THE EVOLUTION OF P2P NETWORKS FOR FILE EXCHANGE: THE INTERACTION BETWEEN SOCIAL CONTROVERSY AND TECHNICAL CHANGE

JAIRO ARES;[1] EDUARD AIBAR[2]*
[1]INTERNET INTERDISCIPLINARY INSTITUTE. UNIVERSITAT OBERTA DE CATALUNYA.
[2]DEPARTMENT OF ARTS AND HUMANITIES. INTERNET INTERDISCIPLINARY INSTITUTE. UNIVERSITAT OBERTA DE CATALUNYA.

Summary: *Since the irruption of Napster in 1999, Peer-to-Peer computer networks for file exchange have been at the heart of a heated debate that has eventually evolved into a wide social controversy across the world, involving legal, economical, and even political issues. This essay analyzes the effects of this controversy on the technical innovations that have shaped the evolution of those systems. It argues that the usual image of a single two-sided conflict does not account for most of the technical changes involved. P2P entrepreneurs and creators show a wide range of motivations and business strategies —if any — and users are not a monolithic group with a common set of goals and values. As a result, the actual historical evolution of those networks does not follow a simple linear path but a more complex and multidirectional development.*

Key words: *P2P networks, file exchange, social shaping of technology, copyright controversy*

* Correspondència: EDUARD AIBAR.
  Department of Arts and Humanities.
  Internet Interdisciplinary Institute
  Universitat Oberta de Catalunya
  AV. Tibidabo, 39-43, 08035 BARCELONA, SPAIN.
  email: *eaibar@uoc.edu*

Introduction[1]

Ever since the irruption of Napster in 1999, Peer-to-Peer (P2P) computer networks for file exchange[2] have been at the heart of a heated debate that has eventually evolved into a wide social controversy in many countries. At stake was —and still is— the model that our societies choose for creation, ownership, modification and distribution of creative material subject to intellectual property rights, such as music, videos or software in general. In the middle of this battle, which came to be known by the media as "copyright wars", even well established civil rights such as privacy of information and communication are being put into question in the name of the fight against piracy. Copyright holders plead for the right —or even the obligation for Internet Service Providers (ISPs)— of analyzing Internet traffic in order to spot illegal downloads, and ISPs —notably Comcast in the US— have tried to slow down the P2P traffic to avoid saturation of their networks and paying excessive transit fees, resulting in another public controversy known as "net neutrality".

Important as these issues are, the underlying debate is also a technological one. P2P networks are typically created on top of the Internet, which was originally a non-centralized and evenly distributed computer network. However, with the advent of the web, the privatization and the increased commercialization of Internet services[3], the net has evolved towards an asymmetric structure where a few servers provide content and a multitude of clients retrieve it. P2P networks constitute an attempt to re-empower the individual computers and hence their users, changing the Internet structure, the traffic patterns, and the balance of power.

Both ISPs and copyright owners have tried to hinder the advancement of these networks with some degree of success. Nevertheless, the P2P community has managed to modify or deploy these technologies in new and changing conditions in order to avoid the legal prosecution of its users. Along the way, P2P users have also proved not to be a monolithic group with clear and shared community oriented goals. Thus, the appearance of downloaders[4],

---

1.   We would like to thank our colleagues Raquel Xalabarder and Joan Arnedo, and two anonymous reviewers, for their useful comments, critiques and suggestions.

2.   The term P2P is used in the field of computer networks to denote a certain architecture and way of interconnecting computers regardless of the applications that run on them. In the present study we will focus exclusively on the networks conceived for the purpose of exchanging files. As in Schoder *et al.* (2005), we consider P2P networks as entities comprising three layers: communication protocols and techniques, client applications running in computers, and the communities of users themselves.

3.   For a detailed historical account of the origins of the Internet see Abbate (1999). For a specific discussion on the privatization process see also Abbate (2010). The commercialization of the Internet may also be framed within the broader issue of the commercialization of scientific research —a hot topic considering present neoliberal approaches to science policy— for an introduction see Mirowski and Sent (2008).

4.   Although in P2P jargon users who do not share are often called *leechers*, we have opted for the term downloaders because we believe it captures the interpretation that this group makes of P2P networks, i.e., networks for "downloading files" rather than for "sharing files".

who did not share files of their own, and entrepreneurs that tried to control these networks for their own profit, forced additional changes in the technology that were not an intrinsic necessity. All together, the picture is one of a public controversy among different social groups and actors, and also of an evolving technology, where both the social and the technological aspects seem to be intertwined with one another.

The main goal of this work is to analyze the controversy around P2P networks for file exchange by explaining why this technology evolved in different forms of distinctive network architectures and use cases as a result of the very dynamics of the controversy. This is indeed a clear case of technology —including technical details— being shaped by a social construction process where different social groups compete to embed their values and visions in the very artifact design (Pinch and Bijker, 1987). Although P2P networks have been studied from other points of view (mainly from the legal one) that take the technological artifact for granted, we will focus here in their technical shaping and evolution. To our best knowledge, such an approach combining both technological and sociological elements has never been done within this controversy, with the exception of limited studies that referred to only one instance of these networks (Spitz and Hunter, 2003).

The predecessor of P2P: IRC networks

Shawn Fanning is well-known for developing Napster, the first P2P network for file exchange, launched in early 1999. When asked in an interview what his motivation was, he replied: "It was rooted out of frustration not only with MP3.com, Lycos, and Scour.net, but also to create a music community. There really was nothing like it at the time" (Varanini, 2000). The word "frustration" is recurrent, and comes in other accounts of the same story (Greenfeld *et al.*, 2000). It points to the fact that there were a growing number of users looking for places from where they could download for free music files in mp3 format after the music industry, personified in the Recording Industry Association of America (RIAA), had started suing music downloading sites forcing them to remove copyright infringing mp3 files.

A second element worth noting in Fanning's answer is the word "community". Ante (2000b) describes Fanning as an assiduous IRC (Internet Relay Chat) user during years, and therefore he must have been well acquainted with online communities and the file exchange that was already taking place in those networks. It can be argued that IRC networks gave him the inspiration for creating Napster.

At the end of 1998 Fanning notices a problem —people cannot get the mp3s they want— and a technology —IRC— which can bring together people from anywhere in the world and enable them to exchange files, albeit in a rudimentary way. His contribution was to understand this environment, and to propose a better technological solution that he thought would solve this social need. As we will see, his solution would create new problems for some of the actors involved, and this would spark a controversy that would ultimately determine how the technology evolved.

IRC was a popular protocol for text communication over the Internet created in 1998 by Jarkko Oikarinen, from the University of Oulu, Finland, who liked tinkering with communication software in his free time[5]. The IRC application he created consisted of two parts: a server —called IRCd or IRC daemon— which was a piece of software he set to run in his Department server, and a client, which needed to run in each of the computers of the users that wanted to chat.

