

# Analyzing BitTorrent: Three Lessons from One Peer-Level View

A. Iosup      P. Garbacki      J.A. Pouwelse      D.H.J. Epema

Faculty of Electrical Engineering, Mathematics, and Computer Science  
Delft University of Technology

P.O. Box 5031, 2600 GA Delft, The Netherlands

{A.Iosup, P.Garbacki, J.A.Pouwelse, D.H.J.Epema}@ewi.tudelft.nl

**Keywords:** P2P analysis, peer-level view, BitTorrent

## Abstract

Peer-to-Peer systems like BitTorrent have recently attracted the interest of the Internet audience with their ability to share content at high speeds, while lowering the burden on the initial data owner. Besides optimizing the transfers, researchers are focusing on robustness and fault tolerance, issues inextricably linked with the location and activity of the network users. In this study we present a peer-level analysis of a large set of BitTorrent measurements and the correlation of three views on the location, time patterns and users activity. We present evidence that BitTorrent content presents non-trivial locality features and show that exploiting these features could greatly benefit the world of P2P systems, and in particular BitTorrent.

## 1 Introduction

Peer-to-Peer file sharing systems are becoming more and more present in the day-to-day life of content consumers. Recent studies show that P2P and especially BitTorrent file sharing traffic occupies as much as one third and over of the global data traffic [4, 5, 7], which makes it currently the dominant data traffic in the world. By comparison, Web surfers produce 10 times less traffic than P2P users [7]. The main challenge with such systems is to efficiently share data in the dynamic, ad-hoc networks of users, a task which relies heavily on the peer characteristics of such networks [9].

In this paper we analyze BitTorrent data covering several months, and over 200,000 unique users. Our analysis offers a unified set of location, temporal, and activity views of the BitTorrent community. We call this analysis a *peer-level analysis* and these views *peer-level views*, because they are characterizing the

peers of the BitTorrent network; we are specifically not characterizing other BitTorrent components (see also Section 2.1).

We show that such a comprehensive study sheds a new and non-trivial light upon the patterns that arise in the world of BitTorrent and, possibly, other P2P networks. Indeed, establishing patterns can result in understanding, and perhaps optimizing such systems. For instance, a P2P system could exploit the possibility that users from a given country would develop communities based on common taste, or that users from the same country would exhibit a similar data access pattern. Similarly, for some ISPs, caching (legal) P2P data transfers would result in huge savings, due to optimized use of the network capacity. Also, it is imperative to study the behavior of users when downloading shared content that is presented under different names (see Section 2.2.2).

The main contributions of our work are:

- The first unified set of location, temporal, and peer activity views of the BitTorrent network, for a large set of files covering different content categories;
- The evidence that BitTorrent shared content presents trivial and non-trivial locality features;
- The first view of BitTorrent's *alias media* phenomenon (see Section 2.2.2) for a very popular file.

## 2 Preliminaries

This section introduces the subject, the BitTorrent P2P network, and the vehicle of our work, the data set.

## 2.1 A brief introduction to BitTorrent

BitTorrent is a second generation peer-to-peer network. As such, its main feature is the high data transfer speed. Data exists in the network in the form of torrents (files or archives with multiple files). BitTorrent employs a *tit-for-tat* mechanism for data sharing – each user has to contribute to the network in order to obtain something from it; this leads in turn to data spreading at much higher speeds. To facilitate the exchange process, files are split in smaller parts, called *chunks*; the whole *tit-for-tat* mechanism is used to exchange *chunks*. Another important BitTorrent design issue is the lack of a contents search system at the peer level.

BitTorrent’s world is composed of peers, trackers, and web sites. As in any other P2P network, *peers* have theoretically equal rights to download files. However, to ensure high speed transfers, users with higher bandwidths are favored by the data transfer algorithm. By definition, for a file, the peers who own all the chunks are called *seeds*, while the peers who own only some or even none of the chunks are called *leeches*. The *trackers* are semi-centralized components (their activity is centralized but their number makes the whole level work as a decentralized network) used to keep track of the network transfers, as such, each file (content) in the network is monitored (*tracked*) by a tracker. By using the services of a tracker, BitTorrent peers can find out who are the other peers that can chunks of the desired file, so trackers also act as *redirector services* for the peers network. *Web sites* are used to locate data. For this, the web sites provide pages with BitTorrent digital content description (movie and music titles, authors, number of user that have the complete file or parts of it, and so on) that a network user may need to retrieve the desired digital data. The most important aspect regarding web sites is that they help providing *moderated* contents – that is, the web site supervisors verify the content that is inserted on the web and do not allow polluted data to be shared with unsuspecting users. Because of this, the BitTorrent network is virtually pollution-free.

