Economics of BitTorrent Communities

Ian A. Kash
Microsoft Research
Cambridge
Cambridge, CB3 0FB, UK
iankash@microsoft.com

John K. Lai
Harvard SEAS
Cambridge, MA 02138
jklai@seas.harvard.edu

Haoqi Zhang
Harvard SEAS
Cambridge, MA 02138
hq@eecs.harvard.edu

Aviv Zohar
Microsoft Research SVC
Mountain View, CA 94043
avivz@microsoft.com

ABSTRACT

Over the years, private file-sharing communities built on the BitTorrent protocol have developed their own policies and mechanisms for motivating members to share content and contribute resources. By requiring members to maintain a minimum ratio between uploads and downloads, private communities effectively establish credit systems, and with them full-fledged economies. We report on a half-year-long measurement study of DIME – a community for sharing live concert recordings – that sheds light on the economic forces affecting users in such communities. A key observation is that while the download of files is priced only according to the size of the file, the rate of return for seeding new files is significantly greater than for seeding old files. We find via a natural experiment that users react to such differences in *resale value* by preferentially consuming older files during a 'free leech' period. We consider implications of these finding on a user's ability to earn credits and meet ratio enforcements, focusing in particular on the relationship between visitation frequency and wealth and on low bandwidth users. We then share details from an interview with DIME moderators, which highlights the goals of the community based on which we make suggestions for possible improvement.

Categories and Subject Descriptors

C.2.4 [**Computer-Communication Networks**]: Distributed Systems; J.4 [**Social and Behavioral Sciences**]: Economics

Keywords

BitTorrent, private communities, peer-to-peer, incentives, share ratio enforcement, resale value

1. INTRODUCTION

Interactions among large numbers of agents on the Internet challenge system designers to not only focus on system-level function, but also to account for user incentives. In systems ranging from eBay to BitTorrent, the designs of reputation systems and sharing protocols pay particular attention to the role of economics in computer systems. In BitTorrent, incentivizing users to contribute by uploading while downloading a file has led to an effective form of file-sharing that now accounts for an estimated 18% of Internet traffic [1].

Despite BitTorrent's success, there is a lack of incentive for peers to continue uploading a file after it is downloaded. *Private BitTorrent communities* are a solution to this problem. Private communities build on the BitTorrent protocol by developing their own policies and mechanisms for motivating members to share content and contribute resources. Communities tend to be organized around a particular interest—e.g., live concert recordings, high definition movies, or the newest TV shows—and registered members acquire files of interest in return for sharing files with like-minded users. There are over 800 active private BitTorrent communities [16], each enforcing its own set of rules that are refined over time to fit the community's goals and needs.

Supported by additions to the original BitTorrent protocol, private communities can track how much each user downloads and uploads. This allows them to require members to upload a certain fraction of the amount they download. This regulation, known as *share ratio enforcement* (SRE), effectively introduces a currency to the system. Users earn credit by uploading files they have or have downloaded, and spend credit by downloading files. For example, a ratio requirement of 0.25 has an uploader earning four credits for every byte uploaded and a downloader spending one credit for every byte downloaded. In accounting for consumption (download) and labor (upload), *private BitTorrent communities are as much economic systems as they are computer systems.*

Anecdotal evidence from discussions among members in private communities points to a rich, multi-faceted set of user behaviors that emerge in response to economic forces. Their stories and shared advice suggest that users often make economic decisions and trade-offs, e.g., by joining new torrents as a way to quickly earn credit that can then be spent on downloading older torrents. If properly directed, economic forces can help to advance a community's goals and lead individuals to make better use of resources, but if misdirected they can lead to skewed incentives and inefficiency.

Previous studies of BitTorrent communities (e.g. [3, 12]) typically emphasize their characteristics as computer systems, focusing on aspects such as the arrival rate of peers to

WWW 2012, April 16–20, 2012, Lyon, France.
ACM 978-1-4503-1229-5/12/04.

a torrent, the quantity of resources available, and the performance experienced by users. A few recent works focus on the economics of private communities, by using theoretical and simulation approaches to examine how ratio enforcement incentivizes contributions and how issues such as lack of credit flow [8] or potential for collusion [11] can create inefficiencies and manipulation opportunities. While the theoretical analysis and simulation results from these works provide some insight, gaining a deeper understanding of the economy in private BitTorrent communities requires rich datasets that can direct our attention to successes and inefficiencies that arise in actual communities for economic reasons.

In this paper, we advance the study of private BitTorrent communities as economic systems by reporting on a half-year-long measurement study of the DIME community for sharing live concert recordings. Using extensive traces of activity on different files and daily snapshots of the activity of all users, we find that:

- There are significant differences between the returns from seeding new and old files, resulting in higher *resale value* for downloading new files.
- Users preferentially consume older files during a 'free leech' period, which provides evidence that users are aware of and react to the resale value of files.
- Given the difference in resale value, frequent visitors to the site have more opportunities to earn credit by downloading and subsequently seeding new files, and on average achieve higher ratios.
- Low bandwidth users do not adjust for their lower earning potential by downloading more new files, and instead just achieve lower ratios.

The paper proceeds as follows. Section 1.1 introduces BitTorrent and related terminology. Section 2 introduces the DIME community, and describes our methodology for obtaining measurements. Section 3 demonstrates the significant difference in resale value between new and old files, and how users react to such differences. In Section 4 we examine how visitation frequency and bandwidth affect user outcomes and behavior. We share details from our interview with site moderators and discuss the implications of our findings in Section 5, with a focus on improving the design of private BitTorrent communities. Section 6 presents related work, and Section 7 concludes.

1.1 BitTorrent and related terminology

BitTorrent [5] is a protocol designed for sharing files via direct peer-to-peer connections between different hosts. A user who wishes to distribute a file to others starts by creating a *torrent* that contains metadata about the file to be distributed. The user then publishes the torrent, typically by posting it on a web site. The torrent, which is downloaded by other users who wish to gain access to the content, points to a centralized server called a *tracker* that is used to coordinate between various peers who are sharing the designated file. Once a peer learns the address of others who are sharing the same file, it directly connects to them and can download and upload pieces of the file. BitTorrent makes a distinction between *seeders*, who are peers that have a full copy of the file (and thus only upload it to others), and *leechers*, who only have a partial copy and engage in both upload and download. Peers start out as leechers while they are downloading the file, and can then serve as seeders once their download is complete.

