Measurements, Analysis and Modeling of Private Trackers

Xiaowei Chen
Computer Science Department
Hong Kong Baptist University
xwchen@comp.hkbu.edu.hk

Xiaowen Chu
Computer Science Department
Hong Kong Baptist University
chxw@comp.hkbu.edu.hk

Abstract—**BitTorrent plays a very important role in the current Internet content distribution. The enormous impact of public and private trackers should not be overlooked. Public trackers are suffering from free-riding problem, but private trackers work very well because of effective Share Ratio Enforcement (SRE) which is an auxiliary incentive mechanism. Therefore, understanding the characteristics of private trackers is essential to BitTorrent content distribution and further development.**

We have crawled and traced fifteen trackers with nearly one million torrents for four months. In this paper, we first provide taxonomy of private trackers, and then present in breadth and depth measurement study on the characteristics of private trackers. We have found that private trackers have apparently different features and statistics from public trackers, ranging from the user viscosity, torrents evolution, user behaviors, content distribution and other metrics to measure the quality of service in private trackers. There are some new features that have not been examined in private trackers by previous measurement studies. Furthermore, we use game theory to study the effectiveness of SRE mechanism. We model the mechanism and propose an improved SRE mechanism to further incent the users and enhance the performance of private trackers.

Keywords-private tracker, incentive mechanism, BitTorrent, content distribution, peer-to-peer networks

I. INTRODUCTION

BitTorrent is currently the dominating P2P file sharing protocol. As one of the core components in BitTorrent protocol, trackers play an irreplaceable role during the distribution process which periodically provides updated peer lists to connected clients. BitTorrent trackers can be divided into two categories, public trackers and private trackers [1].

Public trackers (also known as open trackers) can be used by anyone by adding some tracker addresses to an existing torrent. There are a lot of public trackers such as the Pirate Bay [2], Mininova [3], ISOHunt [4], etc. Although BitTorrent has implemented the Tit-for-Tat (TFT) algorithm as an incentive mechanism, public trackers are still suffering from free-riding problem: first of all, a peer may stop uploading immediately after it finishes the download task; secondly, a peer usually sets a limit on the total upload bandwidth.

In recent years, private trackers become more and more popular and generate a huge amount of Internet traffic [5]. E.g., Torrents.ru alone has more than 1000TB of content and generates more than 46GB/s of Internet traffic. Users in private trackers can usually achieve much faster download speed than users in public trackers can. Based on strict member controlling policy, private trackers adopt Share Ratio Enforcement (SRE) as an auxiliary incentive mechanism to overcome the free-riding issue. SRE forces registered users to maintain a ***share ratio*** (i.e., upload-to-download ratio). In general, a registered user will be banned from private tracker community if his share ratio is lower than a threshold (e.g., 0.5). On the other hand, a user with higher share ratio can receive more benefits. There are two ways to achieve a high share ratio: one is to provide a high upload bandwidth; another one is to prolong the seeding time.

The main differences between public trackers and private trackers are summarized as follows:

- Users & Torrents Scale: Public trackers are open to everyone, but private trackers are available for registered users only. It is remarkable that the number of torrents in private trackers collectively will be much more than that in public trackers in existence [6].
- Incentive Mechanism: Public trackers rely on the BitTorrent TFT algorithm. Private trackers implement an additional SRE mechanism to incent the users to contribute as much as possible.
- Traffic Counting: Public trackers do not count each user's traffic during content distribution. Private trackers accurately record each user's total amount of uploaded data (T_u) and downloaded data (T_d). A user's ***Share Ratio*** is then calculated as T_u/T_d. If a user's Share Ratio is lower than a predefined threshold, he will be banned from this private tracker community.
- Download Performance: The download performance of private trackers are usually much better than public trackers due to the SRE mechanism. This is because a user in private tracker is incented to provide a high upload bandwidth and a long seeding time.

Given the importance and popularity of private trackers, it is essential to understand their characteristics so as to help us design better mechanism and build better sustainable environment. To this end, we conduct a measurement study and provide a thorough analysis on the collected datasets. Our main contributions are summarized as follows:

- We have crawled 13 private trackers and 2 public trackers from September 28, 2009 to February 10, 2009, and have obtained 31 datasets that cover nearly 2.5 million torrents.
- By analyzing the collected datasets, we find that private trackers have apparently different features and statistics

from public trackers, ranging from the user viscosity, torrents evolution, user behaviors, content distribution, et al. We show that SRE is an effective mechanism to incent users to seed as much as possible.

- Furthermore, we use game theory to study the effectiveness of SRE mechanism. We model the mechanism and propose an improved SRE mechanism to further incent the users and enhance the performance of private trackers.

The rest of our paper is organized as follows. Section II provides an overview of private trackers. Section III describes our measurement methodology and the collected datasets. Section IV presents detailed analysis on the measured datasets. The modeling of SRE mechanism based on game theory is presented in section V. Related work is presented in section VI, following by our conclusions in Section VII.

II. OVERVIEW OF PRIVATE TRACKERS

In this section, we provide a taxonomy of private trackers, define certain terms which are widely used in this field, and briefly introduce the operation principle of private trackers.

A. Taxonomy Overview

We begin our taxonomy by classifying the roles and operation processes of private trackers. We break down the roles in private tracker into the two components shown in Table I. In general, private tracker's sysops establish and publish private tracker by using open source codebase, and then operate it by attracting users with high ranking [7]. Users sign up in private trackers by using some approaches [8]. After that, registered users select accepted clients to seed or leech contents above minimum share ratio, and then do contributions if possible (e.g. upload contents, donation, etc.). For each role and operation process we give related terms and definitions in part B.

TABLE I. BREAKDOWN OF PRIVATE TRACKERS

Private Trackers		
Role	***Operation Process***	***Related Terms***
Tracker User	Registration	Invitation Code
	Sharing (Seeding & Leeching)	Share Ratio, Passkey, Slot, Snatched, Freeleech, HnR, Point System
	Contribution	User Class
Tracker Owner	Building & Publishing	Codebase, Seedbox
	Ranking Improving	Scene Release, Pre/Pre'd, Pre-time

Related Terms do not cover all components to operation processes.

B. Terminologies and Definitions

The terms used in the private trackers are not standardized. For the sake of clarity, we define the terms shown in Table I, which are widely used in private trackers. Note that the terms list here are non-exhaustive.

Invitation Code: It is a typical registration way to become a member of closed private tracker communities. Contributing members meet specific requirements are eligible to invite their friends to join this private tracker by sending an invitation code.

