to the gift, such as the widespread diffusion of a particular piece of software or a particular musical subgenre, or might be more general, such as the well-being of a particular community or the good of humanity.

Frey and Jegen survey the evidence for the *crowding-out* effect: the presence of extrinsic motivations results in decreased intrinsic motivation. For example, only a considerable financial reward motivates schoolchildren who collect money door-to-door for charity to be as efficient as volunteer groups.

Frey and Jegen present convincing empirical evidence that the crowding-out effect exists, and underscore the strong crowding-out effect of task-contingent tangible rewards. However, they report that a *crowding-in* effect can also occur: extrinsic and intrinsic motivations can reinforce eachother.

Whether the introduction of an extrinsic incentive results in lower or higher intrinsic motivation depends on the properties of the extrinsic motivation, and is historically and culturally contingent. However, Frey and Jegen conclude that external interventions crowd-out intrinsic motivation if the participants perceive them to be controlling, rather than supportive.

However, even with a crowding-out effect, introducing an external incentive may not reduce overall motivation strength as long as the resulting extrinsic motivation outweighs the crowded-out intrinsic motivation. If there is little or no intrinsic motivation for giving, increasing external incentives for giving should increase the amount of giving, as standard economic theory predicts.

(Benkler, 2004) considers a particular class of resources which exhibit systematic overcapacity. These resources are "large enough so that each unit has systematically more capacity than one person requires over the lifetime of the resource, and small enough that one person can justify putting a unit into operation" (p.357).

One example is seat space in a car, which is shared in car pools. Other examples are desktop computer processing power, bandwidth, and storage. Although Grid computing (Foster and Kesselman, 2004) might change this situation, at present most consumers acquire computing resources by buying a computer, network connection, and storage media with capacities provisioned for their peak needs. Most of the time, however, these resources are not used at full capacity.

Benkler argues that the overcapacity for this class of goods may in some contexts be more effectively harnessed using social sharing rather than markets, because social exchanges have lower transactional costs and can produce and make use of more detailed, textured information than markets can. Benkler also mentions that markets may crowd-out intrinsic motives for sharing.

The existence of the crowding-out effect is a potential issue for the design of file-sharing systems. Common reasons for file-sharing include intrinsic motivations, and some of the mechanisms suggested to promote sharing introduce an external incentive which is intended to be controlling. It is not clear whether or not such incentives will produce a strong enough extrinsic motivation to make up for any crowded-out intrinsic motivation. Therefore, the effect of such incentives needs to be monitored by sharing communities that adopt them, to check the incentives are not counterproductive.

The next section presents a detailed analysis of BitTorrent—one of the most successful file-sharing systems in terms of both adoption and contribution levels. We compare the contribution levels among different BitTorrent communities, one of which has an extrinsic incentive ("sharing ratio enforcement") which is intended to be controlling. As we will see, this community does present significantly higher contribution levels, so any crowding-out effect appears to be compensated for by the increased extrinsic incentive.

3. BitTorrent

File-sharing in BitTorrent works as follows. A user who aims to make a file available to a community divides the file into smaller *chunks*, publishes the file details on a Web server, and creates a *tracker* which records the locations of chunks. To download the file, a new peer uses the tracker to find the locations of the chunks making up the file.

As the peer downloads chunks of the file, it also updates the tracker with their additional locations and may upload to other peers the chunks it has already downloaded. Thus, the burden of bandwidth consumption is moved from the original content originator to all the peers interested in downloading the file.

The distinguishing feature of BitTorrent is its incentive mechanism to encourage cooperation. Peers reciprocate: a peer is most likely to upload (i.e., to serve content) to those peers that served it best providing high download rates. This gives an incentive for peers to cooperate by uploading to the peers they are downloading from.

To search for peers that reciprocate best, each peer periodically initiates uploads to another peer from whom it has not recently received anything (*optimistic unchocking* in BitTorrent parlance). Additionally, this mechanism also allows new peers yet have not acquired any chunk of a file to obtain new content and thus to participate in chunk trading.

In BitTorrent lingo, a *torrent* is a group of peers trading chunks of the same file and using the same tracker. Current BitTorrent implementations allow trading only within the same torrent. This has the advantage of simplifying trading and the accounting mechanism but the significant drawback of segmenting the markets for trading chunks.

The use of a server to publish and locate content makes BitTorrent users gather around Web portals. These portals often specialize in a specific type of content (e.g., creative commons-licensed media or a specific genre of videos) and the community of users of a particular portal may adopt community-specific mechanisms for encouraging contribution. As we will discuss later, some BitTorrent communities have introduced accounting schemes to keep track of contributions across multiple torrents.