Oikarinen, as many other Internet pioneers, could be considered as a *hacker* in the sense defined by Eric Raymond and later used by Pekka Himanen in order to describe his Hacker Ethic (Himanen 2001). In line with this hacker spirit, and as soon as he noticed that his application became popular among his colleagues, Oikarinen provided his IRC to other computer enthusiasts in Finnish universities, who helped him improve it and contributed to making it popular in Finland and beyond. Thus, from its very beginning IRC was an open application, and it became a standard *de facto* when other hackers started to develop new client applications to connect to IRCD using other operating systems than Unix.

A distinctive feature of IRC when compared with other text messaging protocols is its "relay" characteristic, a direct inheritance of BBSs and USENET, two of Oikarinen's sources of inspiration. Relaying allowed for the interconnection of servers running in different machines and for the creation of networks of thousands or even millions of users that could chat with one another in real time. Furthermore, a very interesting protocol strongly related to IRC, and arguably another key inspiration for Shawn Fanning, was the Direct Client Connection (DCC), which allowed direct connection between clients -or peers- without going through a server.

From a functional point of view, two users wanting to establish a DCC connection with one another first need to connect to two servers within the same IRC network, then contact each other in a public chat room or channel, and finally open a DCC connection using their IRC client applications. Once this connection is open, they are free to chat directly as well as to exchange files.

DCC was first implemented by the Australian Roy Trollo as part of IRCII[6] (Rollo), a UNIX client he maintained for some time during the first half of the 1990s, and later integrated in other IRC clients. Some time later, based on DCC, the Windows client mIRC introduced a new functionality called */fserve* that provided a relatively simple way to set up a file server, and together with it came many scripts for creating file bots that automated the file sharing process. Still, file sharing under IRC required some advanced computing skills, so it could not become the kind of mainstream phenomenon that Napster became. Nevertheless, it prepared the road for Internet users that grew used to sharing files through exchange networks that socially could be considered P2P, even if the underlying computer networks were not.

---

5.   http://www.irc.org/history_docs/jarkko.html (last checked 3/10/10)

6.   A description of the DCC protocol can be found at http://www.irchelp.org/irchelp/rfc/dccspec.html (last checked 3/10/10).

Napster

Shawn Fanning created Napster while he was a first year university student. He was inspired when he saw his fellow students struggle with IRC and websites to download music while looking for mp3 files, and he came up with the idea of linking computers directly without going through a centralized file server. It was the year 1999, and the dot-com bubble was very much growing, so in a typical reaction in those times he quit studying and started programming compulsively in order to be the first to hit the market with his application (Greenfeld *et al.*, 2000).

Napster shared many of the characteristics of IRC. To start with, the application came in two parts: client software that the users had to download, and a server[7] which was controlled by Fanning's company, the Napster Corporation. Then, in order to use the service, users had to register a nickname which became their persistent Napster identity, valid for the IRC-like chat rooms available through the program and also visible when downloading and uploading files.

Upon logging in, the Napster client uploaded to the server the list of mp3 files that the user was willing to share. This allowed the Napster server to keep an up-to-date list of all the files being offered for sharing. When users wanted to download a file, they had to submit a query to the server, and received in return a list of the 100 best matching files, sorted according to an estimation of the closest distance for downloading. In the end it was the user who decided from where to download the file and opened a DCC-like direct connection with the peer sharing the file.

The original program, written entirely by Fanning and released in June 1999 was an immediate hit: it reached 10 million users in its first 9 months of activity, and 80 million in the 9 months that followed (Lessig, 2004). However, the program was also highly controversial, and its legality was very quickly put into question by the music industry. The fact that Napster did not store the mp3s and that it was the users who performed the actual file sharing made Fanning and his partners believe that they were not guilty of any copyright infringement. Nonetheless, the industry felt directly attacked, and in December 1999 the RIAA filed a complaint against Napster Corporation, opening a legal battle that lasted more than a year and a half, when Napster was finally forced to shut down their service. Although the legal details of this case fall beyond the scope of this paper, it is interesting to note that the Napster trial exposed some of the conflicting meanings about what P2P networks were. Whereas Napster Corporation tried to present themselves as an ISP and therefore operating under the shelter of the Digital Millennium Copyright Act (DMCA), the court ruled in favor of the music industry, considering it to be a listing service that offered a search engine, directory, index, and links to music files. Furthermore, even if Napster's lawyers managed to demonstrate that

---

7. In practice Napster, had at least 160 servers interconnected with each other (Saroiu *et al.*, 2002) in order to support the heavy traffic generated in its peak moments. Here we consider them as only one centralized server, although this is not strictly true.

there were a number of legal uses of the P2P network, the fact that millions of users engaged in unauthorized exchange of copyrighted material prevailed (Spitz and Hunter, 2003) and reinforced the discourse of the music industry that equated P2P with piracy.

Gnutella

Whereas some hackers tried to copy and reverse engineer Napster[8], others like Justin Frankel decided to take the P2P idea further and improve the system. Frankel had become famous in 1996 at the age of 18 for creating the popular mp3 player application Winamp. Leaving university soon after, he associated with Tom Pepper to create a company called Nullsoft and to develop Shoutcast, an application for setting up Internet radio stations inexpensively. The unique combination of Winamp and Shoutcast caught the attention of AOL, one of the biggest ISPs of the time, who bought Nullsoft in 1999 and kept the founding team on board.

When Frankel came across Napster, as an expert in online music, he was fascinated by it, but he also understood the problem of keeping a centralized database while faced with a combative and piracy-concerned music industry (Kushner, 2004). His response to Napster was a P2P network called Gnutella, which he started together with Pepper as a pet project inside Nullsoft. Besides opening the scheme to non-mp3 files too, the most significant innovation was that Gnutella completely got rid of the centralized server, making it impossible to shut down. In this new network there were to be only nodes that would talk to one another and that would collaboratively find the files requested. It was thus a real P2P network where only one type of software was needed, at the same time client and server. This new architecture matched both his hackeristic views and his knowledge about the music distribution industry:

> "Napster was a company built on people doing things that are illegal. That's wrong (…) I decided to take the wind out of Napster's sails (…) I would not be getting any money from it. I'd be giving power to people, and what can be wrong with that?" (Kushner, 2004).

