The Copyright Crusade

Abstract

During the winter and spring of 2001, the author, chief technology officer in Viant's media and entertainment practice, led an extensive inquiry to assess the potential impact of extant Internet file-sharing capabilities on the business models of copyright owners and holders. During the course of this project, he and his associates explored the tensions that exist, or may soon exist, among peer-to-peer start-ups, "pirates" and "hackers," intellectual property companies, established media channels, and unwitting consumers caught in the middle. This research report gives the context for the battleground that has emerged, and calls upon the players to consider new, productive solutions and business models that support profitable, legal access to intellectual property via digital media.

by
Andrew C. Frank, CTO afrank@viant.com
Viant Media and Entertainment
Reinhold Beutler rbeutler@viant.com
Aaron Markham amarkham@viant.com

assisted by
Bruce Forest

VIANT

1 Call to Arms

Well before the Internet, it was known that PCs connected to two-way public networks posed a problem for copyright holders.

The problem first came to light when the Software Publishers Association (now the Software & Information Industry Association), with the backing of Microsoft and others, took on computer Bulletin Board System (BBS) operators in the late 1980s for facilitating trade in copyrighted computer software, making examples of "sysops" (as system operators were then known) by assisting the FBI in orchestrating raids on their homes, and taking similar legal action against institutional piracy in high profile U.S. businesses and universities.[1] At the same time, the software industry found that it needed to adjust its licensing strategy, shifting from a uniform shrink-wrap "per user" licensing model to more institutionally-focused models such as site licensing, upgrade fees, and service contracts.

In 1996 the Internet drove the BBS into obscurity and created a new movement. The Internet became more than just a ubiquitous computer network: it became a symbol of radical change. Public critics of copyright law such as Richard Stallman of the Free Software Foundation found audiences for their views[2] (as well as their open-source software), summarized by the contentious slogan "information wants to be free."

To briefly review the events that followed,this point of view did not evoke much of a response from intellectual property (IP) organizations for about three years. In 1997, Hotline Communications was founded and began providing the first in a series of public havens for anonymous file trading on the Internet.Also by 1997, MP3 audio format was gaining in popularity (although it had been in the public domain since 1993). As a threat, it was by and large overlooked by the recording industry.[3]

In 1999, the situation changed dramatically. In January, Shawn Fanning, a Computer Science student at Northeastern University, created a file-sharing program called Napster. In May, Napster, Inc. was funded and founded. In August, Napster.com launched its service and became an overnight sensation. And on December 7th of that year, the record industry sued Napster, Inc. for copyright infringement.

That December saw a rapid succession of key events for copyright.About the same time that the RIAA[4] was filing suit against Napster in Northern California, in Toronto a group of television networks was suing icravetv.com for rebroadcasting their shows over the Internet.[5] And in Europe, two hacker groups known as DoD (Drink or Die) and MoRE (Masters of Reverse Engineering) were purportedly looking for ways to play DVDs on the unsupported Linux operating system when they stumbled across a flaw in DVD's copy protection scheme (known as CSS for "Content Scrambling System"). They created an application known as DeCSS that defeated the copy protection scheme, allowing users to "rip" unscrambled DVD contents onto PC hard drives and promptly published their code on the Internet. The code quickly found its way onto thousands of Web sites around the world. Both of these events served to drive home the point that the Internet was a global theater for copyright issues.

The DeCSS incident led the MPAA to sue Eric Corley, the publisher of "2600: The Hacker Quarterly," which operated one of the Web sites on which DeCSS could be found (in defiance of a court order to remove it). The basis of the MPAA's case was a clause in the recently passed Digital Millennium Copyright Act of 1998[6] that specifically prohibited "circumvention of copy protection systems."[7] The lawsuit (in which the MPAA eventually prevailed[8]) set up a bitter debate when organizations such as the American Civil Liberties Union (ACLU), the Electronic Frontier Foundation (EFF), and several prominent scholars took up the issue, claiming that Corley's actions were "fair use" and "free expression" and were protected under the First Amendment. The debate foreshadowed many disputes to come.

The year 2000 saw a near doubling in the adoption of consumer broadband services[9], as well as the continued high speed wiring of U.S. universities and corporations. Along with these developments came the inevitable spread of file trading beyond music to movies and television shows. Hollywood features became particularly popular, with many post-production or guild versions arriving on piracy channels prior to their theatrical release. Similarly, television shows intercepted from satellite feeds through which networks distribute content to affiliates[10] began to appear prior to being aired. (In a recent example, "Endgame," the series finale of **Star Trek: Voyager,** appeared on piracy channels one week before it was broadcast on UPN.)

Today, the RIAA,MPAA,NAB, and other organizations concerned with the protection of intellectual property are striking back in force at Internet file sharing services whose existence is now recognized as a real threat – not necessarily to immediate revenue (where their effects are controversial), but, more importantly, to the ultimate premise of copyright control and enforceability in cyberspace.

The back-and-forth of this legal struggle has been a subject of intense media scrutiny, and we will not review that here. Rather, we will address the more fundamental question: what will it take for media companies to prevail and service a legitimate mass market for content online? In the course of exploring this question, we will study the phenomenon frequently referred to as "Internet piracy" to gain some insight into the often obscure behavior patterns of today's most enthusiastic Internet media audience.

2 Theaters of Conflict

In many ways, the digital copyright war resembles any widespread social struggle that pits organizational interests against individual behaviors. We can see this struggle play out on five distinct fronts: legal, tactical, economic, educational, and architectural.[11] To be clear, we need some definitions.

- **The legal theater.** This is typically the most visible front (since it is public), and its three arenas reflect the U.S. structure of government: in the judicial arena, we have the filing and adjudication of lawsuits and countersuits; in the legislative area we have the passage, debate, and evolution of new laws such as DMCA; and on the executive front we have law enforcement activities, which include the recent practice of sending hundreds of letters to Internet Service Providers (ISPs) citing infringement activities under DMCA.[12] However, the copyright problem on the Internet is of an international scale, which raises considerable additional legal and political complexities. As mentioned, we will refrain from an extensive examination of such conflicts. However, it is worth noting that, at the time of this writing, with the legal defeat of Napster apparently at hand, the RIAA has escalated its legal offensive to include Aimster and Launch. Both of these cases will test the RIAA's ability to prevail against increasingly complex challenges. Aimster, for example, employs encryption in

its technology and does not apparently maintain a centralized database of available content as Napster does.[13] We will focus on such difficulties later in this paper.

• **The tactical theater.** This refers to what security measures IP organizations might take to protect content and what countermeasures hackers might take to unprotect it.What might IP organizations do to track and interdict illegal traffic, and what measures might rogue services take to create protected offshore havens and proxies for untraceable file copying? What countermeasures might IP owners take to censure or disconnect these, etc.? Central to this area of conflict are notions of data security, a highly technical field. We will avoid as much as possible delving into the mechanics of security architectures and instead examine their implications in the design of distribution services.

• **The economic theatre.** As legal and tactical battles run their courses, market forces also serve a regulatory function. The price of content, models for payment, channel conflicts, distribution windows, accounting and reporting practices all have strategic implications for copyright's ultimate value. Also related to these issues is the matter of the cost of pursuing a legal and tactical arms race should it continue to escalate. A recent spate of lawsuits against IP organizations[14] suggests that the cost of defending, as well as prosecuting these cases might continue to rise, and the cost of pursuing individual enforcement measures on a file-by-file, user-by-user basis globally (should it come to this) is incalculable.

• **The educational theater.** One of the critical imperatives of IP organizations is to reinforce public acceptance of copyright norms. The RIAA in particular has made some efforts in this regard with on-campus education programs designed to impart college students with an appreciation for the importance of respecting copyright laws. Education is also a two-way process, and an objective of this paper is to shed light on the current situation to inform strategic thought on both sides of the distribution chain.

• **The architectural theater.** There is a considerable body of work that considers how architecture regulates behavior in physical space. In virtual space, the design of large-scale service architecture is likely to be the ultimate front on which the future of content distribution is decided. It is our basic

intention to suggest some ways of thinking critically about such architectures so that they might ultimately shape the online media consumption experience. We will look at factors that influence consumer behavior such as cost, convenience, choice, and quality. We will examine where legitimate services might gain advantage over peer-based distribution, despite the latter's apparent cost and choice advantages.

Many copyright owners are well aware that demand for content over the Internet may yet create their most profitable distribution channel (being of relatively low cost and broad reach), and that today's peer-to-peer services are by no means ideal solutions for most consumers. There is a strong sense that users might well pay for more convenience and higher quality, more comprehensive services, that subscription and other models hold considerable promise, and that there is great marketing power to be gained from building new, more intimate entertainment-based brand relationships with consumers and communities.

But the path is cloudy and the investments required appear to be large and risky. Digital copyright clearance (at least in music) has shown itself to be a slow and laborious process. Channel conflicts and technical hurdles abound.And, perhaps most importantly, many content companies are still reeling from the bursting of the Internet bubble. We have entered a period of economic slowdown in which Internet-related investments are increasingly – and rightly – subject to scrutiny and bring to mind layoffs and business failures that are the fruits of some overzealous predictions and valuation methods.

There are many who believe the first step should be to finish the job of eliminating the free competition, to drive piracy back underground,and to make it clear who is in control. In effect, many seem to be saying,"Let's win the war before designing the peace."

3 Escalation Scenarios

But can the digital copyright war ultimately be won simply through legal and tactical escalation on the part of IP organizations? Despite the early idealism of "net libertarians," there is momentum building that suggests the answer is "yes."[15]

Several factors weigh in favor of this view:

- **IP organizations have a clear legal advantage.** In the U.S. (and the U.K.), copyright law is fundamentally well established and rests on a solid history of precedents. Although its application in cyberspace may present novel situations, these are rapidly becoming clearer and disputes are generally being resolved conservatively (in U.S. courts). The international situation is a bit shakier, but international copyright law does have some fairly solid roots.
- **IP organizations have a clear economic advantage.** The disappearance of venture funding that many peer-to-peer-based starts-ups need to finance their operations puts them in an extremely weak position to mount any kind of campaign, legal or otherwise. Entertainment start-ups in general are rapidly being either acquired by large media companies (e.g. MP3.com by Vivendi/Universal,myPlay.com by Bertelsmann, etc.) or going bankrupt (e.g.Scour).
- **Content owners are moving in earnest to claim the online space with new services.** Announcements by major media companies suggest that the coming year will see a wealth of new content service offerings. These will test a variety of distribution and payment models and, with proper design and marketing, could generate considerable consumer enthusiasm.
- **IP organizations and content owners are working aggressively with consumer electronics and infrastructure companies to create secure distribution channels and products.** As post-PC Internet-enabled entertainment devices like SONY's Playstation2 and Microsoft's xBox, HDTV, and a new generation of broadband stereo and portable music and video devices become prevalent, along with services like VOD and digital radio, the PC (which is the fundamental component of today's peer-to-peer file trading networks) may lose "share" as an entertainment device. Higher quality/higher bandwidth formats (like HDTV) and killer services (like VOD) could simply raise the stakes of enjoyment. These new products may implement hardware-based security systems that restrict their network operation to secured channels. And established media companies clearly have greater entrée into these channels and product development companies than start-ups or rights organizations like EFF.

Having made note of these factors, let's examine the other side of the story and take note of the risks involved in this set of assumptions.

Returning to the spring of 1999, while Shawn Fanning was busy making it easier for computer users to trade MP3s, another computer scientist in London was working on a more comprehensive strategy to secure freedom of expression on the Internet. His name was Ian Clark and, in June, he completed the initial design of a software architecture known as Freenet.

Although like Napster it is designed to support open file trading, Freenet differs from Napster in some key respects, and represents a class of services we will refer to as **cryptographically distributed databases** (or CDDs). CDDs present a new kind of challenge to copyright owners. In essence, rather than supporting copying among individual users, CDDs create an encrypted database that aggregates all of the contributions of its constituents and spreads it across all of their machines. (See Figure 1.)

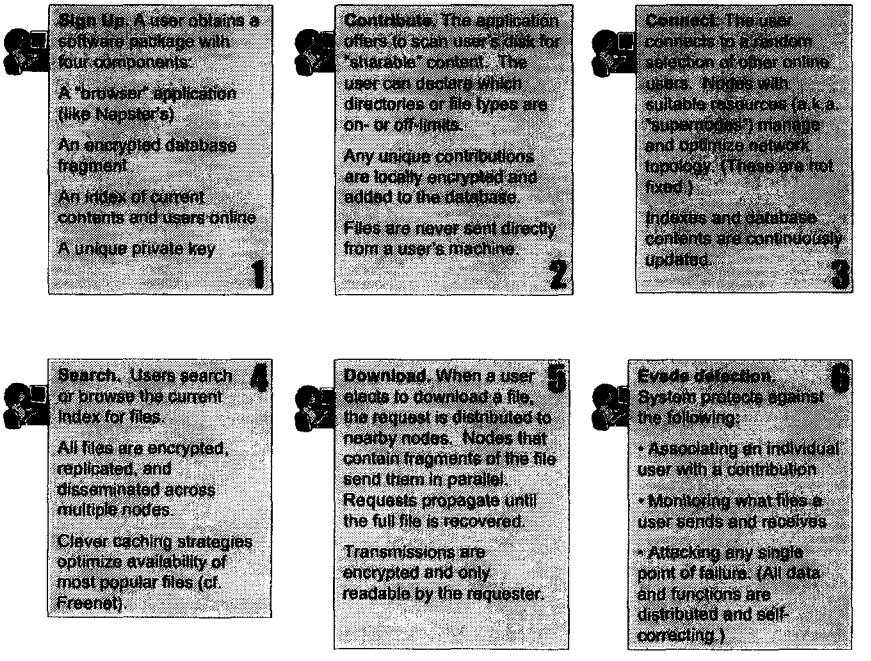


Figure 1 – Anatomy of a CDD

Such architectures suggest that the ultimate threat to copyright owners comes not from rogue companies like Napster and Scour (who can demonstrably be dealt with through legal measures), but from a new kind of "doomsday weapon" for the opponents of IP (or those for whom the undermining of copyright is seen as "collateral damage" in a greater struggle).Let's examine some of the implications of this threat.

- **CDDs present no target of opportunity.** Because such a service is decentralized, assuming there were a clear legal basis from which to target it, there is unlikely to be any controlling entity to hold accountable. This situation exists with Gnutella[16] today, whose network is not under any single entity's control. Going after the creators of such services does little good: first, these systems are generally being developed using Open Source collaboration,[17] so hundreds of anonymous developers worldwide may contribute to the final product; second, even if these alleged perpetrators could all be identified and prosecuted,they would not be able to stop the service once it was unleashed, as it would be running on the individual computers of its users.
- **CDDs have a non-infringing use.** Another recourse IP organizations might take on discovering an infringing CDD would be to do as they have done with Gnutella, and issue DMCA compliance letters to the ISPs of participating users (assuming these users could be identified, which in itself is doubtful). Legally, however, because their content is encrypted, CDDs have a substantial non-infringing objective: to protect freedom of speech from censorship. They are thus presumably entitled to First Amendment protection, offering ISPs (which could make a safe-harbor claim under the rules of DMCA) a strong disincentive to go after their subscribers. Since it is clearly not in the interest of IP organizations to see copyright protection portrayed as indistinguishable from free speech in a hypothetical court case, the remaining option would be to directly go after individuals who use it for infringement. Putting aside the notion that this could be a great many individual across a great many nations, they may be simply unidentifiable.
- **CDDs prevent the identification of their content sources.**
By design, and because content is disseminated, the origin of any particular file is cryptographically protected, and bits of the file are assembled from the storage of many different users, none of whom is even aware of what the data on

their systems might represent (since it's encrypted). In a hypothetical infiltration scenario, an enforcement agent, identifying a suspicious file through its search key, would create a direct peer-to-peer connection with an individual suspect who appears to be hosting the file, verify that he is actually offering copyrighted material illegally, and identify him (through his ISP[18]) by his IP address (which the peer-to-peer client must know if it is to make a direct connection; if the connection is intermediated by a proxy service that provides anonymity, presumably liability transfers to the proxy service in this scenario). In the case of disseminated data, however, the receiver of content is dealing with a cloud, rather than an individual, when acquiring a file. The cloud's actual physical contents are spread across individual machines all over the world, and the actual source of the content is untraceable, leading back to the problem of attacking the service itself.[19]

- **CDDs can be built from existing content and infrastructure.** The final recourse may be to protect all new content, storage, and consumption devices (as we will explore shortly). Nonetheless, the dependence on insecure legacy distribution formats and hardware (CDs, DVDs, PCs, etc.) makes a clean transition to a new environment unlikely, as SDMI (Secure Digital Music Initiative, a consortium formed in 1999 to create secure standards for music distribution) has demonstrated.[20]

As of this writing, CDDs still exist predominantly as a theoretical threat.Many systems and proposals have appeared that have some of the characteristics of a CDD (see Appendix B): Freenet, Publius, Aimster, Filetopia, Morpheus, WinMX, DirectConnect, Bearshare and many others. But all of these have yet to deliver a full-fledged CDD implementation.