We are interested in comparing the collaboration patterns in different BitTorrent communities. To this end we investigate two opposite behaviors: freeriding and seeding.

- ***Freeriding***. A *freerider* downloads but does not upload any data. Freeriders aim to avoid the bandwidth and computational cost of file uploading (Feldman et al., 2003), or, they see uploading as more immoral or riskier than downloading in the case of copyright-infringing content. Additionally, firewalls impact peers' ability to contribute: if two peers are behind firewalls, they cannot exchange data and they might appear to freeride.

  To characterize the freeriding at the community level we define the *freeriding ratio* as the percentage of peers that are freeriders. The low freeriding ratios we see in existing BitTorrent communities show that BitTorrent is successful in boosting cooperation (Section 3.2).

- ***Seeding***: A *seeder* is a peer that has finished downloading but is still connected to the torrent and might upload to other peers. A torrent benefits from seeders as they generally increase content availability and download rates. BitTorrent sites and client software exhort users to seed; however, no incentive for seeding is present in the protocol, thus for BitTorrent communities without additional incentives, seeding is a form of gifting.

  To characterize the seeding behavior for an entire community we define the *seeding ratio* at a given time as the percentage of all peers that are seeding at the time. We compute this ratio for snapshots taken of different BitTorrent communities, and suggest reasons for the significant differences we find (Section 3.30).

Our study quantifies freeriding and seeding levels across six BitTorrent communities and detects which extensions to the protocol successfully discourage freeriding and boost seeding.

3.1. Methodology

Our data comes from six BitTorrent communities: *bt.etree.org*, *piratebay.org*, *torrentportal.com*, *easytree.org*, *btefnet.net* and *alluvion.org*. For the rest of this paper we refer to these as *etree*, *piratebay*, *torrentportal*, *easytree, btefnet*, and *alluvion*. Table 1 summarizes the main characteristics of these communities.

	Content type	# of torrents	# of peers	Data about individual peers
etree	Music	567	4,492	Yes
easytree	Music	2,586	25,687	Yes
piratebay	Films, music	13,054	320,900	No
torrentportal	Films, music, etc	10,115	357,428	No
btefnet	TV episodes	476	78,897	No
alluvion	User forum content	204	2,807	Yes

Table 1: Characteristics of the communities analyzed

We collect data by crawling the public torrent report pages for each community. Each crawling provides a snapshot of the community at a given moment. Since we obtain data from a large number of torrents of different ages, we expect to have a representative sample from the different stages of a torrent's life.

We collect data on all the torrents that are active at the time of sampling, have at least three peers with upload or download greater than zero, and have at least one seeder, so that there is at least some sharing going on and peers are able to upload. At each site, for each torrent we collect its corresponding file size, torrent age in days, number of participating peers, and the total volumes uploaded and downloaded.

In addition, as Table 1 shows, for *etree, easytree* and *alluvion* we have obtained data about the state of each participating peer: the amount downloaded and uploaded and whether the peer is a seeder. In the case of *etree* and *easytree* we also know whether the peer is *connectable*, that is, whether the tracker is able to open a connection to it. The *etree* and *alluvion* data is public, while the *easytree* data was provided by the system administrators.

3.2. Freeriding Behavior in BitTorrent

As Table 2 shows, freeriders and low-sharing peers are much rarer in BitTorrent than in other peer-to-peer communities such as Gnutella or eDonkey where freeriders can amount to as much as 85 percent of the peer population (Adar and Huberman, 2000; Saroiu et al., 2002; Fessant et al., 2004; Hughes et al., 2005; Stutzbach and Rejaie, 2006). It appears therefore that the design of the BitTorrent protocol does result in increased cooperative behavior.

	etree	*easytree*	*alluvion*
Characteristics of overall peer population:			
▪ Share of connectable peers	49%	60%	not available
▪ Peer age	Median: 10.6h	65% less than 24h old	Median: 6.4h
Characteristics of *freerider* population			
▪ Share of freeriders (peers with no upload)	6%	5%	2%
▪ Share of connectable freeriders	19%	41%	not available
▪ Freerider age	Median 3.3h	72% percent less than 24h old	Median 1.0h
Low-sharing peers (0.25 threshold)	28%	24%	23%

Table 2: Characteristics of freerider and overall peer population at *etree, easytree*, and *alluvion*.

Table 2 summarizes the freeriding characteristics we measure at the three sites for which we have individual peer data: *etree, easytree* and *alluvion*. The key fact is that for each of the three sites the level of freeriding is extremely low: at most 6 percent of peers do not upload anything. (The differences in freeriding ratios between the sites are not statistically significant at the 0.01 significance level).