File sharing communities provide a set of services: they host the website on which torrent files are posted, host the trackers used to coordinate the sharing of each file, and keep track of updates that are sent by the various peers about the upload and download that they have performed on each file. Each community thus serves as a center for coordinating the sharing of files among members in that community, and for keeping records of each member's contributions.

2. OVERVIEW AND METHODOLOGY

In this section we present an overview of DIME and its economy, discuss our methodology for obtaining measurements, and share results on user contribution and consumption.

2.1 DIME

DIME (`www.dimeadozen.org`) is a private BitTorrent community in which users share live concert recordings (bootlegs) in lossless audio format. Sharing concert recordings has a rich history prior to BitTorrent, as music enthusiasts would trade tape and CD recordings of their favorite bands. DIME provides a community in which to continue this tradition of bootleg trading, but with the convenience afforded by its website, forum system, and BitTorrent trackers. Shows uploaded on DIME cover a wide range of music genres, and include new shows from currently touring bands as well as older shows recorded decades ago. DIME prohibits the posting of any 'official material,' and maintains lists of artists, venues, and shows that are disallowed on its tracker.[1] According to DIME's FAQ, this helps to avoid legal troubles, and aims to respect artists' rights.

DIME allows open registration, but restricts the maximum number of accounts so as to reduce server load and work for moderators.[2] While the site is typically full, new accounts open up frequently, as existing accounts that are inactive for long periods of time are periodically removed from the system.

2.2 DIME's economy

By tracking the upload and download of members beyond a single torrent, communities are able to require that members perform some minimal amount of work. DIME enforces a *share ratio* of 0.25, which requires members to upload at least a quarter of the amount they download (in bytes).[3] We define the amount of *credit* or *wealth* each user has on DIME as:

$$\text{Credit} = 4 \times \text{upload} - \text{download}$$

which is the amount a user can download (in bytes) without uploading and still satisfy DIME's share ratio requirement.

Note that every transfer of data adds credit to the system. Because DIME requires a share ratio of 0.25, if a byte is sent

[1] `http://wiki.dimeadozen.org/index.php/Main_Page`

[2] During the course of this study the maximum number of accounts was approximately 110,000. As of February 2012 this number has increased to approximately 130,000.

[3] The minimum share ratio allowed on DIME is much lower than that allowed in other communities with ratio requirements.

from peer A to peer B, then B loses a unit of credit, but A gains four units of credit. This creation of credit counters loss from users with positive balances who exit the system or otherwise "hoard" credit that will never be spent [10].

DIME does not constantly enforce the share ratio, but rather does so in *share ratio enforcement (SRE) cycles*, which define particular download amounts after which the system will check a user's ratio against the minimum requirement. One can view the enforcement cycles as a form of loaning; set at every 5GB, they allow a user to have a negative amount of credit as long as the balance is positive by the next enforcement cycle. For example, this helps new users, who begin with no credit, to download their first files. Users who fail to meet the requirement at an enforcement cycle are not allowed to download additional files until their wealth becomes positive via uploading. Donating to DIME extends one's enforcement cycles, which effectively increases how much one can borrow without adding to one's wealth.

The *price* of a file is the amount of credit deducted from the account of the downloader, which is simply the size of the file (in bytes). The price per byte is thus the same across all files on DIME. The *resale value* of a file is the amount of credit earned by an uploader, which is four times the amount he uploads (in bytes). This resale value depends on the upload rate achieved (the rate of return), which depends on the uploader's bandwidth and may also change over time as seeders and leechers join and leave a torrent, and on the duration of seeding, which is up to the user. For example, suppose that files A and B have the same size, but file A has few seeders and many leechers while file B has many seeders and few leechers. All other things equal, file A promises a higher upload rate and thus a higher resale value for the same duration of seeding. To the extent that users are constrained by their ability to earn credit or simply want to maintain higher ratios, the resale value of a file is important and can influence user decisions.

Occasionally, DIME has a *free leech* period, during which users do not spend any credits when downloading files. In other words, the prices of all files are fixed to zero during free leech. Users still receive credit for uploading, so files retain their resale value. Our data covers one such period that lasted three days.

2.3 Methodology

DIME's servers collect information that is reported periodically by the BitTorrent clients of its members, which it tracks and displays in the form of HTML pages available to all members. We obtained the following information by performing periodic crawls of the website:[4]

- **Account profile snapshots**: We took periodic snapshots of the profile pages of all user accounts in the system. These profile pages included static information such as the user's join date and dynamically updated information such as the user's ratio, and upload/download amounts and rates.[5]

- **Torrent traces**: We recorded traces of torrent detail pages from the time a torrent was posted for a number of torrents. These pages included information about the seeders and leechers on the torrent and their current upload and download amounts for the torrent. We downloaded the torrent details pages every five minutes for the torrents being tracked.[6]

- **Torrent snapshots**: In late 2010, we also started to take snapshots of all active torrents in the system. For these snapshots we crawled the same pages as the torrent traces, but did not track individual torrents and instead took less frequent snapshots of all torrents.

While most of the statistics we collect are precise, two require some care. One is the maximum upload bandwidth available to a user. DIME tracks the maximum upload bandwidth it has ever observed for a user, but the actual maximum bandwidth of a user varies over time. While at an individual level the reported value may be a noisy signal of how much bandwidth a user can typically provide, in aggregate our results suggest it provides a reasonable signal. For example, on average the upload rate of a peer is roughly linear in this quantity (see Figure 3(a)).

The other statistic is the current upload of a peer when tracking a torrent. We did not perform peer-level measurements, so we only have access to the data that peers reported to the tracker. Though we crawled each tracked torrent every five minutes, empirically we observe that a peer's reported upload updates every 20 to 30 minutes. We can derive upper and lower bounds on the peer's upload during these 20 to 30 minute intervals, but do not have finer grained information. When computing statistics such as upload rates, we assume the upload is distributed equally across these intervals, and aggregate data from many users to mitigate errors due to this assumption.