## 2.2 The data set

This section describes the data set that is used in this study. We used 100 nodes from DAS-2<sup>1</sup> to record/follow peers at a one minute resolution. For this study, we discarded the nodes that could not be followed (e.g., firewalled nodes). It is important for this study that the data set covers over 95% of the existing peers for this data set [8]. For the complete description of the data acquisition methods we point the reader to [8].

<sup>1</sup>The Distributed ASCI Supercomputer, `cs.vu.nl/das2`

Group name	Group size (# of Files)	# of Records (in M)
All	120	42.81
Detailed	12	34.83
Varied	108	7.98

Table 1: The three groups of files.

### 2.2.1 Data set structure

The data set tracks the peer behavior during the download of 120 files using the BitTorrent network. Tracking observations are stored in *trace files*, in the form of time-stamped records, with each record consisting of an (*IP address, port, number of chunks*) tuple.

The data set was acquired between December 2003 and March 2004. A number of 12 *big* trace files have been acquired in the period December 2003-January 2004. Another 108 *small* trace files have been acquired during March 2004. The terms *big* and *small* refer to the number of time-stamped observations, with *small/emphbig* files containing under/over 500,000 time-stamped observations.

### 2.2.2 Alias Media

We have observed that the same contents (e.g., movie *X*) can be available in the same P2P network in several forms, in very similarly, but not identically, named files (e.g., movie *X* may appear as *X.by.YYY*, *X.created.on.03.2004* and so on). We call the whole set of files regarding the same contents *alias media*.

### 2.2.3 Data set distribution

The 120 *trace files* in the data set cover a wide set of file types, from games to movies and from music to computer applications. There are three general groups (see Table 1): group *All*, with all the trace files available in the data set, group *Detailed*, containing all big trace files, and group *Varied*, consisting of all the small trace files. The *Detailed* group contains trace files with sufficient observations to provide a detailed peer-level information. The *Varied* group contains trace files from a variety of content types (e.g., movies, games, music, and applications).

We have also split the data set into nine specific categories (see Table 2). For each category of files we have selected from group *All* the file with contents belonging to that category and for which we have made the most observations. In the case the selected file is part of an *alias media* group, we have selected all the traces corresponding to its *alias media*.

## 3 The Methods

This section presents the methods we used to analyze the peer-level of BitTorrent.

Category name	Category size (# of Files)	# of Records (in M)
Movie, popular	10	19.73
Movie, themed	1	2.12
Movie, anim�	2	0.01
Game, popular	1	13.14
Game, themed	1	0.17
Music, popular	1	0.15
Music, album	1	0.05
Application, popular	1	0.02
App, educational	1	0.03

Table 2: The nine categories of files.

### 3.1 Location analysis

Like any other overlay network, BitTorrent needs a proper peer location analysis – we need to discover the continents, countries, and cities where the users are located, and which are the Internet Service Providers (ISPs) which provide the underlying data transport network.

There is also a clear need to compare the number of users to their traffic. We define a user’s *weight* as the amount of data transferred by that user. Using the *number of chunks* field in the observation record, we were also able to measure the user weights for each country and continent.

#### 3.1.1 Geographical analysis

The geographical analysis provides information about the distribution of users and user weights per continent, country, and city. The distribution of users per continent provides interesting insights into the coarse—by continent—locality of the data. The distribution of users per country is critical for establishing whether there are countries with a clear preference for specific types of content. Beside the trivial correlation country/language, there are countries having strong preferences for just some types of contents (e.g. Japan being an important consumer only for animes). Finally, identifying shared content preferences for distinct cities could help optimizing the immediate (local) data sharing. It is noteworthy that these features that cannot be derived from any current model of a P2P network, as they vary with the type of contents being shared.