To a large degree this is explained by the difficulty of contacting these peers from outside: 81 percent of the freeriders at *etree* and respectively 59 percent at *easytree* are behind firewalls. In addition, at all three sites the group of peers that have not uploaded anything includes disproportionately many peers of low age. If a peer has not downloaded any chunks needed by other peers in the torrent, then it will be unable to upload data to them.

Thus peers that have not uploaded data are not necessarily trying to freeride: they may be willing to upload but unable to at present. Indeed, some peers that had uploaded nothing at the time of our first *etree* sample, uploaded data later.

Most popular BitTorrent client implementations allow users to set the maximum upload rate, and, while they do not allow setting it to zero they do allow setting it to a low value. As a result, a peer who wishes to minimize its contribution to the system may attempt to use this setting to upload only a small amount of data relative to the amount it downloads.

To investigate the frequency of this behavior we introduce the *sharing ratio*, defined as ratio between the upload and download volumes of a peer in a torrent. Obviously, the sharing ratio is related to freeriding: a freerider has a zero sharing ratio.

A number of BitTorrent community administrators have decided to tackle what they see as unacceptably low contribution levels and to enforce lower limits on sharing ratios. In these communities, members with sharing ratio lower than a specific threshold cannot join new torrents. For example, *easytree* uses a fixed threshold of 0.25 to define low-sharing peers subject to sharing ratio enforcement.

In such communities the member's sharing ratio is calculated by considering the total upload and download made in all the torrents the member has participated in, whereas our measurements only consider the upload and download made in a single torrent. However, our measurements give some idea of likely sharing ratios over multiple torrents.

Since a fixed low-sharing threshold can be seen as arbitrary, Figure 1 presents the distribution of peers in the three communities according to their sharing ratios. We believe the slightly higher levels of cooperation at *easytree* are explained by the use of sharing-ratio enforcement against low-sharing peers. As we will show later, *easytree* also has significantly higher levels of seeding.

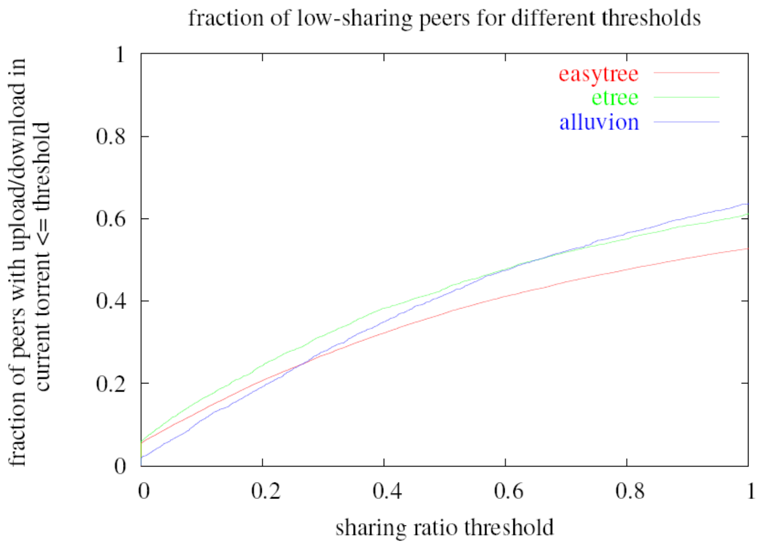

Figure 1: Cumulative distribution of sharing ratio for the three sites where we are able to collect individual peer data. To make the plot readable in the region of interest we limit we plotting these distribution only for sharing ratios lower than 1.

In (Andrade et al., 2005) we present empirical evidence that, under certain conditions, largely when there are a large number of seeders in a torrent, freeriders have faster access to content than well behaved participants. However, in the communities we survey, we find low levels of freeriding even in these conditions.

We attribute this situation to two factors: a psychological effect and the economics of effort. Firstly, the general user perception is that BitTorrent economic incentives to penalize freeriders are always working, thus the psychological barrier to attempt to freeride is high. Secondly, most popular BitTorrent clients cannot be configured for freeriding and it takes considerable effort and skill to modify the source code of a client to change it to a freerider.

3.3. BitTorrent and Seeding

BitTorrent files are usually published on Web sites that consist of listings of torrents and HTML links to their associated trackers. An important distinction between BitTorrent and other file-sharing systems is the use of this centralized location mechanism. This centralized architecture has obvious drawbacks: a single point of failure/attack, potential bottlenecks, and high operational costs. However, it allows the implementation, at the site level, of new features to compensate for the lack of incentives for seeding in the BitTorrent protocol.