In March 2000 Frankel and Pepper posted an early version of the program on the web page of Nullsoft, but they were required almost immediately by their management to remove it. At that time AOL was in merger talks with Time Warner, a music and media corporation that had sued Napster for Copyright infringement and that did not appreciate this new piece of software. Thus, the original Gnutella was only allowed to live for a few hours in Nullsoft's servers, although enough time to spread the idea and raise the interest of other hackers that would continue with the project within the open source community and re-

---

8.   By May 2000, 7 months after the launch of Napster, there were in the SourceForge website up to 19 open source projects to develop various Napster clients and also a Napster like server http://web.archive.org/web/20000511171541/http://open-nap.sourceforge.net/ (last checked 3/10/10).

lease a protocol specification known as Gnutella 0.4, which became the official reference for those who started to develop Gnutella applications.

The specification stated that the nodes of the Gnutella network were to be called *servents* (SERVers + cliENTS), and that there would be no hierarchy among them. A servent would join the network by connecting to one or more existing servents. Then, when a servent wanted to look for specific content in the network, it would send a query to all the servents it was connected to and who, in their turn, would also send it to all servents they were connected to, etc. If a servent recognized that it had the content requested available, it would send a message back with its contact details, and this message would be routed backwards in the opposite way it had been transmitted in first place. File exchange between a servent that had launched the request and the one that had replied positively would take place out of the Gnutella network using a direct one-to-one http connection.

The choice to route the query messages back and forth through the network was a way to send the queries in an anonymous way, without sender identification, and thus protected from eavesdropping and potential legal action. Nevertheless, this technological choice was rather controversial. As soon as Gnutella started to be deployed, users realized that it was much slower than Napster. As Ripeanu and Foster (2001) noticed in their measurements of Gnutella network, the combination of queries and overhead traffic mushroomed as soon as the number of hosts grew, slowing down the overall performance and user experience, and even blocking those peers who were accessing the network using a dial-up modem and did not have enough bandwidth available.

Delio (2000) describes the atmosphere of those days, where several developers disagreed on the implementation choices for Gnutella. Despite the interest from the hacker community for a completely decentralized network, the technical problems and the entrepreneurs looking for the next big commercial success drove the development towards a more hierarchical structure. For instance, a now defunct company named Clip2 built an application which was a Gnutella "super peer" that hid traffic from some servents. Clip2 ceased operations in mid-2001, but the idea of two types of nodes persisted and was taken by version 0.6 of the protocol that was introduced in 2002 (Klingberg & Manfredi, 2002). Thus, starting from this version, Gnutella distinguished between two kinds of peers: *leaf nodes* and *ultra peers*. Leaf nodes were just connected to one or more ultra peers, whereas ultra peers had connections to both leaf nodes and other ultra peers, and were responsible for routing queries and keeping the network alive. The decision for a peer to operate in one or the other mode was taken dynamically, and therefore the same software client was used for both types of nodes.

In either version of the protocol the network lacked central control, as its creators had carefully engineered in order to avoid legal action from the music industry and other copyright holders. Their approach was successful against the first wave of lawsuits like the one that ended up closing Napster. However, it would not stand up against the industry's change

of tactics, which moved from targeting P2P companies to the actual users. In 2002 the RIAA started asking the ISPs for the names of the subscribers behind file exchanges, and even filed a suit against the American ISP Verizon that had refused to collaborate alleging a breach of confidentiality. In a parallel action, several lobbies of the industry jointly addressed more than 2,300 higher education institutions in the United States urging them to collaborate in stopping file sharing across their networks and eventually, the first law suits against individuals were filed in September 2003.

KazAa and the FastTrack network

Despite the prominence of American companies and individuals, P2P was not only an American phenomenon. Niklas Zennström and Janus Friis, two young Europeans and former employees from the Swedish low cost telecommunication services provider Tele2, were responsible for writing an important chapter in P2P history.

In 1999, while living in Amsterdam and looking for ideas to start up their own company, they heard about Napster. Roth (2004), in an article after an interview with them, recalls that Zennström had been frustrated about having to buy network capacity for his ISP in order to cope with the traffic imbalance caused by too many Europeans downloading and streaming content from servers in the United States. Thus, the idea of P2P and users sharing content among themselves caught on in their minds, and they decided to create a network where users could exchange any kind of content.

Zennström and Friis invested their own money and hired a team of programmers in Estonia with whom they had had relations through Tele2. In barely 4 months the team developed a protocol which they called FastTrack that had three major advantages over Napster. Firstly, it supported exchanging all kind of files and not only mp3s; secondly it did not require a big farm of computers that grew exponentially with the number of users; and thirdly, it could resume file downloads that were interrupted by locating another peer that had the same file.

Although the details of the protocol are a company secret, some parts of it have been described (Ding *et al*., 2005). FastTrack can be classified as somewhere between Napster with its centralized search and the completely decentralized Gnutella. The protocol is based on two types of nodes: *standard nodes* and *supernodes*, which were the equivalent of ultra peers in Gnutella version 0.6. In the same fashion as Gnutella, the network is difficult to shut down, as the peers work independently from a central server.

Zennström and Friis licensed the protocol to two companies that generated their own clients and P2P networks, Grokster and iMesh. Additionally, the FastTrack creators also launched their own network, called KaZaA, with its own client program. Each of these networks worked according to the same protocol, but were incompatible with one another. As for their business model, all three used client software with targeted advertising, which proved controversial among users and triggered some reverse engineering hacker projects to deactivate this feature.

FastTrack networks became the most popular P2P networks during 2003 and 2004. In a longitudinal measurement study, Karagiannis et al. (2003) observed the dominance of FastTrack over all the other P2P protocols, and Liang et al. (2004) reported that on any given day in 2004 there were 3 million users in the KaZaA network -the most popular of the three- and refer to sources that measured FastTrack traffic to be 76% of all P2P traffic in the Internet.

FastTrack's success was another blow for the media industry that initiated a new set of legal actions against the different parties involved in these networks. Even if the lack of a central server made FastTrack in theory less vulnerable than Napster, the three companies Grokster, iMesh and KaZaA were sued by the media industry in the United States. Of the three, Grokster decided to fight and lost, being forced to close by the Supreme Court in 2005. iMesh reached a settlement, which allowed it to continue operating by distributing files protected with Digital Rights Management (DRM) technologies[9], and KaZaA decided to ignore the legal actions for a certain time, on the basis that they were not operating under United States' jurisdiction. Thus, the industry also sued the KaZaA companies in their home countries, but the European side of the case proved more difficult for them, as the legal texts were different than in the United States, and neither the Dutch nor the Estonian courts found any of the accused parties guilty. However, by the time the sentences were ready, Zennström and Friis had already moved on to their new venture, the P2P telephony operator Skype, after selling KaZaA to an obscure Australian company incorporated in Vanuatu, Sharman Networks, that would eventually sign a settlement with the industry.