Share Ratio: The private tracker calculates the share ratio for each user, which is the total amount of data the user has uploaded, divided by the total amount it has downloaded.

Passkey: It is a unique identity that private trackers assign each registered user. Passkey is usually a hexadecimal string. It is appended to the announce URL in the .torrent file which is dynamically generated for each member. This is to prevent private torrents from being uploaded to public websites. Well behaved users should not leak the passkey and announce URL to other members or public trackers; otherwise, they will be banned from the private trackers.

Slot: The slot system is used to limit the concurrent downloads for members that have ratio below a minimum value. If all the download slots are filled, the system will deny any connection before validating.

Snatched: It indicates that how many times a torrent file is completely downloaded.

Freeleech: When a torrent is flagged with freeleech, it means leeching that torrent will neither affect member's download amount nor decrease his share ratio. Meanwhile seeding that torrent will continue to increase upload amount.

HnR: It refers to the behavior that a peer stops uploading as soon as it completes the downloading. HnRs are highly prohibited, as they are counter-productive to sharing and the health of the torrent. Different trackers have different rules about HnR.

Point System: Many private trackers incorporate a "point/credit system" with SRE mechanism. Points can be earned maintaining a good share ratio; uploading torrents, etc. Points can be spent to improve member's account authority.

User Class: Most private trackers deploy a ranking system to categorize their users based on each user's contribution to the community. For example, the users in CHDBits are categorized as shown in Table II [9].

Codebase: Typically private trackers do not develop themselves, but instead use or modify many open source codebases, such as TBDev [10], TS SE [11] , etc.

Seedbox: A seedbox is a private dedicated server used explicitly for torrent transfers or seeding at high rate.

Scene Release: Basically it means that a DVD or software which has been cracked by someone or some groups.

Pre/Pre'd: When a release group pres a release (distribution content), it will be available for other people and the distribution will start.

Pre-time: Since proper Scene Releases are not directly pre'd on trackers, normally they are done on IRC channel, there exists a waiting period before it arrives the private trackers.

C. Operation Principle of Private Tracker

Private trackers implement a strict set of rules to control member eligibility and content quality. Users sign up to be a member of private tracker through an invitation system. Under a passkey system, each member is given a unique announce URL to perform content distribution. In a predefined User Class system which adopts SRE mechanism, members can be automatically or manually promoted or demoted. The typical operation principle of private tracker is illustrated in Fig. 1.

TABLE II. USER CLASS IN CHDBITS TRACKER

Different User Classes		
User Class	***Explanation***	***How to Work***
Peasant	Demoted users who must improve their ratio within 20 days or they will be banned.	Demoted to this class if: Downloaded>10GB & ratio<0.3
User	Default class of new members. Can upload subtitles and delete subtitles. Cannot view NFO file.	Downloaded>50GB & ratio<0.4 Downloaded>100GB & ratio<0.5 Downloaded>200GB & ratio<0.5 Downloaded>400GB & ratio<0.7
Power User	Can view NFO file. Can view user list, Ask for reseed, Send invitation, View Top 10 and View other's torrents	Be a member ⩾ 5 weeks; Downloaded ⩾ 50GB And ratio ⩾ 1.05; Auto demoted if ratio < 0.95.
Elite User	Elite User or above would never be deleted if parked.	Be a member ⩾ 10 weeks; Downloaded ⩾ 120GB And ratio⩾ 1.55; Auto demoted if ratio < 1.45.
Crazy User	Can view other users' history of comments and forum posts. Can be anonymous when seeding/leeching.	Be a member ⩾ 15 weeks; Downloaded ⩾ 300GB And ratio⩾ 2.05; Auto demoted if ratio < 1.95.
Insane User	Can upload torrents without going through the Offer section.	Be a member ⩾ 20 weeks; Downloaded ⩾ 500GB And ratio⩾ 2.55; Auto demoted if ratio < 2.45.
Veteran User	Veteran User or above would never be deleted whether parked or not.	Be a member ⩾ 25 weeks; Downloaded ⩾ 750GB And ratio⩾ 3.05; Auto demoted if ratio < 2.95.
Extreme User	Can update outdated external information.	Be a member ⩾ 25 weeks; Downloaded ⩾ 1TB And ratio⩾ 3.55; Auto demoted if ratio < 3.45.
Ultimate User		Be a member ⩾ 30 weeks; Downloaded ⩾ 1.5TB And ratio⩾ 4.55; Auto demoted if ratio < 3.95.
Nexus Master		Be a member ⩾ 30 weeks; Downloaded ⩾ 3TB And ratio⩾ 4.55; Auto demoted if ratio < 4.45.

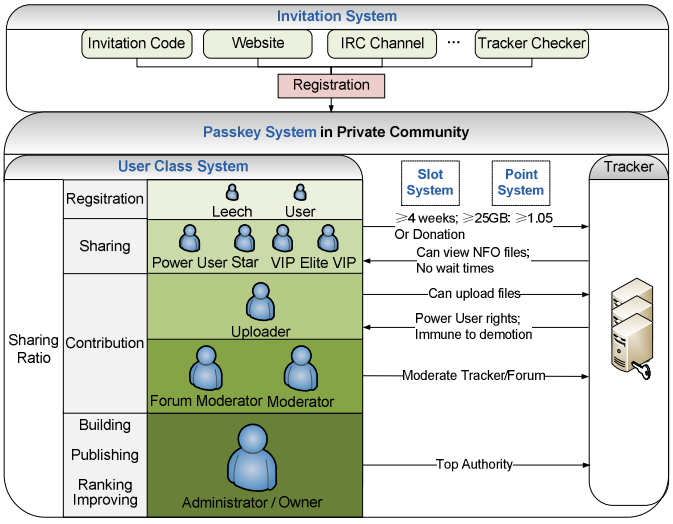

Figure 1. The Typical Operation Principle of Private Tracker

III. MEASUREMENT SETUP

In this section, we present the details of our measurement setup, methodology, and an overview of our collected datasets.

A. Selection of Private Trackers

There are a huge amount of private trackers on the Internet [12]. Most of them are closed systems, and it is difficult for an outsider to receive an invitation code. Nevertheless, based on the list and categories of [7, 12], we successfully joined 13 representative private trackers to perform our measurement study. For comparison purpose, we also crawled 2 well-known public trackers. We summarize all these trackers in Table III.