The three principal weaknesses we see today are:

- **Lack of true dissemination.** Of the services mentioned above, only Freenet and Publius (which is conceived as a document publishing system) truly disseminate their contents into an amorphous cloud distinct from its users. The remainder require a direct peer-to-peer connection, which might be infiltrated.[21] As of this writing,the Freenet community has yet to release a usable Windows client and demonstrate its real-world scalability.

- **Lack of true decentralization.** Services like Aimster require corporate entities to operate centralized servers (of either content or users). These entities are relatively clear targets for enforcement (although,as mentioned in the previous section, each will present a new set of legal challenges).
- **Lack of scalable design.** As Gnutella and Freenet have each shown in their own ways, designing an effective, self-organizing distributed system with acceptable performance and usability characteristics is not a trivial matter, particularly in the absence of a funded corporate infrastructure. Although these are the conditions under which many peer-to-peer file trading systems have been developed thus far, it is not clear that such development environments will necessarily produce or widely distribute an effective CDD.[22]

This final point is not very comforting. If Napster, Gnutella, MP3, deCSS, and divx have taught us anything, it is to respect the Internet's viral effect: Small bits of code that emerge from obscurity even for a brief period[23] can spread almost instantly and have tremendous network effects under the right conditions. Demand for content over the Internet is a powerful force; legal and tactical measures might only be able to go so far in curbing it.

How serious a threat are CDDs? The answer depends on two factors. The first is an understanding of the current environment for file trading: its magnitude, trends, supply chain,and migratory abilities. The second is an understanding of what tactics may be available to legitimate sources to co-opt such a movement before it becomes widespread and out-of-control. We will discuss legitimizing countermeasures in section 5, but for now, three questions about CDDs to ponder are:

- How might files get deleted from a CDD?
- How do conflicting claims to keys (asset names) get resolved?
- How do users know if a piece of content is trustworthy?

4 Guerrilla Insurgence

What is the current state of affairs in file trading? This data is not easy to come by for a variety of reasons. First, there are multiple channels to consider, many of which are either private or require active participation to gain entry and thus can't be monitored (e.g. private ftp sites, virtual private networks (VPNs), invitation-only chat

rooms, campus and corporate LANs, certain peer-to-peer networks, etc.). Second, of the channels that can be monitored, the data revealed varies considerably. And third, much of the data is self-reported and unverifiable.

Nonetheless, we believe that some systematic observations can give some insight into what's going on. (See Appendix A for more details on the methodology employed in arriving at the following figures.)

The Internet is Not the Web

It still seems as though the press and the public are often confused about the distinction between the Web and the Internet, and the two terms are frequently used interchangeably. The Web – and the network of Web sites of which it's composed – is but one of the Internet's many channels. Each of these channels is defined by a protocol, or set of communication standards, by which content is openly exchanged. The Web is the set of content destinations that can be viewed with a browser that sends and receives content (generally HTML or XML) in HTTP[24] protocol. Other popular Internet channels include email, newsgroups, chat, instant messaging, and FTP (File Transfer Protocol).

Any of these channels is quite capable of file transfer. Some are more conducive to it than others, but all are used to some extent to send and receive media files. Napster and its various clones are Internet applications (although they generally have affiliated Web sites where the apps can be downloaded; see Appendix B). They use various protocols (including HTTP) to transmit search requests and results and establish file transfer connections.

Further details here are irrelevant, except to note the one common factor that ties all of these protocols together: they all share a common underlying protocol, which is IP (Internet Protocol,not to be confused with Intellectual Property). This is important because it establishes the universal notion of an IP address, which requires registration to be part of the Internet. Thus anyone operating on the Internet can – in theory – be traced by his or her IP address, setting a limit to absolute anonymity for legal purposes.[25]

There are various distinctions among file trading applications in both the area of how content is discovered (i.e. searched or browsed), as well as in how it is transferred. We can divide trading applications into seven categories:

File Trading Architectures

Type	Examples	Description	Copyright Enforcement Potential
1. Central Server	FTP sites, Web sites	A central server or server cluster with a fixed IP address.	Relatively simple once located (although requires ISP to comply with DMCA, which could create jurisdictional issues outside U.S.). If private (i.e. password protected with no public invitation), may be difficult, but threat is correspondingly less widespread.
2. Web + Net Application	iMesh, Scour	A central server with a custom Web application that mediates transfers among multiple clients.	Relatively simple. iMesh and Scour (which has recently emerged from bankruptcy following its failure to defend a lawsuit by the MPAA) have installed filters and are apparently attempting to comply with IP restrictions.
3. Multiple Server	Hotline, Usenet	A distributed collection of servers mirroring the same or similar content.Users connect to random servers and trade with them directly (Usenet), or with each other through them (Hotline).	Moderately difficult. "Cancelbots" exist which can automate the deletion of content. Services themselves are presumably legally protected by substantial non-infringing use. (Hotline, for example, was originally established as a backup service for user data.And Usenet has substantial non-infringing history and content.)

4. Central Server + Peer-to-peer transfer	Napster, Aimster	A central server farm mediates connections; transfer is directly from peer-to-peer. Controlling authority provides legal target.	Legal standing still unclear. Although operating under injunction, Napster has yet to hear a judgment, and its attempts at compliance through filtering have exposed the complex issue of burden of notification.[26] Aimster presents even more difficulties, because it claims no central directory of assets, only an open "buddy list"that one can join which enables direct searching of all of the "buddies'" assets.[27] Aimster has also added encryption[28] and a carefully designed copyright infringement claim form[29] that highlights the burden of notification issue as further defenses.
5. Multiple Server + Peer-to-peer transfer	IRC, Napigator (OpenNap et. al.)	A decentralized collection of servers mediating peer-to-peer transfers (no controlling authority)	Moderately difficult. Napigator, which navigates among "OpenNap" (and other) Napster clones (many outside of the U.S.),continues to run in a diminished capacity as of this writing, following the departure of MusicCity from the network (more on this shortly). IRC itself has a long history of substantial non-infringing use and is almost certainly immune from shutdown, but individuals trading files on IRC (or their offshore proxy intermediaries[30]) are exposed by IP. An automated interdiction process is feasible for IRC (assuming the proxy issue could be dealt with).

6. Pure Peer-to-peer	Gnutella (Bearshare, Limewire),Morpheus[31] (KaZaA)	All servers are clients; no independent operation of any kind. Searches proceed from node to node in a "tree"structure.	Difficult. The service itself can't be attacked as there is no controlling authority. The effectiveness of coercing ISPs to censure users of these applications without specific evidence remains to be seen. Ultimately, however, traders must know each other's IP address (or proxy) to exchange files.
7. CDD	Freenet	See section 3	Almost impossible. See next section.

Of these seven, we will focus primarily on numbers 4, 5, and 6 in this section. Numbers 1 and 2 either do not show significant traffic (Web, Scour) or are not visible (ftp, private Web site). iMesh,however, has been considered in the study.

Underground Markets

To illustrate how things work, we will focus on the area of movies and television shows. Most readers are probably familiar with Napster, but music is somewhat different, as its acquisition through CD-ripping has been popularized, while high quality video assets are still relatively difficult to convert into a tradeable form. Thus a more evolved supply chain has developed.

We will begin our brief tour of the file trading underground on IRC.While not the most widely known of services (its population is restricted to fairly hardcore traders), like a wholesale market its customers tend to engage in higher per-user trading volumes than other services, and its selection features more up-to-date content (e.g. current run movies and television shows).

We are seated at a home PC with a cable modem connection. We have loaded "miRC " (a popular IRC shareware client for Windows) and managed to configure it to connect to an IRC network, which we choose at random. The software presents us with some popular chat rooms to join, but we'll have none of that. We clear the box of offerings and click on "search channels." Here, we can search for keywords in channel descriptions. We type the word "movies." Instantly a box appears and

fills with hundreds of names and cryptic (often lurid) descriptions of channels we might join.

We choose one of the more popular ones (they are conveniently sorted by numbers of current users, which are displayed next to the names) and a window opens on a garishly colored scroll of geeky-looking text that goes jerking up the screen.

```
<[vU]SocialDistortion> My server is up at [/ctcp [vU]SocialDistortion !Movies inc. Scary Movie & X-Men] [Serves: 0/5] [Sends: 2/2] [Queues:
  10/30] [Current speed: 28173 cps] [Top speed: 48182 cps by Brit] [Sent: 57.28Gb & 189 files] [People served: 5280] [Serving: The Patriot,
  The Perfect Storm, Rocky And Bullwinkle, Chicken Run Boys & G
<[vU]trixter> My server is up at [/ctcp [vU]trixter see tee see pee me] [Serves: 0/999] [Sends: 2/2] [Queues: 5/5] [Min CPS: 9.8 Kb/s]
  [Current speed: 53.7 Kb/s] [Top speed: 83.3 Kb/s by [vU]Omniserv[vU]] [Sent: 87.99Gb & 208 files] [People served: 2521] [Serving: i am a
  bot. DO NOT MSG ME!]  -( NVCDVault Script )-
<Freshness> can anyone resume [vU]scary.movie.tfe.fs.divx.avi at 52%?
<Alpha187> My serve is up at [/ctcp Alpha187 Warez To the Millions!] [Serves: 0/5] [Sends: 1/1] [Queues: 4/10] [Min CPS: 1000 cps] [Current
  speed: 40336 cps] [Top speed: 97126 cps by scoobydoo] [Sent: 36.40Gb & 118 files] [Serving: The Newest [vU] releases..]  -( NVCDVault Script
  )-
<[vU]Muffin|Away> My server is up at [/ctcp [vU]Muffin|Away !100% Pure Entertainment With Rating System] [Serves: 0/3] [Sends: 1/1] [Queues:
  15/25] [Min CPS: 7.8 Kb/s] [Current speed: 17.2 Kb/s] [Top speed: 0 b/s by n/a] [People served: 65] [Serving: [vU] Releases - X-Men -
  Rocky&Bullwinkle - ScaryMovie - ThePatriot - ThePerfectStorm - Many Many More
<Freshness> can anyone resume [vU]scary.movie.tfe.fs.divx.avi at 52%?
```

Figure 2 – An IRC Channel

Amidst the incomprehensible banter, we recognize announcements known as triggers scrolling up the screen. Servers (who are running what's known in the trade as "fserves") are announcing their wares like mongers at a fish market. Clients (who pay nothing) attempt to connect to these fserves directly by typing their triggers to browse their wares and initiate downloads (which, due to the popularity of such services, usually results in becoming "queued": put at the end of a long line of waiting customers). Transfers themselves are peer-to-peer (which, in IRC-speak, is known as "DCC" for "Direct Client-to-Client").

Often we'll encounter what's known as a "net split." When this happens, everyone on a segment of the network is disconnected and has to begin again. Fortunately, resuming a file transfer can be configured to restart from where it left off. Unfortunately, all queue spots are forgotten and must be re-established.

We'll often get kicked off or banned from a channel for some unknown infraction.As we develop our skills we may learn to use "bots" (automated scripts) to streamline some of our activities. Or, more likely, we will get tired of this and seek an easier way.

Magnitude of Traffic

IRC is not for the time-impoverished. It is fiercely non-intuitive, adolescent, seamy, and cultish. Yet we estimate at least 87,000 feature length movies a day are traded through its channels,[33] which represents about 22% of the estimated 300,000 - 500,000 features movies traded daily on the Internet.

It is possible to estimate this because the servers proudly announce their traffic statistics. Observing a single movie piracy channel continuously for a week, for example, recently revealed the following activity:

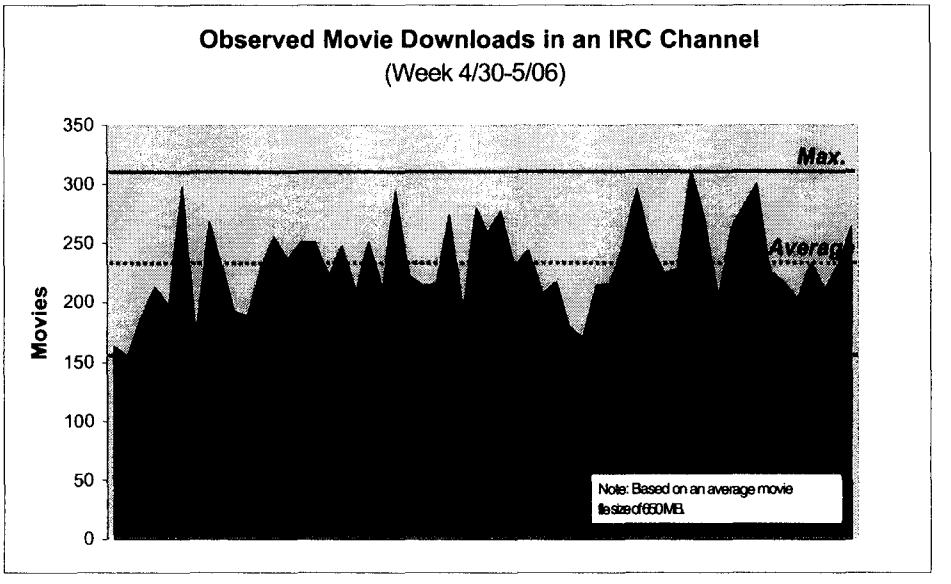


Figure 3 – IRC Movie Piracy Channel

There are approximately 1500 channels on IRC devoted to movie piracy,[34] (the one shown above is among the more popular.) They vary in simultaneous users (SU) from about 500 on the most popular to fewer than 50 on the lowest tier. As mentioned, at any given time most of the servers on IRC have full queues. Thus, a gating factor for the amount of traffic is the ratio of servers to users on a given channel.Another indicator of channel efficiency is the ratio of observed SU to total download traffic (derived by tallying the reported serving bandwidth for all active servers on a channel). These factors are plotted below, for a variety of channel sizes.

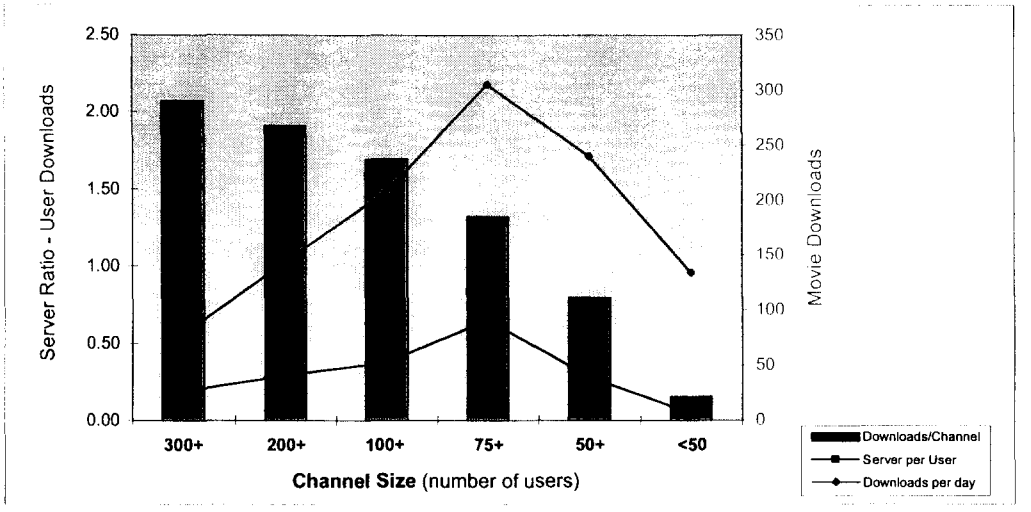


Figure 4 – IRC Channel Profile

These channels play host to an audience of perhaps 70,000 total SU, of which about 17,000 are running fserves. An average fserve delivers just about 5 movies a day, but this can vary widely, with the fastest among them delivering perhaps 1,000 a day (on a T3 connection),and the slowest (or most unpopular, or possibly misconfigured) delivering no significant content at all. (See Appendix A for more details.) As suggested earlier, the IP addresses of these servers are available (by typing "whois" followed by the user's "nick," or nickname, as it appears on the screen); however, increasingly these are being disguised through the use of offshore proxies.

