

Perceptions & Use of BitTorrent P2P File Sharing by Dartmouth College Students*

Alex Gerstein
Dartmouth College
Computer Science, Film Studies
`alex.gerstein@dartmouth.edu`

Scott Gladstone
Dartmouth College
Computer Science, Economics
`scott.gladstone@dartmouth.edu`

November 12, 2013

Abstract

In our final project, we hope to investigate peer-to-peer (P2P) file sharing network protocols, specifically those related to torrenting (BitTorrent). After a general overview of the system and security flaws present, we plan to examine P2P in the context of Dartmouth College. By surveying students and speaking to Dartmouth College Computing Services, we hope to understand the disparity between the perception and actual use by students of torrent networks. If available, we also hope to obtain and analyze data from the College on student bandwidth habits, download frequencies, or other metrics that the College records. The outline below summarizes the questions we hope to answer and the individuals we hope to speak with.

*This paper constitutes the final project for the Fall 2013 iteration of CS 55: Security and Privacy, taught by Charles C. Palmer, Adjunct Professor of Computer Science, Dartmouth College; CTO Security and Privacy, IBM Research.

1 Introduction

BitTorrent is a protocol supporting peer-to-peer (P2P) file sharing that facilitates the distribution of large data files over the Internet.¹ BitTorrent is one of the most common protocols for transferring large files and is often used for distribution of very popular files and files available for free, such as literary texts, audio files, movies, and applications. “Torrenting” is the process by which a user engages with a BitTorrent client in order to send and receive those desired files. To understand the uniqueness – and security implications – of the BitTorrent protocol, one must first develop a working understanding of P2P file sharing. From there, the legal implications of P2P file sharing and BitTorrent can be made clear. With this foundation, the authors then describe a study in which data about student perceptions of torrenting at Dartmouth College are compared with data collected by Dartmouth College network administrators describing student incidences of torrenting. Conclusion and policy recommendations are then made with the goal of aligning student perceptions and administrative desires.

1.1 Peer-to-Peer (P2P) File Sharing

According to one Internet dictionary, P2P is a “type of Internet network that allows users with the same program to connect with each other and access files on one another’s hard drives without the intervention of a server computer.”² This decentralized network architecture consists of individual nodes called *peers* that act as both suppliers and consumers of resources. In P2P networks, tasks are shared among multiple, interconnected peers who each make a portion of their resources – processing power, disk storage, or network bandwidth – available to all network participants, without the centralized coordination of a server-client model.³ A comparison of a decentralized-P2P network and a centralized server-based network are presented in Figure 1 and Figure 2 below.

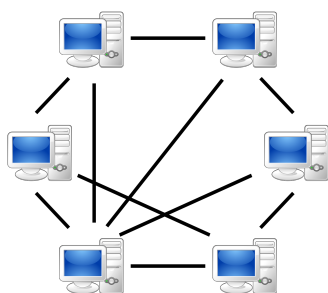


Figure 1: A peer-to-peer (P2P) network with interconnected nodes (“peers”) sharing resources. Source: Wikipedia.

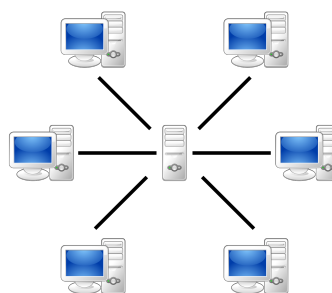


Figure 2: A client-server model network, where individual clients request services from centralized servers. Source: Wikipedia.

P2P networks are often found in residential home networks, allowing users to configure their computers in peer workgroups to allow sharing of files, printers, and other resources among all devices; with a number of computers running similar network protocols, P2P is a convenient way to access shared resources.⁴ However, the term “P2P” used today generally does not refer to the network architecture from which its name is derived, but to P2P file sharing systems. Using P2P software applications such as Napster or Kazaa, users are able to search for, transfer, and download data files over the Internet with any user on the same P2P file sharing network. The implications of such a system were huge: if *any* user on the system had a file that another user demanded, the second user could – via the network – obtain the file. The second user would then also be a source of the file in the P2P network, making it faster and easier for more users to download the files.