TABLE III. TRACKER LIST

Category		Tracker Information		
		Name	***# of Active Torrents***	***# of Registered Users***
Public	General	thePirateBay	19,931,758	4,196,330
		TorrentPortal	2,345,113	1,197,133
Private	General	TorrentLeech	24,347	N/A
		RevolutionTT	29,760	N/A
		Demonoid	236,983	N/A
		Bitsoup	13,474	N/A
		ILoveTorrents	9,313	N/A
	HD (High Definition)	CHDBits	12,093	18,939
		HDStar	6,805	8.397
		HD-Torrents	8,356	35,924
	Foreign	Torrents.ru	723,798	3,804,600
	Music	DimeaDozen	36,615	109,990
	TV	TheBox.bz	51,999	N/A
	DVD	AsianDVDClub	26,144	43,043
	Adult	PureTNA	66,883	1,306,501

Data updated until. 16:00, February 4, 2010. N/A indicates the tracker does not provide it.

B. Data Crawling and Collection

We use passive method to collect private tracker traces by using a crawler developed by ourselves. Data collection is recorded periodically by crawling web pages provided by each private tracker. These datasets are divided into three aspects, as shown in Table IV. "Tracker statistics" covers the information of registered users, torrents, seeders, leechers, and seeder-to-leecher ratio, etc. "Top 10" includes information about top members based on traffic, speed, share ratio, etc. "Torrent list" contains detailed torrent information, such as torrent type, name, size, added time, snatched times, seeders, leechers, etc. Notice that trace duration includes tracker maintenance periods during which our crawling cannot be done.

TABLE IV. CRAWLING DATASETS

Tracker Name	Tracker Statistics	Top 10	Torrent List	Trace Duration (days, mm/dd/yy)
thePirateBay	√	N/A	(Part)	**31**, 12/29/09 – 01/28/10
TorrentPortal	√	N/A	√	9, 02/02/10 – 02/10/10
TorrentLeech	N/A	√	√	27, 01/15/10 – 02/10/10
RevolutionTT	N/A	N/A	√	**37**, 12/26/09 – 02/10/10
Demonoid	N/A	N/A	√	10, 01/17/10 – 01/26/10
Bitsoup	N/A	N/A	√	**103**, 10/16/09 – 01/26/10
ILoveTorrents	N/A	N/A	√	**119**, 09/28/09 – 01/26/10
CHDBits	√	Top 250	√	**119**, 09/28/09 – 01/26/10
HDStar	√	√	√	**112**, 10/05/09 – 01/26/10
HD-Torrents	√	√	√	19, 01/23/10 – 02/10/10
Torrents.ru	√	N/A	N/A	**55**, 10/01/09 – 11/25/09
DimeaDozen	√	√	√	9, 02/02/10 – 02/10/10
TheBox.bz	N/A	N/A	√	9, 02/02/10 – 02/10/10
AsianDVDClub	√	√	√	9, 02/02/10 – 02/10/10
PureTNA	√	√	√	9, 02/02/10 – 02/10/10

N/A indicates the tracker does not provide it or we cannot acquire it.

There are three challenging issues we need to address: incomplete torrent list of public trackers, information loss caused by crawling time intervals, and limited trace durations.

ThePirateBay only allows displaying recent 50 pages for each torrent category and 300 pages for all torrents, so we cannot depict the whole picture of thePirateBay, but we have crawled as many torrents as we can. On the other side, we have succesfully crawled the complete torrent lists from all private trackers.

The frequency of data collection should be carefully tailored to achieve a smooth crawling. Many trackers enforce a minimum time interval between two page visits. To solve this issue, we estimate the relation between information loss ratio with time interval. We configure our crawlers to execute every 4 hours for tracker statistics and torrent list, every day for Top 10 because its change is not frequent. Complete torrent list can reflect the overall population of torrents in the tracker. Furthermore, we have crawled two popular torrents every 10 minutes to study torrent life span.

Another issue is the limited trace duration because we try to draw unbiased conclusion from these complete torrent lists. To avoid bias when studying complete torrent lists, we refer to the method in [13]. Those torrent lists whose trace durations are longer than 30 days will be studied in our paper. We also take other samples into study but they can be viewed as preliminary results reference.

IV. MEASUREMENT AND ANALYSIS

In this section, we present our detailed measurement results and analysis. We discuss the user viscosity, single torrent evolution, user behaviors, distribution of private trackers, and the comparison with public trackers.

A. *User Viscosity of Private Trackers*

According to the traffic ranking (updated data on Feb. 6, 2010) from Alexa [14], we select 6 popular public trackers (thePirateBay.org, Torrentz.com, ISOHunt.com, BTJunkie.org, TorrentReactor.net, Mininova.org) and 6 most famous private trackers (Torrents.ru, Demonoid.com, TorrentLeech.org, PureTNA.com, Bitsoup.org, TheBox.bz) from our crawling list.

Page views measure the number of pages viewed by site visitors. Multiple page views of the same page made by the same user on the same day are counted only once. Page views per user is the average number of unique pages viewed per user per day. Fig. 2 shows the current page views per user of each tracker. It is obvious that private trackers have higher page views per user than public trackers. This implies that users in private trackers have more interest than public trackers. User viscosity of private trackers is better than public trackers.

Next, we take a close look at this viscosity of users by analyzing the number of users in each class of CHDBits. Fig. 3 shows our measured data of CHDBits in a 3-month period[1]. The total number of users keeps relatively stable[2]. In general, stability of total number of users is applicable to other private trackers. It is mainly dependent on two factors: the capacity of the private tracker and the maintenance of the tracker's administrator.

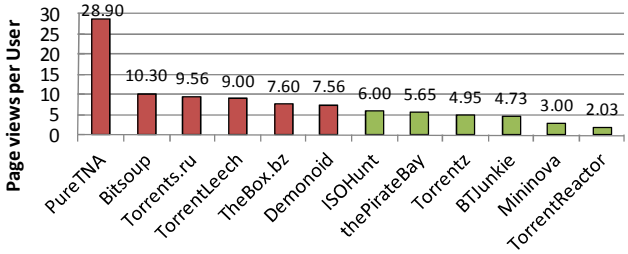

Figure 2. Page views per User in Popular Trackers

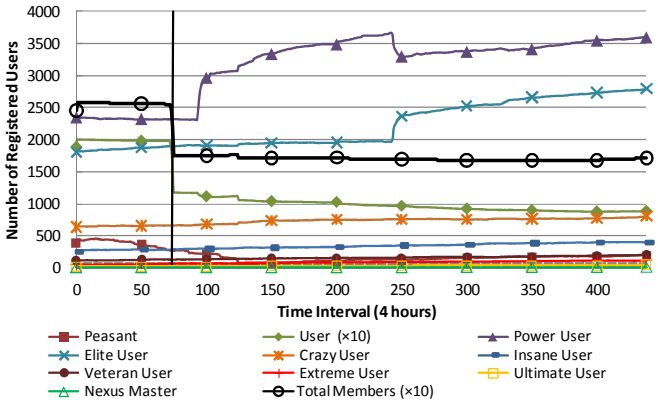

Figure 3. Number of Registered Users of each User Class in CHDBits

Different private trackers adopt different user class policy, but their internal structures are very similar. In Fig. 3, most of the users belong to the User class (i.e., share ratio is lower than 0.9). We observe that at the number of users in User class in decreasing with time, while the numbers of users in Power User and Elite User classes are increasing. It shows the effectiveness of the ranking system in private trackers as users are incented to promote themselves to a higher user class.