2.4 User Contribution and Consumption

Figure 1 shows a snapshot of the historical upload and download amounts of all users on February 20th, 2010. There are 109,891 users in the system at the time of the snapshot, of whom 7.4% have donated money to the site. Note that almost all (non-donating) users who download more than 10GB and are still in the system have a ratio above 0.25 and that many users have a ratio above 1. This shows that many users choose to behave "altruistically" and upload more than the minimally required amount. DIME and other private communities promote such behavior by encouraging users to upload at least as much as they download,[7] and by issuing social rewards to users with high ratio. For example, users earn special badges for attaining specific levels of activity, are often more respected in the community, and are given additional privileges on the site.[8] These factors inspire many users to upload more than what is required by the minimum share ratio, and suggest that even users with ratios

[4]Our study is conducted with permission from DIME moderators, and with approval from Harvard University's Institutional Review Board.

[5]We performed daily scrapes between April 28, 2010 and September 27, 2010, and multiple scrapes per day between December 23, 2010 and January 21, 2011. Out of 153 possible days between 4/28/10 and 9/27/10, we are missing 32 days due to scrape failures.

[6]Our first batch of traces tracked 173 torrents posted after April 29, 2010 until June 26, 2010. Our second batch of traces tracked 176 torrents posted after June 27, 2010 until September 7, 2010.

[7]`http://wiki.dimeadozen.org/index.php/DimeFAQ:DIME_Ratio_Primer`

[8]`http://wiki.dimeadozen.org/index.php/EzTorrent:VIP_Perquisites`

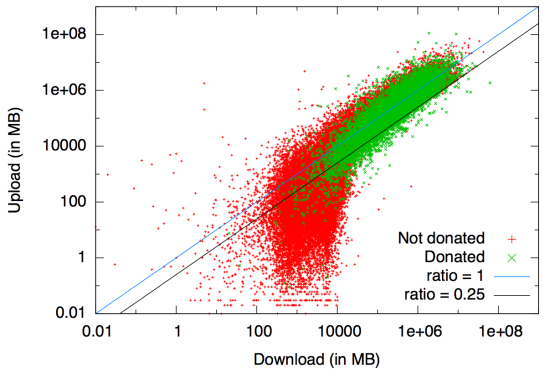


Figure 1: Snapshot showing all users' upload and download amounts. Users marked in green donated money to the site; users marked in red (including those covered by the thick green) did not.

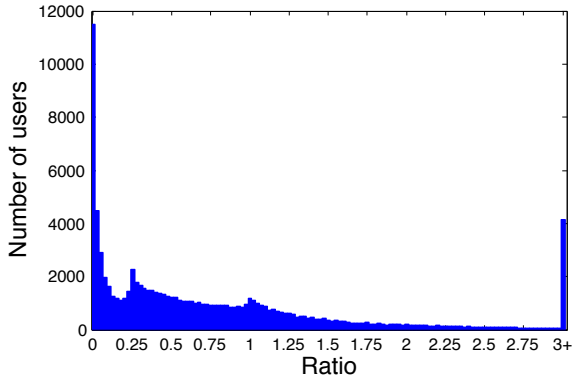


Figure 2: A Histogram of user ratios from a snapshot taken on May 1st, 2010 with bin sizes of 0.025.

significantly above 0.25 may care about the resale value of a file.

We also see in Figure 1 that most users who have uploaded at least 1GB are above the 0.25 threshold, and virtually all users who have uploaded at least 10GB are above the threshold. This suggests that users who are restricted by the first enforcement cycles may be free riders who do not intend to become long term members of the community. Such users are initially given the benefit of the doubt, but the extent to which they can free ride is limited. We also see that there are a number of users with 100GB or more of download who are significantly below the ratio of 0.25. Presumably, these users make repeated donations to periodically extend their SRE cycles. From the system's perspective, this is not particularly bad: there aren't enough of these users to cause a problem with the functioning of the economy, and donations collected can be used to cover server and other operating costs.

In analyzing user ratios, we find that 50 percent of users have a ratio of at least 0.5 and 30 percent of users have a ratio of at least 1. Of the users with ratios less than 0.25, only 6.5% (or around 2000 users) downloaded more than 20GB, indicating again that most users with low ratios are free riders who will either donate or leave the system.

Figure 2 shows a histogram of the share ratios of users in the system. We observe distinct increases around ratios 0.25 and 1. The spike at 0.25 is consistent with a group of users performing the minimum amount of work required to remain active in the system due to share ratio enforcement. The bump around 1 shows some users attempting to contribute at least as much as they receive from the system, which is consistent with what DIME recommends that users do.

3. RESALE VALUE

Share ratio enforcement and users' desire to maintain particular ratios require users to earn credits through uploading to keep up with credits spent through downloading. In this section, we examine the difference in resale value between new and old files, and show how such differences may affect users' decision-making in terms of which files to consume.

3.1 Resale Value and Torrent Age

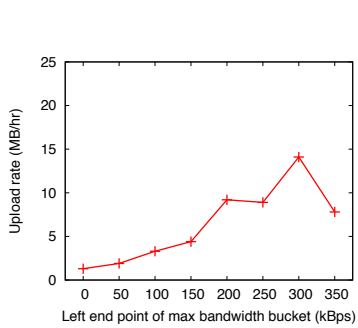
In order to examine the relationship between user behavior and resale value, we first consider factors that affect the resale value of torrents. An analysis of our collected data shows that the rate of return from seeding is highly correlated with the *age of a torrent*, i.e. the time elapsed since the torrent was first posted. A priori, it is unclear whether new torrents or old torrents will result in the highest returns to seeding as there are competing effects at play. Early in the life of a torrent there are more leechers who wish to download the file, suggesting a higher return to seeding. However, there are also more seeders around, suggesting that users may face more competition with other users for upload. By tracking the activity on individual torrents on DIME, we find that *earning potential is significantly higher during the early lifetime of a torrent and decays as the torrent ages.*

We use the first batch of torrent traces (173 torrents) to obtain an aggregate estimate of the upload per period of time seeding over the age of the torrent. For each torrent, we compute an estimate of the upload rate as follows. For each seeder on a torrent that is not its original uploader, we construct a sequence of (upload, (start time, end time)) pairs which gives an estimate of how much the user uploaded in (start time, end time). We then bucket these observations by time, so that for each bucket of five hours, we have the total upload as well as the total time spent seeding. From here, we divide total upload by total time to get an estimate of the upload rate in the time bucket. We then take the average of these upload rates across all torrents in our set of traces. Torrents that had no seeding activity in a time bucket are included with a rate of zero.