In the rest of this section we present two community level solutions to encourage higher seeding ratios and present empirical evidence of their effectiveness.

- **Sharing-ratio enforcement.** As mentioned above, some BitTorrent communities, including *easytree*, *empornium.us*, and *pwtorrents.net*, periodically enforce that the sharing ratios of participating members are above a minimum value. These sites keep a long-term history of contributions, and prevent members whose sharing ratio is below a certain threshold from gaining access to new content. In some cases this decision mechanism also takes member “age” into account. Since a member can efficiently increase her sharing ratio by uploading content as a seeder, the mechanism provides an indirect incentive for seeding as well as a direct incentive not to freeride.

- **Broadcatching.** Several BitTorrent Web sites use Really Simple Syndication (RSS) feeds (Bond, 2005) to advertise newly published files. *Broadcatching* is the use of BitTorrent clients to automatically download files advertised through RSS feeds. For example, a user may subscribe to an RSS feed of a site that publishes past episodes of television series, and state interest in any new episode from a particular series. Whenever the RSS feed announces matching content, the user's client will download it automatically. In the time between finishing a download and the user checking to see whether new files have arrived, the client remains connected as a seeder. Thus, as a side effect, broadcatching results in increased seeding and sharing because users maintain their clients running for longer time.

	Share of seeders
piratebay	28%
torrentportal	31%
alluvion	42%
etree	52%
btefnet	55%
easytree	59%

Table 3: Seeding ratios for the six BitTorrent communities we investigate

Table 3 shows the seeding ratios for the six communities we investigate. The differences in seeding ratios can be attributed to a combination of social and technical factors:

- *Communities with undeniably legal content see higher seeding ratios*. Among the sites that do not use extra mechanisms to increase cooperation (*etree, piratebay, alluvion,* and *torrentportal*), the sites that enable distribution of only legal content (*etree* and *alluvion*) have larger seeding ratios.
- *Broadcatching increases the seeding ratio. Btefnet* torrents have significantly more seeding than those in *piratebay* and *torrentportal*. The use of broadcatching may explain this.
- *Sharing ratio enforcement increases seeding ratios. Easytree* torrents have significantly more seeding than those at any other sites. We attribute this to the sharing ratio enforcement: this is an extrinsic incentive which does not appear to crowd out other motivations.
- *The social characteristics of different communities strongly influence sharing behavior*. The moderator of *easytree* attributes the relatively high amount of cooperation to the sharing culture among offline consumers of the type of content (bootleg recordings) distributed via *easytree*. When a large group of new users who did not share this culture joined *easytree* the sharing ratio went down, and site administrators successfully introduced sharing-ratio enforcement to reverse its decline.

These differences in seeding between sites are all significant at the 0.01 significance level. We have verified using partial regressions that they are not explained by differences in torrent age or in file size.

At all six sites we investigate, torrent seeding ratio is negatively correlated with the file size file (i.e., torrents sharing small files display higher seeding ratios). We believe this is explained by the fact that the time peers remain online after they finish downloading is independent of the file size. Small files take less time to download. Thus, the fraction of their total time connected to the torrent that peers spend seeding is larger in torrents that share smaller files.

Economics of effort may influence seeding behavior, just as it influences freeriding. Unlike Gnutella and other file-sharing client software, at the time of our study, most BitTorrent clients do not have an option to leave the file-sharing network automatically as soon as a download has finished. Even after a download is completed the client continues to seed until the user manually disconnects it. It is likely that this increases participation and seeding.

3.4. Contributions in the same community

Besides comparing peer behavior across communities, we look at how contribution is distributed over peers in the same community. We examine the distribution of content contributions for *etree* and *piratebay*, and the distribution of upload volumes for a 21-day trace of *etree*. A user contributes content by publishing a new torrent at a site—a gifting behavior.

For *etree*, the 16 percent of users who contributed most are responsible for the majority of content, while for *piratebay* this value is 9 percent. Figure 2 shows the cumulative distribution of content contribution for users which have contributed something at two sites. As the Figure shows, in both communities around 95 percent of the contributors publish at most 10 new torrents, while a minority of contributors publishes a large number of torrents.