B. *Torrents of Private Trackers*

In this subsection we investigate the behavior of a single torrent and compare with a torrent in public tracker TorrentPortal. Fig. 4 presents the evolution of the number of seeders, leechers, snatched times and the sum of the former three of a single popular torrent in CHDBits. The main finding is that the number of seeders is significantly larger than the number of lechers most of the time. It will be further discussed in subsection C.

Fig. 5 shows the number of peers in an active torrent (i.e., a torrent with at least one seeder) in DimeaDozen and TorrentPortal. Torrents are ordered from the largest to smallest based on the number of peers joined in each torrent. It is denoted as Torrent Rank in the label of X-axis. In Fig. 5, we can see there exist "giant" torrents not only in public trackers, but in private trackers which contain more than 1,600 seeders. Besides, we can see there are more peers in a torrent in private tracker than in public tracker. That guarantees the high download speed in private tracker. Of course, due to the limitation of server capacity, the "giant" scale in private tracker cannot reach the level of public tracker. E.g. in TorrentPortal, there are ten torrents with 8,388,607 seeders,

[1] Notice that the unit of curves "User" and "Total Members" is 10.

[2] The reason of the big drop of total members and User class around time 70 is because the administrator deleted many inactive members.

and some of them even have 8,388,607 leechers. These torrents are usually popular TV series and movies.

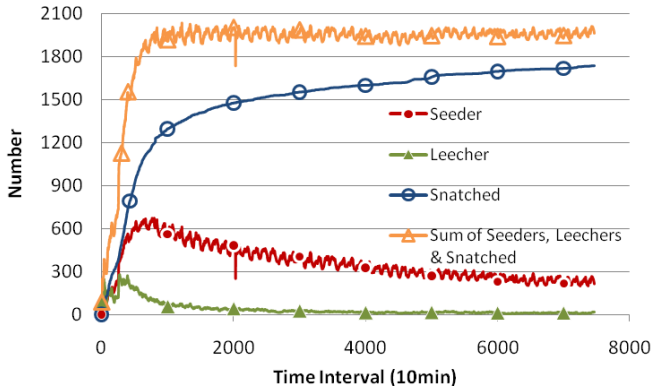

Figure 4. Seeder, Leecher, Snatched & Sum of Single Torrent in CHDBits

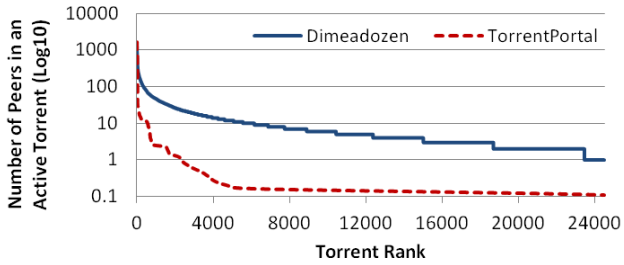

Figure 5. Number of Peers in Active Torrents

Fig. 6 and Fig. 7 show the age distribution and transfer speed of torrents respectively in DimeaDozen (music private tracker). In Fig. 6, about 40% torrents are uploaded in 3 months and 82% of torrents are uploaded within one year. Surprising observation is that there are torrents nearly 3 years that are still active. In Fig. 7, though there are a lot of long live torrents with seeders in private tracker, more than 98% torrents with small leechers have no more than 100KBps in this music private tracker, only those popular torrents attract many leechers and have very high speed. It implies private tracker has a starvation condition that the number of seeders is significantly larger than the number of leechers. Furthermore, because the supplies from seeders exceed the demands of leechers, so the leechers' download speed can be very high in private trackers.

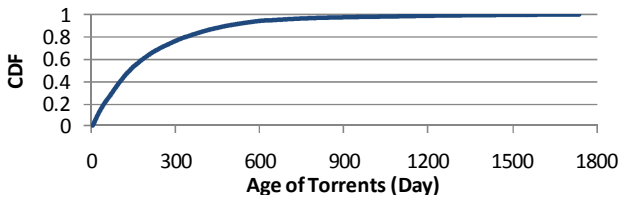

Figure 6. Age Distribution of Torrents in DimeaDozen

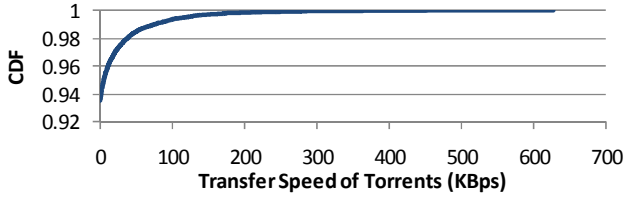

Figure 7. Transfer Speed Distribution of Torrents in DimeaDozen

C. *User Behaviors of Private Trackers*

We have shown that private trackers have better user viscosity, longer seeding time and more active seeding peers, etc. In some degree, they reflect the activity of members in private tracker communities. This subsection presents our findings about user behaviors in private trackers.

1) *Activity of Users*

Active user is a set which includes two parts: active tracker user who is actually participating the content distribution, and active browsing user who is browsing the private tracker web site. Fig. 8 shows the active user ratio (active users to total users) statistics per day and per week in CHDBits and HDStar. In Fig. 8, though these two trackers have different number of users, their active user ratios are surprisingly very close to each other: the per day ratio keeps around 50% and per week ratio keeps around 80%. From Fig. 8 we can see that, compared with active user ratio per week, there still exists nearly 30% members who are inactive per day.