Structured Overlay Networks

The networks described so far, even when functionally very different, share a common approach to solving the problem of sharing files in a large network. They all start from the point of view of the downloader, and proceed as follows: first they look for the content they want to download, and then they contact the peer that hosts the file and request it. The underlying assumption is that files are available in an unstructured way in the network, and that they first need to be found before they can be downloaded.

Nevertheless, this approach is not the only one to solving the downloading problem. Soon after P2P gained interest from the public, different groups at universities and other research institutions started to make proposals for new network architectures. A distinctive set of solutions correspond to *structured networks* like Pastry (Rowstron and Druschel, 2001), Tapestry (Zhao *et al.*, 2001), CAN (Ratnasamy *et al.*, 2001) or Chord (Stoica *et al.*,

---

9.   DRMs are a set of technologies that impose limitations on the use of digital content, for example by restricting reproduction to one specific device or during a limited period of time. Their use is controversial, and is highly criticized by organizations like the Free Software Foundation or the Electronic Frontier Foundation. Content distribution companies like Apple Corporation have now abandoned these techniques for music files, although they keep using them for other types of files.

2001). All these implementations share in common the fact that files are stored in a predictable way in the network, and that every node that contains files is identified with a certain key that relates directly to the file it contains. The correspondence between files and locations is kept through a Distributed Hash Table (DHT). As follows from this term, these algorithms are distributed, and thus do not have a central server that can be switched off. They are also based in hash functions, which can generate unique signatures per files.

Due to their complexity and limitations in handling queries and the dynamic behavior of users that join and leave the network (Taylor & Harrison, 2009), these networks were never very popular as a stand alone. Nevertheless, their techniques were later incorporated by other P2P networks, reaching some success as part of bigger systems. The most successful of these was Kademlia, a network proposed by two researchers at New York University, Maymounkov and Mazières (2002), which was later to be included in variants of eMule and BitTorrent protocols.

Mojo Nation

Mojo Nation is a failed P2P network that we are including here both as an example of the multidirectional development of the technology and also because of the influence it had on the future of P2P networks. It was the creation of Jim McCoy, a veteran of the Internet that left his job at Yahoo in May 2000 to start a company that he called Autonomous Zone Industries —the name being a reference to a novel by the anarchist author Hakim Bey— and develop his own P2P network. With a powerful libertarian inspiration (Cave 2000), McCoy defined his software as "a cross between Napster and eBay" (McCullagh, 2000), although it also compares to Freenet[10], which seems to have started as a simultaneous development with no connection between the two projects. He used his own money to finance his venture, but failed to raise new rounds of capital, and eventually went out of business in the year 2002 in what seemed a combination of lack of funding, not enough users, and bigger than anticipated technical issues.

As a network, Mojo Nation represented an important milestone in the development of P2P technologies, with distinctive features arising directly from McCoy's ideology[11]. One of the most important differences with other P2P networks was that files were not directly shared from the publishing peer's computer, but rather split, distributed and replicated through other computers of the network. In this way, when a P2P user wanted to retrieve one file from the network, the file would not just come from only one place but from several, and it would need to be reassembled before it could be used. This technique was known as *swarming*, and was conceived in order to make content available even after the origina-

---

10.   We will discuss Freenet later on.

11.   A technical description of the protocol was available at the Mojo Nation site: http://web.archive.org/web/20020127125928/www.mojonation.net/docs/technical_overview.shtml (last checked 3/10/10).

tors turned their machine off, providing the possibility of round-the-clock trading and market liquidity. Nevertheless, swarming also had a positive side effect with a dramatic improvement in download speed, allowing the network to get over the upload throughput limitations. In a traditional point to point connection between two residential users, the uploader was very likely to have more severe limitations in the uploading speed, either because of the asymmetry of ADSL technology, or simply because of ISP-imposed limitations. By using swarming, the downloader could combine several uploaders at their maximum speed until the combination reached the maximum download throughput.

As regards the conflict with music and media industries, McCoy also foresaw several mechanisms which he thought would keep him safe from being sued. Firstly, the search function was outsourced to users that were willing to run file trackers and get credit for that. So, at least in theory, Autonomous Zone Industries could not be liable for providing a service like Napster. Secondly, in the event that a content creator reported a violation of copyright, McCoy's company could mark the blocks in the network related to that file as bad blocks, and effectively stop the sharing of that file. And thirdly, he also foresaw a "tipping" mechanism by which users could make a donation to the content creators. As part of his vision, he expected that the music labels would publish using his network and be financed with the donations from downloaders (Cave, 2000).

With all its complexity, Mojo Nation attracted quite a lot of interest of the media in the year 2000, but the network never really took off. Nevertheless, some of its ideas were taken up by Bram Cohen, an employee of Autonomous Zone Industries that left the company in 2001 to start his personal project, the BitTorrent P2P client.

BitTorrent

BitTorrent constitutes another important landmark in P2P evolution, both due to its widespread success and to the acceptance (albeit partial) of the music and movie industries. Its protocol was created by Bram Cohen, a hacker who in the same tradition as Fanning and Frankel also dropped out of college during the 1990's , although he only reached notoriety at the age of 26 when he created BitTorrent (Berfield, 2008).

Cohen left Mojo Nation not with the idea of starting a company, but rather to develop a project that would give him personal satisfaction and that would suit his hackeristic values. As he expressed in an interview in 2005 after working for several start-ups that went bankrupt, he just wanted to write something for himself in his own way and give it away for free: "You get so tired of having your work die. I just wanted to make something that people would actually use" (quoted in Thompson, 2005).

BitTorrent initially became popular among Linux fans, who used the program to share and download Linux distributions, and it progressively gained acceptance among other types of users. CacheLogic, a consultancy firm that provided services for ISPs, estimated based on real Internet traffic measurements that between January and June 2004 BitTor-

rent's share of P2P traffic rose from 26% to 53% worldwide, whereas Fastrack descended from 46% to 19% in the same period[12]. In a second study a few months later they reported that BitTorrent was responsible for 30% of all Internet traffic at the end of 2004[13].

From its beginnings, the project was completely open source, non-commercial, coordinated by Cohen himself, and allowed for alternative software clients written by other parties. However, by the end of 2004, Cohen decided to start a company using venture capital to exploit the protocol success. Faithful to his hackeristic approach, Cohen then separated the source code availability from his own commercial enterprise, which meant moving the BitTorrent project from his personal web page to two new ones: www.bittorrent.org for the protocol development and evolution, and www.bittorrent.com, for BitTorrent Inc.