Where Do They Get the goods?

As mentioned, IRC represents a kind of a wholesale market for content.Assets often appear on IRC before they arrive on more intuitive "retail" channels like WinMX.Assets move into other channels primarily through the shared directories of users of multiple systems.

The following table summarizes the most popular movie formats, and their sources. The code letters used are usually included as part of the filename, so that interested users have an indication of source and quality.

Pirate Movie Sources			
Source	**Type**	**Method**	**Quality**
Home Video	dvd – DVD rip	Most likely a deCSS[35] base utility or set of tools was used to extract the movie from the DVD and re-encode to avi or divx.	High (in some cases difficult to distinguish from original when seen on TV).
Post-production	sc – screener	Sometimes sourced from a trade/guild pre-release copy of the movie or a leak in post-production.Often available pre-theatrical release.	Usually high, but may differ from theatrical release (missing scenes, rough cuts and transitions, different edits, etc.).
Theater	ts – telesync	Recorded in the projection room of a theatre directly from the reel as it plays; audio taken directly from projector.	Medium to high. Often include Asian subtitles (Chinese, Thai).
Theater	cam – screen cam	Video camera pointed at the screen in a theater.	Low to Very Low.
Home Video, Broadcast, Station Feeds	VHS rip, TV rip	Digital recording of an ana-log VHS or broadcast signal. Often represents recordings of popular television shows off the air (or intercepted from wild satellite feeds[36]).	Medium to Low. Scan lines often prominent. Image crawl and other video artifacts common.

As indicated, users can discern the asset type by looking closely at the filename. Here are some examples:

- [vv]exit.wounds.hbo.cam.divx.avi
- [vv]almost.famous.teg.ts.sub.divx.avi
- [vv]the.Mexican.dui.cam.divx.avi

In the first example, Exit Wounds, [vv] is the identifier for the source (encoder) of the movie: not where it was originally recorded, but who encoded it. In this case it's identified to be encoded in divx.[37] The cam identifier defines the type of movie as described above. It is a camera copy. The origin of the movie is from hbo, not Home Box Office, but a ring of pirates who process camcorder tapes recorded in theaters.

In the second example, Almost Famous, you see sub, which indicates that the movie has subtitles, often in one or two Asian languages.

In the third example, The Mexican, you see dui, which could indicate another pirate group, but more likely refers to the unsteady hand of the camera operator – "driving under the influence."

As a general rule, when a movie first hits the silver screen, there will be various cam versions available – most of which are of poor quality. As different pirate groups manage to infiltrate the projection room of the theater, telesync (ts) versions begin to spring up and replace the shoddy cam versions of the movie. File sizes of the movies also go up as the quality improves. Telesync and Screener (sc) versions tend to be 25% - 40% larger in file size than cam versions, which average about 400 MB.

VCD (Video Compact Disc) is a popular format in Asia that never caught on in the U.S.,[38] and that, on pirate channels, is synonymous with pirated movies. When a movie is released on VCD in Asia,usually much earlier than on DVDs in the U.S., copies begin to appear, followed by DVD rips once the DVD is released. The VCD and DVD rips tend to be 650 - 700 MB; however, lengthy movies such as Gladiator are much larger, and are broken into two or more parts – one part for each VCD or DVD.

Despite the widespread availability of DeCSS and other hacker tools, creating a usable compressed file from a DVD or other source is a time-consuming and technical process. Thus, prior to hitting the distribution channels such as IRC, most movies are encoded by a host of independent groups that specialize in such activities. Many of these groups distribute solely on IRC and can be identified not only by their file naming conventions, but also by their often elaborate intro sequence inserts (often replacing the standard FBI warning) advertising and crediting their group for the encoding of the movie.

There is some crossover between the encoding/distribution realm and the actual sourcing of the movie. Some of the listed groups such as SmR claim full credit for the movie, while others, such as VCDVault,acknowledge the source groups such as hbo or mgm.

The prevalence of Cantonese and Thai subtitles, along with conversational evidence, suggests that Asia is the source for a large number of these movies. Users

and administrators of the channels on IRC will refer to FTP dumps or "distro" servers as the original source of the movie and often these have yet to be encoded (i.e. compressed into a distributable form). Since FTP dumps are generally private and require not only passwords, but also explicit knowledge of their existence, it is difficult to estimate the volume of traffic these servers receive. It is safe to assume, however, that the majority of available content arrives from organized upstream distribution channels of IRC and other file sharing apps, rather than end-users (as is the case with music, where CD ripping has become a consumer activity). This is rather puzzling since there is no obvious compensation model that would presumably be necessary to garner the interest of true black marketers in these channels. We speculate that there is a barter mentality that motivates such individuals to contribute so that they can in turn receive similar goods that they can convert to physical product to be vended on street corners. Access to these primary servers by IRC distributors (if acknowledged at all) may simply be regarded as a negligible toll of the hacker-supported networking service.

In any case, once a movie is encoded and available somewhere on IRC, fserve operators get access to private, high bandwidth distribution servers. This speeds the availability of the latest movies to the regular users who wait in queues when the movie is first released and is advertised only on a few nodes of a channel's server network. Since anyone with broadband access can become an fserve operator through a little "social engineering"(a hacker term for conversational manipulation, often using a false pretext), it doesn't take long before they are also granted access to the IRC distribution servers, where they can get movies earlier and more reliably than regular users.

Not surprisingly, many IRC users appear to be college students on campus networks, where they must overcome bandwidth restrictions set by the campus network administrators. Where restrictions might prevent a student from running an fserve, instead they might become a private node of distribution on their school LAN by setting up their own FTP dumpsite for a circle of friends. They arrange for other IRC users to deposit as many files as possible on their system so that by the time the campus administrators temporarily disable their Internet connection for exceeding their bandwidth quota, they are fully stocked. They continue to act as a

repository while someone else in the group takes a turn at running the dumpsite for newer content.

The following chart illustrates the supply chain:

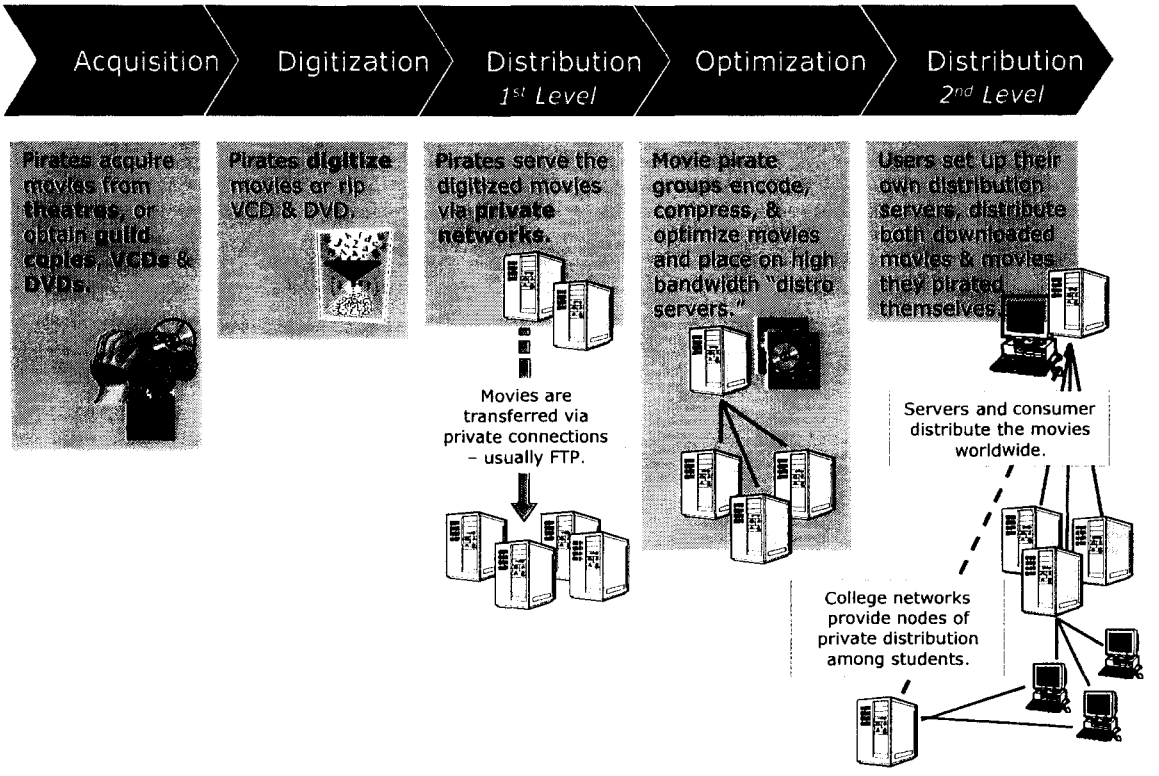


Figure 5 – Movie Piracy Supply Chain

Where Do They Go?

As mentioned, our estimate of the total number of daily movie downloads over the Internet falls in the range between 300,000 to 500,000 per day. We have arrived at this number by creating an estimation model that extrapolates our IRC observations to other systems that announce less data (see Appendix B) but, being more intuitive, are more widespread.

Since other systems do not supply a way of gauging what percentage of their traffic is in movies vs. television, music, software or other assets, we have used the overall percentage of broadband access[39] as a starting point (i.e. upper bound) for our estimation of percentage of movie downloads. We have also applied a discount to the efficiency factor (ratio of SU to downloads), based on anecdotal evidence that suggests large file downloads are less frequently successful on peer-to-peer applications than IRC.Our efficiency factor (that is, the percentage of bandwidth that can be productively used for downloading content) is about 40% for IRC and 20% for other channels.

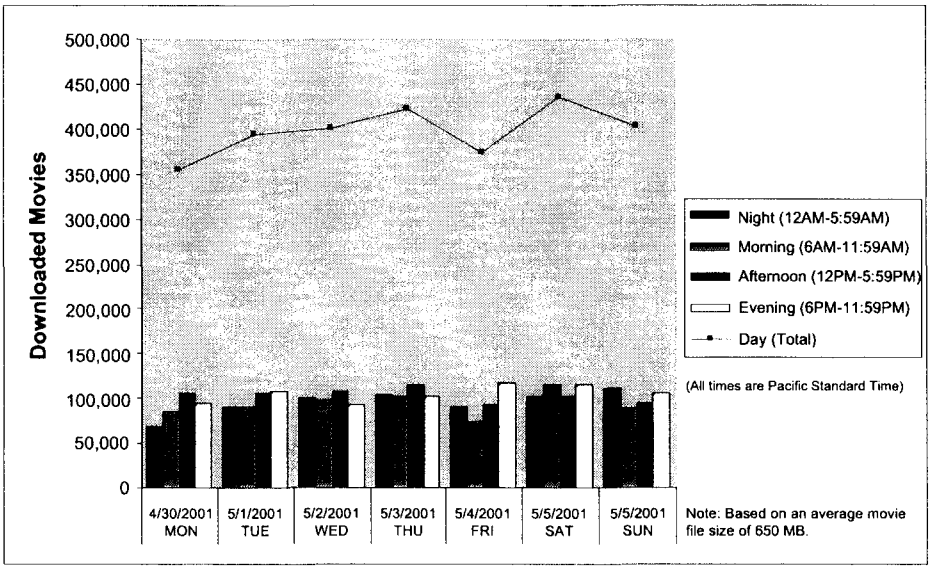


Figure 6 – Estimation Model Results for Movies

Another relevant area of data revealed by analyzing IRC logs is the composition of domains of the IP addresses of the participating users, which give some indication of geographical distribution.While this is not a perfect indicator (.com domains, for example, can originate from outside the U.S.), this does give a picture of an emerging global phenomenon.

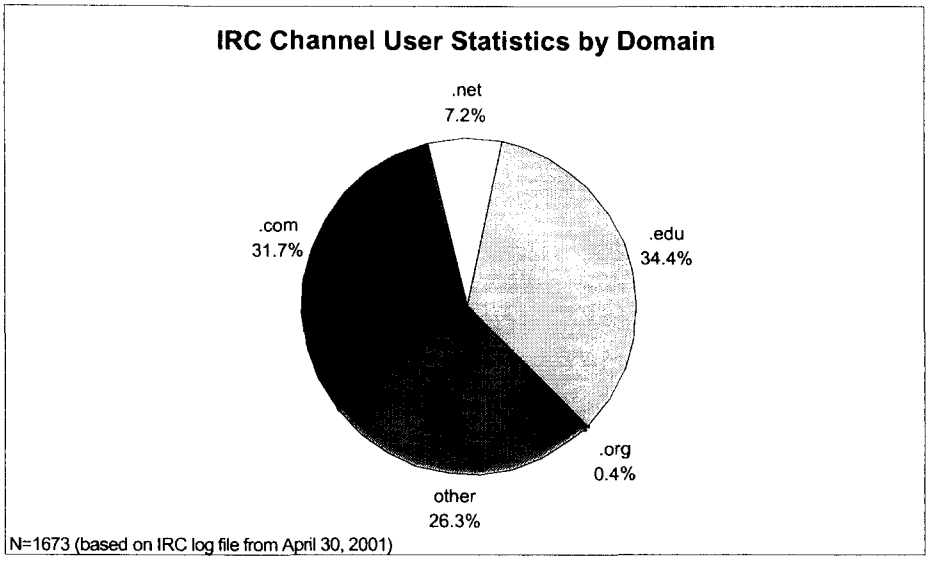


Figure 7 – IRC User Statistics by Domain

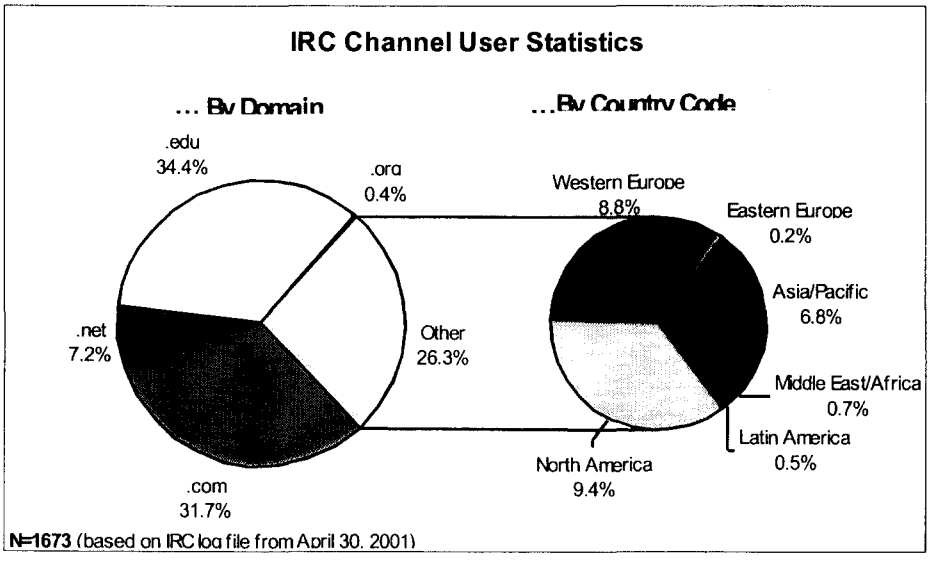


Figure 8 – IRC User Statistics for International Domain Suffixes

Much of the criticism leveled at the RIAA over their treatment of Napster had to do with the perception that a tremendous audience was squandered.At its peak, Napster had 64 million users, more than any other Internet service with the excep-

tion of giants like AOL. Moreover, they were all known to be music fans. In the words of John Taplin of Intertainer, "They had an audience! That's the hard part!"