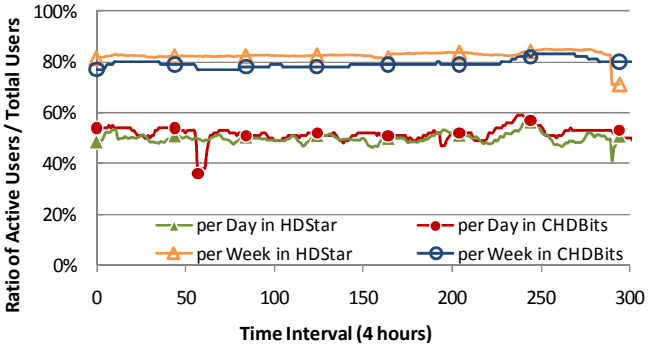

Figure 8. Ratio of Active Users/Total Users per Day & per Week

Then we select one week period (Oct. 13, 2009 to Oct. 20, 2009) in CHDBits and HDStar to compare the change of active browsing user and active users connecting to the tracker (for short, active tracker users). In Fig. 9, the active browsing users in two trackers present the same trend with the change of time. The wave peak happens at 22:00PM every day and the sub-peak is at 10:00AM every day, which means that there are most of browsing users around these two time points. The wave hollow is located at 6:00AM every day.

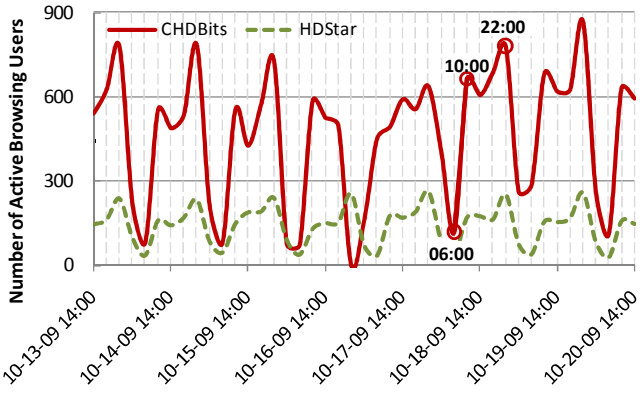

Figure 9. Number of Active Browsing Users

When adding the factor of active tracker users, the result of ratio (active tracker users to active browsing users) in Fig. 10 shows that the wave peak becomes 6:00AM, while the wave

hollows are at 10:00AM and 22:00PM. That means most users in private trackers keep staying in the BitTorrent swarm. For example, at 6:00AM, most users are not browsing, but their clients keep connecting to the tracker, so the ratio of active tracker users to active browsing users presents peak status. The reason of the CHDBits irregular wave changes (around 22:00PM, Oct. 17, 2009 to 10:00AM, Oct. 18, 2009) in Fig. 10 is that the tracker suffered non-periodical DDoS attack, which also affects the tracker in the following 2 days, so members' clients have difficulty in connecting to the tracker.

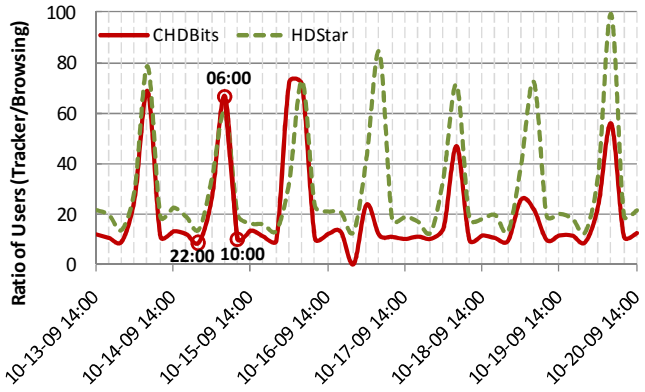

Figure 10. Ratio of Tracker Active Users/Active Browsing Users

2) *Activity of Peers*

In order to further investigate users' behaviors in private trackers, we track the ratio of seeding peers to leeching peers and show the result in Fig. 11. Generally, the number of seeding peers in private trackers is 5 to 15 times of the number of leeching peers[3]. It indicates that users in private trackers are seeding more than leeching. The wave properties of three trackers also follow the interchange pattern of daytime and nighttime.

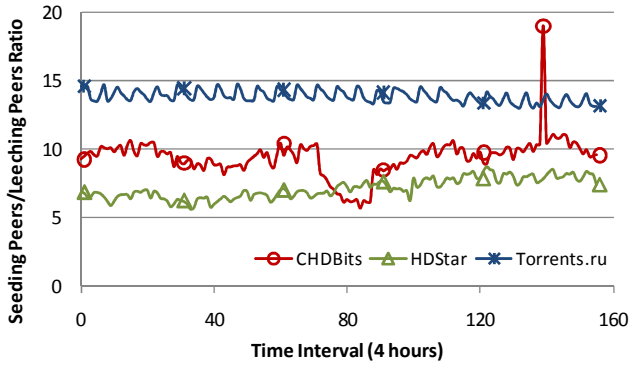

Figure 11. Ratio of Seeding Peers/Leeching Peers in three Pirvate Trackers

3) *Traffic of Private Trackers*

This subsection shows that statistics of the traffic of private trackers. Traffic here indicates the sum of uploaded and downloaded traffic. In Fig. 12, on one hand, we can see that the total traffics per user of CHDBits and HDStar trackers keep increasing from 430TB, 496TB to around 955TB, 851TB, respectively. Fig. 13 shows that total traffic in Torrents.ru is surprising 15-35GBps with nearly 700,000 living torrents. The tremendous traffic may bring heavy burden to our ISPs. On the other hand, the ratios (total uploaded to total downloaded) of two trackers maintain the same value (around 5). This further proved that users in private trackers are doing more seeding than leeching. It reflects a kind of imbalance: supply is more than demand, which also proves that the high download speed in private tracker is partly because of the enormous seeders in the swarm.