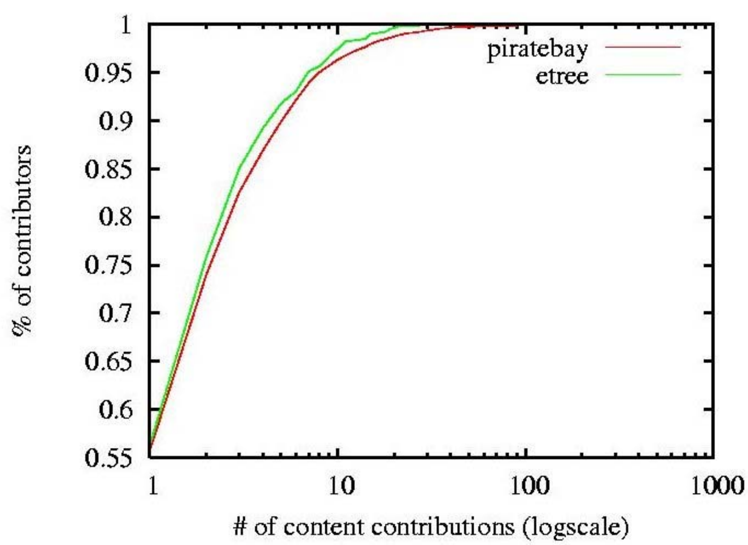

Figure 2: Cumulative distribution of the number of content contributions among users who have contributed some content in *etree* and *piratebay*. The x axis is logscale.

We now turn to investigating the distribution of upload volumes among peers. *Etree* uses semi-anonymized IP addresses as identifiers. As a result, it is not possible to identify a peer across all torrents that it took part in and measure the total upload volume over all these torrents. We therefore look at how the upload volume is distributed in each torrent.

To this end, we select from our 21-day trace the torrents with file sizes between 600 and 800MB, so as to eliminate the effect of the correlation between file sizes and seeding levels. Next, we select the torrents with at least 30 peers, out of which at least 10 are seeders and at least 10 are peers that download content (leechers in BitTorrent parlance). This results in 181 selected torrents.

Figure 3 presents the total upload volume contributed by the 10 percent of most active uploaders considering only seeders, only leechers, and all peers together.

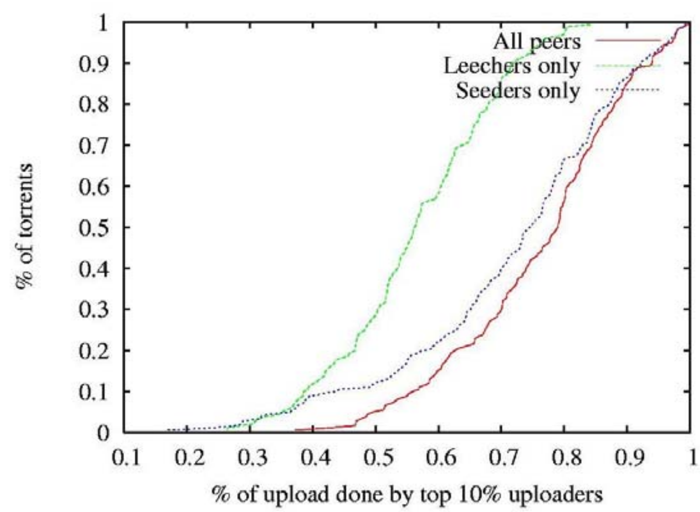

Figure 3: Percentage of torrents in which the traffic volume uploaded by the top 10 percent of uploaders is less than a given percentage of the entire contribution volume**.** Data is from a 21-day trace of *etree*. All torrents are for files with sizes between 600 and 800 MB.

Figure 3 shows that the top 10 percent of uploaders upload less content than the total uploaded by the rest of the peers in the torrent in only 5 percent of the torrents we examined. For half of the torrents, these same 10 percent of uploading peers contribute more than 79 percent of all the bandwidth used in the torrent.

The data for seeders only reveals a similar behavior; in most of the torrents, a small percentage of seeders uploads more than all other seeders put together. When we consider only leechers—who are trading their bandwidth—we see that while a minority of leechers dominates uploads by the leechers, the distribution is less extreme.

We attribute these non-uniform distributions primarily to heterogeneity of users' bandwidth. However, the distribution of bandwidth available for contribution by leechers and by seeders should be similar. Thus, if it were the only factor determining the amount of contribution, we should see similar distributions of upload volumes for both types of peers. We attribute the more extreme distribution for seeders to the effect of a large variation in the time that users leave their client online after having finished downloading.

Note that the skewed distribution of uploading and content contribution happens both within a group of peers (the leechers) who earn a direct reward for their contribution, and within a group (the seeders) who do not. In *etree* and *piratebay* there is no incentive mechanism for content contribution or upload by seeders. Upload by leechers is an example of trading.

This is evidence that incentives provided by the sharing mechanisms are not the only constraints on the distribution of contribution levels for trading and gifting in BitTorrent communities.