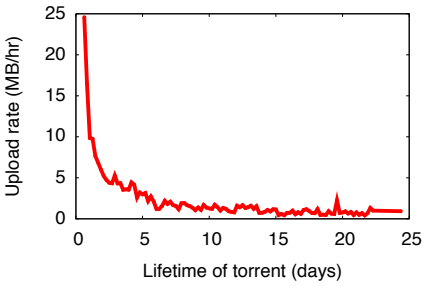
Figure 3(b) shows that the average upload rate on a torrent is extremely high in the hours immediately following its posting, and that there is a severe drop in rate of return over the course of the first few days. After five days, the decrease in upload rate slows, but continues for the lifetime of the torrent.[9] The large discrepancy between the returns from seeding early and seeding late shows that when a user downloads the file may be more important than how long the user plans to seed it.

While Figure 3(b) shows that the upload rate is higher for seeders who join a torrent early, it could be that the pop-

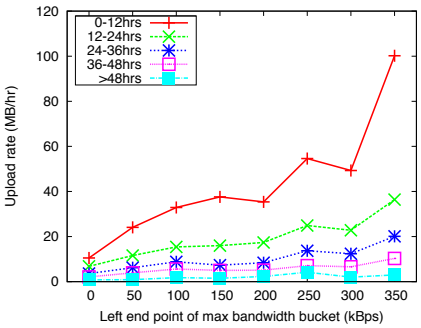
[9]The slow decline in the tail may be an artifact of torrents dying and our measurements recording a rate of zero for these torrents that are inactive.



(a) Upload rate vs Max bandwidth.



(b) Upload rate vs. Torrent age.



(c) Upload rate vs. Max bandwidth for different torrent age buckets.

Figure 3: Effects on Upload Rate.

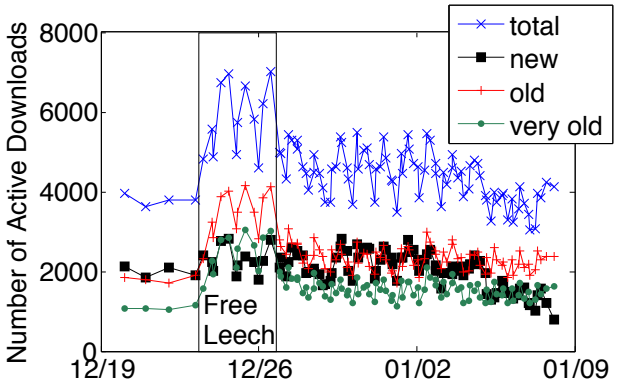


Figure 4: Leeching activity before, during, and after a free leech period.

ulation of seeders who join a torrent early is different than the overall population. For instance, it could be that those who join a torrent early tend to have higher upload bandwidth, and that is accounting for the observed discrepancies in upload rate. Figure 3(c) shows that even after controlling for the effect of upload bandwidth, the average upload rate is higher earlier in the life of a torrent. We see that higher bandwidth leads to higher upload rates as expected, while earlier join times magnify this effect by changing the slope of the plotted relationship. Figure 3(c) also suggests that an effective way to compensate for connection speed is to join torrents earlier. For example, while we refrain from giving precise numbers due to measurement noise, the figure suggests that joining in the first 0-12 hours as a low bandwidth user (50-150 kBps) may yield higher upload rates than joining in the first 12-24 hours as a higher bandwidth user (150-250 kBps).

These observations show that all else equal, newer torrents have a higher resale value. For each unit of time spent seeding, a user can gain more credit seeding a new torrent than an old torrent.

3.2 Resale Value and Decision Making

The significant difference in resale value between new and old files suggests that users can often earn credit by downloading files soon after they are posted, but will have to spend accumulated credit to acquire older files of interest. Given this, we expect users to preferentially download newer files, and predict that users would be more willing to download older files if their prices were lowered. Through a natural experiment that occurred during our study, we are able to confirm these hypotheses.

From December 23, 2010 to December 26, 2010, DIME had a free leech period, during which downloading did not count against a user's credit but uploading still provided credit. Figure 4 shows the number of active downloads during a three week period that includes the free leech period. We observe significantly more active downloads during the free leech period than during the days before and after free leech, where the amount of download activity during free leech is 50% to 75% higher than during the days following free leech.[10] In the days before and after free leech, we observe that the number of active downloads of files uploaded within the last week (new files) is nearly identical to the number of active downloads of files older than a week (old files). But during free leech, demand for old files increased 60% to 70% while the demand for new files did not change significantly. Given that there are approximately 25 times more old files than new files at any given time, these findings imply that users are typically downloading significantly more copies of newer files than older files, but that during free leech users react to the change in prices by consuming many more older files. For "very old" files (those that are more than sixty days old), the demand nearly doubled during the free leech period.

From an economic perspective, users can download newer files of interest without worrying much about credit (since these files can actually increase their wealth), but have to download older files discriminately if constrained by their wealth. Free leech provides a significant opportunity to acquire these files for free, during which users are able to download files they want (old or new) without worrying about impacting their wealth.

There is no particular bonus for seeding during free leech, but the increase in download activity allows seeders to earn

[10]Note that prior to the free leech period our data has only a single observation each day, while during and after free leech we have multiple observations per day. The results during the latter period captures some of the daily fluctuations in usage that are typical of private communities [7].

more credit per unit of time spent seeding. Interestingly, there was essentially no increase in the number of seeders during free leech, either overall or among those with low share ratios. Given the increase in the number of active downloads during free leech, more downloads are supported by the same number of seeders during this period. Assuming that the characteristics of the population of seeders (e.g., their bandwidth distribution) are more or less the same during free leech and at other points in time, this provides evidence that there is typically a supply of upload bandwidth that is not being used because no one is currently leeching the files those users are seeding. While this finding points to a potential inefficiency of the system, having a supply of seeders with available bandwidth on older torrents does allow these files to remain available to users who choose to download them. It also highlights why the resale value of files is important: there is a pool of users who are willing to seed more but their content may not be of interest to others, so their efforts could be spared or better directed to files that are in greater demand to increase download speeds. Alternatively, since users do appear to value older files (as evidenced by the demand during the free leech period), we could attempt to make downloading these files more attractive to increase the total welfare produced by the system. We discuss this issue further in Section 5.2.

4. GENERATING WEALTH

The difference in resale value between new and old files have implications on the ability of users to acquire wealth on DIME. We examine in this section how the frequency of site visits may affect earning potential, and how low bandwidth users handle the economic forces within DIME.