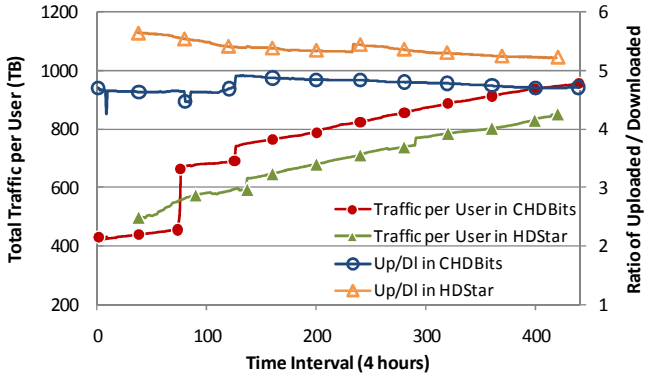

Figure 12. Total Traffic/User & Ratio of Total Uploaded/ Downloaded

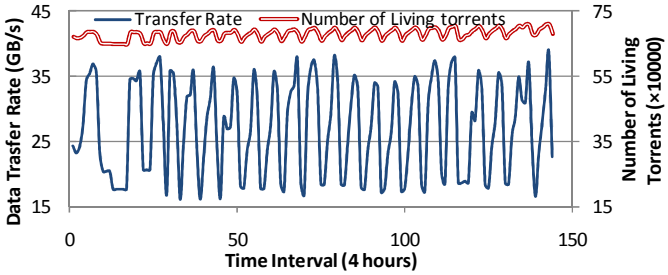

Figure 13. Data Transfer Rate in Torrents.ru

4) *Top 250 of Private Tracker*

CHDBits provides us lists of Top 250 from many aspects, such as best downloaders, best uploader, best share ratios, etc. Fig. 14 shows the change of top 250 members with best share ratios. Though CHDBits have over 15,000 members, there are no more than 250 members who have share ratios of over 10. Actually, most of members only have the share ratio of no larger than 1. E.g., in current DimeaDozen, there are 39% members whose share ratios are lower than 0.25, 75% members whose share ratios are lower than 1.

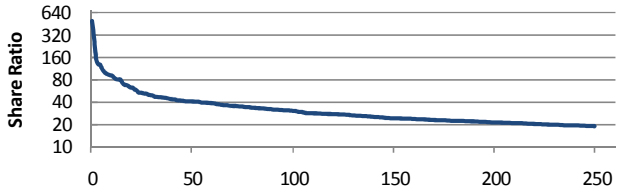

Figure 14. Top 250 Members with Best Share Ratio in CHDBits

Fig. 15 presents the fastest uploading and downloading speed in top 250 members. Combined the results of subsection B and C, we may conclude that some members with better share ratios are not very active. Besides, there are many seeding peers in private trackers, so members can easily download with high speed. This will induce low share ratios,

[3] The abnormal change in CHDBits curve is because the relocation of the server at that time, so the users can hardly connect to the tracker.

and keeping with low share ratios is definitely not safe in private trackers. Therefore, active members have to seed for a long time (shown in Fig. 6), and hungrily aim for safe share ratio to avoid being banned from the trackers.

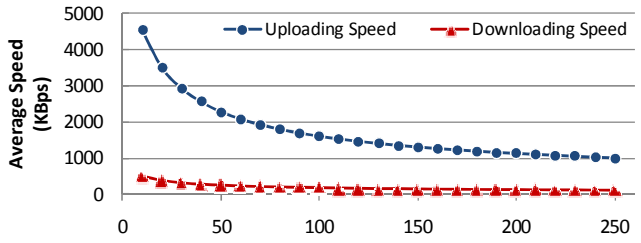

Figure 15. Top 250 Members with Fastest Transfer Speed in CHDBits

5) Active Torrents Rate of Private Trackers

Active torrents rate is the ratio of the number of torrents which have at least one seed to the total number of torrents in the tracker. Along with the increasing torrents scale, maintaining the active torrents rate is one of KPIs (Key Important Factor) to evaluate the active level and ranking of trackers. Fig. 16 compares the active seeder rate of 14 trackers. Apparently, most private trackers are far better than public trackers.

We use $N(s)$ indicates the number of seeders in a active torrent, and $N(l)$ indicates the number of leechers in a active torrent. $N(s\_l)$ is the sum of torrents which $N(s) > N(l)$, and $N(l\_s)$ is the sum of torrents which $N(l) > N(s)$. Fig. 17 compares the ratio of $N(s\_l)$ to $N(l\_s)$. Except for AsianDVDClub, all private trackers have more $N(s\_l)$, which means they are more active than public trackers.

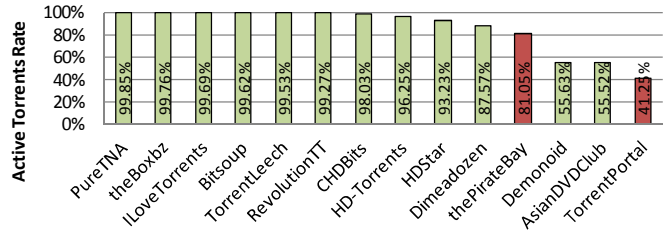

Figure 16. Rank of Active Torrents Rate in Trackers

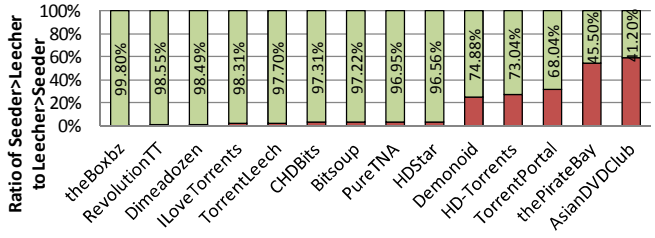

Figure 17. Rank of Seeder > Leecher / Leecher>Seeder Ratio

D. Content Distribution of Private Trackers

This subsection shows the content size distribution and category distribution in private trackers. In order to be representative, we choose to show results of one music tracker (DimeaDozen), one HD tracker (CHDBits) and one general tracker (RevolutionTT).

1) Content Size Distribution

Fig. 18 shows the classification of content size distribution base on the number of contents in two private trackers, respectively. We can see the most number of contents are small contents with 0-1GB, followed by the contents which are large than 100GB.

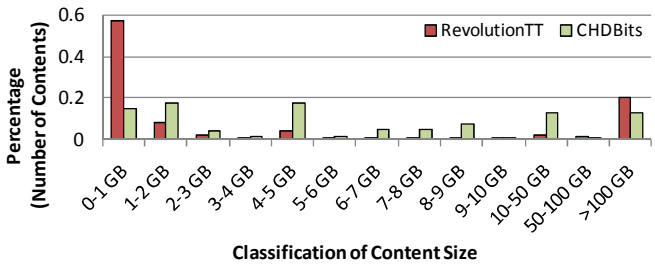

Figure 18. Distribtuion of the Number of Contents

2) Category Distribution

Fig. 19 and Fig. 20 show the category of content distribution based on the number of contents in two private trackers, compared with peer distribution respectively. We can see that popularity of content category basically matches the peer distribution accordingly. The DimeaDozen provides detailed categories of music, that is, 61 categories, which shows a long tail distribution. We list the top 5 category percentage in Table V. Fig. 5 shows that TV series, movie and adult are the leading types which occupy 81% share in private tracker. There is a great diversity in the content distribution based on category.