The first P2P file sharing networks, such as the purely-mp3 sharing network Napster, relied on a central index server to assist with the transfer of files. When someone searched for a file, the central index server – which contained an index of all of the users and their shared content – searched for all available copies of the file and presented them to the user; the file would then be transferred directly between the two private computers.⁵ Because the file sharing occurred over a central network, Napster was held liable for copyright infringement over the sharing of mp3 files and was shut down in 2001.⁶ New protocols such as BitTorrent represent a technological evolution of P2P file sharing networks.

1.2 BitTorrent Protocol

BitTorrent is a P2P file sharing protocol that makes use of unique design to reduce the bandwidth and download time for more frequently requested data. The protocol allows a single user to download files quickly by allowing users to join a “swarm” of hosts to download and upload from each other simultaneously. Thus, instead of attempting to transfer a large file from one user to another like other P2P network protocols – a slow, bandwidth-intensive process –, BitTorrent allows a user to download small segments of the file from a variety of users, significantly reducing computational energy and download time.¹ Because of this design, BitTorrent is often used for distribution of very large files, very popular files, and files available for free, as it is more efficient to distribute these files using BitTorrent than a regular download.

When a user “torrents” a file, meaning that the user is attempting to download a file over a P2P network using a BitTorrent client implementing the BitTorrent protocol, they begin by loading a .torrent file into a BitTorrent client. A .torrent file is a computer file that contains metadata about files to be distributed, including file names, sizes, checksums of all individual pieces, and a list of the network locations of “trackers” – a server that facilitates communication between peers on a P2P network – to help participants in the system seeking to download the same data find one another.⁷ The tracker shares the peers’ IP addresses with other BitTorrent clients in the “swarm,” the group of all peers sharing a torrent. Each piece of a file specified in a torrent is protected by a cryptographic hash contained in the .torrent file; this ensures that any modification of the piece can be easily detected, preventing

a user from downloading both accidental and malicious modifications of any pieces of a file.¹

Once connected, a BitTorrent client downloads bits of the files specified in the torrent in small pieces, downloading different segments from data from different computers in the swarm. Once the BitTorrent client has some data, it can then begin to upload that data to other BitTorrent clients in the swarm; by using both downloading and uploading bandwidth and dividing the files into pieces, the time required to download the file is significantly reduces.¹ Thus, it is worth noting the following concept: “Everyone downloading a torrent is also uploading the same torrent. If 10,000 people are downloading the same file, it doesn’t put a lot of stress on a central server. Instead, each downloader contributes upload bandwidth to other downloaders, ensuring the torrent stays fast,”.⁷ From this protocol, clients never actually download files from the tracker itself; the tracker keeps track of the clients connected in the swarm, but does not actually participate in the process of downloading or uploading data.⁷

The process of sharing data via BitTorrent starts with a *seed*, or the initial machine possession 100% of the data. Then *leechers*, also known as *downloaders*, which refers to peers that do not have the entire file or are in the process of downloading the file. As the number of downloaders increases, the reliance of new leechers on old leechers decreases, as each successful downloader becomes a source of the file for future downloaders.⁸ From that point, the process repeats and continues, allowing peers on the network to share any and all files to which any other peer has access.

1.3 BitTorrent Security Concerns

What are security issues related to using BitTorrent? torrent poisoning (wiki)

2 Legal Issues

Torrenting is not inherently illegal.

2.1 Implications for Individuals and Network Hosts

What are the legal implications for an individual user for torrenting copyrighted information?

2.2 Dartmouth Policy and Practice

What are the legal implications for a university like Dartmouth with students that do illegal torrenting? What are some past cases and punishments relating to torrenting?