4.1 Visitation Frequencies

Since newer torrents provide higher resale values, we expect users who can regularly manage to download files early in a torrent's lifetime to have the best earning potential. Assuming that files of interest to a user are uploaded more or less uniformly across time, the likelihood that a file of interest will be new when a user sees it on DIME is directly correlated with how frequently the user visits the site. To study the relationship between visitation frequency and earnings, we use contiguous daily snapshots of all DIME users from April 28th, 2010 to July 20th, 2010 to obtain records of each user's upload and download amounts during this period, as well as the number of days on which they were seen on the site. After filtering out users who only visited once or never, and users for whom we did not have at least 60 days of data (e.g., new members) in this measurement period, we are left with a dataset containing records for 38,583 users.

We plot in Figure 5 the distribution of upload to download ratios obtained by users with particular visitation frequencies during the measurement period. We see from the graph that very few of the most frequent visitors (2%) obtain a ratio of less than 0.1, and over 50% earn a ratio above 1. On the contrary, nearly 40% of the most infrequent visitors earn a ratio of less than 0.1, and only 20% earn a ratio above 1. While this suggests that frequent visitors may be presented with more earning opportunities, we also see that not all frequent visitors earn a high ratio. This is likely due to choice: as users who visit often can more easily join newer torrents of interest to earn credit, they can also respond to their higher earning potential by consuming more older files,

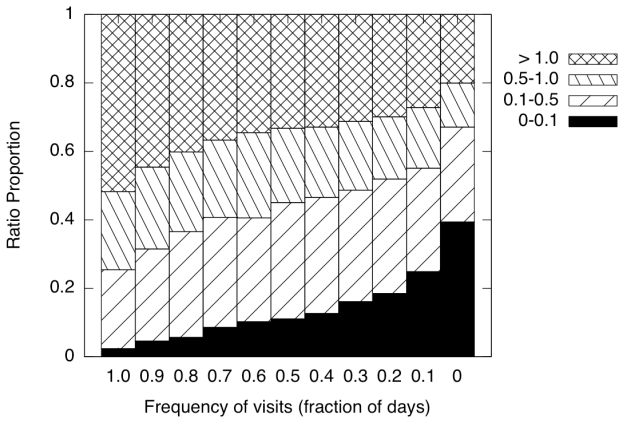


Figure 5: Distribution of users by upload to download ratio, bucketed by visitation frequency.

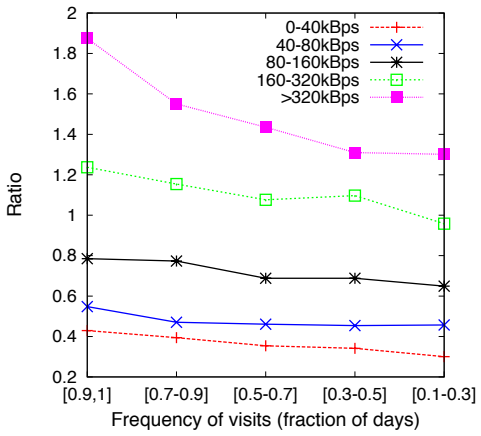


Figure 6: Aggregate ratio vs. Visitation frequency, bucketed by upload bandwidth.

or seeding for shorter periods of time (e.g., disconnect soon after the download completes).

To ensure that the observed difference is not due to differences in the distribution of bandwidth among users at different visitation frequencies, we plot in Figure 6 the aggregate ratio achieved by users at different visitation frequencies, separated out by bandwidth. We see that on average, regardless of the bandwidth group, users who visit the site most frequently earn higher ratios, with the lowest bandwidth and highest bandwidth users earning significantly higher ratios when they arrive earlier. This provides further evidence that the observed effect is due to differences in earning potential, as caused by being able to join files earlier.

From Figure 7, we see that frequent visitors not only earn higher ratios, but also download and upload more on average than infrequent visitors. In addition to having higher in-period consumption, the proportionally higher upload implies that frequent visitors are also acquiring significantly more wealth that can be utilized for future spending. While some of the difference in wealth earned can be explained by higher rate of earning on newer files, it may also be a representation of actual demand, wherein users with lower

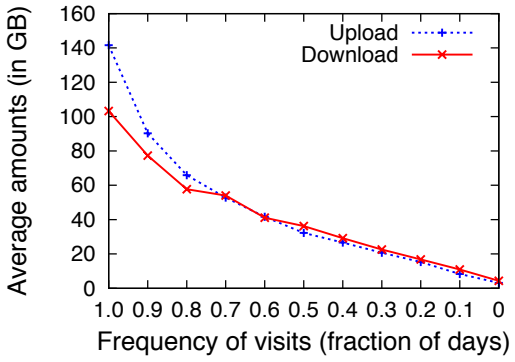


Figure 7: Average download and upload amounts (in GB), bucketed by visitation frequency.

demand (e.g., interested only in a few particular bands) may choose to visit the site less frequently and may also be less concerned with maintaining a high ratio. It is also worth noting that while one can download files of interest while their resale value is high by visiting the site frequently, doing so takes effort and should be thought of as a costly action. We examine how users make economic decisions in response to such trade-offs below.

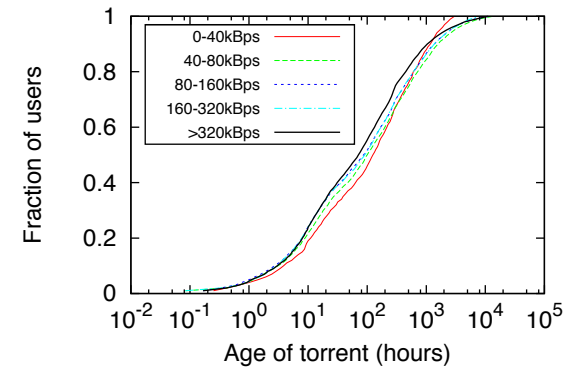
4.2 Low-bandwidth users

To examine how users respond to the forces within the DIME economy, we consider the behavior of low bandwidth users, who are particularly susceptible to economic pressures within the system.