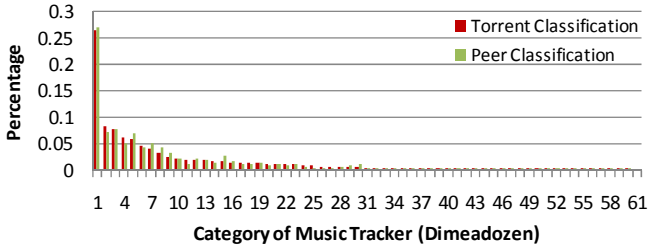

Figure 19. Classification of Peers and Torrents in Dimeadozen

TABLE V. TOP 5 CATEGORIES IN DIMEADOZEN

Category	Rock	Alternate	Jazz	Singer or Songwriter	Progressive Rock
Percentage	27.03%	7.17%	7.63%	5.15%	6.92%

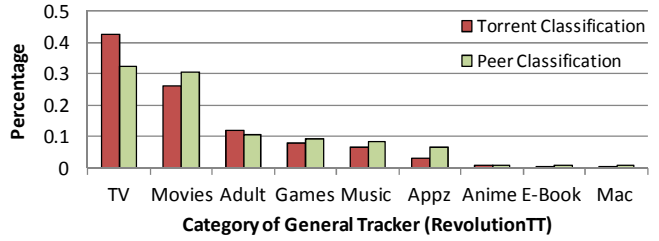

Figure 20. Classification of Peers and Torrents in RevolutionTT

Except for the above measurement, there are other metrics which can measure the performance of private trackers, such as pre-time, firewall rate, freeleech, download slots, etc. We will give more detailed in our technical report.

V. ANALYSIS AND IMPROVEMENT OF SRE MECHANISM

In this section, we will analyze the effectiveness of private tracker's SRE mechanism based on game theory. Furthermore, we model the mechanism and propose an improved SRE mechanism to incent the users' activity and enhance the quality of service of private trackers.

A. Analysis of SRE Mechanism based on Game Theory

From a macroscopic view, the total download bandwidth is equal to the total upload bandwidth in a BitTorrent swarm. The reason of free-riding problem is that some users contribute a lot but receive unfair return. BitTorrent adopts TFT algorithm as a kind of incentive mechanism, but it has been shown that increased upload contribution only marginally improves download rates [15]. Peers lack the motivation to seed and have no reason to contribute once they have satisfied their immediate demands. In private trackers, we have previously shown that SRE (Share Ratio Enforcement) is a very effective auxiliary incentive mechanism through our measurement study. Here we use a general game-theoretic framework to analyze SRE and answer the following question: Why does the TFT mechanism alone fail to achieve the same level of performance as private trackers that use SRE?

In order to simplify the game, we assume that each peer has two strategy sets, upload (*UL*) and download (*DL*). *N* is the set of natural number.

$UL = \{(ul_1, ul_2, ..., ul_p) \mid p \in N\}$, denotes p different upload levels sorted from lowest to highest;

$DL = \{(dl_1, dl_2, ..., dl_q) \mid q \in N\}$, denotes q different download levels sorted from lowest to highest.

Each peer selects a pair of (ul_i, dl_j) from *UL* and *DL*, and builds its own strategy space S_k:

$$S_k = \{(ul_i, dl_j) \mid i, j, k \in N,\ 1 \le i \le p,\ 1 \le j \le q\}$$

Then *n* peers in P2P system construct a strategy space set *S*:

$$S = (S_1, S_2, ..., S_n), \quad n \in N$$

In general, peer wants to download by fulfilling its bandwidth configuration or maximum bandwidth. But it is difficult to download as pre-selected in *DL* set. In P2P content distribution network, the actual download performance is decided by the number of active peers joined in the swarm. Here we use the actual download performance of peer i to define the utility u_i according to its own strategy s_i. The set of u is defined as:

$$u = (u_1, u_2, ..., u_n), \quad n \in N$$

Each peer has an optimal strategy s_i^* from its strategy space S_k. We define the optimal strategy set as s^*:

$$s^* = (s_1^*, s_2^*, ..., s_n^*), \quad n \in N$$

Then the Nash equilibrium can be defined as:

$$u_i(s_1^*, s_2^*, ..., s_i^*, ..., s_n^*) \ge u_i(s_1^*, s_2^*, ..., s_i^{'}, ..., s_n^*),$$
$$\forall s_i^{'} \in S_i,\ s_i^{'} \ne s_i^*,\ \forall i \in N$$

If peers are rational, they will choose to upload with minimum speed and download with maximum speed (ul_1, dl_q). If all peers upload with maximum speed, then peers' utility will become high. If peer i chooses $(ul_1^{\ i}, dl_q^{\ i})$ and peer j chooses $(ul_p^{\ j}, dl_q^{\ j})$, then peer i's utility will become intermediate, peer j's utility will become low. Therefore, two peers' utilities can be represented in Table VI.

TABLE VI. UTILITY FUNCTION

	Peer *j* (ul_{high})	**Peer *j*** (ul_{low})
Peer *i* (ul_{high})	$(\underline{u_{high}}, \underline{u_{high}})$	$(u_{low}, u_{intermmediate})$
Peer *i* (ul_{low})	$(u_{intermmediate}, u_{low})$	$(\underline{u_{low}}, \underline{u_{low}})$

Based on game theory, we know there are two Nash equilibriums in Table VI. (u_{low}, u_{low}) is strict Nash equilibrium, and (u_{high}, u_{high}) is weak Nash equilibrium. Because peers are rational, so both sides will choose upload with minimum speed to reach the strict Nash equilibrium, and hence Tragedy of Common will occur. Though BitTorrent adopts TFT as a conflict resolution scheme, notice that TFT is based on repeated game and it can only take effect in infinite repeated game. However, there does not exist actual infinite repeated game but only finite repeated game in P2P content distribution. We will get the same result in finite repeated game and in one time game [16]. So the final result of using TFT algorithm induce the strict Nash equilibrium (u_{low}, u_{low}).

But if we introduce SRE mechanism in P2P content distribution system, peers cannot just be rational, and they are enforced to upload/seed to reach certain share ratio under a giving SRE policy. Even if a peer still chooses to upload with minimum speed, it must reach the required minimum share ratio by seeding for a long time. Therefore, the strict Nash equilibrium is broken, and the equilibrium will transit to the weak Nash equilibrium (u_{high}, u_{high}). This is an optimal equilibrium. All peers download and upload with high speed. This will best utilize the potential of all peers and achieve the maximum efficiency of content distribution.

B. Modeling of SRE Mechanism

Here we choose one general tracker (RevolutionTT) and one HD tracker (CHDBits) to model their SRE mechanisms. Their SRE mechanism is defined as Table VII.