Be that as it may, after Napster installed its filters in March, usage dropped from over a million average concurrent users to under 200,000.As was widely predicted, many of these users migrated to unregulated "OpenNap" servers, access to which was granted by a program called Napigator. Napigator was soon subsumed into another client, called File Navigator 2.0, which in turn was superceded by File Share 2.0. Both of these programs had the added effect of removing Napster's restriction of only supporting the trading of MP3 files, opening the OpenNap network to movies, television, and computer software as well.

Case Study: Morpheus

Behind much of this migration was a company called MusicCity.com of Nashville, TN. MusicCity operated over 30 OpenNap servers out of its domain.

Then, on April 22nd, MusicCity shut down its OpenNap services. Users attempting to connect to its servers instead received a pop-up message asking them to download a new MusicCity application called Morpheus (presumably named after Laurence Fishburne's master hacker character in The Matrix). The application instantly created a new, proprietary peer-to-peer network.About 27,000 users connected on the day it launched.

Morpheus (at the time of this writing) is a licensed version of a program called KaZaA Media Desktop, created by Amsterdam-based FastTrack[40] – although MusicCity has announced plans to release a new client application in July.[41] Morpheus/KaZaA is of the pure peer-to-peer class of architectures, and addresses some of Gnutella's performance issues[42] by implementing a system of "supernodes"[43] that are run by individual users with superior resources (e.g. bandwidth). MusicCity also claims to be "copyright friendly" by offering content providers the capability to deploy "third-party digital rights management schemes" to protect the copyrights of their digital content distributed through the MusicCity network. The exact nature of this offer is unclear as of this writing, however.[44]

It is also notable that the Morpheus/KaZaA client user interface has a number of bugs and usability problems that have made it the subject of some pointed criticism by the user community (see alt.music.* newsgroups). It is also restricted to the Windows platform (whereas Linux is considered to be the system of choice for hackers, and Napster supports Macintosh™ systems as well).

Nonetheless, as the OpenNap network choked under the sudden removal of its main source of service, the usage of Morpheus grew rapidly, as shown by the table below. It continues to grow as we go to print, having recently surpassed Napster itself in number of simultaneous users.

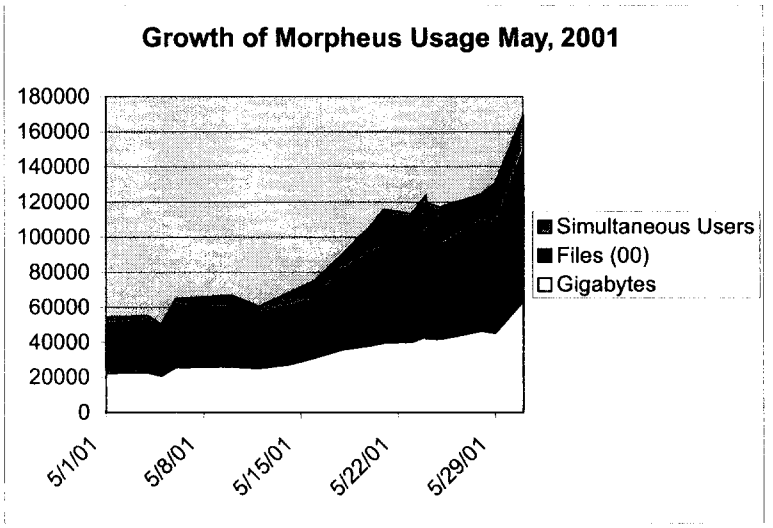


Figure 9 – Morpheus Rising

This case illustrates how, despite user interface woes and a hostile legal climate, a new peer-to-peer service can still double its usage in the span of a month and emerge from nothing to become a hit. Not only does this illustrate that switching costs for this audience are low, but also that there remain ample opportunities for established brands to launch winning applications, if they can deliver the right experiences.

As we look at the tactics of legitimacy, there is one additional lesson here to be considered.

Open access to peer-to-peer systems is a level playing field. Content owners can put their protected product in these channels and even establish new and better channels just as easily as pirates. The trick is in the subtlety and creativity of the approach. To understand what might be possible, we must first review the tactics of secure packaging and Digital Rights Management (DRM).

5 The Cryptographer's Arrows

Many have compared the war on piracy to the war on drugs, in the sense that a policy of pure enforcement is unlikely to be successful.[45] However, to the extent that some measure of active enforcement must be part of any solution that preserves notion of copyright enforceability, data security is a key battleground. Here there are considerable arsenals of weaponry arrayed on both sides.

On the side of IP organizations is a detailed (but sometimes confusing) array of DRM and copy protection schemes for data. There is considerable controversy over all aspects of these: their ultimate security, usability, practicality, and even in some cases desirability. However, DRM remains a cornerstone of secure digital distribution, so its implications need to be examined (although again the subject is very deep and we are only touching the surface).[46]

On the other side we have already seen something of the potential of CDDs, but how secure could they really be? If content protection schemes can ultimately be cracked (as is the knee-jerk claim of any hacker, with some substantial historical support), why can't encrypted databases and file transmissions also be cracked? Or, why can't offending files simply be deleted, filtered, or overwritten? Or, might the channels themselves simply be filled with protected content, so that pirated wares are the rare exception? We will look at these questions, too, in some detail.

The Aim of Digital Rights Management

Real-world DRM applications are still in their infancy. In theory, DRM would allow content owners to embrace the widespread distribution infrastructures we have been studying with protected content that could be licensed on demand. If this were possible, distribution costs would fall dramatically, with most of the bandwidth and storage paid for by consumers, potentially freeing more resources for develop-

ment and promotion while driving broadband adoption for ISPs. In practice, only a handful of experiments have been done with this,[47] and there is considerable doubt about its ultimate appropriateness, even on the part of DRM designers![48]

To better understand the challenges and opportunities presented by DRM, we will review its security and design implications.

First You Need a License

Core to the concept of DRM is the notion of a digital license. In practical terms, a license is a data record that defines specific usage rights for a protected content file or group of files granted to an individual ("licensee") on a transactional basis. It is part of a trio of software components that include an encrypted content file, a DRM-enabled player, and the license itself.

DRM on the Internet allows a copyrighted file to be freely distributed in encrypted form such that anyone can possess or copy it, but in order to use it (i.e., play it), the prospective user must acquire a license. Without the license, the file appears to be a random jumble of bits. The license itself cannot (theoretically) be copied or transferred in any way – it is somehow bound to the user's system or personal identity. The license also contains various restrictions on usage. These can get extremely arcane, as they intend to reflect the minutiae of contract language associated with content, but they generally fall into three categories:[49]

- **Render rights.** Restrictions as to where content can be viewed, printed, or played, as well as limits on time or numbers of plays before expiration (i.e. rental or pay-per-play) models. Render rights can also be used to enforce geographical and channel-based release windows.
- **Transport rights.** Limit on the number of copies that can be made (i.e. one for backup purposes), or a limit on how the file might be transferred (i.e. moved) to another device (such as a portable music player or CD burner), or loaned (i.e. removed to another user and returned later).
- **Derivative works rights.** These include restrictions on content that may be extracted, edited, or embedded in another piece of content.

Designing Experiences with DRM

The most common objection to the use of DRM,particularly in B2C scenarios,is the complexity of the user experience.This is especially relevant in the case of musical assets,which tend to be single song files,for which much extra overhead (e.g. authorizing a separate credit card transaction for each song downloaded) is unlikely to be tolerated by users (especially in light of free alternatives). Movies offerings are likely to present a more acceptable scenario,since consumers are already used to the overhead required to rent a movie either at a video store or on a pay-per-view television service.

To understand this objection, we will consider a typical[50] DRM user experience scenario, as compared to a typical peer-to-peer file copy.

DRM-Enabled Web-based Content	Peer-to-peer (Napster, etc.)
Step 1: Locate content (~3 minutes)	**Step 1:** Locate content (~1 minute)
• Search for content by title, artist, or keyword	• Search for content by title, artist, or keyword
• Add to shopping cart	
• Check out and authorize purchase with credit card	
Step 2: Download content (typically ~5k cps analog; ~13 minutes for 4mb MP3 file: ~80k cps DSL/cable modem; ~135 minutes for a 650mb MPEG4 movie file)	**Step 2:** Download content
• Initiate file transfer	• Initiate file transfer
• Locate file on computer	(Media management built into app)
• Click on content file[51]	• Click on content file
Step 3: Obtain license (~2 minutes)	(Media player typically built into app)
• Get directed to license provider's registration site	[Note: Many unprotected formats support viewing or listening as file downloads (like streaming)]
• Enter more data (name, email,password, credit card)	[Additional note: Unprotected files can also be enjoyed on a portable MP3 player (in the case of music) or burned on a CD in VCD format for TV viewing (in the case of video)]
• Download license	
• Click again to play file	
Total overhead (excluding download): 5 minutes	**Total overhead (excluding download): 1 minute**

Although DRM in this form does not seem to improve on the convenience of the experience, for large assets such as movies or video serials it may not be too much to ask,especially if, instead of comparing with a Napster-like application, we use a platform like IRC as our benchmark.

Moreover, this is not the only way that DRM can be used effectively. We will explore some more creative alternatives in the next section.

Why Use DRM?

As much trouble as DRM may be for consumers, it is probably worse for content owners. This is because, to be used effectively, DRM really has to be incorporated as part of an overall modularized media production and distribution process.

In many cases, this means creating metadata standards and workflows among media asset and content management systems that may exist in disparate parts of a media organization, or even across different organizations.

Although we have started from consumer-centric view of it, DRM's role is at least as important in license syndication and intermediated distribution scenarios. Use of DRM in a third-party licensing scenario will allow content owners to implement standard licensing arrangements to online intermediaries without having to own or manage all of the channels through which this packaged content might flow (although Direct-to-Consumer is sure to provide an important distribution model).

This has an important benefit of addressing two key issues of consumer choice: first, the issue that consumers are generally not accustomed to navigating their content choices based on the origin (label, studio) of the content they seek. If channels are unable to easily distribute all – or at least most – of the content their audience might be seeking (regardless of origin) because of license restrictions, then users have a continued incentive to seek alternatives that offer more complete selections.[52] Second, the design of today's peer-to-peer applications has erased the traditional distinctions that separated destinations by media type and put all media in one place. Thus it is natural for these consumers to think of a single application for managing all of their media – their own personal media libraries, if you will –

that organizes content both owned and sought, whether music, movies, television, games, books, photographs or anything else for that matter. As branded service channels develop to serve the cross-media choices of various communities with recommendation engines and custom programming, they will become trusted custodians of personalization data and marketing access, and they will be natural targets for the distribution of protected media packages whose DRM-encoded terms will make flexible licensing arrangements possible.

While content creators and their conglomerates – as well as new start-ups – will continue to develop and operate such channels, for these to be successful, a system of universal open licensing across media types for all protected content would serve to eliminate a key advantage of rogue services: the advantage of ubiquitous choice offered by the aggregation of all the users' personal collections.

Considerable work has already begun in regard to standardized DRM metadata languages on the part of organizations like MPEG, SMPTE, ISO, and OMG.[53] The success of these efforts will go a long way to creating a favorable environment for protected content.

New Distribution Fortresses

Returning to the security of DRM,one important feature of any DRM scheme is that content is packaged in a secure container.[54] Secure containers are designed to keep content encrypted at all times except when it is within a so-called "secure perimeter."

This notion raises a serious difficulty when one is considering DRM's potential application to today's home PCs. The trouble is, no matter how secure the encryption itself may be, once the content is consumed on an insecure hardware platform, it can no longer truly be considered secure. Even proposed CD and DVD schemes that allegedly prevent these media from being ripped to a PC (while still maintaining backward compatibility)[55] must still be capable of being recorded in analog, after which compression is likely to obscure any detectable difference in capture formats. And content from secure files must at some point appear unencrypted in the buffer of a video or audio card, where a dedicated hacker might grab it.

These limitations have led to many proposals, including most notably the CPRM proposal to advance copy protection to the hardware level by implementing content recognition standards in hard drives and other recordable media.

This initiative has raised considerable objections among consumer advocates,[56] and is likely to create consumer backlash against such products, as they are likely to be perceived as restricting activities in the gray areas of "fair use" as the interaction of protected hardware with unprotected software from any number of sources (some of which may be legitimate) creates confusion and frustration. Moreover, it makes the objective of co-opting peer distribution for legitimate use a far more complex goal to achieve.

In fact, perhaps DRM's most effective use will be to educate people as to what their legal rights and restrictions actually are so that they can voluntarily comply with the law if they are so inclined.In any case,when it comes to copy protection,it seems more practical to consider DRM a deterrent rather than a prohibitive technology.

Moreover, DRM has the greatest odds of being effective if it is introduced subtly, so as not to present an obvious target of opportunity for hackers (who you know will crack it), but rather to convince non-partisan consumers who are uninterested in conflict or complexity that legitimate content is desirable and easy, rather than challenging and restrictive. What's needed is a "sneak attack."

Content Missiles

How might such a co-opt plan actually be accomplished? Isn't the process of attempting to extract payment for something that might be obtained for free bound to be a painful one?

The key lies in adopting a creative approach.DRM is perhaps more flexible than has been widely portrayed,or even conceived by its designers.Recall that DRM packages can contain any usage rules desired, or even no rules at all.The fundamental notion is that, when an unlicensed piece of content is opened in a media player, there is an opportunity to issue a license under any terms that can reasonably be designed. In particular,this provides a triggering event that can load a Web site and

thus be utilized in any number of ways beyond the simple extraction of payment or registration information.We will examine some of these in the next section.

DRM is a Content Management Discipline

Core to the idea of DRM is the wholesale embrace of digital content management techniques throughout all phases of the media production and distribution process.This is the infrastructure that allows the capture of key metadata that not only encapsulates rights restrictions, but all other information about a content element as well. The overall goal of such an effort is to realize the potential of modular media.

The goal of modular media is to maximize the latent value in any production process by supporting the most flexible possible range of distribution and consumption scenarios. In addition to allowing for more freedom of distribution, modular media has several other distinct advantages.

- **Modular media differs from a Web site in key respects.** The Web is a static, fragile, cross-linked structure of pages.The movement or removal of one page can cause errors in any number of pages that link to it. Even dynamic pages have static locations and contexts in which they can be experienced. In contrast, modular media is self-contained and mobile. It travels across the Internet as a file that anyone can send, receive, or open. This enables multi-level distribution, and maximum contextual adaptability, including cross-platform adaptability. In theory, the same content package can be sent to any Internet-enabled device and format itself appropriately whether it is opened on a PC, in a TV set-top box (STB), a personal video recorder (PVR), a kiosk, a mobile PDA or a even screen phone.This implies that great economies of scope for each piece of content can be achieved. Moreover, unlike a Web site that must service each request from a central server, modular media only checks back when necessary (e.g. to report data or request a new license).
- **Modular media can be streamed or downloaded.** Controversies over streaming vs. download distribution strategies are tangential to the adoption of modular media. These are questions of delivery infrastructure. Modular media and DRM can support usage rights for both modes of delivery.

• **Successful content management adopts comprehensive metadata standards upstream in content production and archival processes.** As mentioned, DRM relies on, and is to some degree a goal of, content management. Digital content management emphasizes the use of metadata to facilitate the storage, search and retrieval of all relevant information related to an asset or class of assets. The key is to ensure that metadata and assets co-exist in self-contained packages. As mentioned, this can create problems when assets are handled in disparate systems and environments. It is not always practical to maintain content in a single package, especially when considering derivative or related works and formats. In general, the solution to this is to separate the concepts of physical and logical packages A logical package can consist of any number of related physical or logical elements, in various locations and states. An effective content management system and process will always maintain referential integrity of object metadata across instances and versions of the underlying content. To accomplish this requires a single point of organizational authority for metadata design and standardization.

• **Producers might consider cross-platform rights management implications in content and experience design.** To coin a phrase, it's not the bits – it's the experiences. Sponsorship opportunities and tie-ins – enabled by self-contained license policies – trigger interoperability, and are a source of potential revenue.

• **Information collection and relationship "ownership" are also license considerations.** Ownership of the wealth of consumer data created through the consumption of media assets needs to be assigned, and can be a bargaining chip in license negotiations. They can also be enabled by the mechanics of DRM.

6 Architecting the Peace

In the previous section we saw that DRM-encapsulated assets,distributed through peer-to-peer networks as well as Web sites and discs,represent new opportunities to monetize assets.But the question remains as to what will be necessary to gain acceptance of such practices among users who may be predisposed to reject such notions as downloading encrypted files. How might content owners create an environment in which consumers feel obliged to pay for experiences that are not easily replicated?