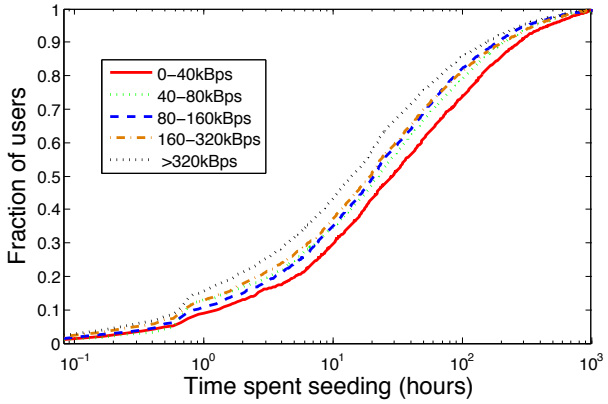
We have seen that returns to seeding depend on bandwidth and torrent age. In particular, upload rates are significantly higher if a user arrives early to a torrent, and tend to increase linearly with bandwidth over a given period. This suggests that, all else being equal, low bandwidth users will earn less credit than high bandwidth users. In this section, we examine four actions low bandwidth users could take to compensate for this: join torrents earlier, seed longer, download less, and aim for lower ratios.[11]

Our results show that low bandwidth users do not join torrents earlier than other groups. Figure 8(a) shows, on a log scale, the CDF of the time after a torrent's creation at which users arrive. Each line represents a class of users within a particular bandwidth bucket. We see that most of the lines are quite similar, with low bandwidth users joining slightly later than high bandwidth users. This suggests that many such users are unable or unwilling to change their behavior in order to join torrents earlier. One possible explanation is that, while checking the site more frequently can allow users to join desired torrents earlier, doing so requires manual effort and may be costly or infeasible for many users.

[11]In this section, we assume that bandwidth limitations are independent of user demand. Since we can only measure the amount of bandwidth users make available for BitTorrent rather than their true upload capacity, we cannot rule out the possibility that behavior we observe is due to users who choose to contribute low bandwidth due to lower demand despite having high upload capacity. While this is a limitation of the study, we mitigate potential effects by using the highest upload rate DIME has ever recorded for each user.



(a)



(b)

Figure 8: (a): CDF of time of first appearance on torrents by users' upload bandwidth. (b): CDF of time spent seeding by users' upload bandwidth.

While low bandwidth users do not appear to join torrents earlier than other users, we do find that low bandwidth users seed longer than high bandwidth users. Figure 8(b) shows, on a log scale, the CDF of seeding time, again grouped by bandwidth. Here the ordering of lines is consistent, with higher bandwidth users spending less time seeding than lower bandwidth users. For example, the median user with a bandwidth between 0-40kBps spends 1.4 times as long seeding as the median user with a bandwidth between 80-160kBps, who in turn spend nearly 1.5 times as long seeding as the users in the highest bandwidth buckets.

A third way to compensate for low bandwidth is to download less. Our findings from observing the change in users' download amounts between April 28th, 2010 and July 20th, 2010 show that this is indeed the case. Figure 9 buckets users by bandwidth, and shows that on average low bandwidth users download less than high bandwidth users. The difference in download amounts thus suggests that the price of files in the system may be preventing lower bandwidth users from being able to fulfill their demand.

We also see from Figure 9 that low bandwidth users upload less in proportion to their download, thus earning lower ratios. Users in the lowest bandwidth bucket earned an aggregate ratio of 0.4 during this measurement period, while users in the highest bandwidth bucket earned an aggregate ratio of 1.7. Since DIME only requires users to maintain

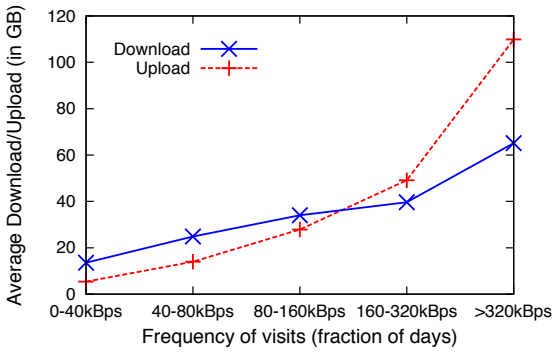


Figure 9: Average download and upload amounts of users bucketed by bandwidth.

a ratio of 0.25, aiming for lower ratios seem like a reasonable strategy for low bandwidth users acting in this economy. That said, their lower upload volume and earned ratio imply that they acquire significantly less wealth than high bandwidth users, leaving them with less savings for future consumption. This in some sense also limits their ability to strategically join early: while the average new torrent provides significant earning opportunities, any particular new torrent may not. Unlike high bandwidth users with higher savings, low bandwidth users may be less able to account for uncertainty in earnings while attempting to maintain a ratio above 0.25.

In summary, we find that low bandwidth users do not join torrents earlier, but they do seed for longer, download less, and earn lower ratios. If we assume that low bandwidth users have the same demand as users in other buckets, the combination of seeding for longer periods yet downloading less suggests that the increase in seeding is unable to fully compensate for the low bandwidth of the users. This result agrees with the observation by Andrade et al. [3], that while users were generally willing to seed longer, doing so did not seem to make them much more successful as uploaders. The insights from Section 3.1 provide an explanation: as the rate of return for seeding a file drops significantly over time, seeding for longer does not result in a significant increase in credits earned.

5. DISCUSSION

Our ultimate goal is not only to understand the economic factors driven by DIME's current policies, but to find ways to improve the community. While one can establish a number of desirable properties, understanding which improvements to focus on and what tradeoffs to make depend on the particular goals of the community. While decisions can sometimes be made with respect to general user preference (e.g., trading off higher download speeds with the amount of seeding required), most decisions on resource distribution and availability will force us to weigh the conflicting preferences of different users, and to draw on the political and moral viewpoints of the community (e.g., what kind of users are wanted in the system, who should get to consume more, and what kind of files should be allowed). In this section we discuss our desiderata as motivated by our interview with DIME's moderators, and then discuss specific changes with potential for improving DIME.

5.1 Desiderata

We conducted an email interview with DIME moderators in January, 2011.[12] From the interview, we learn that decisions on site policies are often community-driven. For example, DIME's ratio requirement of 0.25 is the result of a vote among community members in 2004. This ratio requirement has not been changed since.[13] The choice of this relatively low sharing ratio indicates that DIME is more open than other communities to the "less fortunate," who may have slower Internet connections or cannot visit the site frequently. The fact that every user is granted 5GBs of download before ratio enforcement begins strikes a balance between giving new users a leg up and allowing in some hit-and-run leechers to download for free. The moderators noted that the enforcement cycle may need to be extended to account for the increase in file size over the years, but were concerned that "hit-and-run leechers would be able to take even more without giving."