We performed the IP address—continent, country, city association for the groups and categories presented in Tables 1 and 2. We used MaxMind’s<sup>2</sup> GeoIP libraries and databases and on WebLog Expert’s<sup>3</sup> databases to map IP addresses to their corresponding continents, countries, and cities. Having a local copy of an IP address lookup database allowed

<sup>2</sup>MaxMind, <http://www.maxmind.com>

<sup>3</sup>WebLog Expert, <http://www.weblogexpert.com>

us to completely map our large data set in a reasonable amount of time. We did not encounter any problems when mapping IP addresses to their respective continents and countries, as at most 1% of the observations could not be matched. However, over 30% of the IP addresses could not be associated with any city, mainly because the IP address-city name mapping is not always publicly available.

#### 3.1.2 Organizational analysis

We follow the MaxMind’s definition of organizations: corporate networks, or ISPs for home users. The organizational analysis is essential for establishing the impact of caching file transfers at the ISP level, with potentially huge savings in the transferring costs. Since caching may raise important legal issues when being applied to illegal content, it is noteworthy that P2P networks are used now also for transferring legal content (for a typical example see [3]).

Again, we were able to establish users and user weights distributions for each ISP identified during the analysis, for all groups and categories of files. At most 1% of the IP addresses could not be matched to their ISPs.

### 3.2 Temporal analysis

We define the *number of available chunks* for a file or a group of files as the sum of owned number of chunks reported by all the observed users for that file or group of files.

The temporal analysis reveals the time-patterns at country level, in seeds/leeches ratio, and of the distribution of the number of available chunks per hour.

#### 3.2.1 Time patterns at the country level

We define *peak intervals* to be periods of time when the time-number of users graph presents spikes, meaning that there are significantly more users than usual.

We have analyzed the distribution of the number of peers per country, given an hourly sampling rate. The results can also be correlated with the *download robustness of the network* (see below).

#### 3.2.2 Time patterns in the seeds to leeches ratio

We have measured the seeds-to-leeches (see also Section 2.1) ratio evolution with an hourly sampling rate, which gives a good insight on the evolution of transfers within the network.

#### 3.2.3 Downloads robustness analysis

The *download robustness of the network* characterizes the difficulty one would encounter to break apart the

network of peers sharing a file or a group of files. We link the number of chunks available at any moment in the network to the download robustness of the network: the higher the total number of available chunks in the network, the higher the chance that the network would survive any given node's failure; hence, the robustness of the network scales at least linearly with the number of total available chunks in the network. A further study of the correlation between these two measures goes beyond the purpose of this work.

### 3.3 Peer activity analysis

There are various ways of analyzing peer activity/behavior in P2P networks. In this section we are focusing on three important issues: identifying collector peers, identifying super-peers, and analyzing the downloading behavior of the BitTorrent users.

#### 3.3.1 Identifying super-peers

We define the nodes who own at least two complete files, and are also visible in the network, as *super-peers*, thus being *seeds* for at least two files. These nodes are acting like repositories of data; their failure could greatly affect the network, while their presence greatly benefits the network, for example because their presence guarantees the existence of shared content back-ups.

#### 3.3.2 Identifying collector peers

We define the nodes who attempt to download at least two files as *collector* peers. BitTorrent could greatly benefit from the existence of *collector* peers, for instance because they could be queried for various files without using the web pages/trackers systems.

Collectors should eventually become super-peers, when they finish downloading several files. This makes worth the effort of detecting the collector nodes as soon as possible, as these may later contribute greatly to the robustness of the network.

#### 3.3.3 Downloading behavior

We call *downloading behavior* the correlation between starting a number of downloads and finishing them or part of them. We are especially interested to analyze the downloading behavior for group *Varied* and for category *Movie, popular* (an alias media set).

We have designed a simple scoring mechanism for the download process, very similar to the cycles mechanism in [6]. Every 1% downloaded from any file is scored with 1 *point*. Downloading a complete file or downloading 75% of a file and 25% of another both score 100 *points*.