Experience Design

Consumption model design at some stage cannot be divorced from media type. Films, TV, music, and games all present diverse usage and economic models that must inform experience design.

The music industry is probably most evolved in its thinking about these matters, and has perhaps the most acute challenge. To address this challenge, the players are already forming distribution constellations and alliances, although few specifics have been revealed about what new experiences the next generation of digital music distribution services might deliver, other than that they seem to be embracing a tiered subscription model for digital downloads.

The film industry has also announced a number of Internet distribution initiatives, albeit fewer than music.

As we await such services, let's contemplate how some DRM packaging alternatives might work by examining some simple speculative scenarios. In particular, it is time to examine how a media file might engage users in a commercial way when they click to play.

- **Start with a handshake.** In the simplest case, a license might be issued at the point that a server receives notification that a certain content file has been opened on a certain user's machine. This is more information than was previously available (and the handshake might be quite transparent to the user). Moreover, licenses can maintain (server-side) records of usage by individual – to allow, say, the first three movie downloads (from an affiliated service) to be free. This is perhaps a loss-leader type of approach.
- **Can we talk?** More interestingly, a licensing event represents an opportunity to present a service's privacy policy to a user and force them to acknowledge it. This opens the door to the Internet model that allows consumers to trade personal data for permission to use content. A simple question might be the next step (name, age, gender) in the personalization process. Ultimately, real addresses may be garnered through promotional give-aways of appropriate content-related merchandise.
- **I know what you need.** Targeted rich-media direct-response ads and

promotions are the next step. A license can be issued contingent on viewing an advertisement or group of segments, or a user can be given a choice between watching the ad or paying for the license outright.Ads can also create "droplet"-type applications containing a coupon or other promotional offer, which sits on a user's desktop until exercised, expired, or deleted. If the user is online, a new (and sufficiently bandwidth-friendly) ad can be streamed or downloaded for future viewing each time the content is opened. Moreover, the selection of ads can be highly targeted based on information accumulated in prior licensing exchanges (as well as the nature of the current content), or by self-selected interest, raising the theoretical CPM rates for such opportunities.

- **Extras.** To further improve the experience, other content-related promotional materials can be included in the package such as browser skins, screen savers, etc. These extras, like the extras in a DVD, can be made difficult to include in counterfeit packages.
- **Getting to the next level.** Finally, up-selling consumers to subscriptions, event tickets, fixed media versions of content (which at some point are likely to become collectors items) or premium service packages offered by their ISPs (who might be affiliates in this model) is also an opportunity afforded by this type of licensing process.

The list goes on.While none of these ideas is particularly new, little has been said about their applicability outside of a Web site, particularly in the form of a modular media file, which, as we've seen, has many advantages over a Web site attempting the same things.

A package of desirable content is perhaps a more suitable location for some of these notions than a Web site, as we return to the notion that the Web (as opposed to the Internet) may be best leveraged to provide informational and transactional support to content services that themselves can be decoupled from any specific site. Being site-independent, these might gain widespread viral distribution through any number of channels, including those now considered illegitimate.

Controlling Namespaces

Now that we've seen, at a high level, how DRM might help to co-opt and legitimize content channels in addition to simply protecting content, let's return to the

subject of CDDs. It should be clear that creators of protected files could benefit from the distribution capabilities of a well-established CDD. There is no more security risk here than with any other form of distribution (unless one conceives a threat from a greater prevalence of hackers using CDDs), and a content owner need not care how a file arrives as long as the package has not been tampered with.

Assuming a set of rules and experiences could be designed that would be attractive to users, the question becomes, how can legitimate files compete with illegitimate copies in an unregulated database?

Recall that a CDD differs from a "standard" peer-to-peer file trading architecture in that, instead of "peering" into the aggregated, poorly kept file stores of countless semi-anonymous individuals, the CDD structure unifies the database so that, in essence, only one asset per name or identifier exists (although it may be replicated and distributed across thousands of nodes). When a user (passively or actively) adds a file to the pool, if the filename (or key) he is using already exists, the file cannot be added under that name. (If it is identical to the one in the pool it need not be added.)

As noted, this situation begs certain questions. First, how would files ever get deleted from such a system? Unlike a peer-to-peer system, which can filter or delete keys, there is no clear way to delete a key or a file from a CDD. In fact, in Freenet's implementation, deliberate deletion is intentionally made impossible – or as impossible as possible! Instead, deletion occurs when a file falls into disuse for a determined period of time.

Other design options, of course, are possible. There could be a special key that gives super-users deletion privileges. But this notion is in conflict with the intent of CDDs, and, moreover, it would trivialize copyright enforcement.

Note that the "no-deletion" feature also implies that files cannot be overwritten (except by the original author, with implications we will see presently). This scenario begs the second question: how are competing claims to a given filename resolved?

This is a familiar problem.It was exactly the problem the Internet engineers had when they designed the system of URL (Uniform Resource Locator) names known as DNS (Data Name Services). DNS has been an area of controversy since it was created, and is no less so today as ICANN[57] struggles (under much criticism) to balance a variety of conflicting agendas.[58]

There are generally only two possible answers to this problem: either names are claimed on a first-come-first-served basis, or there is some controlling authority that arbitrates and is presumably accountable for trademark restrictions, infringements, etc. The second case makes some sort of enforcement feasible (although, as Napster's name games have revealed,claiming a name doesn't necessarily stop the traffic in the underlying asset), although the first scenario seems much more consistent with CDD philosophy, and hence more likely to be implemented.

If they are lucky, content owners might simply claim their names first. Since they already have trademark management processes in place for their titles, adding such a step would not add significant cost to the existing process, as name claiming could be a highly automated procedure across any number of CDD services. Moreover, although obscure new names may not be discovered, the contents of the associated file (if it existed) could contain any kind of teaser or viral marketing concepts.

This brings us to the third question, that of trust and authorship. In our earlier discussion of CDDs, we assumed that files would be added anonymously, with no way to trace the origin,and this is the case. However, as the designers of Freenet have recognized, branding is no less important in pirate environments than commercial ones. As we've seen, pirates build loyal followings by adding their marks to an asset (both its name and its contents). These brand statements (and the followings they create) appear to be at least as important as the security afforded by anonymity.

Thus,a system of pseudonym signatures is proposed.These will identify that similarly signed assets came from the same source,although the source's actual identity will be cryptographically protected.In this way,contributors will be able to build reputations for quality on which users will rely when searching or browsing or selecting content.

This final observation plays to content creation's greatest advantage: quality control. The ability to guarantee a quality experience is likely to more than justify some form of payment for most consumers, whose time is of value to them.

Conclusion: Peace on Demand

Many media conglomerates have made much of their recognition that broadband convergence has laid the foundation for a future of on-demand delivery of all media. AOL Time Warner, for example, has articulated their vision of a world where subscription services can offer unlimited access to any media, any time. The trouble, as we've seen, is that on-demand distribution of modular media has arrived before many of the companies producing the content are ready for it, creating an opportunity for hackers and start-ups to satisfy global demand where legitimate content is scarce, and traditional distribution practices (such as release windows) irrelevant.

As the IP organizations representing content industries attempt to stem the flow of free goods, we fear that without substituting competitive legitimate services, their actions may have the ironic effect of imposing a Darwinian force on free file trading systems, under which only the most legally and technologically resilient survive. We've seen how CDDs, for example, represent a new, virulent strain of such services.

These observations should reinforce the motivation for content companies to accelerate their plans to mobilize for digital delivery, rather than scale back in a wave of complacency.

We have also seen how embracing digital delivery is not merely a matter of designing new Web sites or piecemeal subscription services,but implementing unified content management processes that consider intellectual property rights management and standardization from end to end in the production and distribution units throughout media companies. Implementing such systems and processes may be challenging,as they require content companies to reconfigure and align many disparate,even siloed, processes. Nonetheless, we believe that these kinds of cross-business initiatives,combined with bringing creativity and vision to the details of online consumer experiences on the Web and beyond,present a far better alternative to betting the future on simply escalating a struggle that may never be won.

Appendix A: Notes on Traffic Estimation Methodology

Although the subject is notorious, data on the traffic of copyrighted files through peer-to-peer networks on the Internet are surprisingly scarce.

An optimal wide-scale investigative framework would aggregate observations from a large number of geographically and topologically disperse monitoring posts, including high-traffic routers on the Internet backbone, last-mile broadband service provider facilities, and college campuses as well as representative Network Operation Centers (NOCs) in major countries where traffic is evident. Such an examination is well beyond the scope of this work, which considers only data from a small set of consumer end-user premises. Nonetheless, we believe that some systematic observations can produce more insight than is generally available to the public, and that certain observations are location independent and indicative of traffic conditions as a whole on various representative services. Thus we have undertaken to provide an analysis that is at least supported if not proven by our evidence.

Objectives of Methodology

In order to both quantify and qualify peer-to-peer file-sharing trends, we are specifically interested in the following metrics:

- Audience size and growth
  - o Total
  - o By service (Napster, IRC, etc.)
  - o By bandwidth/connection type (modem, DSL,Cable, T1,T3)
  - o By country/region
  - o By domain type (.edu,rr.com, athome.com, co.com)
- Supply size
  - o Black market (capturers, by method)
  - o Servers
  - o Capacity (simultaneous users)
  - o Capacity (simultaneous transfers)
  - o Upload/download ratios
- Traffic
  - o Raw (tbps)
  - o Titles/day

- o Titles/userday
- o By File type
- o For film: by window

- Usability (qualitative assessment)
  - o Time to locate (by asset class)
  - o Time to download (by bandwidth)
  - o Download reliability (chance of completion)
  - o Quality metrics
    - Accuracy of labeling
    - Image quality
    - Sound quality
    - Completeness
    - Graffiti (extraneous added content)
    - Search and browse capabilities
- Business models (where applicable)

Analysis Framework

The overall framework for this analysis is shown in Figure 10 below.

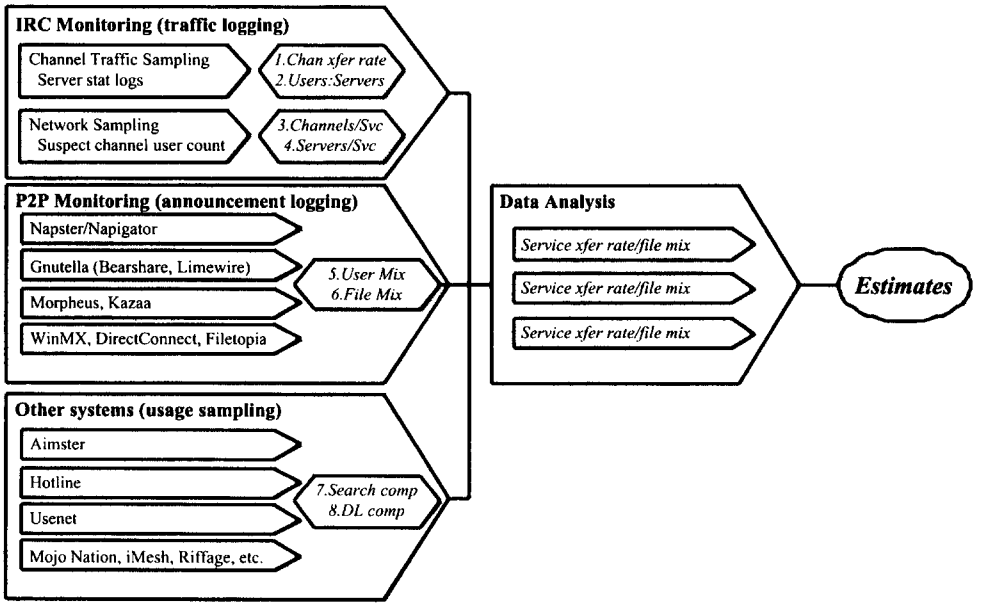


Figure 10 – Traffic Analysis Framework

As the above diagram shows, we have divided file-sharing services into three categories. The first is Internet Relay Chat (IRC), which for our purposes has two key benefits as a starting point: it presents a rich amount of traffic data, and it is easily scriptable to capture this data methodically.

The second is the more user-friendly (but in some ways less efficient) category of peer-to-peer applications typified by Napster and Gnutella and its spin-offs.

These systems, whose interfaces are generally similar to Napster's, all have some degree of traffic reporting, typically: users, files, and gigabytes of data online.

Using observations derived from IRC we can make some assumptions about traffic based on these figures which we regularly monitor. (The ratios among these figures, particularly users to files, also say something about the characteristics of a service, especially as related to the so-called "free rider problem" of peer-to-peer design.)

The final category contains services that, while perhaps among the most popular (i.e. Aimster), announce little if anything about their traffic. (Note that popularity can be inferred from both their own press announcements and from various sites and newsgroups that monitor, rate, and discuss peer-to-peer file sharing applications. Examples include Jupiter/MediaMetrix, Webnoize NEWS, p2ptracker.com, clip2.com, and many others. Although these sources do not generally provide enough statistics to draw any definite conclusions about traffic, we can draw some qualified conclusions about the scale of their contribution.)

Limitation of Approach

In addition to restricting the scale of monitoring to a small number of consumer broadband locales, it should be noted that we observed an additional restriction regarding our own participation in file trading activities.

A study undertaken by law enforcement or with support of various content owners may benefit from a methodology based on active participation in file trading activities. Such a methodology could not only accumulate more data more quickly,

but might also further the cause of infiltration into widespread non-public areas that otherwise remain hidden.

The methodology presented here, and pursued in the course of this study, is based only on passive observation.

IRC basis

In pursuing this analysis, we have begun with the most detailed and usable source of traffic information available, Internet Relay Chat.[59] IRC servers announce statistics on shared channels, many of which are publicly accessible. (See Figure 11.)

```
<[vV]BluDeViLz> My server is up at [/ctcp [vV]BluDeViLz
  Good vV releases] [Serves: 0/5] [Sends: 2/2] [Queues:
  4/10] [Min CPS: 9.8 Kb/s] [Current speed: 83.6 Kb/s]
  [Top speed: 84.0 Kb/s] [Sent: 62.53Gb & 344 Files]
  [Serving: All the newest releases...for a txt list of
  what vV and I have to offer, type: !MoviesList]  =(
```

Figure 11 – An IRC Server Announcement

In monitoring popular piracy channels it is possible to parse such announcements and accumulate figures into a database (and/or spreadsheet format). This forms a basis for gauging activity on other services that provide less information.

By capturing and analyzing IRC traffic logs (using customized IRC scripts running under MIRC, a popular IRC client for Microsoft Windows), we can obtain (from both server announcements and by logging into the servers themselves and requesting "stats") useful statistics for a channel.

Channel Transfer Rate

An important indicator of total traffic is a channel's total reported instantaneous transfer rate for all servers. We obtain this by multiplying each server's reported (average) bandwidth (in characters per second, or cps) by their reported number of simultaneous transfers and adding these figures together for all servers on a channel.

To translate the channel's transfer rate into an estimate of number of movies (or other assets), three additional figures are needed:

1. The channel's "file mix," which is obtained by sampling the "sends"(reported transfers) and breaking them down by asset type. (Even though a channel may be dedicated to movies, music, or "warez" [computer software], this does not preclude other uses, although in practice we observe such deviations to be small);
2. A target asset's average size (e.g. 650MB for a feature-length movie of "high" quality, 4MB for an average MP3 file);
3. The overall success rate of the transfers (in other words, percentage of abandoned downloads).

User to Server Ratio

Another statistic that is useful in gauging the general dynamics of self-organizing peer-to-peer file-sharing networks is the number of active file servers that can be inferred from a given size user base. Although by definition peer-to-peer clients have the capacity to serve and are encouraged to do so, unequal distribution of bandwidth and storage resources upsets this symmetry in practice. (Note that we must make a distinction in a peer-to-peer environment between a network server, such as a Napster or OpenNap server that provides search and directory services, and a content server that is operated by an end-user with suitable content, bandwidth, and storage.)

Here the size and consequent bandwidth distinction between MP3 and movie trading becomes important.Serving movies effectively requires a broadband connection. Moreover, most consumer broadband connections provide limited capabilities in this regard, as they tend to restrict upstream bandwidth to around 100kbps.[60] (To illustrate, a single institutional server operating on a T3 can potentially serve more content in an hour than a dedicated T1-based node might serve in a day, or a cable modem might serve in a week of continuous operation.)