Bittorrent.org defines BitTorrent as a "free speech tool" that enables content to be published at low cost using cooperative distribution, and uses the motto "give and ye shall receive!"[14]. From this declaration we can extract two key ideas that are embedded in the protocol. First, BitTorrent was designed to share and to publish, rather than to download, and whatever didn't fit in this use was stripped out of the basic form of the protocol. This even goes to the extreme of not providing any centralized or distributed content search mechanism that could appeal to downloaders. Peers need to contact a tracker —the original publishing peer— for a specific file and they will be instructed on where to find other peers that are downloading the same file. The location of the tracker and of the content it hosts is handled outside the protocol, typically through a simple search with an Internet browser or sometimes through more sophisticated methods like DHTs established by the P2P application.

The second key idea is the motto "Give and ye shall receive". Being designed as a tool for publishers, it follows that those who need to pay a price in terms of computing power and bandwidth are the receiving peers. So, the protocol forces the downloaders to share among themselves and makes free riding virtually impossible.

The general behavior of the protocol is described in a paper by Cohen (2003). The most significant architectural element is the splitting of files into smaller blocks, an idea he borrowed from Mojo Nation's swarming distribution. Thus, as a tracker begins publishing a file, it sends different blocks to the peers that approach it and then, since each peer receives a different block, peers can start sharing the pieces they already have without the need to contact the tracker. Combining this feature with the forced upload results in faster file distribution as more users try to download the same file. This particular feature makes the protocol particularly well suited for distributing popular files.

---

12.   http://web.archive.org/web/20061104200853/www.cachelogic.com/home/pages/studies/2004_09.php (last checked 3/10/10).

13.   http://web.archive.org/web/20061018024808/www.cachelogic.com/home/pages/studies/2005_06.php (last checked 3/10/10).

14.   http://www.bittorrent.org/introduction.html (last checked 3/10/10).

BitTorrent seems to provide an elegant technological solution to the debate for all the different parties. Coming from the world of hackers, it suits those who want to share. "Give and ye shall receive" is part of the hacker culture, and therefore this group feels at ease with this implementation. But BitTorrent also suits users more focused on downloading than in sharing. The improved download speeds are an advantage that arguably moved users away from FastTrack and other networks to BitTorrent. Finally, the separation of the search function from the file sharing itself is an architectural change that suits both the content providers and the P2P entrepreneurs. For the content providers, the technology now becomes neutral. Thus, they will not sue the technology developers this time, but the sharers that make unauthorized content available. Regarding the P2P entrepreneurs, they were liberated from the tedious task of developing a closed-garden proprietary software. By using commoditized standard software, a new generation of entrepreneurs focused on launching websites that offer search services like The Pirate Bay or Mininova, financed by advertising. These sites became the new target of the media industry, that tried to have the sites closed or banned using their legal arm.

eDonkey, eMule and the eDonkey2000 network

eDonkey was the original creation of Jed McCaleb, a UC Berkeley dropout who combined working as a programmer in the Bay area with his own software projects, often released as shareware. When he discovered Napster, he was so taken by the idea of summing up individual computers to make a "massive hard drive" (Healey, 2005) that he quit his job and decided to start writing his own competing application.

McCaleb's network was based on clients and servers, like the original Napster, although he proposed replacing Napster's central server with independent servers that could specialize in specific types of content, and would thus allow for faster and more targeted searches. However, what he did not foresee was that many operators would link their servers to create one large server network that was to be known as eDonkey2000 (ed2k), and that transformed the architecture of the network into something similar to Gnutella 0.6 or FastTrack. Another improvement he devised was a simple swarming mechanism, which he thought would be a way to accelerate the download of big (heavy) files -hence the name Donkey.

In 2001 McCaleb incorporated a company called Metamachine, and a few months later was joined by Sam Yagan, a Harvard graduate and young entrepreneur. McCaleb continued in charge of the technical developments while Yagan took care of the business side.

As eDonkey started to gain in popularity, its story was similar to that of KaZaA and the other FastTrack companies. As a way of getting some revenue they offered two P2P client applications, one free with advertising, and one premium that was advertising-free but had to be purchased. However, this did not please a number of users, and a team led by a hacker named Hendrik Breitkreuz decided to create an open source and improved client for Windows, which they called eMule.

In 2005, ed2k was reported by Sandvine[15] as the most popular P2P network in some European countries, including 72% of all file sharing in Germany and 80% in France, while BitTorrent was already more popular in the United States and the United Kingdom. However, Metamachine eventually got into trouble with the copyright industry, and despite attempts to reach an agreement for distributing copyrighted material, they were forced out of business in 2006 and stopped maintaining the eDonkey client.

When Metamachine ceased operations, eMule was already a mature and popular project able to work in the e2dk network. Furthermore, it had been extended with support for a DHT based network, which means that although some content was still residing in dedicated servers, many files resided in a decentralized structured network. eMule continues to be used to this date, and its web page reports that the program has been downloaded more than 417 million times[16]. The protocol is slower than BitTorrent, though it is still preferred by many users who find a wider selection of older material for downloading that is not easily found in many Torrent sites[17].

Content distribution P2P platforms

Although the music, movie and gaming industries have traditionally shown strong opposition to P2P networks, they too have experimented with this technology at different stages. As part of the judiciary or extra-judiciary agreements with some networks like iMesh or KaZaA, they asked the latter to start distributing DRM-protected files, but this didn't succeed. In a development of this strategy, the main Hollywood studios cut a deal in 2006 with BitTorrent Inc in order to create the Torrent Entertainment Network, a content distribution network that combined free files with for-rent and for-sale premium ones, all distributed using the BitTorrent protocol. This network was not a commercial success, and ended up closing in December 2008.

Other experiments were carried out by broadcasters like the UK based British Broadcasting Corporation (BBC) and British Sky Broadcasting (BSkyB) using proprietary P2P technology from a company called Kontiki in combination with DRM. This technology creates managed networks that only allow authorized content to be injected by the broadcaster and move control away from the users. BBC's iPlayer, running on this technology, was maintained between 2005 and the end of 2008, when BBC replaced it with http-based video streaming software after complaints about excessive bandwidth consumption both from users and ISPs[18]. In 2010 Sky also replaced its Sky Player application with a browser based service.

---

15.   http://www.sandvine.com/news/pr_detail.asp?ID=88 (last checked, 3/10/10). Sandvine is a company that provides services for managing Internet traffic to ISPs, and therefore has access to real life data of its customers.

16.   http://www.emule-project.net/home/perl/general.cgi?l=1&rm=download (last checked 3/10/10).

17.   http://sharereactor.com/faq/ (last checked 3/10/10).

18.   See http://crave.cnet.co.uk/software/0,39029471,49291924,00.htm (last checked 3/10/10) and http://news.bbc.co.uk/2/hi/technology/7336940.stm (last checked 3/10/10).