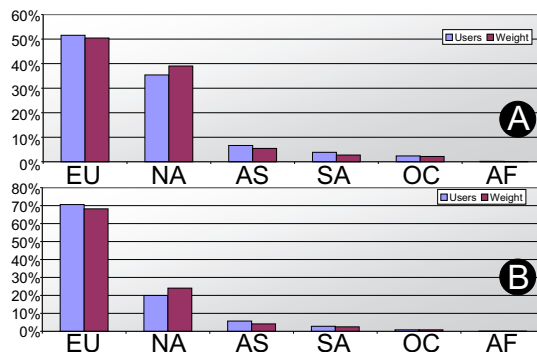


Figure 1: The distribution of users per continent: (a) for group Varied; (b) for group Detailed.

## 4 The Results

This section presents our results, as well as a number of potential BitTorrent improvements, based on these results.

### 4.1 Location analysis

This section presents the results of the location analysis.

#### 4.1.1 Geographical analysis

In Figure 1 we show the number of users and the total user weight per continent. The most important information is that Europe is the most important contributor to the BitTorrent network, both for the Detailed and the Varied groups.

Figure 1 clearly indicates that, for a given continent, the percentage of users (from the total number of users in the world) is very close to the percentage of their weight (from the total weight in the world). The weight percentage is higher up to 5 percentage points than the number of users percentage for the continents that have a good network infrastructure, like North America, and lower below 5 percentage points for countries with a worse network infrastructure, like Asia.

In Figure 2 we show the distribution of number of users and user weight per continent, for two content categories: applications and games. The figure shows that the distribution of users per continent varies with the file category. In Figure 2 (a), North America dominates the transfers of the *themed application* group, while in Figure 2 (b), Asia overcomes North America for the *themed game* group. This shows the existence of a coarse locality property for certain files or groups of files.

In Figure 3 we show the number of users and the total user weight per country. Unlike the situation from the distribution of users per continents, the distribution of users per country has no clear winner: US

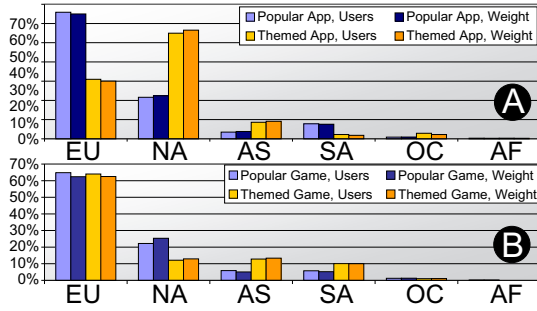


Figure 2: The distribution of users per continent: (a) for Applications; (b) for Games.

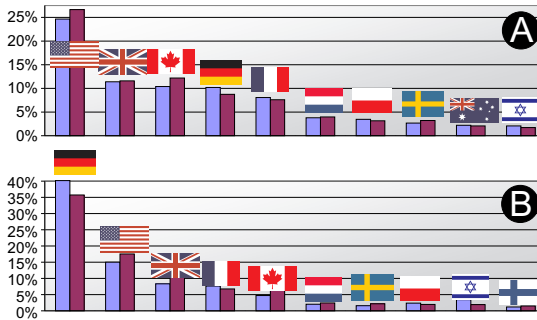


Figure 3: The distribution of users per country: (a) for group Varied; (b) for group Detailed.

is the most active for group *Varied*, while Germany has the same position for group *Detailed*. In both cases, Great Britain, Canada and France are in the *Top5*.

We move now to the locality properties of the shared data. Figures 4 and 5 show the number of users and the total user weight per country for the categories Applications and Games.

Figure 4 shows that Israel is the most important contributor to the sharing of a themed movie, while Japan is only making marks in downloading Animé Movies. Figure 5 shows that Hong Kong and Chile are surprisingly important contributors to the sharing of a themed game. We attribute (very speculatively) Hong Kong’s contribution to the fact that the themed game was a soccer management simulator of the English and other European soccer competitions, which are subject to intense gambling in Hong Kong [2]. In the case of all these surprise countries, their interest for other similar content (movies for Israel and Japan and games for Chile and Hong Kong) was at best reduced. The cumulated interest for other contents in the same category was at most 20% of the interest for the media which presented a distinctive locality feature.

The same discussion as in the case of Figure 1 holds for Figures 2, 3, 4 and 5: the percentage of users is very close to the percentage of their weight.