Consequently, as services grow popular, server queues tend to grow quite long, so in order to gauge overall transfer rates it is useful to gain an understanding of the behavior of the ratio of servers and server bandwidths to users for various sizes of user base. (In short, we demonstrate that user populations will typically grow

much more quickly than server capacity on a peer-to-peer network, and that therefore bandwidth and file transfer rates have a logarithmic relationship to users.) In services like Morpheus that announce only numbers of users, files, and gigabytes online, knowing this approximation gives us an indication (along with file mix and other data) of what traffic may be assumed from such a mass.[61]

Channels per Service

Since each major IRC network consists of tens of thousands of channels, it is impractical (at least on the scale of this effort) to monitor them all. It is, however, possible to monitor their identities (keywords) and the average number of servers present (at least for those that publish).

Servers per Service

Aggregating channel server figures across networks gives a total figure for the number of servers on IRC, which can be used (extrapolating from our sample set) to estimate total bandwidth and file transfer figures for IRC.[62]

To summarize, using in-depth IRC channel observations over time, we derive approximations for bandwidth per server and servers per channel size (numbers of users). Next, we sample each major IRC network for the number of target asset (e.g. movie) trading channels and its respective attending users, apply our server-bandwidth-channel size approximation, and accumulate the results.

Napster/Gnutella Clone Composition

Moving to the "Napster/Gnutella clone" category, we have two sources of data: the overall channel stats (users, files, and gigabytes of data) and search results. Search results (which include self-reported[63] user bandwidth) give a weak indication of the broadband and content mixes.

These observations are systematically logged over a period of time to produce a profile for each service.

Non-reporting Services

Services such as Aimster, Freenet, Hotline, Usenet, ftp, etc. provide the least amount

of information. In some cases (e.g.,Aimster),estimations have been made based on usage reports from the services themselves. In other cases (e.g., Usenet, Freenet), qualitative review of the experience compared with other alternatives has led us to discount their relative contribution.

While this practice, along with the noted existence of many private channels to which we have no access, may cause us to understate file-sharing traffic, this supports our intention to provide a conservative but realistic estimate that we may safely prefix with "at least."

Data Model

The following is a data model for file sharing services like IRC.Each service maintains its own database through logging activities. This structure is used to perform analysis on the data.

For the purposes of this study, a SQL Server database was used. SQL Server has the advantage of being able to export directly to Microsoft Excel for numeric analysis, charting, etc.

Channel
Name (Text)
Description (Text) Note: we will tag keywords that identify channel to be of interest to study (e.g.VCD, Movie, Warez, etc.)
Number of users (CUME, Real)
Number of servers (CUME, Real)
Average traffic (Real, mbps)
Average volume (Real, gb/day)
Server
Name (Text,Nick)
Channel (ID)
Domain (ID, see below)
Asset profile (by type – see below)
Max speed (Real, mbps)
Average speed (Real, mbps)

Average volume (Real, gb/day)
Average files (Real, files/day)
Average uptime (Time)
Minimum cps (Integer)
Avg Sends (Real)
File
CUME volume (Real, gb/day, for computing breakdown)
User
Domain type (.com, .edu, .xx country code, etc.)
ID:T3,T1,Cable, DSL, ISDN, Modem
CUME volume
CUME servers
CUME clients

Conclusion

This section has presented the primary research methodology used in this paper, which seeks to portray objectively the ways in which individuals today are using personal computers to distribute, obtain,and consume media.

It has been our intention to present a balanced and educated view of these phenomena, not to rigorously prove any specific metrics for piracy.

We hope that by contributing to the overall understanding of these matters that those whose business or livelihood they may affect will be better equipped to construct a well-informed and effective strategic response.

Appendix B – Sharing Services

Aimster

Statistics[64]

Classification	General File sharing
Architecture	Pure P2P, Gnutella network
Est. Number of Users	4.5 Million
Supported Platforms	Windows 95,98,NT, 2000
File Size	2.7MB

Target Audience

Aimster is primarily directed at novice users who want to share files, such as family photos, files from work, or private information with their friends or relatives.

Service Description

Aimster is designed to latch on to AOL's popular messaging client,AIM, and the company has discussed similar plans to glom onto MSN's Messenger and Yahoo!'s chat client. The central concept of the original Aimster was buddy sharing. Aimster integrates its technology with instant messenger applications to provide users with a file-sharing tool.Aimster allows users to share files with buddies specified on a buddy list, accounting for much of its popularity, or to a limited extent with users from other networks. The obvious downside of this is that even with a decent number of friends, the amount of searchable media will be much more limited than a non-buddy focused network like Napster. The most current release of the software changes this approach somewhat in favor of the more Internet-wide search and sharing. In addition, the new software also supports additional media types, including video.

Aimster business focus is on forming portal partnerships. Aimster has been in talks with Intel and Capitol Records.[65] The latter yielded an experimental and limited deal to distribute promotional media related to the release of Radiohead's new album, kid a.

Features Overview[66]

File sharing	✔	Chat	✔
Voice chat	✗	Search	✔
Search by media type	✔	Buddy list	✔
Download manager	✔	File tracking	✗
Audio player	✗	Video player	✗
Image viewer	✗	Proxy support	✔
Client update	✗	Friends notification	✔
Connection speed setting	✔	Concurrent upload	✔
Concurrent download	✔	User directory	✔
Login	✔	Upload port restrictions	✗
Resource usage tracking	✗	Recommended files	✗
Skins	✔	General stats	✔
Buddy restricted sharing	✔	Forums	✔
Other media views	✗	White board	✗
Collaborative search	✗	Shared document updates	✗
Streaming	✗	Hide IP	✗
Auto-resume	✗	Dynamic ping	✗
Instant messaging	✔	Ignore user	✗
Ban user	✗		

BearShare

Statistics

Classification	General File sharing
Architecture	Pure P2P
Est. Number of Users	Unknown
Supported Platforms	Windows 95,98, NT, 2000
File Size	1.36MB

Target Audience

BearShare is designed to let users share files with everyone on the Internet and download files that other people are sharing.

Service Description

BearShare is a client application built on the Gnutella protocol. The basic operation of this protocol allows sending searches for files to computers that are connected to the same network. The actual downloading of files takes place outside of the Gnutella network. Users do not need to be connected to other hosts before they begin downloading a file (unless the host is behind a firewall). Similar to other sharing services, only the files that users specify in a shared folder are visible to other users who are searching.

Like other Gnutella enabled programs, BearShare does not depend on a central server to allow users to search or download files. If an individual host goes down, user ability to search for files remains unaffected. Users have complete control over the number of hosts, if any, that are allowed to connect to their computer, whether or not they can download shared files, and which files will be shared. The program works with all types of files and lets users search for files that contain specific words or phrases. The program allows multiple active searches, it resumes incomplete up-/downloads, offers customizable file filters, and provides comprehensive statistical reporting on a separate page. It also offers automatic connection maintenance, file previewing, and displays server type, user agent and versions used by other users.

BearShare uses the FreePeers Agent, a communications module, which abstracts the handling and maintenance of data without the use of a centralized server. Free Peers, Inc., the maker of BearShare provides peer-to-peer software for exchanging information over the Internet and local area networks.

Features Overview

File sharing	✔	Chat	✗
Voice chat	✗	Search	✔
Search by media type	✔	Buddy list	✗
Download manager	✔	File tracking	✗
Audio player	✗	Video player	✗
Image viewer	✗	Proxy support	✗
Client update	✔	Friends notification	✗

Connection speed setting	✔	Concurrent upload	✔
Concurrent download	✔	User directory	✗
Login	✔	Upload port restrictions	✔
Resource usage tracking	✔	Recommended files	✗
Skins	✗	General stats	✔
Buddy restricted sharing	✗	Forums	✗
Other media views	✗	White board	✗
Collaborative search	✗	Shared document updates	✗
Streaming	✗	Hide IP	✗
Auto-resume	✔	Dynamic ping	✗
Instant messaging	✗	Ignore user	✗
Ban user	✗		

Hotline

Statistics

Classification	General File sharing
Architecture	Hybrid P2P
Est. Number of Users	3 Million
Supported Platforms	Windows, Mac, Linux (in development)
File Size	6MB (client), 3.4MB (server)

Target Audience

More Internet experienced and technically versatile users.

Service Description

Hotline Connect is a suite of two free applications that enables Internet users to communicate and share files and information over the Internet. The service offers real time chat, conferencing, messaging, data warehousing, file transfer and streaming capabilities.

The software allows the user to explore online communities, preview multimedia files, post news and to participate in online chats. It offers a full range of personalization, including bookmark deletion for the client, server-side folder upload/download configuration, and the ability to disable cookies for complete client-side privacy control.

In addition,the Hotline software supports cross platform operation, accessible to both Mac and PC users and Linux versions are currently being developed.Setting up your personal server is slightly more complex then typical P2P applications, which seems almost intentional. But there are limitations to Hotline's technology: Hotline has a significantly lower number of running servers than clients and the application's search feature returns "sites" rather than actual files. Finally, the separation of server and client software, resulting in two separate installation processes, seems archaic.

Hotline is marketed as an application to help build community around existing businesses or organizations (e.g., a file-sharing extension of an existing Web site). The business model focuses on in-client advertising and developing network affiliations.

Features Overview

File sharing	✔	Chat	✔
Voice chat	✗	Search	✔
Search by media type	✔	Buddy list	✔
Download manager	✔	File tracking	✗
Audio player	✔	Video player	✔
Image viewer	✔	Proxy support	✔
Client update	✗	Friends notification	✔
Connection speed setting	✗	Concurrent upload	✔
Concurrent download	✔	User directory	✔
Login	✔	Upload port restrictions	✗
Resource usage tracking	✗	Recommended files	✗
Skins	✗	General stats	✔
Buddy restricted sharing	✗	Forums	✔
Other media views	✔	White board	✗
Collaborative search	✗	Shared document updates	✗
Streaming	✔	Hide IP	✗
Auto-resume	✗	Dynamic ping	✗
Instant messaging	✔	Ignore user	✗
Ban user	✔		

Napster

Statistics

Classification	Music Specific File sharing
Architecture	Central Server for Search Library
Est. Number of Users	6 Million
Supported Platforms	Windows 95,98,NT and 2000, Mac
File Size	1.5MB

Target Audience

Internet users in general and music lovers.

Service Description

The most popular MP3 sharing service which serves the largest community of music lovers on the Internet. It offers instant messaging, chat rooms, hot list tracking online buddies, and news about up-and-coming artists.

User music libraries are automatically added to the Napster's "library"each time users launch the software. A user's search for an artist or song is therefore a global search of all the hard drives of users currently logged in.Although Napster's library is centralized, there is no central storage location for the files. Search results display file names, size, and type of Internet connection of the file's owner, allowing the user to select the fastest location from which to download.Adding others to a user's "hot list" enables the user to see the rest of their collection.

Unfortunately, Napster's greatest strength – its huge base of users – can also cause problems. Since so many people are using the service at any given time, others downloading his or her songs can consume much of bandwidth available to an individual user who uses only a dial-up connection to access the service. Currently, the company is undergoing significant changes to address a variety of copyright violation charges. Several court decisions have called into question Napster's contention that since it does not distribute the files itself, it can't be in violation of copyright laws. This effectively forced Napster to implement filters that prohibit the exchange of copyright protected files.

Features Overview

File sharing	✔	Chat	✔
Voice chat	✗	Search	✔
Search by media type	✔	Buddy list	✗
Download manager	✔	File tracking	✗
Audio player	✗	Video player	✗
Image viewer	✗	Proxy support	✗
Client update	✗	Friends notification	✔
Connection speed setting	✔	Concurrent upload	✔
Concurrent download	✔	User directory	✔
Login	✔	Upload port restrictions	✗
Resource usage tracking	✗	Recommended files	✗
Skins	✗	General stats	✔
Buddy restricted sharing	✗	Forums	✔
Other media views	✗	White board	✗
Collaborative search	✗	Shared document updates	✗
Streaming	✗	Hide IP	✗
Auto-resume	✗	Dynamic ping	✗
Instant messaging	✗	Ignore user	✗
Ban user	✗		

Gnutella

Statistics

Classification	General File sharing
Architecture	Pure P2P
Est. Number of Users	N/A
Supported Platforms	Depending on the specific client, generally support for Win 9x, ME,NT, 2000,Linux, MAC,BEOS
File Size	Client-specific

Target Audience

Depending on the specific client, it ranges from novice Internet users to more technical and computer-savvy individuals.

Service Description

The first Gnutella software was created and posted on the Internet by programmers at Nullsoft, an America Online subsidiary. AOL immediately denounced it as "an unauthorized freelance project," and the software was removed, but Gnutella and its concepts continued to spread and gain significant traction over past years. Since then, many companies and affiliations of programmers have released a variety of Gnutella-based software and clients.

With the emergence of popular Gnutella clients such as LimeWire and BearShare it has become much easier to join the Gnutella network and use it to share files. These software releases have also helped to reduce some early problems with the network, such as traffic overloads and freeloaders who download files without sharing their own. They also helped the Gnutella network grow tenfold in just a few months. Currently, an estimated 31,000 unique hosts are recorded on the network, according to data compiled by LimeWire.

Gnutella itself is a protocol designed for sharing files in a distributed network, a standard format that allows two pieces of software to communicate, allowing the exchange of any type of file through connecting every computer in the network to every other machine, with no single server responsible for distributing all of the content (the essence of P2P). It is important to note that Gnutella is an open standard. No single organization owns it or has control over it, so the survival and success of Gnutella does not depend on any single company.

Features Overview (initial Gnutella client)

File sharing	✔	Chat	✗
Voice chat	✗	Search	✔
Search by media type	✔	Buddy list	✗
Download manager	✔	File tracking	✗
Audio player	✗	Video player	✗
Image viewer	✗	Proxy support	✗
Client update	✗	Friends notification	✗
Connection speed setting	✗	Concurrent upload	✗
Concurrent download	✗	User directory	✗

Login	✗	Upload port restrictions	✗
Resource usage tracking	✗	Recommended files	✗
Skins	✗	General stats	✗
Buddy restricted sharing	✗	Forums	✗
Other media views	✗	White board	✗
Collaborative search	✗	Shared document updates	✗
Streaming	✗	Hide IP	✗
Auto-resume	✗	Dynamic ping	✗
Instant messaging	✗	Ignore user	✗
Ban user	✗		

Morpheus

Statistics

Classification	General File sharing
Architecture	Hybrid P2P
Est. Number of Users	Limited,but steadily growing
Supported Platforms	Win 9x, ME, NT, 2000
File Size	1.44MB

Target Audience

Primarily users of MusicCity's Web site and services

Service Description

Morpheus is a slightly altered version of KaZaA (a distributed,self-organizing network and one of Europe's premiere P2P services) modified to work with MusicCity's array of servers. Morpheus is a full-featured P2P file-sharing application that enables users to search for all types of digital media across the MusicCity Network. Morpheus is neither central server-based, nor solely based on the Gnutella file-sharing protocol. Instead,the program uses a proprietary peer-to-peer protocol to share files among users on a single network.

Morpheus offers advanced features such as play list creation and extended media management. It also has an embedded Microsoft media player for audio and video playback, instant messaging, and CPU utilization throttling to limit peer

computer utilization. The program automatically resumes broken content downloads, using SmartStream™ by finding another source and monitoring the network until the requested content stream becomes available again. FastStream™ allows simultaneous transfers of content files from multiple sources for fast downloads of large files, even from users with slower connection speeds. Morpheus file transfer claims to be fully encrypted to protect user privacy, transmissions, and unauthorized intrusions. In addition, the Morpheus application claims to support third-party digital rights management technology allowing content providers to protect the copyrights of their digital content distributed through the MusicCity's network.

KaZaA

Service Description

The developers of KaZaA describe the software as the first expression of a commercial framework for a new content delivery system – direct delivery of a media from artist to consumer.[67] The application is built on the KaZaALib, which contains the foundation for a direct payment or fund-based media market.KaZaA's search engine produces fast and relevant returns using a proprietary content-description technology. It features an intuitive UI that allows even the novice user to learn it quickly. But KaZaA suffers like so many other P2P clients from a limited user base and offers only limited resources.