DIME moderators view themselves as a "user-help-desk", and are willing to help users in share ratio violation to get back on track by providing advice and temporarily extending the enforcement cycle. One desideratum may thus be to provide additional mechanisms for helping "less unfortunate" users and poor decision-makers to earn credit, while keeping abusers out of the system.

When asked about the significant difference in resale value between new and old files, moderators simply responded that "this is the nature of BitTorrent," and that it encourages users to arrive which helps to share and distribute files. The moderators also had no issues with users visiting the site more often and downloading newer torrents as a means to earn credit, and believe that this helps DIME be "a vital community." While forming an active community in which users can download files of interest as long as they put in the effort to contribute (e.g., by joining new torrents regardless of interest) is important, alternative mechanisms for rewarding contributions of users can potentially offset the unmet demand on older files due to the lack of earning potential. A second desideratum is thus to increase demand for older files, while balancing the goal of maintaining a vital community.

Finally, while system performance is important to DIME moderators, they are also looking to reduce server costs when possible. Moreover, moderators do get very busy, and given that they serve an important function in the community and their time is a scarce resource, policies and mechanisms to reduce their own workload can benefit the system. Reductions in the load each user inflicts on the servers and on the moderators will allow, for example, for an increase in the size of the community.

5.2 Potential Changes

5.2.1 *Restricting access of new users to older files*

Our measurements reveal that new users have an increased tendency to download old files. Figure 10 shows that the median user who registered within the last 0-14 days initiated download 96 hours into the torrent's lifetime, while the median veteran user (who had an account for more than 50

[12]A transcript of the interview is available at `http://eecs.harvard.edu/~hq/DIMEinterview/`

[13]There were a number of attempts spurred by particular community members, but they "weren't successful because of the community's resistance."

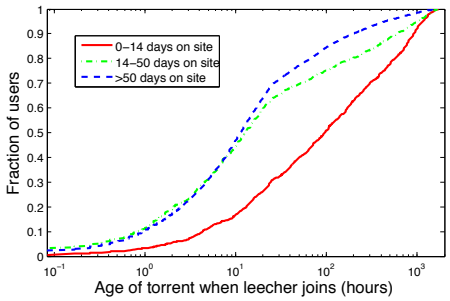


Figure 10: A CDF of times leechers begin to download torrents. Users are grouped by the age of their account.

days) tended to join a torrent after only 11.3 hours. This effect may have several causes. First, users who have just joined may find old files appealing, and were not around to download them when they were new. Second, these users may be less aware of the pitfalls of downloading old files, which can quickly result in them having negative credit. Finally, some users may be joining the site to get a particular file, and may not be interested in staying for the long run. These "free riders" know in advance that they will not need to regain their lost credits and will not upload the file; they may thus place no value on the gains from potential future resale of the files and just download indiscriminately.

Newcomers who unwittingly end up with a negative amount of credit may be driven to create new accounts, or may turn to moderators for a temporary extension of the SRE cycle (a temporary loan of credits). An approach requiring less manual intervention would be to limit the access of users to older files, e.g., until they gain more experience on the site, or only when they have enough absolute wealth to cover the entire cost of downloading the old file. This would both help new users avoid the potential mistake of getting into debt for downloading a file they cannot later upload, and at the same time would also make free-riding less appealing as more effort would be needed to access many files.

A possible pitfall of this approach is that new users may be dissuaded from joining the site if they cannot initially access some material they desire. While this is something to be wary of, DIME is currently running at capacity and new users need to wait for accounts to become available. If this is a concern, an alternative would be to caution users with a warning, or apply softer limit based on their current ratios.

5.2.2 *Increasing demand for files*

In conventional markets, the price of services that have too much supply and too little demand naturally drops. But on DIME, all transfers are credited equally, so prices remain fixed. One can imagine adopting a credit system in which uploads and downloads convert to credit based on the prices of files. In such a system, one can attempt to adjust the price of torrents by slowly lowering the price over time, by making all files beyond a certain age cheaper, or by making the price depend on the seeder to leecher ratio in the torrent. This would attract more reluctant downloaders, and give additional hints to seeders about how to best direct their efforts. Related approaches to helping match supply and demand across torrents are considered in Antfarm [13] and PACE [4].

Price alterations should be done very carefully. If the cost decreases too much, too many users will wait to download files and too few will seed them which will amount to a stagnation in the economy. We also need to be careful not to make it too easy to earn money. Theoretical models [10] show that if it is too easy to earn money, rational users feel "rich" and decrease their willingness to work, leading to a vicious cycle where fewer and fewer users contribute.

6. RELATED WORK

A number of papers empirically study private BitTorrent communities, generally concluding that private communities exhibit higher download speeds and availability than public trackers. While our study tracks information similar to that of earlier studies, we conduct a series of long traces and can thus examine how user behavior changes over time. Additionally, our torrent level traces allow us to study how activity on individual torrents varies over time, leading to our novel study of resale value and its implications.

In a series of papers, Andrade et al. [2, 15, 3] study traces from seven BitTorrent communities, some of which use SRE. They find that peers contribute significantly more, particularly by seeding for longer periods of time, in communities with SRE. They also study the arrival rate of peers to torrents, showing that it is initially high, but rapidly drops and then has a long, slowly-decaying tail. This arrival pattern is consistent with our observation that the greatest opportunities to gain upload as a seeder are early in the life of a torrent.

Liu et al. [11] study a user snapshot of HDChina, which uses a variable SRE depending on download amount, and show that seeder / leecher ratio is significantly higher in HDChina than in public torrents. The authors also develop a model of incentive mechanisms in BitTorrent communities and show that a ratio mechanism provides good incentives. They argue that collusion is an inherent problem in private communities and propose an entropy-based method for detecting collusion.

Hales et al. [8] report some basic statistics from a seven day trace of a community using SRE at a ratio of 0.67. They show that a majority of the uploading each day is contributed by ten percent of peers, possibly starving others of the opportunity to maintain an acceptable ratio while downloading desired files. Using a theoretical model and simulations, they demonstrate conditions under which this occurs. Rahman et al. [14] build on this through additional modeling and simulations and show how an adaptive policy can help avoid credit crunches by instituting free leech periods when many peers are "stuck" at a low ratio.