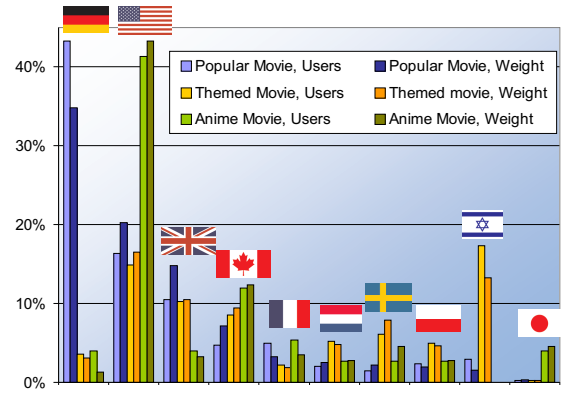


Figure 4: The distribution of users per country for category Movies.

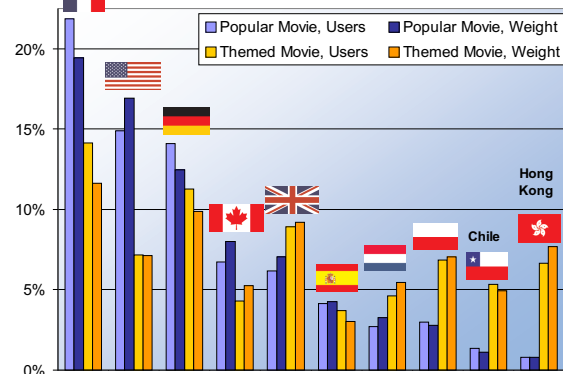


Figure 5: The distribution of users per country for category Games.

The city-based analysis did not result into meaningful results for our goals. Indeed, many of the users gather around several important cities, like Madrid, Paris, or Atlanta, but the fact that over 30% of the IP addresses could not be matched to real cities prevents us from generalizing our conclusion.

Another aspect has become an obstacle to creating a conclusive users/city distribution: the existence of *buffer cities*—small cities that host important network conjunction nodes, like Oldenburg and Eschborn (DE), and Herndon (US). The *buffer cities* cover another 15% of the total IP addresses.

#### 4.1.2 Organizational analysis

Table 3 shows the top 10 ISPs with regard to the weight of their users.

We find that that the top 20 organizations are attracting over 37% of the world-wide traffic for the Varied group and over 54% of the world-wide traffic for the Detailed group.

Figure 6 shows that, for some files, only a few ISPs cover the vast majority of the users. For instance, for the German version of a popular movie, over 60% of

Organization (Country)	Weight (%)
Deutsche Telekom AG (DE)	5.51
Comcast Cable (US)	4.85
America Online (US)	2.30
Road Runner (US)	3.28
British Telecommunications (GB)	2.09
Pac Bell Internet Services (US)	2.12
Proxad Free SAS (FR)	1.90
Telewest HSD Platform (GB)	1.86
Telia Network Services (SE)	1.63
Neostrada Plus (PL)	1.23
Total	26.77

Table 3: The Top10 organizations by users weight for group Varied.

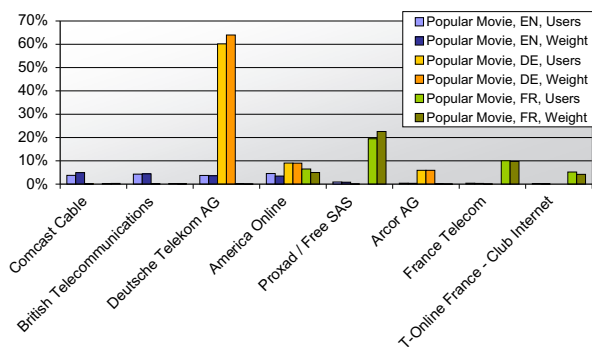


Figure 6: The distribution of users per organization for the Popular Movie sub-categories English(EN), German(DE), and French(FR).

the users use the services of only one ISP, *Deutsche Telekom AG*. However, for the French version of the same movie, the top 4 ISPs, *Proxad/Free SAS*, *France Telecom*, *T-Online France - Club Internet*, and *America Online* cover only approximately 50%. It is clear that, when only a few ISPs serve the majority of the users, optimizations are possible.

A similar discussion as in the case of Figure 1 holds for Figure 6: the percentage of users per organization is very close to the percentage of their weight.