3 Methods, Data & Analysis

The goals of this research were three-fold: first, to determine the student perception of the occurrence of torrenting at Dartmouth; second, to investigate the actual occurrence of torrenting at Dartmouth based on data provided by Dartmouth College Computing Services; and third, to examine the relationship between student perceptions and actual incidences of torrenting and make policy recommendations with the purpose of aligning student perceptions and student network activity. The methods employed to carry out these goals and the data collected are described in the sections 3.1 to 3.3.

3.1 Student Perception Survey

3.1.1 Participants

The participants in the survey were 130 students at Dartmouth College. The population consisted of second-year, third-year, and fourth-year students at the College. Participation was anonymous and voluntary. All prospective participants were sent a link to a Google Form that was used to create the survey; all participants had completely voluntary choice in clicking on the link, examining the questions, and completing and submitting the survey. The population examined in this study may be biased in the sense that many individuals who completed the survey could have personally known the researchers or were members of CS 55: Security & Privacy, the class for which this research was completed.

3.1.2 Survey Design

The survey was designed via Google Forms and was administered by providing a link to the survey to all prospective participants. The survey consisted of seven questions, six of which were mandatory. The mandatory questions asked the participant for his/her person experience with torrenting, whether they believed other students at Dartmouth torrented, the number of files that they believed the average student torrented, whether they believed Dartmouth recorded information about student torrenting, and whether their knowledge of Dartmouth Copyright Violation Policy impacted the amount that they torrented. The seventh, optional question asked whether the participant has any thoughts as to how Dartmouth College could reduce the incidence of torrenting on campus. All surveys were timestamped when submitted and resubmissions/changing of answers were not permitted. See Appendix A for a copy of the survey administered.

3.1.3 Perception Data

Of the 130 students, 78 (60.0%) stated that they had previously downloaded a file via torrenting in general, but only 52 (40.0%) stated that they had downloaded files via torrenting

while at Dartmouth. This indicates that 52 out of 78 individuals (74.3%) who had torrented in general had also torrented while at Dartmouth. Of the 52 students who designated that they torrented at Dartmouth, 34 (65.4%) claimed to download one to five files per month, on average; 5 (9.6%) claimed to download zero files per month on average; 5 (9.6%) claimed to download six to ten files per month on average; and 9 (17.3%) claimed to download 11 or more files per month on average.

Considering all students (both torrenters and non-torrenters), 82 of 130 students (63.1%) claimed to download zero files per month on average. This was highly significantly different ($p < .00001$) from the 16 (12.3%) that those participants believed that the average Dartmouth downloaded via torrenting on average. The median self-reported number of files downloaded via torrenting on average was zero, while the median perceived number of files downloaded by the average student was one to five; based on the distributions presented in Figures 3 and 4 below, there is a clear right shift in Figure 4 as compared with Figure 3, indicating that there is a significantly higher perception of student torrenting than what actually occurs.

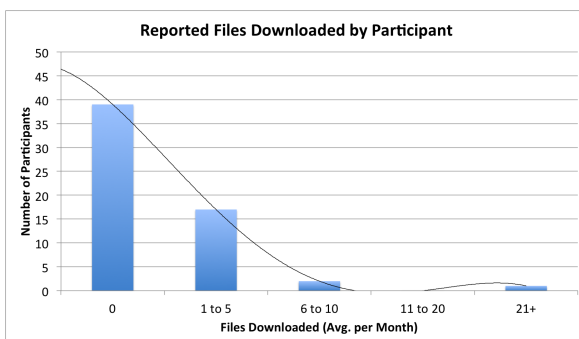


Figure 3: Bar distribution of reported number of files downloaded via torrenting by the participant.