Features Overview

File sharing	✔	Chat	✔
Voice chat	✗	Search	✔
Search by media type	✔	Buddy list	✗
Download manager	✔	File tracking	✗
Audio player	✗	Video player	✔
Image viewer	✗	Proxy support	✗
Client update	✗	Friends notification	✔
Connection speed setting	✔	Concurrent upload	✔
Concurrent download	✔	User directory	✗
Login	✔	Upload port restrictions	✗
Resource usage tracking	✗	Recommended files	✔
Skins	✗	General stats	✗
Buddy restricted sharing	✗	Forums	✔

Other media views	✗	White board	✗
Collaborative search	✗	Shared document updates	✗
Streaming	✗	Hide IP	✗
Auto-resume	✗	Dynamic ping	✔
Instant messaging	✗	Ignore user	✗
Ban user	✗		

WinMX

Statistics

Classification	General File sharing
Architecture	Hybrid P2P
Est. Number of Users	Unknown
Supported Platforms	Win 9x,ME, NT, 2000
File Size	1.59MB

Target Audience

Broad spectrum of Internet users, intermediate to expert users.

Service Description

WinMX is another OpenNapster clone. The client allows the user to connect to multiple networks simultaneously and share and download any file type. Users can search in "Napster-style" for different digital media types by title or author. In addition,WinMX supports multiple simultaneous searches. The software can track and resume broken transfers. It has the ability to limit upload and download bandwidth, features upload and download bandwidth graphs, and per-user and overall upload and download queuing.WinMX also has full chat capabilities. The program includes a close program or shutdown computer (for Win9x only) option when current transfers are completed.

Features Overview

File sharing	✔	Chat	✔
Voice chat	✗	Search	✔
Search by media type	✔	Buddy list	✔
Download manager	✔	File tracking	✗

Audio player	✗	Video player	✗
Image viewer	✗	Proxy support	✗
Client update	✔	Friends notification	✗
Connection speed setting	✔	Concurrent upload	✔
Concurrent download	✔	User directory	✔
Login	✔	Upload port restrictions	✔
Resource usage tracking	✗	Recommended files	✗
Skins	✗	General stats	✔
Buddy restricted sharing	✗	Forums	✗
Other media views	✗	White board	✗
Collaborative search	✗	Shared document updates	✗
Streaming	✗	Hide IP	✗
Auto-resume	✔	Dynamic ping	✗
Instant messaging	✗	Ignore user	✗
Ban user	✗		

Mojo Nation

Statistics

Classification	Resource sharing
Architecture	Other
Est. Number of Users	Tens of Thousands
Supported Platforms	Win 9x,ME, NT, 2000, Linux
File Size	5.8MB

Target Audience

Users cognizant of Internet community, free-rider phenomena and "publisher rights."

Service Description[68]

Mojo Nation enables users to publish and share different forms of data, including text, movies, and other binary files. Unlike other online file-sharing systems that are plagued by freeloaders who use the service but don't ever contribute, Mojo Nation charges for every transaction internal tokens called "Mojos." Once a user has reached the preset credit limit, he or she must contribute "something" – whether resources or "cash" – to the Mojo community. The company uses a system of

brokers, agents, relays and trackers to monitor resource usage across the network. The distributed data service built on top of this micro-payment system provides a reliable and scaleable method for peer-to-peer content distribution.

Mojo Nation is de-centralized and secure; once data is published it cannot be deleted or controlled. Users are authenticated and tracked across the network based on a generated and encrypted key that represents their account. Content publishers and consumers can respectively publish or retrieve data with as much anonymity as they desire.

Points of criticism are a slow search and limited protocol support of the proxy.

Features Overview

File sharing	✔	Chat	✗
Voice chat	✗	Search	✔
Search by media type	✔	Buddy list	✗
Download manager	✔	File tracking	✗
Audio player	✗	Video player	✗
Image viewer	✗	Proxy support	✔
Client update	✗	Friends notification	✗
Connection speed setting	✗	Concurrent upload	✗
Concurrent download	✗	User directory	✗
Login	✗	Upload port restrictions	✗
Resource usage tracking	✔	Recommended files	✗
Skins	✗	General stats	✗
Buddy restricted sharing	✗	Forums	✗
Other media views	✗	White board	✗
Collaborative search	✗	Shared document updates	✗
Streaming	✗	Hide IP	✗
Auto-resume	✗	Dynamic ping	✗
Instant messaging	✗	Ignore user	✗
Ban user	✗		

IRC

Statistics

Classification	General File sharing
Architecture	Multi-server
Est. Number of Users	Tens of Thousands
Supported Platforms	Win 9x, ME, NT, 2000
File Size	5.8MB

Target Audience

Advanced Internet users and programmers.

Service Description[69]

Originally, Jarkko Oikarinen from Finland designed Internet Relay Chat in 1988. Since then, IRC has developed into an increasingly popular Internet activity and is used in over 60 countries around the world. IRC is a multi-user chat system that operates on a client/server basis. The most popular IRC clients for Windows are currently mIRC and PIRCH.

IRC Networks

IRC servers can be linked together into IRC networks, with each public channel accessible to any user connected to any server on the same network. The largest IRC network, called EFnet (Eris Free net), usually serves over 15,000 users at any given moment. Smaller ones, like Undernet and Dalnet are significantly less populated but provide more stability and convenience for the user. Since IRC servers are connected in a linear fashion they struggle with performance issues. Unlike Web transactions, in which data packets are re-routed whenever one node becomes unavailable, IRC networks experience frequent broken connections between servers. This usually takes place when network traffic becomes heavy and leads to a significant number of interrupted file transfers.

IRC Channels

People meet on "channels" (a virtual place, usually with a certain topic of conversation) to talk in groups, or privately. Channels on IRC are dynamic in the sense that anyone can create a new channel, and a channel automatically disappears when the

last person on it leaves. There is no restriction to the number of people that can participate in a given discussion, or the number of channels that can be formed on IRC. Channel windows inform the user about participants and channel operators. Posted messages also appear in the channel's window along with server announcement, specifying server access information, current file sharing activity, server bandwidth, etc.

File Sharing

Similar to other sharing services, IRC is designed to let users share files with everyone on the Internet and download files that other people have designated for sharing. The application supports all types of files but offers no central search feature to find files. Many of the channels are almost exclusively dedicated to the exchange of digital music or movie files. One of the most popular channels for digital movies is VCD Vault with an average number of about 120 servers.

IRC Scripts

Scripts allow users to automate specific tasks that are routinely performed while using IRC. Many of the scripts are designed to compensate for the lack of features that are common to other popular file sharing services. For example, server scripts and bots effectively substitute the central search feature handling searches in a predefined manner (e.g., servers running the VCD Vault script will respond to the user typing the command "@find <search param>" allowing a rudimentary file searching capability). Scripts can also overcome some of the aforementioned issues such as incomplete downloads by providing an auto-resume feature. Some scripts have been coalesced into bots, which tend to be rich with IRC features and can help automate much of the administrative tasks of running a channel on IRC. "Eggdrop" is a well-known bot on IRC and is used to help moderate a channel by keeping lists of banned users, channel operators (admins), and those with extra privileges (voiced), as well as provide extra features such as games and file servers (fserve). Piggybacking on top of DCC, the standard for direct chat on IRC, the user is able to make a direct connection with the server (or other user) and either send or get binary files such as movies.

Features Overview

File sharing	✔	Chat	✗
Voice chat	✗	Search	✗
Search by mediatype	✗	Buddy list	✗
Download manager	✔	File tracking	✗
Audio player	✗	Video player	✗
Image viewer	✗	Proxy support	✔
Client update	✗	Friends notification	✗
Connection speed setting	✔	Concurrent upload	✔
Concurrent download	✔	User directory	✗
Login	✔	Upload port restrictions	✔
Resource usage tracking	✗	Recommended files	✗
Skins	✗	General stats	✗
Buddy restricted sharing	✔	Forums	✔
Other media views	✗	White board	✗
Collaborative search	✗	Shared document updates	✗
Streaming	✗	Hide IP	✗
Auto-resume	✔	Dynamic ping	✔
Instant messaging	✗	Ignore user	✗
Ban user	✗		

iMesh

Statistics

Classification	General File sharing
Architecture	Hybrid P2P
Est. Number of Users	2 Million
Supported Platforms	Win 9x, NT, 2000
File Size	1.2MB
Key Technologies	Plug-ins for Media Browsers

Target Audience

Novice Internet users.

Service Description

iMesh is a file-sharing tool that allows users to search for files on other computers

also running the iMesh client. One of iMesh's noteworthy features is its ability to simultaneously download a file from up to five other computers, reducing the vulnerability to incomplete data transfers due to interrupted connections during downloads. Other features are customizable client skins, preview of files while downloading, and the ability to choose directories available for sharing through what is iMesh's "Share Wizard."

iMesh offers a sizable and growing community and is currently developing a business model (advertising and sponsorship hybrid) that embraces copyright holders most pressing concerns. iMesh has incorporated a new file-identification technology into its software that could be used to track non-authorized material and remove it from the network.[70] However, the technology's success depends on iMesh's ability to develop it into a secure subscription service platform.

Features Overview

File sharing	✔	Chat	✗
Voice chat	✗	Search	✔
Search by media type	✔	Buddy list	✔
Download manager	✔	File tracking	✗
Audio player	✗	Video player	✗
Image viewer	✗	Proxy support	✔
Client update	✔	Friends notification	✗
Connection speed setting	✔	Concurrent upload	✔
Concurrent download	✔	User directory	✔
Login	✔	Upload port restrictions	✔
Resource usage tracking	✗	Recommended files	✗
Skins	✔	General stats	✔
Buddy restricted sharing	✔	Forums	✗
Other media views	✗	White board	✗
Collaborative search	✗	Shared document updates	✗
Streaming	✗	Hide IP	✗
Auto-resume	✔	Dynamic ping	✗
Instant messaging	✗	Ignore user	✗
Ban user	✗		

Filetopia

Statistics

Classification	General File sharing
Architecture	Hybrid P2P
Est. Number of Users	Unknown
Supported Platforms	Win 9x, NT, 2000,Linux (third party support)
File Size	1.1MB

Target Audience

Users concerned about their privacy and Internet security.

Service Description

According to the developers of Filetopia, its objective is to create a productive, secure and easy to use file-sharing environment for the Internet. Filetopia offers an integrated File Client/Server. It has the ability to recover from errors, implement dynamic block sizes, retry or resume downloads,and restore files automatically. The Collection Manager allows the user to create a list of all files and manage the actual disk files directly from that list. Users who don't want to share all of their files freely can resort to Filetopia's Trade Mode protection. The Chat Client helps users to identify people with similar interests and the messaging system allows them to participate in subject-related channels or to send private messages to one another. Filetopia addresses security concerns with advanced encryption algorithms protecting users privacy using 10 different algorithms with a default of 256 bits and IP bouncers.[71]

Filetopia is a product developed by Bitmap Multimedia,a small company based in Spain, Europe. Bitmap Multimedia started as Software Development Company in 1990.In 1995, it began to focus on the Internet where it concentrated on providing valued-added services to other companies in the space.[72]

Features Overview

File sharing	✔	Chat	✔
Voice chat	✔	Search	✔
Search by media type	✔	Buddy list	✔
Download manager	✔	File tracking	✔

Audio player	✗	Video player	✗
Image viewer	✗	Proxy support	✔
Client update	✗	Friends notification	✔
Connection speed setting	✔	Concurrent upload	✔
Concurrent download	✔	User directory	✔
Login	✔	Upload port restrictions	✗
Resource usage tracking	✗	Recommended files	✗
Skins	✗	General stats	✗
Buddy restricted sharing	✔	Forums	✔
Other media views	✗	White board	✗
Collaborative search	✗	Shared document updates	✗
Streaming	✗	Hide IP	✔
Auto-resume	✗	Dynamic ping	✔
Instant messaging	✔	Ignore user	✔
Ban user	✔		

Appendix C – Selected References

Anon."CEA Digital Download Conference Features Hot Debates Over Copyright and Consumers' Home Recording and Fair Use Rights." CE News March 7,2001.

Anon."Global-Scale Distributed Storage Systems." Bluesky mailing list, http://www.transarc.ibm.com/~ota/bluesky/ February, 2001.

Anon."Life After Napster?" Yippee! http://www.yippee.net/html/win/afternapster.htm

Anon."Potlatch Protocol – a decentralized architecture for gift economies." http://www.potlatch.net/ March 2, 2001.

Anon."Publius Censorship Resistant Publishing System." http://www.cs.nyu.edu/%7ewaldman/publius/ 2000.

Auvinen,Mikko. 'Ave',mIRCStats. http://gamma.nic.fi/~mauvinen/mircstats/ 1998-2001.

Business Software Association. Sixth Annual BSA Global Software Piracy Study May, 2001.

Catidian."Web Sees No Shortage Of Napster Alternatives." news:alt.music.mp3.napster January 15, 2001.

Clark, Ian et. al. The Freenet Project. freenet.sourceforge.net 1999-2001.

Dingledine, Roger. The Free Haven Project: Design and Deployment of an Anonymous Secure Data Haven."MIT Master's Thesis," http://www.freehaven.net/papers.html June 2000.

Forman, Peter and Saint John, Robert W. "The Future of Digital Entertainment." Scientific American November, 2000.

Free Haven Project: Links to Anonymous Publication Projects. http://www.freehaven.net/links.html

InterTrust. What is Digital Rights Management? The Need for a Rights Managed/Enabled World. http://www.intertrust.com/main/metatrust/whatsdrm.html

Koepsell, David R. The Ontology of Cyberspace: Philosophy, Law, and the Future of Intellectual PropertyAugust 30, 2000.

Slashdot: News for Nerds. http://slashdot.org

Lessig, Lawrence .Commons: The Life and Death of Innovation on the Net November, 2001.

Lessig,Lawrence. Code and Other Laws of Cyberspace June 2000.

Litman, Jessica. Digital Copyright: Protecting Intellectual Property on the Internet March, 2001.

Madore, David A. "A Method of Free Speech on the Internet: random pads." http://www.eleves.ens.fr:8080/home/madore/misc/freespeech.html 2000.

Mann, Charles C. "A Whole New Vision for Napster." Yahoo! Internet Life June, 2001.

Mann, Charles C."Who Will Own Your Next Good Idea?" The Atlantic Online September, 1998.

P2Ptracker. http://www.p2ptracker.com/research/summary

PC Pitstop. Napster and the File-Sharing Revolution: Application Penetration and Effects of the PC and Internet Community April, 2001.

Shapiro, Andrew. The Control Revolution: How the Internet is Putting Individuals in Charge and Changing the World We Know May 15, 2000.

Stallman,Richard."Re-evaluating Copyright: The Public Must Prevail." Free Software Foundation. http://www.fsf.org/philosophy/reevaluating-copyright.html Spring, 1996.

University of Utah, General Piracy Information (Reference site). http://u.cc.utah.edu/~bac2/piracy/general.html

Van Tassel, Joan. Digital Content Management: Creating and Distributing Media Assets by Broadcasters NAB, 2001.

Woodhead,Robert."Tipping – a method for optimizing compensation for intellectual property." http://tipping.selfpromotion.com/ February 26, 2001.

FOOTNOTES

1. See http://www.spa.org

2. See, for example, Richard Stallman, "Reevaluating Copyright: The Public Must Prevail;" Published in Oregon Law Review, Spring 1996, http://www.fsf.org/philosophy/reevaluating-copyright.html

3. For example, The New York Times, March, 1997 quotes RIAA head Hillary Rosen as saying: "The low quality of [MP3] files should prevent this format from threatening control of our intellectual property. Why would anyone listen to a sub-CD quality song when they can easily buy the CD at the local Tower Records?"

4. To be accurate, it was the labels themselves that filed suit.

5. This incident notably raised specific international issues surrounding the Internet and copyright law, as such rebroadcasts were legal in Canada. Nonetheless, the broadcasters prevailed forc - ing icravetv to suspend its broadcast activities and opening a technological front known as geofiltering by which Internet services might restrict reception on a geographical basis.

6. A.k.a . DMCA, October 8, 1998, available at http://www.loc.gov/copyright/legislation/hr2281.pdf

7. Section 1201. Circumvention of copyright protection systems. "...no person shall manufac - ture, import, offer to the public, provide, or otherwise traffic in any technology, product, service, device, component, or part thereof, that is primarily designed or produced for the pur - pose of circumventing a technological measure that effectively control access to a work protected under this title."