TABLE VII. SRE MECHANISM

	Downloaded Data	**Share Ratio (no less than)**
RevolutionTT	3.0GB	0.3
	6.0GB	0.50
	10.0GB	0.80
	>10.0GB	>0.8
CHDBits	10GB	0.3
	50GB	0.4
	100GB	0.5
	200GB	0.6
	400GB	0.7
	>400GB	0.95

In general, SRE mechanism is related with two variables: share ratio (r) and the volume of downloaded data (d). RevolutionTT and CHDBits have different SRE mechanisms. According to Table VII, we found that SRE in both trackers

can be modeled as S-Curve: it first starts with a rapid growth stage and then enters into a stable and smooth stage. We model the S-Curve by the following equation:

$$\ln(r) = b_0 - \frac{b_1}{d}, \quad r \geq 0, d > 0, r, d \in R \tag{1}$$

Through our regression analysis, b_0 is 0.179 and b_1 is -4.326 in RevolutionTT. In CHDBits, b_0 is 0.023 and b_1 is -11.81.

C. Improvement of SRE Mechanism

Though the SRE mechanism achieves a big success in private trackers, we still notice that there are a lot of inactive users. Although administrators can delete inactive users periodically, it is better to improve the activity of users. We will consider the following factors:

- *R*: It is the SRE mechanism used in private trackers.
- *A*: It denotes the Activity of user in terms of seeding time ($\Omega(t_s)$) and the number of seeding torrents ($\Psi(n_i)$).
- *P*: It denotes the Period of time to the last seeding ($\mathrm{P}(t_l)$), the longer to the present, the larger the *P* is.

Our improved SRE mechanism (RAP) is shown as Eq. (2) to calculate the final ratio:

$$Ratio_{final} = r - \mathrm{P}(t_l) + \sum_{i=1}^{n} \Omega(t_s) \times \Psi(n_i), \tag{2}$$
$$r, t_l, t_s, n_i \geq 0, r, t_l, t_s \in R, n_i \in Z$$

Eq. (2) enforces that a user can not be inactive for a long time; otherwise his final ratio will be decreased by the inactive time period. The final ratio can also be increased by the seeding time and number of seeding torrents.

VI. Related Work

To the best of our knowledge, there are little studies based on measurements, analysis and modeling with BitTorrent private trackers. In 2005, Nazareno et al firstly studied the Share Ratio Enforcement (SRE) in private trackers [17]. They used the same cooperation metrics to compare several BitTorrent communities, including one small scale private tracker (easytree.org). They viewed the SRE as extrinsic gifting which is giving not motivated by a direct, immediate benefit [18]. Then Nazareno et al deeply studied another private tracker (bitsoup.org) with other two BitTorrent communities based on resource demand, supply and their relationship [13]. It seems that private trackers do not attract enough researchers' attention at that time. Until 2009, Tribler team proposed that if a peer uploads more than it downloads in a swarm means the peer should have more credit. Private trackers possibly exist "credit squeeze" which indicates that the lacking of credit leads to significantly reducing of system efficiency [19].

Though the previous investigations revealed some characteristics of private trackers and obtained some primitive research results to SRE mechanism, the noticeable limitation is samples they studied are one or two private trackers, which provides a limited view of current various private trackers. It makes some ideas they proposed stayed in hypothesis stage but not conclusive. Though Chao et al. investigated 800+ private trackers, depicted a broad and clear picture of private trackers landscape [1], there is no clear picture of private trackers inside the box. They did not focus on SRE mechanism, and lacked in-depth studies on some metrics to measure the quality of service provided in a relatively wide range of private trackers.

VII. Conclusion

In this paper, we provide a useful taxonomy of private trackers. In breadth, we have crawled and compared system behaviors among 13 private trackers and 2 public trackers for a 4-month period. We find that private trackers have apparently different features and statistics compared to public trackers, ranging from the user viscosity, single torrent evolution, user behaviors, content distribution and other metrics to measure the quality of service in private trackers. In depth, we study the effectiveness of SRE mechanism based on game theory. Besides, we model the mechanism and proposed an improved SRE mechanism to further incentive the activity of users and enhance the quality of service of private trackers.

References

[1] C. Zhang, P. Dhungel, Di Wu, Z. Liu,and K.W. Ross, "BitTorrent Darknets," Infocom 2010, San Diego, 2010.

[2] "The Pirate Bay," http://thepiratebay.org/.

[3] "Mininova," http://mininova.org/.

[4] "isoHunt," http://isohunt.com/.

[5] "The 100 Most Popular Private Trackers Of 2008," http://filesharefreak.com/2008/12/31/the-100-most-popular-private-trackers-of-2008/, December 31, 2008.

[6] "Private Trackers with the Most torrents," http://filesharefreak.com/2009/09/19/private-trackers-with-the-most-torrents/, September 19, 2009.

[7] "Torrentking.org," http://www.torrentking.org/?order=rating.

[8] "The Essential Guide to Getting Into Private Trackers," http://filesharefreak.com/2009/06/18/the-essential-guide-to-getting-into-private-trackers/.

[9] "CHDBits," http://chdbits.org/faq.php.

[10] "TBDEV.net,"http://sourceforge.net/projects/tbdevnet/.

[11] "TS SE," http://templateshares.net/.

[12] "Trackers List," http://filesharefreak.com/trackers-list/.

[13] Nazareno Andrade, Elizeu Santos-Neto, Francisco Brasileiro, et al. "Resource demand and supply in BitTorrent content-sharing communities," Computer Networks Journal, December 2008.

[14] "Alexa," http://www.alexa.com/.

[15] M. Piatek, T. Isdal, A. Krishnamurthy, T. Anderson, "One hop reputations for peer to peer file sharing workloads," in NSDI08.

[16] Gibbons, R, "A Primer in Game Theory," Harvester, New York, 1992.

[17] Nazareno Andrade, Miranda Mowbray, Aliandro Lima, et al. "Influences on cooperation in BitTorrent communities." Third Workshop on the Economics of Peer-to-peer Systems, SIGCOMM 2005, August 2005.

[18] Matei Ripeanu, Miranda Mowbray, Nazareno Andrade, and Aliandro Lima, "Gifting Technologies: A BitTorrent Case Study," First Monday, vol. 11(11), November 2006.

[19] D Hales, R Rahman, B Zhang, M Meulpolder, et al. "BitTorrent or BitCrunch: Evidence of a credit squeeze in BitTorrent?" In the Proceedings of the 5th Collaborative Peer-to-Peer Systems (COPS) Workshop, June 29 - July 1, 2009.