3.5 Suggestions for BitTorrent

In this section we summarize BitTorrent-specific suggestions to increase contribution levels inspired by our own BitTorrent experience as well as by a literature survey. Section 4 looks at general lessons that can be applied at most peer-to-peer communities to enhance gifting.

Since gifters are sensitive to the efficiency with which their contribution is used (Kerr, 1996), a number of suggestions aimed at improving the efficiency of underlying BitTorrent mechanisms are likely to impact gifting levels. These suggestions include the use of chunk ordering to solve the 'last block problem' (Gkantsidis and Rodriguez, 2005), and minimizing network overheads by intelligently pairing up the peers that trade chunks (Sherwood et al., 2004; Bindal et al., 2006).

The time peers spend online is one key characteristic influencing the level of gifting (seeding in this case) in a torrent. If peers spend a long time online, this also has a positive effect on long-term content preservation and availability (Adar, 2005; Guo et al., 2005).

(Guo et al., 2005) point out that client software that allows participation in multiple torrents concurrently will increase overall participation time in all torrents. This is because all torrents share the same, limited capacity network link; thus file downloading in each torrent will take longer. The resulting increased participation time will improve content availability and seeding in the torrents corresponding to small files.

To increase average online peer time, (Adar, 2005) suggests, counter-intuitively, that the original content source and the possible seeders reduce their upload rates according to their estimate of how frequently new peers join the torrent, so as to ensure a minimum level of participation in the torrent. Peers already participating in the torrent will need to spend more time online to download files if data sources are slower, thus providing improved content availability.

In the communities we study, torrents corresponding to smaller files have higher seeding ratios. Our explanation of this is that the time after a download completes and until the user decides to take the client offline is not closely related to the size of file distributed by the torrent. Thus, we suggest two ways to increase seeding time: first, modify the default client behavior to remain active online, seeding, for a time proportional to the size of the file that has been downloaded. Second, community administrators can recommend a maximum file size to be published.

4. Suggestions for Gifting Technologies

While this paper has so far focused on BitTorrent, our findings have general implications for gifting technologies. In the rest of this section we suggest ways to increase the level of gifting in a community, arising from our study of BitTorrent.

- *Promote both gifting and trading*. In BitTorrent communities, gifting and trading coexist. At the BitTorrent sites in our study other than *easytree*, there is no incentive to seed. Nevertheless seeding, a good example of gifting, occurs in these sites, together with trading in the form of uploading of content in order to download. *Easytree*'s extrinsic incentives do not appear to crowd out other motivations, and both types of motivation can be present for a single peer.

  Designers of gifting technologies should therefore not assume that users are exclusively interested in gifting or in trading, and should aim, while monitoring for possible motivation crowding, to concurrently encourage both types of contribution.

- *Support cultural norms for gifting*. According to *easytree* moderators, the presence (or absence) of a cultural norm for 'sharing' has a noticeable effect on the levels of altruistic contributions.

  More generally, designers of gifting technologies should consider adding features to support social and cultural norms that promote gifting. For instance, there are notices on BitTorrent sites and clients which try to create a social norm by exhorting users to keep their clients running after they have finished downloading. It is plausible that these have a positive effect on the amount of seeding.

- *Take prolific gifters into account.* Our measurements of *etree* revealed that even in the presence of an incentive mechanism, a small fraction of all peers is responsible for the majority of resources provided. Among gifters, this concentration was even more pronounced.

  This observation should be considered by designers. It is important not to assume that all peers will be able to contribute equally, and the efficient use of the resources owned by the small fraction of gifters who contribute most is likely to be a key issue in the efficiency of the system. Similarly, encouraging these gifters to keep participating in the system is important to assure a sustained level of contributions. Of course, any incentive provided to such users should be carefully designed so as not to crowd out intrinsic motivations for gifting.

Our study of BitTorrent communities supports the idea of "economics of effort" applied to gifting: that the easier it is for users to gift, the more likely they are to do so. The rest of our suggestions arise from this idea.

- *Make gifting the default*. Economics has frequent examples of the power of the default choice. For instance, the Economist magazine reported that when an American firm changed their employee pension plan from opt-in to opt-out the level of enrolment jumped from 49 percent to 86 percent (*Economist* staff writer, 2005). Similarly, in communities based on gifting, one way to reduce the effort needed to gift is to make gifting the default, so that it takes some effort not to gift.

For BitTorrent clients, uploading while downloading is the default; changing this behavior requires changes to the client source code. We believe that this is an important factor in keeping freeriding in BitTorrent to a low level.