Darknets and anonymous P2P networks

Darknets and *anonymous* P2P networks represent another attempt by file sharers to go underground in the face of growing opposition to P2P usage. We can define a darknet as a private computer network where all the users know and trust each other and can be used to exchange files and messages confidentially inside the group. Biddle *et al.* (2002) connect these networks with the tradition of copying music tapes and computer programs with family and friends, and argue that darknet-based P2P networks will play an important role in the future as an easy way to exchange copyrighted material with no risk of being prosecuted.

These networks are typically designed for groups of up to 50 users, all of them sharing the same public encryption key and being able to access all the shareable content of all the other users. They do not need a central server, and due to their size they are good at locating popular content of common interest for its users, but not rare files outside the small world they represent.

Perhaps the most popular darknet to date is WASTE, an open source project created in 2003 by the P2P hacker Justin Frankel, also creator of the original Gnutella. As in the case of Gnutella, he chose to publish his program in the webpage of Nullsoft, once again infuriating its parent company AOL Time Warner, which requested the program to be removed immediately. The program as well as some variants of it continues to live in the open source community Sourceforge.

Whereas darknets are useful in small world scenarios, they are not scalable and therefore lack some of the advantages of the more conventional P2P networks. As an alternative, there is a growing number of networks that are being developed with the ambition of becoming a large, anonymous P2P network. A number of these derive from an anonymous protocol called Onion Routing that was originally developed by a group of researchers in the US Navy (Goldschlag *et al.*, 1999). Examples of them are Tor and I2P.

An alternative network is Freenet, a project that grew out of a paper written by the Irishman Ian Clarke while he was a final year student at the University of Edinburgh (Clarke, 1999). Freenet is a decentralized network that allows sharing, not only of files, but also of web pages, e-mail and notice board messages in a completely anonymous way. After graduating in 1999, Clarke continued the development as an open source project, inviting other hackers to collaborate.

Freenet presents striking technical similarities with Mojo Nation, although it is generally slow and the contents shared can get lost and are not guaranteed to be found again. Freenet was conceived as a non-commercial tool to promote freedom of speech and defense of democracy by ensuring that "the government cannot control its population's ability to share information, to communicate"[19]. The project is still under development, but the network is fully usable and its creators have made efforts to make it available to the peoples of

---

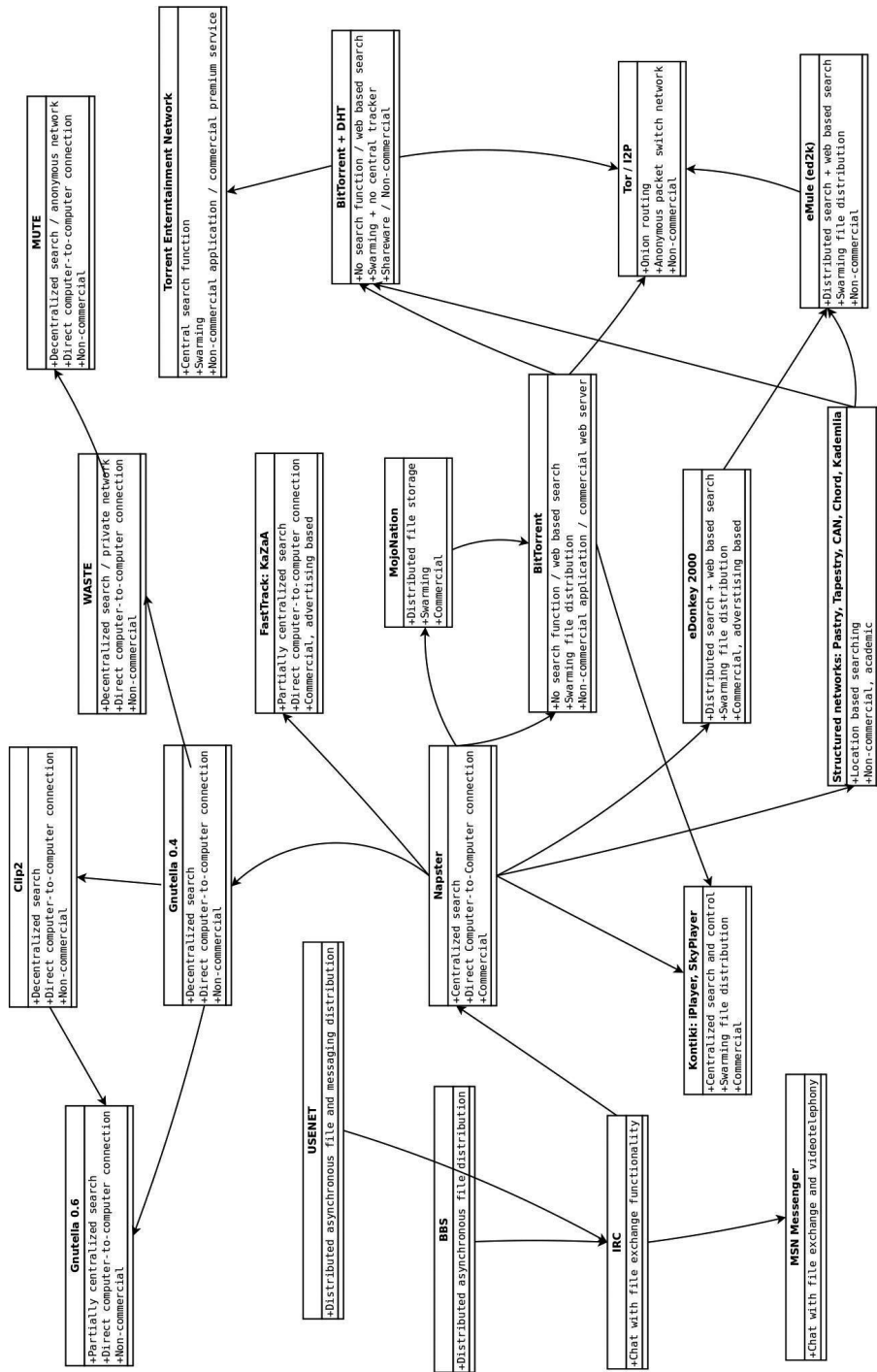19.   http://freenetproject.org/philosophy.html (last checked 3/10/10).

Figure 1.   Evolution of P2P networks

China, Iran, and other countries notorious for government Internet monitoring practices. As per the copyright conflict, Clarke is very assertive: "You cannot guarantee freedom of speech and enforce copyright law," and therefore, Freenet "must prevent enforcement of copyright"[20].

Another network that reached some notoriety was MUTE, an open source project developed in 2004 by the independent programmer Jason Roher. Roher conceived his network as a form of social activism against the tactics of the RIAA, "encouraging people to break an unjust law as a form of social protest" (Wen, 2004).