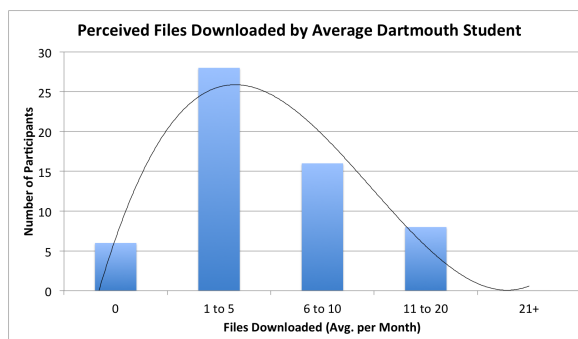


Figure 4: Bar distribution of perceived number of files downloaded via torrenting by the “average Dartmouth student.”

With regard to whether participants believed that Dartmouth recorded student torrenting information or whether knowledge of Dartmouth’s Copyright Violation Policy affects how often the participant torrents, the results were less conclusive. Of the 130 students asked if Dartmouth records when someone torrents a file, 47 (36.2%) said Yes; 28 (21.5%) said No; and 55 (42.3%) said they were Unsure. Attempting to break down the results further by number of files participants reported downloading or by the number of files that participants perceived that other students download were also inconclusive; more data would need to be collected to determine this result. Regarding whether knowledge of Dartmouth’s Copyright Violation Policy affected personal torrenting practices, there was no significant difference ($p > .05$) between the number of participants who said Yes (N=63, 48.5%) and who said No (N=67, 51.5%).

3.2 Dartmouth Network Patterns

Adam Goldstein, PaloAlto, types of metrics

3.3 Relationship between Student Perception and Dartmouth Network Data

Draw some conclusions and relationships between analytics (e.g. perception ? practice)

4 Conclusion

Summary, Questions, and Policy Recommendations

References

- ¹ “BitTorrent: Beginner’s Guide.” *BitTorrent Inc.* Accessed November 4, 2013. Available at <http://www.bittorrent.com/help/guides/beginners-guide>.
- ² Shannon, Victoria. “The End User: P2P starts to mature.” *New York Times*. July 9, 2005. Available at <http://www.nytimes.com/2005/07/08/technology/08iht-ptend09.html>.
- ³ Rdiger Schollmeier, A Definition of Peer-to-Peer Networking for the Classification of Peer-to-Peer Architectures and Applications, Proceedings of the First International Conference on Peer-to-Peer Computing, IEEE (2002).
- ⁴ <http://compnetworking.about.com/od/p2ppeertopeer/a/p2pintroduction.htm>
- ⁵ “Peer-to-peer file sharing.” *Wikipedia*. Accessed November 4, 2013. Available at http://en.wikipedia.org/wiki/Peer-to-peer_file_sharing#History.
- ⁶ Douglas, Guy. “Copyright and Peer-To-Peer Music File Sharing: The Napster Case and the Argument Against Legislative Reform.” *Murdoch University Electronic Journal of Law*. Vol. 11, No. 1 (March 2004). Available at <http://www.murdoch.edu.au/elaw/issues/v11n1/douglas111.html>.
- ⁷ <http://www.howtogeek.com/141257/htg-explains-how-does-bittorrent-work/>
- ⁸ http://en.wikipedia.org/wiki/Glossary_of_BitTorrent_terms#Leech

** statistics on P2P we may want to use: <http://www.ipoque.com/sites/default/files/mediafiles/documentstudy-2008-2009.pdf>

Appendix A

Student Perception Survey

1. Have you ever downloaded files (music, movies, applications, etc.) via torrenting?
 - Yes
 - No
2. Have you ever downloaded files (music, movies, applications, etc.) via torrenting while at Dartmouth?
 - Yes
 - No
3. On average, how many files do you download via torrenting per month at Dartmouth?
 - 0
 - 1-5
 - 6-10
 - 11-20
 - 21+
4. How many files do you think an average student at Dartmouth downloads via torrenting per month?
 - 0
 - 1-5
 - 6-10
 - 11-20
 - 21+
5. Do you think that Dartmouth records when you or someone else torrents a file?
 - Yes
 - No
 - Unsure
6. Does your knowledge of Dartmouth's copyright policy affect how often you download files via torrenting? Dartmouth policy states that, "If your computer begins to consume excessive network resources, Computing Services will investigate your network activities in order to keep the network operating smoothly?.Sanctions [for illegal downloading] may include suspension of network access (meaning loss of e-mail and course web site access) and formal college disciplinary action."
 - Yes
 - No
7. Do you have any thoughts on how Dartmouth could reduce student torrenting on campus?