Meulpolder et al. [12] study five communities, three of which use SRE. They find that more stringent ratio requirements lead to higher download speeds, longer seeding time, and fewer firewalled peers.

Zhang et al. [16] study the landscape of private BitTorrent communities and estimate that over 800 private communities combine to have approximately the same number of torrents as publicly available trackers and have significantly more active users at any time.

In a pair of papers, Chen et al. [7, 6] study 17 communities, including a 68 day trace of DIME, and note that those that use SRE have significantly greater user activity and seeding. Their study of DIME is more limited and focuses primarily on the characteristics of users. They model the tendency of

peers to be starved for opportunities to upload and discuss mechanisms such as free leech periods that communities use to ameliorate the problem.

Huberman and Wu [9] propose an incentive mechanism for peer-to-peer exchange that credits servers for seeding files, much like the SRE mechanism for private BitTorrent communities. They conclude that such a mechanism creates an incentive for servers to provision the long tail of files that may not be accessed very often. Indeed, we observe that many older files are still actively seeded.

7. CONCLUSION

We have presented a study of DIME's complex economy. In it, we have shown that older files are much harder to upload to others and thus have low resale value, but are equally priced per byte. This effect skews the everyday consumption of files toward newer, more popular files—an effect that is reflected in the increased desire to download old files when the price is lowered during a free leech period. While the difference in resale value makes it difficult for infrequent and low bandwidth users to earn credit by downloading files of interest, DIME's relatively low minimum share ratio nevertheless allows such users to participate in the community. These results show how economic policies in private BitTorrent communities can reflect social norms, and affect the composition of users and behaviors accordingly.

Based on these insights, we have suggested a few possible changes that may improve the efficiency of DIME's economy. However, it is important to note that, even without further intervention, DIME's survival despite changing conditions, such as increases in bandwidth and file size and having a dynamic user population, is a tribute to both the community spirit that it maintains, and to the robustness of its economy. Changes should thus be introduced with great care.

While we have focused on DIME, our observations on the difference in resale value between new and old files and its effect on user behavior and wealth should hold in other communities for which the pricing of files is age independent. An interesting direction for future work is to study a diverse set of BitTorrent communities to gain a fuller understanding of how different economic policies adopted by individual communities can affect system performance, influence the type of users they attract, and advance community goals.

Acknowledgements

We thank DIME moderators for allowing us to conduct our study, and in particular to jupiter2101 and erwe for the interview. We thank Jane Calhoun for her assistance with the IRB process, and David Parkes for helpful discussions. John Lai and Haoqi Zhang are generously funded by a NDSEG Fellowship and NSF Graduate Research Fellowship respectively.

8. REFERENCES

[1] An estimate of infringing use of the internet. Technical report, Envisional, January 2011. `http://documents.envisional.com/docs/Envisional-Internet_Usage-Jan2011.pdf`.

[2] N. Andrade, M. Mowbray, A. Lima, G. Wagner, and M. Ripeanu. Influences on cooperation in BitTorrent communities. In *Proc. of the ACM SIGCOMM Third Workshop on the Economics of Peer-to-Peer Systems (P2PEcon)*, pages 111–115, 2005.

[3] N. Andrade, E. Santos-Neto, F. V. Brasileiro, and M. Ripeanu. Resource demand and supply in BitTorrent content-sharing communities. *Computer Networks*, 53(4):515–527, 2009.

[4] C. Aperjis, M. J. Freedman, and R. Johari. Peer-assisted content distribution with prices. In *2008 ACM Conference on Emerging Network Experiment and Technology (CoNEXT 2008)*, page 17, 2008.

[5] BitTorrent Inc. BitTorrent web site. http://www.bittorent.com.

[6] X. Chen, X. Chu, and J. Liu. Unveiling popularity of BitTorrent darknets. In *Proc. of the Global Communications Conf. (GLOBECOM)*, pages 1–5, 2010.

[7] X. Chen, Y. Jiang, and X. Chu. Measurements, analysis and modeling of private trackers. In *Proc. of the IEEE Tenth Int. Conf. on Peer-to-Peer Computing (P2P)*, pages 1–10, 2010.

[8] D. Hales, R. Rahman, B. Zhang, M. Meulpolder, and J. A. Pouwelse. BitTorrent or BitCrunch: Evidence of a credit squeeze in BitTorrent? In *Proc. of the 18th IEEE Int. Workshops on Enabling Technologies: Infrastructures for Collaborative Enterprises (WETICE)*, pages 99–104, 2009.

[9] B. A. Huberman and F. Wu. Bootstrapping the long tail in peer to peer systems. In *Proceedings of the First Workshop on Economics of Networked Systems (NetEcon'06), Ann Arbor*, 2006.

[10] I. A. Kash, E. J. Friedman, and J. Y. Halpern. Optimizing scrip systems: Efficiency, crashes, hoarders and altruists. In *Eighth ACM Conference on Electronic Commerce (EC)*, pages 305–315, 2007.

[11] Z. Liu, P. Dhungel, D. Wu, C. Zhang, and K. W. Ross. Understanding and improving ratio incentives in private communities. In *Proc. of the 30th Int. Conf. on Distributed Computing Systems (ICDCS)*, pages 610–621, 2010.

[12] M. Meulpolder, L. D'Acunto, M. Capota, W. Wojciechowski, J. A. Pouwelse, D. H. J. Epema, and H. J. Sips. Public and private BitTorrent communities: A measurement study. In *Proc. of the 9th Int. Workshop on Peer-to-Peer Systems (IPTPS)*, 2010.

[13] R. S. Peterson and E. G. Sirer. Antfarm: Efficient content distribution with managed swarms. In *Networked Systems Design and Implementation (NSDI)*, 2009.

[14] R. Rahman, D. Hales, T. Vinko, J. A. Pouwelse, and H. J. Sips. No more crash or crunch: Sustainable credit dynamics in a p2p community. In *Proc. of the 2010 Int. Conf. on High Performance Computing & Simulation (HPCS)*, pages 332–340, 2010.

[15] M. Ripeanu, M. Mowbray, N. Andrade, and A. Lima. Gifting technologies: A BitTorrent case study. *First Monday*, 11(11), 2006.

[16] C. Zhang, P. Dhungel, D. Wu, Z. Liu, and K. W. Ross. BitTorrent darknets. In *Proc. 29th IEEE Int. Conf. on Computer Communications (INFOCOM)*, pages 1460–1468, 2010.