Roher's P2P client program created a decentralized anonymous network based on the Gnutella 0.4 protocol. Besides encrypting all the file communications to avoid eavesdropping, his main contribution was masking the IP addresses of the participant computers and avoiding direct computer-to-computer connections for download. His rationale was that RIAA lawsuits started with the IP address of a computer, which was enough to discover the name and address of the sharer.

Conclusion

The historical evolution of P2P we have outlined does not follow a simple linear path towards more efficient or horizontal ways of sharing files through the Internet: it resembles more the so-called multidirectional model proposed by constructivist accounts of technology (Pinch & Bijker, 1987). Many forces and actors with different and sometimes conflicting interests have had a role in shaping P2P networks. In particular, the controversy that has developed around music and film downloading and its very evolution from the so-called copyright wars to a broader societal debate on cultural production and access to it has been one of the key features in triggering most of the technical innovations we have analyzed.

Designers of most P2P networks share a very similar profile: young, visionary and technologically savvy people who understood the potential uses of P2P networks and in a very entrepreneurial way tried to create new forms of business with them. Most of them showed an ambiguous ideological stand fluctuating between utilitarian and romantic individualism; a trend that has also shaped the evolving, and sometimes messy, ethos of the Internet (Streeter, 2011, 113). But as entrepreneurs, they also had different approaches to creation of value and business models. In general, the early experiences suffered from the dot com bubble illness that Porter (2001) describes as "applying creative accounting in the form of dubious performance metrics such as number of visitors or unique users". This was the case of Napster, which was offered completely free, with no advertising in the client application, and never managed to articulate a way in which it could make any profit, yet still managed to raise several rounds of funding (Ante, 2000a). Other entrepreneurs tried to be a bit more explicit about their plans for creating a profit. For example, KaZaA, whose founders already

---

20.   http://freenetproject.org/philosophy.html (last checked 3/10/10).

had significant business experience, created not only a complex legal structure engineered to escape legal prosecution, but also included advertising in the client application based on analyzing users' search patterns, technology licensing, and support for downloading paying content with embedded DRM mechanisms.

When the tasks of searching and sharing were separated, P2P entrepreneurs focused on providing their services around the search function, whereas the sharing and the protocol part became a commodity which did not allow for economic value creation. However, this shift did not change the problems with the copyright holders, who continued to sue the entrepreneurs behind sites like Demonoid, Mininova or The Pirate Bay.

In fact, the dramatic court battle and eventual closure of Napster signaled very clearly to the creators-to-be of P2P networks that a centralized network was not legally viable anymore. Thus, with a few early exceptions, all the other subsequent networks moved towards various decentralized architectures. The early Gnutella represented a swing towards an egalitarian and completely flat network, which corresponded as well to the ideological views of hackers like Frankel, who believed that everybody should be sharing information on equal terms. However, some sort of centralized management was still required by entrepreneurs that wanted to control the technology in order to obtain revenues to capitalize their investments. The first proposal to include advertising in the P2P client, though, proved not to be a lasting one. Even if KaZaA and eDonkey worked with advertising for some time, both saw alternative advertising-free applications get developed by hackers that eventually replaced the official ones.

All in all, the controversy has seen an extraordinary use of very diverse strategies by the different sides and actors, on a global level and unprecedented scales, in order to impose their views. These strategies have been aimed at establishing particular interpretations of the technology involved and thus forcing specific uses (or non-uses) of it. Some of the actions have been addressed to change the context of use —as in the legal prosecution of users or P2P companies and the involvement of political institutions; others have taken a semiotic character in order to impose negative meanings on opponents —as in the use of concepts such as 'pirate', 'mafia'[21] or 'terrorist'[22]— in order to discredit them. However, most notably, some strategies have been aimed at building particular meanings and values into the very technical design in order to compel users to follow them in a more implicit and subtle way —as in the changes introduced in BitTorrent, the DRM systems, or the swarming mechanisms in Mojo Nation. This move echoes the *inscription* process described by actor-network scholars (Latour, 1992) and proves to be a more solid and lasting strategy to achieve change. Whether the next chapters of this controversy will be written in legal or software code remains to be seen.

---

21.   The terms mafia and MAFIAA, a parody acronym for a hypothetical Music And Film Industries Association of America, are used widely on the Internet by supporters of free sharing of media files.

22.   Jack Valenti, president of the MPAA: "We're fighting our own terrorist war" (Harmon 2002).

References

ABBATE, J. (1999). *Inventing th e Internet*. Cambridge: MIT Press.

ABBATE, J. (2010). Privatizing the Internet: Competing Visions and Chaotic Events, 1987-1995. *IEEE Annals of the History of Computing,* 1058-6180, pp. 10-22.

ANTE, S. E. (2000a). Inside Napster. *Business Week Online*. Available at <http://www.businessweek.com/2000/00_33/b3694001.htm> [Last checked 2/2/11].

ANTE, S. E. (2000b). Napster's Shawn Fanning: The teen who woke up web music. *Business Week Online*. Available at <http://www.businessweek.com/ebiz/0004/em0412.htm> [Last checked 2/2/11].

BERFIELD, S. (2008). BitTorrent's Bram Cohen isn't limited by Asperger's. *Business Week Online.* Available at <http://www.businessweek.com/magazine/content/08_43/b4105046863317.htm> [Last checked 2/2/11].

BIDDLE, P.; ENGLAND, P.; PEINADO, M.; WILLMAN, B. (2002). The darknet and the future of content distribution. *Proceedings of the 2002 ACM Workshop on Digital Rights Management*. Washington, DC, Nov. 18.

CAVE, D. (2000). The Mojo Solution. *Salon*. Oct 9. <http://www.salon.com/technology/view/2000/10/09/mojo_nation/index.html> [Last checked 2/2/11].

COHEN, B. (2003). Incentives build robustness in BitTorrent. *Proceedings of the 1st Workshop on Economics of Peer-to-Peer Systems.* Berkeley, CA. <http://www.bittorrent.org/bittorrentecon.pdf> [Last checked 2/2/11].

CLARKE, I. (1999). *A Distributed Decentralized Information Storage and Retrieval System*. Undergraduate thesis. University of Edinburgh. <http://freenetproject.org/papers/ddisrs.pdf> [Last checked 2/2/11].

DELIO, M. (2000). Gnutella development gnotted. *Wired News*. 09 Nov.

DING, C. H.; NUTANONG, S.; BUYYA, R. (2005). Peer-to-peer networks for content sharing. In *Peer-to-Peer Computing: The Evolution of a Disruptive Technology*, eds. R. Subramanian and B. Goodman, pp. 28-65, Idea Group Publisher: Hershey, PA.