Appendix B

Email Notice of Unauthorized Use of Copyrighted Property

Sent: Friday, October 28, 2011 4:01 PM

To: *****

Subject: FW: Notice ID: ***** Notice of Unauthorized Use of ***** Corporation Property

Dartmouth College has received a notice claiming that an IP address registered to you has been used to offer for download via the Internet, without permission, a copyrighted work, cited in the complaint below.

Infringement is actionable under federal copyright law and can require the payment of damages of up to 0,000 per violation, as well as carrying the potential for criminal penalties. Respect for intellectual property ownership is a fundamental value in academia. Copyright infringement is also a violation of College policy. Additional information about the College's copyright policy - including peer-to-peer file sharing - can be found online at <http://www.dartmouth.edu/copyright/>.

By return e-mail, please indicate that (1) you have removed from your computer any programs that may be used to offer copyrighted material via the Internet for download by others without the copyright owner's permission (examples of such programs include BitTorrent and Limewire) and (2) that you will desist from activities that infringe copyright. If you need assistance in removing file-sharing programs from your computer, the College's Computing Services department would be glad to assist you. For this purpose, please contact the Student Help Desk or your IT Consultant.

If you fail to respond promptly, the College will block your IP address. The College must take this step in order to avail itself of the legal protection afforded service providers under the Digital Millennium Copyright Act. Please note, in most cases this notice comes as a result of the uploading or downloading of music, video, audio, movie or television files that, as a functionality of the file-sharing software used for downloading, is available for uploading by Internet requests without your awareness or permission anytime that your computer is turned on. Scans by copyright owners (such as the company that made the complaint in this case) capture both uploading and downloading of these files.

If you believe this notice is in error, or choose to challenge it, please contact me and I will explain the procedure. This procedure entails the filing on your part of a counter-notice asserting your right to the material and your right to distribute it, and satisfying all of the requisite legal formalities. Only after you have filed a proper counter-notice can a request to have your IP address unblocked be granted.

Thank you for your cooperation. I look forward to hearing from you.

- Chief Information Security Officer, Dartmouth College, Hanover, NH 03755

Appendix C

File Sharing Program Gigabyte Usage¹

24 Hour

App Sub Category	Application Name	Gigabytes	Percentage
file-sharing	bittorrent	177.09	70.50
file-sharing	dropbox	44.56	17.73
file-sharing	ftp	20.25	8.06
file-sharing	azureus	3.42	1.36
file-sharing	pando	1.94	0.77
file-sharing	google-drive-web	1.53	0.61
file-sharing	skydrive-base	1.49	0.60
file-sharing	boxnet-base	0.89	0.36

7 Day

App Sub Category	Application Name	Gigabytes	Percentage
file-sharing	bittorrent	1180.17	60.10
file-sharing	dropbox	394.77	20.10
file-sharing	ftp	313.54	15.97
file-sharing	pando	36.40	1.85
file-sharing	skydrive-base	12.94	0.66
file-sharing	boxnet-base	10.02	0.51
file-sharing	azureus	7.69	0.39
file-sharing	google-drive-web	7.62	0.39

30 Day

App Sub Category	Application Name	Gigabytes	Percentage
file-sharing	bittorrent	4757.79	60.89
file-sharing	dropbox	1697.71	21.73
file-sharing	ftp	1046.01	13.39
file-sharing	pando	137.98	1.77
file-sharing	azureus	53.67	0.69
file-sharing	skydrive-base	52.94	0.68
file-sharing	boxnet-base	33.32	0.43
file-sharing	google-drive-web	29.01	0.37

¹Provided by Adam Goldstein, Dartmouth College Computing Services.