## 4.2 Temporal analysis

This section presents our results regarding the analysis of the time patterns at country level, in seeds:leeches ratio, and in the distribution of the number of available chunks.

### 4.2.1 Time patterns at the country level

In Figure 7 we show the time patterns in the number of users, at the country level and for group *Varied*. We found four peak intervals. Sorted by importance, they are: 8 : 30 AM, 1 PM, 6 – 9 PM, and 12 – 1 AM. Given that the majority of users are in Europe, the peak intervals mean that the majority of users join the

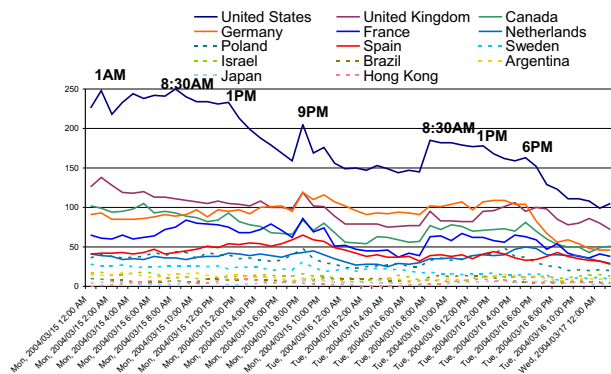


Figure 7: Time patterns at the country level for group Varied.

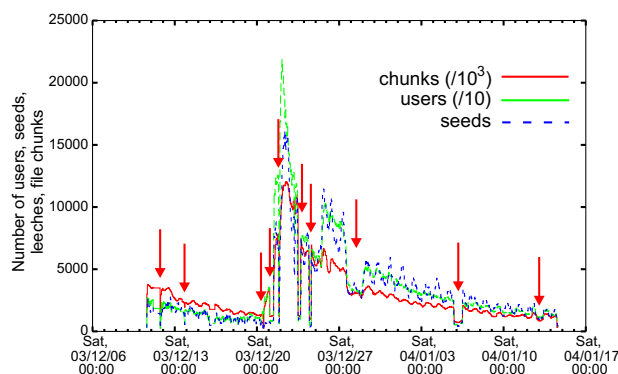


Figure 8: The normalized number of users, seeds, and number of file chunks for group Detailed.

network during *slow* working hours (early morning and during the lunch break), soon after they return home, or before sleep time. The last two intervals confirm the behavior of P2P networks users, as identified by [9, 8]. However, the first two peak intervals also indicate a new bias towards *slow* working hours, which was not previously observed. Unfortunately, our data set size requires future measurements to validate this conclusion.

### 4.2.2 Time patterns in the seeds to leeches ratio

In Figure 8 we show the normalized number of users, seeds, leeches and number of file chunks for group *Detailed*.

As expected, for a file or an *alias media* group, the ratio between the number of seeds and leeches tends to remain constant (see Figure 8). This means that the time does not influence the seeds:leeches ratio.

The ratio between the number of seeds and leeches is not the same for different files or groups of files, so improvements of the BitTorrent protocol based on the seeds:leeches ratio are unlikely.

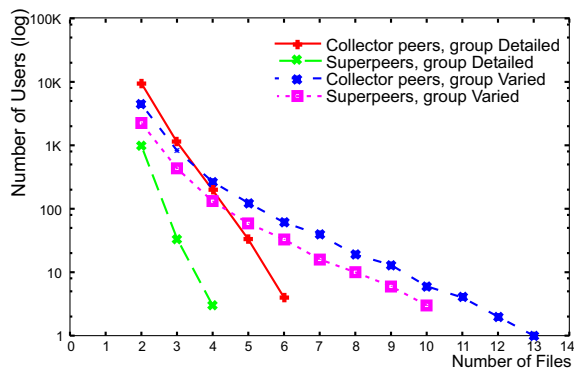


Figure 9: The number of superpeers and collector peers.

#### 4.2.3 Downloads robustness analysis

In Figure 8 we indicate with arrows the falls in the total number of available chunks for group *Detailed*.