- *Increase the effectiveness of gifting*. Some of the suggestions for improving the BitTorrent protocol in Section 04 are concerned with reducing protocol BitTorrent overheads, and, as a result, increasing the amount of useful gifting that takes place for a fixed amount of contributed resources. Since community members are generally more likely to gift if they perceive that their gifts are efficiently used and have a strong positive effect on the recipient, increasing effectiveness may also increase peers' motivation to gift.
- *Make gifting safe*. Our data shows that users can be reluctant to seed files with illegal content. The burden of legal, ethical, or security-related risks as a consequence of gifting can be seen as a particular type of effort required for gifting.
- *Make gifting a byproduct of actions carried out by users for their own benefit*. Users' contribution to the Napster index occurred as a natural byproduct of their actions carried out for their own benefit, e.g., when downloading songs to a portable music player or burning a CD (Bricklin, 2001). Thus, in Napster, gifting occurred with no effort from the gifter at all.

  Similarly, the seeding that occurs in the time between a *btefnet* client finishes downloading new content and the user noticing that the content has been downloaded is another example of gifting as a byproduct.
- *Allow different kinds of gifting with different effort levels*. Effortless gifting, or gifting by default, is not always possible. However, designers of gifting technologies generally have the option of enabling different types of gifting, some of which require limited effort, so that users can choose their gifting levels.

  This can be seen in BitTorrent: most peers participate in chunk trading, a smaller proportion of them are seeders, and an even smaller number provide original content. Another example of a gifting technology with differing effort levels is online newsgroups: Bradley Horowitz of Yahoo! speculates that only about one percent of Yahoo! user population might create a new group while ten percent might add content to an existing group (Horowitz, 2006).

  Naturally the rewards and incentives that motivate contributions should be commensurate with the effort involved. *Easytree's* sharing ratio enforcement mechanism has, in some cases, the side effect of creating a motivation to upload original content, a type of contribution that requires more effort than simply uploading existing content. A member whose sharing ratio is below the low sharing threshold cannot join new torrents. In this situation, one way for the member to raise its sharing ratio is to contribute new content. In this case there is a substantial reward for the extra effort: regaining the ability to join new torrents.

5. Summary

In BitTorrent file-sharing communities users contribute by introducing new content to their community or by serving content downloaded from other users. Reciprocation and gifting coexist as motivations to contribute.

Our analysis of traces from six BitTorrent communities show that BitTorrent chunk trading mechanisms are efficient in discouraging freeriding and ensuring a minimal level of participation from each peer. BitTorrent communities have significantly lower levels of freeriding compared to other communities that have similar content and user demographics but are supported by other peer-to-peer file-sharing software, e.g., Gnutella or eDonkey.

Additionally, in the BitTorrent communities we study, we detect significant levels of gifting. The variation in gifting levels across communities is linked to community-specific technical and social factors, including incentive mechanisms deployed by some of the communities. Our data for the distribution of contributions within a single community shows that a minority of gifters in a community are responsible for most of the gifting.

We have used this analysis to suggest changes to the protocol and to the design of associated BitTorrent Web sites to promote gifting, and we have also presented some suggestions for gifting technologies in general.

6. About the Authors

Matei Ripeanu is an Assistant Professor with the Computer Engineering Department of University of British Columbia. He is broadly interested in large-scale distributed systems: peer-to-peer and Grid systems. Matei received his PhD in Computer Science from The University of Chicago.

Miranda Mowbray works at Hewlett Packard Laboratories, where she researches online communities. She co-founded *e-mint*, the association of online community professionals. Her PhD is in Algebra, from London University.

Nazareno Andrade is a PhD student at Universidade Federal de Campina Grande, researching large-scale resource sharing. He is interested in content distribution mechanisms, online communities and peer-to-peer systems, and works in the OurGrid project.

Aliandro Lima is a Research Assistant in the Distributed Systems Laboratory at Universidade Federal de Campina Grande. He holds a MSc degree in Informatics from the same University, and is currently working on an asynchronous communication infrastructure for distributed systems, within the OurGrid project.

7. Acknowledgements

The authors would like to thank the *easytree.org* administrators Jupiter and Erwe for their helpful discussions and data. This work was partially developed in collaboration with HP Brazil R&D.

8. References

Economist staff writer, The Economist, 2005, "Pensions by Default," August 25, 2005, http://www.economist.com/finance/displayStory.cfm?story_id=4316190.

Eytan Adar, 2005. "Drawing Crowds and bit welfare," *ACM SIGecom Exchanges,* volume 5, number 4, pp. 31-40.

Eytan Adar and Bernardo A. Huberman, 2000. "Free Riding on Gnutella," *First Monday,* volume 5, number 10, pp. 1-20.