GOLDSCHLAG, D.; REED, M.; SYVERSON, P. (1999). Onion routing. *Communications of the ACM*, 42(2):39-41.

GREENFELD, K. T.; TAYLOR, C.; THIGPEN, D. E. "(2000). Meet the Napster. *Time Magazine*. 02 Oct. <http://www.time.com/time/magazine/article/0,9171,998068,00.html> [Last checked 2/2/11].

HEALEY, J. (2005). The piracy arms race. *Los Angeles Times*. 2 Oct. <http://articles.latimes.com/2005/oct/02/business/fi-piracytwo2> [Last checked 2/2/11].

HARMON, A. (2002). Black Hawk download: Pirated Videos Thrive Online. *The New York Times.* 17 Jan. <http://www.nytimes.com/2002/01/17/technology/circuits/17VIDE.html> [Last checked 2/2/11].

HIMANEN, P. (2001). *The Hacker Ethic and the Spirit of the Information Age*. London: Secker & Warburg.

KARAGIANNIS, T.; BROIDO, A.; BROWNLEE, N.; CLAFFY, K. C.; FALOUTSOS, M. (2003). *File-sharing in the internet: A characterization of p2p traffic in the backbone*. Technical report. University of California, Riverside, USA.

KLINGBERG, T.; MANFREDI, R. (2002). Gnutella 0.6. *The Gnutella RFC*. June. <http://rfcgnutella.sourceforge.net/src/rfc-0_6draft.html> [Last checked 2/2/11].

KUSHNER, D. (2004). The world's most dangerous geek. *Rolling Stone*. <http://www.rollingstone.com/news/story/5938320/the_worlds_most_dangerous_geek> [Last checked 2/2/11].

LATOUR, B. (1992). "Where Are the Missing Masses? The Sociology of a Few Mundane Artifacts". In: Bijker, W. and Law, J. (eds). *Shaping Technology/Building Society. Studies in Sociotechnical Change*. Cambridge (MA): MIT Press; 225-258.

LESSIG, L. (2004). *Free Culture. How Big Media uses Technology and the Law to lock down Culture and Control Creativity*. New York (USA): Penguin Press.

LIANG, J.; KUMAR, R.; ROSS; K. W. (2004). *The kazaa overlay: A measurement study*. Technical report. Department of Computer and Information Science. Polytechnic University, Brooklyn, NY, USA. September.

MAYMOUNKOV, P.; MAZIÈRES, D. (2002). Kademlia: A peer-to-peer information system based on the xor metric. *Proceedings of the 1st International Workshop on Peer-to-Peer Systems (IPTPS '02)*: 53-65.

MCCULLAGH, D. (2000). Get Your Music Mojo Working. *Wired*. 29 Jul.<http://www.wired.com/science/discoveries/news/2000/07/37892> [Last checked 2/2/11].

MIROWSKI, P.; SENT, E-M. (2008). "The Commercialization of Science, and the Response of STS". In: E.J. HACKETT, O. AMSTERDAMSKA, M. LYNCH & J. WAJCMAN (eds). *The Handbook of Science & Technology Studies*, Third Edition, Cambridge, MA: MIT Press; 635-90.

PINCH, T.; BIJKER, W. (1987). "The Social Construction of Facts and Artefacts: Or How the Sociology of Science and the Sociology of Technology Might Benefit Each Other". In: BIJKER, W.; THOMAS, E.; HUGHES, P.; PINCH, T. (eds.). *The Social Construction of Technological Systems: New Directions in the Sociology and History of Technology*. Cambridge (MA): MIT Press; 17-50.

PORTER, M. E. (2001). Strategy and the internet. *Harvard Business Review* 79(3):62-79.

RATNASAMY, S.; FRANCIS, P.; HANDLEY, M.; KARP, R.; SCHENKER, S. (2001). A scalable content-addressable network. *Proceedings of the 2001 SIGCOMM conference* 31(4):161-172.

RIPEANU, M.; FOSTER, I. (2001). *Peer-to-peer architecture case study: Gnutella network*. Technical report, University of Chicago.

ROTH, D. (2004). Catch us if you can. *Fortune Magazine*. 9 Sept. <http://money.cnn.com/magazines/fortune/fortune_archive/2004/02/09/360106/index.ht> [Last checked 2/2/11].

ROWSTRON, A.; DRUSCHEL, P. (2001). Pastry: Scalable, decentralized object location, and routing for large scale peer-to-peer systems. *Lecture Notes in Computer Science*, 2218: 329-350.

SAROIU, S.; GUMMADI, P. K.; GRIBBLE, S. D. (2002). A measurement study of peer-to-peer file sharing systems. *Proceedings of Multimedia Computing and Networking (MMCN '02)*; San Jose, CA, USA; January.

SCHODER, D.; FISCHBACH, K.; SCHMITT, C. (2005). Core Concepts in Peer-to-Peer Networking. In *Peer-to-Peer Computing: The Evolution of a Disruptive Technology*, eds. R. Subramanian and B. Goodman, pp. 1-27. Idea Group Publisher: Hershey, PA.

SPITZ, D.; HUNTER, S. (2003). *The social construction of Napster*. MIT Sloan Working Paper No. 4445-03. <http://ssrn.com/abstract=469283> [Last checked 2/2/11].

STOICA, I.; MORRIS, R.; KARGER, D.; KAASHOEK, M. F.; BALAKRISHNAN, H. (2001). Chord: A scalable peer-to-peer lookup service for internet applications. *Proceedings of the 2001 conference on Applications, technologies, architectures, and protocols for computer communications*. San Diego, CA, USA; Jan.; 149-160.

STREETER, T. (2011). *The Net Effect: Romanticism, Capitalism, and the Internet.* NY University Press: New York.

TAYLOR, I. J.; HARRISON, A. (2009). *From P2P to Web Services and Grids: Evolving Distributed Communities*. Springer-Verlag New York Inc: New York (USA).

THOMPSON, C. (2005). The bittorrent effect. *Wired*. 13 Jan.<http://www.wired.com/wired/archive/13.01/bittorrent.html> [Last checked 2/2/11].

VARANINI, G. (2000). *Q&A: Napster creator Shawn Fanning*. Available at <http://news.zdnet.com/2100-9595_22-96066.html> [Last checked 2/2/11].

WEN, H. (2004). *Anonymous, open source p2p with mute*. Available at <http://www.onlamp.com/pub/a/onlamp/2004/08/12/mute.html?page=1> [Last checked 2/2/11].

ZHAO, B. Y.; KUBIATOWICZ, J.; JOSEPH, A. D. (2001). *Tapestry: An infrastructure for fault-tolerant wide-area location and routing*. Technical Report CSD-01-1141, U. C. Berkeley.