The existence of such falls indicates that the network is not robust at all times. We attribute these sudden falls to failures of the higher levels of the BitTorrent networks, usually the web sites aiding the location of trackers. This assumption is also supported by the fact that most of the sudden falls in the number of chunks available in the network occur during the Christmas period (20-27 December 2003), when a number of web sites failes [8], and by the fact that the seeds:leeches ratio tends to remain constant over time.

### 4.3 Peer activity analysis

In this section we present our results regarding the identification of collector peers and super-peers, and the analysis of the BitTorrent users downloading behavior.

#### 4.3.1 Identifying super-peers

In Figure 9 we show the number of superpeers, for groups *Detailed* and *Varied*.

There are very few BitTorrent peers that display the super-peer behavior. The probability that a node would have a certain number of files decreases supra-exponentially with the number of files. The supra-exponential coefficient is more accentuated for *alias media* than for the *varied* group of files, hinting that BitTorrent users rarely get files that they do not need (they get only one file from an *alias media* set).

#### 4.3.2 Identifying collector peers

In Figure 9 we show the number of superpeers and collector peers, for groups *Detailed* and *Varied*.

The analysis of collector behavior yields similar results to the analysis of super-peers behavior. There are, again, very few peers that have the collectors property, but significantly more than super-peers. For

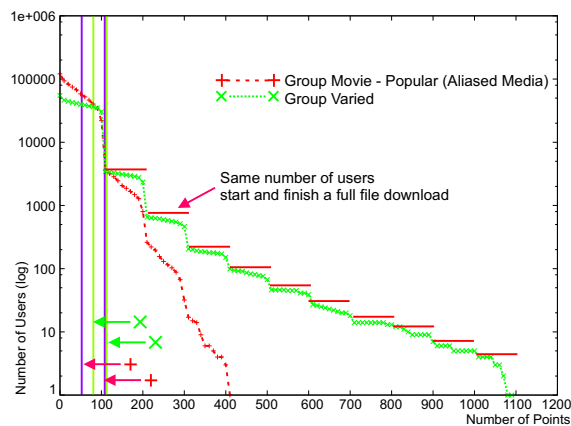


Figure 10: The distribution of the number of points per user.

group *Varied*, and for a given number of files, the number of collectors is approximately twice as the number of super-peers. The ratio is even bigger for the *alias media* set.

#### 4.3.3 Downloading behavior

In Figure 10 we present the distribution of the number of points per user, for the *Varied* group and for an *alias media* set.

The analysis based on the scoring mechanism described in Section 3.3.3 revealed that users tend to download the equivalent of full files. The dotted lines in Figure 10 show that the number of users that start and end downloading at least the equivalent of  $N$  files is almost constant for group *Varied*. Users downloading files from *alias media* sets, however, present a totally different behavior: as shown by the abrupt fall depicted in Figure 10, they tend to stop after downloading one of the many files from the same *alias media* set.

The average number of points is also an indicator of the typical behavior of BitTorrent users. For group *Varied*, users have an average of 81 *points*, which indicate that many of the users drop before actually getting complete files from the network. A special *end-of-download* extension should probably be added to the network to help users overcome this problem. The average number of points for users having downloaded at least one file, for the same group, is 113 *points*, which means that users quit the network after downloading one file. This information also supports our conclusion regarding the number of super-peers and collectors in BitTorrent network. The *alias media* group users have scored an average of 52 *points*, which supports the idea that users quit the network after downloading one file from an *alias media* set.

## 5 Related work

Our work is similar in goal to a number of recent studies on Gnutella [9, 10, 12], KaZaa [6, 1], BitTorrent [3, 8], and on several P2P networks [11]. Our work complements and builds upon these previous efforts by providing and correlating information for two layers of characterization: the locality layer (geographical and organizational views) and the behavioral layer (temporal, network robustness, the downloads consistency, and peer activity views). This work is also the first to analyze the *alias media* phenomenon and its implications in users downloading behaviors.

A comprehensive study of BitTorrent's characteristics has been performed by the very authors of this work in [8]. Their work tries to characterize all BitTorrent components, providing a limited insight into the low-level characteristics of network. To our knowledge, the only complete study of the geographical aspects of a P2P network is the study of Gnutella's global geographical and organizational views by Yazti et al. [12]. A previous study ended with the first mapping of Gnutella peers per continents [10], but did not analyze the organizational level.