Nazareno Andrade, Miranda Mowbray, Aliandro Lima, Gustavo Wagner and Matei Ripeanu, 2005, "Influences on Cooperation in BitTorrent Communities," in proceedings of 3rd ACM Workshop on Economics of Peer-to-Peer Systems (P2P Econ), Philadelphia, PA, pp. 111-115.

Yochai Benkler, 2004. "Sharing Nicely: On shareable goods and the emergence of sharing as a modality of economic production," *Yale Law Journal,* volume 114, number 273,

Ruchir Bindal, Pei Cao, William Chan, Jan Medval, George Suwala, Tony Bates and Amy Zhang, 2006. "Improving traffic locality in BitTorrent via biased neighbor selection," in proceedings of 26th IEEE International Conference on Distributed Computing Systems (ICDCS 2006), Lisbon, Portugal, pp. 66-80.

John Bond, "RSS Frequently Asked Questions", www.rss-specifications.com/rss-faqs.htm, accessed 8/8/2006.

Andrew Chien (2004). Massively Distributed Computing: Virtual Screening on a Desktop Grid, in: The Grid2. I. Foster and C. Kesselman editors. San Francisco, CA, Morgan Kauffman.

Michal Feldman, Kevin Lai, John Chuang and Ion Stoica, 2003. "Quantifying Disincentives in Peer-to-Peer Networks," in proceedings of 1st ACM Workshop on Economics of Peer-to-Peer Systems, Berkeley, CA, pp. 67-72.

Fabrice Le Fessant, Samir Handurukande, Anne-Marie Kermarrec and Laurent Massoulie, 2004. "Clustering in Peer-to-Peer File Sharing Workloads," in proceedings of 3rd International Workshop on Peer-to-Peer Systems (IPTPS), San Diego, CA, pp. 90-95.

Ian Foster and Carl Kesselman (2004). *The Grid: Blueprint for a New Computing Infrastructure (Second Edition)*. New York: Morgan-Kaufmann.

Bruno Frey and Reto Jegen, 2001. "Motivation Crowding Theory: A Survey of Empirical Evidence," *Journal of Economic Surveys,* volume 5, number 5, pp. 589-611.

Christos Gkantsidis and Pablo R. Rodriguez, 2005. "Network coding for large scale content distribution," in proceedings of 24th Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM 2005), Miami, FL, pp. 2235-2245.

Lei Guo, Songqing Chen, Zhen Xiao, Enghua Tan, Xiaoning Ding and Xiaodong Zhang, 2005. "Measurements, Analysis and modeling of BitTorrent-like systems," in proceedings of ACM SIGCOMM Internet Measurement Conference, New Orleans, LA, pp. 670-685.

Bradley Horowitz, "Creators, Synthesizers, and Consumers", accessed 8/8/2006, http://www.elatable.com/blog/.

Daniel Hughes, Geoff Coulson and James Walkerdine, 2005. "Freeriding on Gnutella Revisited: the Bell Tolls?," *IEEE Distributed Systems Online,* volume 6, number 6, pp. 90-97.

Miguel Izal, Gillaume Urvoy-Keller, Ernst W. Biersack, Pascal Felber, A. Al Hamra and L. Garces-Erice, 2004. "Dissecting BitTorrent: Five Months in a Torrent's Lifetime," in proceedings of Passive and Active Network Measurement 5th International Workshop, Antibes Juan-les-Pins, France, pp. 1-11.

Norbert L. Kerr, 1996. "Does My Contribution Really Matter? -- Efficacy in Social Dilemmas," *European Review of Social Psychology,* volume 7, number 8, pp. 209-240

Kevin McGee and Jorgen Skågeby, 2005. "Gifting Technologies," *First Monday,* volume 9, number 12, pp. 1-15.

Andy Oram (2001). *Peer-to-Peer: Harnessing the Power of Disruptive Technologies*. Sebastopol, CA: O'Reilly & Associates.

Stefan Saroiu, P. Krishna Gummadi and Steven D. Gribble, 2002. "A Measurement Study of Peer-to-Peer File Sharing Systems," in proceedings of Multimedia Computing and Networking Conference (MMCN), San Jose, CA, USA, pp. 673-695.

Rob Sherwood, Ryan Braud and Bobby Bhattacharjee, 2004. "Slurpie: A Cooperative Bulk Data Transfer Protocol," in proceedings of The 23rd Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM 2004), Hong Kong, pp. 234-250.

Daniel Stutzbach and Reza Rejaie, 2006. "Characterizing Files in the Modern Gnutella Network: A Measurement Study," in proceedings of ACM Multimedia Computing and Networking (MMCN '06), San Jose, CA, pp. 567-581.