One of the first studies regarding the characterization of P2P file-sharing systems is the work of Saroiu et al. [9]. They characterize the one-point-to-target latency, the bottleneck bandwidth, the user connection/disconnection frequency, and the number of files in the network, and correlate this data. In [3], the authors present the life of a file shared using BitTorrent. The file comes from the operating systems domain and is particularly large, thus being outside the targets of a common P2P file-sharing system user. The analysis presented in [11] contains an analysis of traffic patterns for a number of P2P networks, during just one day.

## 6 Conclusions and future work

In this paper we have presented a manifold view of the geographical, organizational, temporal, network robustness, and peer activity analysis of the BitTorrent file-sharing system.

Three important lessons surface from our study. The most important lesson is that BitTorrent shared contents displays various levels of locality, both trivial and non-trivial. The most obvious types of locality are the communities of users speaking the same language sharing localized versions of some data. The non-trivial examples include the preference of some countries for particular categories of contents (e.g. Hong Kong downloading a soccer management simulator) and the presence of single ISPs that serve over 50% of some media's users. The second lesson is that BitTorrent users are strongly influenced by the presence

of *alias media*—the same contents presented under different names and/or languages. The final lesson is that global views over P2P networks do not yield conclusive characterizations of the network under survey; instead, detailed views organized per categories of files and *alias media* should be used.

For the future, we plan to improve this study with a more recent and diverse set of data, and to add temporal and peer activity analysis. This would help us conclude on several issues that have arisen in this work, but which could not be proved because of the nature of the study. The continuation of this work would most likely also lead to a number of possible improvements of the BitTorrent network.

## References

- [1] K. P. Gummadi, R. J. Dunn, S. Saroiu, S. D. Gribble, H. M. Levy, and J. Zahorjan. Measurement, modeling, and analysis of a peer-to-peer file-sharing workload. In *SOSP '03: Proceedings of the nineteenth ACM symposium on Operating systems principles*, pages 314–329. ACM Press, 2003.
- [2] Hong Kong Govt. Gambling review: A consultation paper, Jul 2001. <http://www.info.gov.hk/archive/consult/2001/gambling-e.pdf>.
- [3] M. Izal, G. Urvoy-Keller, E. Biersack, P. Felber, A. Al Hamra, and L. Garcés-Erice. Dissecting BitTorrent: Five months in a torrent's lifetime. In *Passive and Active Measurements (PAM 2004)*, April 2004.
- [4] JoltID. PeerCache. 2005.
- [5] T. Karagiannis, A. Broido, N. Brownlee, k. claffy, and M. Faloutsos. Is P2P dying or just hiding? In *Global Internet and Next Generation Networks (Globecom 2004)*, Dallas, Texas, US, Dec 2004.
- [6] N. Leibowitz, M. Ripeanu, and A. Wierzbicki. Deconstructing the Kazaa Network. In *Proceedings of The Third IEEE Workshop on Internet Applications*, page 112, June 23-24 2003. San Jose, California.
- [7] A. Parker. The true picture of peer-to-peer file-sharing. CacheLogic Presentation, July 2004.
- [8] J. Pouwelse, P. Garbacki, D. Epema, and H. Sips. The BitTorrent p2p file-sharing system: Measurements and analysis. In *Proceedings of the 4th International Workshop on Peer-To-Peer Systems (IPTPS'05)*, Ithaca, New York, USA, February 2005.
- [9] S. Saroiu, P. Gummadi, and S. Gribble. A measurement study of peer-to-peer file sharing systems. In *Proceedings of Multimedia Computing and Networking*, 2002.
- [10] R. Schollmeier and G. Kunzmann. GnuViz - Mapping the Gnutella Networks to its Geographical Locations. *Praxis der Informationsverarbeitung und Kommunikation (PIK)*, 26(2):74–79, 2003.
- [11] S. Sen and J. Wong. Analyzing peer-to-peer traffic across large networks. In *Second Annual ACM Internet Measurement Workshop*, November 2002.
- [12] D. Zeinalipour-Yazti and T. Foliass. A quantitative analysis of the gnutella network traffic. Course Project for "Advanced Topics in Networks", with M. Faloutsos at the University of California - Riverside, Dpt. of CS, April 2002.