8. Although the case remains active in appeal: for recent coverage see Lewis, Mark, "Hacker Publication, Studios File Free Speech Answers in DeCSS Case," Webnoise News, June 4, 2001. Also, Scarlet Pruitt, "Hacker Mag's Court Case Hinges on Fair Use," TheStandard.com, May 2, 2001, http://biz.yahoo.com/st/010502/24156.html

9. Forrester: 5.01M to 10.97M (119%), Jupiter Research: 4.8M in 1999 to 8.6M in 2000 (78%). In Q1, 2001 there were 68.5M U.S. homes connected to the Internet according to Telecommunications Reports International.

10. See John Borland, "Net Pirates nab TV episodes from the sky," CNET News, May 24, 2001, http://dailynews.yahoo.com/h/cn/20010524/tc/net_pirates_nab_tv_episodes_from_the_sky_1.html

11. Larry Lessig (http://www.lessig.org/) in several related publications invokes a similar frame - work of four regulatory fources: law, norms, markets, and architecture, which in cyberspace he refers to as "code." (cf. Code and Other Laws of Cyberspace, http://code-is-law.org/). We find it useful to distinguish "code" into its combative "tactical" elements (e.g. cryptographic mea - sures) and its cooperative "design" elements (e.g. usability) in our analysis.

12. See Sue Zeidler, "Researcher Cites Drop in Use of Napster Clones," Reuters, May 17, 2001, http://dailynews.yahoo.com/h/nm/20010517/wr/online_opennap_dc_1.html, also, regarding MPAA's letters to ISPs, http://biz.yahoo.com/st/010502/24156.html, http://www.newsbytes.com/news/01/164614.html

13. See Matt Richtel, "Aimster Heads Down a Path Already Taken by Napster," The New York Times, June 1, 2001. The article quotes Mark Radcliffe, an intellectual property lawyer at Gray Ware & Freidenrich: "Without a central database [of available files], the R.I.A.A. has a tougher case to make."

14. Aimster v. RIAA, Launch Media et. al. vs RIAA, and Professor, Felton vs. RIAA, SDMI, and Verance provide some timely examples. See, for instance, Lewis, Mark, "Princeton Researcher Sues RIAA, SDMI, and Verance, Issuing New Challenge to DMCA," Webnoize News, June 6, 2001.

15. See, for example, John Borland, "Music Trading Heads Back Underground," CNET News, May 8, 2001, http://news.cnet.com/news/0-1005-200-5862906.html?tag=tp_pr, also Dawn C. Chmielewski, "Digital music outlets pop up as major labels fight on," The Mercury News, June 3, 2001, http://www0.mercurycenter.com/partners/docs1/056568.htm

16. Technically, Gnutella is not a service or an application but an infrastructure on which other services and applications (such as BearShare and Limewire – see Appendix B) are built.

17. See http://www.sourceforge.net

18. Note that this step is also highly problematic in the case of so-called "havens" and BNC's (see http://havenco.com for one example) that provide anonymous proxies for users through offshore facilities that are not easily subject to legal jurisdiction. However, since such facilities still themselves need to be connected to the Internet through some service provider, there remains the recourse of forcing their disconnection through pressure on their upstream providers. Here, escalation appears to favor enforcement, although the point is disputable.

19. There has been some speculation as to the degree of vulnerability such a system might have to a denial of service (DoS) type of attack, or virus infiltration, but such hacker tactics themselves are generally illegal.

20. See, for example, "Is SDMI Dead or Just Sleeping?," Webnoize News, May 29, 2001. Summary quote: "There's been no official announcement that the Secure Digital Music Initiative has called it quits, but for practical purposes it's dead. Consortium representatives say there is no consensus among members on technologies that can do the job without creating problems for content creators, device makers and consumers alike."

21. Although many have raised the objection that any large-scale infiltration of a network like Gnutella would be prohibitively expensive, a possible lesson of the SPA's is that a relatively small number of high-profile enforcement actions can serve to set an "example" that has large-scale behavioral effects. (On the other hand, these were accompanied by a fundamental shift in business model by the software industry.) We will not enter this debate, but to note again that escalation appears to favor enforcement in this case.

22. Note, however, that scalability is typically an attribute associated with peer-to-peer architectures. Cf O'Reilly, op. cit.

23. For example, Gnutella, which was developed by the Nullsoft group at AOL, was only available for download for a few hours before it was pulled off of the Internet. That was enough time for millions of copies to eventually propagate into the hands of users.

24. HyperText Transfer Protocol. See http://www.w3.org or http://www.ietf.org for formal definitions.

25. In practice, international issues and hacker tricks (e.g. "IP spoofing," whereby a hacker substitutes his address in outgoing packets with a forgery) make this system less than perfect. Nonetheless, IP addresses provide considerable evidence for escalation of enforcement measures.

26. See Deborah Asbrand, "Patel to RIAA: You Are the Weakest Link," TheStandard.com, April 30, 2001, http://www.thestandard.com/article/0,1902,24092,00.html

27. Johnny Deep, CEO and founder of Aimster has made much of the buddy list distinction. This appears to appeal to the language of the 1992 Audio Home Recording Act's definition of "publication" for the purposes of exempting "fair use." (Section 101 defines "...publicly" to mean "...at a place open to the public or at any place where a substantial number of persons outside of a normal circle of a family and its social acquaintances is gathered..." ...thus begging the definition of what legally constitutes "social acquaintances" for the purposes of fair use. Although Aimster's "find buddies" function is clearly not what the law intended, legal precision on this terminology may cause difficulties (see also Section 107: Limitations on exclusive rights: Fair use). AHRA can be found at http://www.loc.gov/copyright/title17/

28. ...and sued RIAA for attempting to monitor usage in another ironic interpretation, this time of the DMCA "Circumvention of Protection" clause (see Note 7), claiming their encryption was

in essence of form of protection.

29. See http://www.aimster.com/copyright.phtml; the form appears to be provided by Aimster's ISP, Abovepeer.

30. See Note 18.

31. We have actually classified Morpheus/KaZaA has hybrid peer-to-peer. Although they are pure peer-to-peer in the sense that they claim autonomous network operation, not all nodes are equal. This is described in the next Section.

32. http://www.mirc.com

33. Nor are we including pornography channels in this estimation (whose numbers must be far greater but have not been measured by us).

34. These are identified by certain keywords in their names and descriptions, including "movie," "divx," and "vcd."

35. See Section 1.

36. See Section 1.

37. See http://www.projectmayo.com/

38. It is of some significance to U.S. movie downloaders that most DVD players sold in the U.S. can play VCD format discs. Thus aficionados can create VCD versions of their movie down - loads, burn them on a CD, and enjoy them on their televisions.

39. We have considered broadband a minimum requirement for downloading movies, ignoring numerous perplexing observations of individuals attempting to download movie files through analog modems, a process that, if successful, would take perhaps a week of telephone time and hence could hardly represent a significant contribution, although it is certainly a powerful indicator of pent-up demand.

40. http://www.fasttrack.nu

41. See Mark Lewis, "Secretive P2P MusicCity Reveals Revenue Plans," Webnoize News, May 30, 2001, http://news.webnoize.com/item.rs?eID=20010607&ID=13214 (payment required).

42. Gnutella suffers from narrowband bottlenecks since each client shares an undifferentiated bur - den of the network traffic, meaning a single modem user can slow down an entire portion of the network.

43. For this reason we technically classify it as "hybrid."

44. This offer, which MusicCity has promised to clarify, seems suspicious for two reasons: first, one can simply introduce DRM-protected content into any peer-to-peer network (as has been done with Gnutella, see Note 47).

45. For example, Esther Dyson, The New York Times, September 20, 2000: "A policy of all-out prosecution sounds to me like the current war on drugs: not likely to be successful and likely to create more problems than it solves."

46. For a good summary of DRM technology as it applies to Digital Content Management for the broadcast industry, see Van Tassel, op. cit.

47. Notably, sightsound.com early last year distributed content licensed from Miramax Films over Gnutella, the results of which have not been released, although the experiment has not to our knowledge been repeated.

48. Scott Moskowitz and Peter Cassidy of Blue Spike, a pioneering media security company, have been quoted thus: "The bottom line is that a great many solutions with potential for construc - tively animating authentic markets for digital media are within reach, but can't get on the agenda because of the institutionalized limitation of the technical imaginations being brought to bear on the problem." As quoted in Van Tassel, op.cit.

49. See Van Tassel, op. cit., p. 60.

50. Note to the use of the word "typical": Many variants are possible, including the prospect of pre-licensing content, etc.

51. At this point, if the user doesn't already have the right DRM-compliant player installed, a detour is necessary, during which the new player must be downloaded, installed, and, often, the machine rebooted (after closing saving any unsaved files), possibly reconnected to the Internet, and the file again located and opened, at a cost of about 17 minutes on an analog modem line, or 5 minutes on a broadband connection.

52. Moreover, if anti-trust issues are to be avoided, presumably content owners cannot all simply co-own the only legitimate distribution channels for all of their content.

53. Object Management Group, http://www.omg.org

54. For an example of how this works from InterTrust, a leading DRM provider, see http://www.digibox.com

55. For example, SunnComm's MediaCloQ is a proprietary CD audio format that was recently used by Music City Records (no relation to MusicCity.com) to release a copy-protected CD from Charley Pride, "A Tribute to Jim Reeves." Results have been inconclusive. For some interesting alternative approaches (designed to drive traffic to fan Web sites), see http://www.sun - ncomm.com/news_cannotbecopied.html and http://biz.yahoo.com/bw/010207/az_sunncom_3.html

56. In particular, John Gilmore, "What's Wrong With Copy Protection", February 16, 2001, http://www.toad.com/gnu/whatswrong.html...John Gilmore is member of the Electronic Frontier Foundation (http://www.eff.org).

57. Internet Corporation for Assigned Names and Numbers, http://www.icann.org

58. See Bob Francis, "Domain Naming Policies Increasingly Under Fire," Newsfactor.com, June 4, 2001, http://dailynews.yahoo.com/h/nf/20010604/tc/10253_1.html

59. For the following discussion some knowledge of IRC structure will be necessary. IRC, which we refer to as a "service," is (like the Internet) a collection of networks running a common proto - col. IRC consists of four major public networks and hundreds of minor ones (both public and private). Each network consists of a number of channels, which are essentially chat rooms that also recognize special commands. The major networks consist of tens of thousands of chan - nels, each of which is operated by a loose affiliation of channel operators. Channels are identified by name and description. A network's Channel list can be searched by keyword, which is how users seeking certain types of files can find them. For instance, searching a net - work's Channel list for the word "movie" will turn up hundreds of channels devoted to trading movies. A user seeking movies can join a number of such channels at the same time and moni - tor them (either manually, manually with the aid of scripts, or by using a script-program known as a "bot"). Some channels may have obscure names or may be secret so that only by referral would a user know that the channel serves movies. Popular channels will typically have dozens of servers online announcing themselves to the chat room in the fashion shown above. When a user logs onto a server she connects via DCC chat directly with to a file server (a.k.a. fserve – an IRC script running on the server), can type instructions like "dir" to see the files available, and can then initiate a "Direct Client-to-Client" (DCC) transfer. Users may also trade directly via DCC with other users, including users that are running their own servers. See Section 0, or http://www.irc.org for more information on IRC.

60. Note that some cable modem users are able use a software patch to illegally modify their modems' firmware to "uncap" their upstream bandwidth and can reportedly obtain upstream speeds of up to 4mbps using this method (which is specific to a cable modem standard known as DOCSIS).

61. Although the design and use of different services can surely impact this ratio, we can demonstrate its applicability (at least to large file transfers) by showing the results of certain counter-assumptions and design innovations (e.g. "Swarmcasting," which can serve a single file from multiple sources, and persistent transfers, which can resume across sessions).

62. Note that we must be careful of certain details here. IRC channels recognize operators (ops) who are permissioned to run fserves on the channel. Ops may optionally be "voiced," meaning they show up as visible in the channel as opposed to being available by private invitation only. Joining a channel reveals a list of voiced ops currently online, however not all voiced ops necessarily run fserves. Nonetheless, empirical research leads us to believe that relatively few servers are not voiced, and relatively few voiced ops are not serving, therefore these marginal figures tend to cancel each other out and we can use the number of ops as a reasonable proxy for the number of servers.

63. Self-reported bandwidth statistics, which are configured by a user when installing a piece of software, are individually unreliable. Nonetheless, according to sources such as Real Networks which monitor the accuracy of self-reporting in their own client software, the errors on aggregate tend to cancel each other out (i.e. the degree of over-reported and under-reported bandwidth is roughly the same).

64. Statistics based on information from P2Ptracker (http://www.p2ptracker.com/research/companies), individual analysis of download sites and company information published at promotional sites.

65. Seattle P-I.com (http://seattlep-i.nwsource.com/business/aims29.shtml)

66. Features Overview adapted from the P2Ptracker site (http://www.p2ptracker.com/research/companies.asp)

67. See http://www.kazaa.com/

68. See http://www.mojonation.net/intro.shtml and http://www.zeropaid.com/mojo/

69. See http://www.mirc.com/faq.html

70. See http://www.p2ptracker.com/research/company.asp?ProductName=iMesh

71. See http://www.p2ptracker.com/research/company.asp?ProductName=Filetopia

72. Company information can be found at: http://www.filetopia.org/index6.htm

NOTES

NOTES

The Copyright Crusade II

Abstract

In June, 2001, Viant released The Copyright Crusade, a research report in which we examined the phenomenon of peer-to-peer (P2P) file transfer of copyrighted works – in particular music and feature-length films – over the Internet. At the time we conducted detailed observations of this traffic and concluded that, on average, between 300,000 and 500,000 films per day were being transferred over Internet channels such as usenet, IRC, Gnutella, and FastTrack.

Since that time, this research has been widely cited and interest has remained high. Both Jack Valenti (CEO and Chairman, MPAA) and Michael Eisner (CEO, Disney) cited the study in U.S. Senate testimony. News sources continue to inquire as to whether these numbers have changed over the last year.

by
Andrew C. Frank, CTO
Viant Media and Entertainment
afrank@viant.com

VIANT

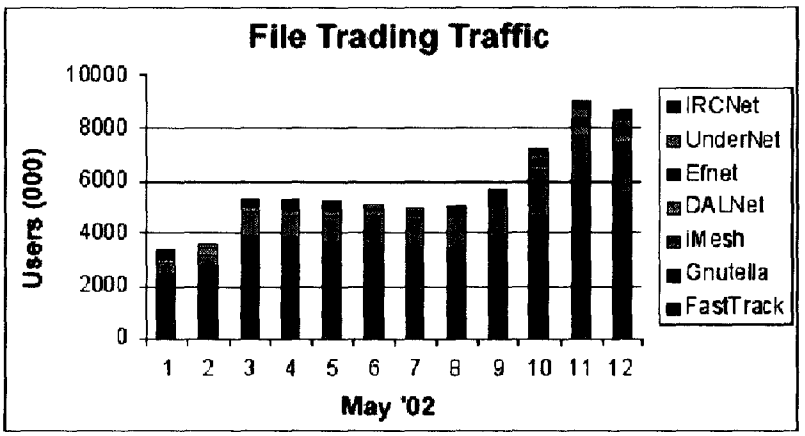
The Attack Begins

The beginning of May, 2002 established a new high-water mark for online film piracy activities, and may have marked the beginning of a new period of expansion. The release of two highly promoted movies, Sony Pictures' **"Spider-Man"** on May 3, and Lucasfilm/Fox's **"Star Wars: Episode II – Attack of the Clones"** on May 16, precipitated an unprecedented frenzy of illicit trading activity when both films appeared, on or before their public release dates, on the Internet.

A "cam" version of "Spider-Man" (captured by a video camera in the audience of a pre-release screening) was first detected the evening of May 2 and quickly spread on May 3, the day the film opened in theaters. "Star Wars," also in cam form, first appeared on May 8,and trading became widespread on May 9, six days before the film's opening in theaters.

"Star Wars" was distributed in four separate files of 70 to 180 megabytes each, while the original "Spider-Man" capture was encoded in a single file. The "Spider-Man" cam has since been superseded by a "telesync" version (captured from within a projection booth with high-quality sound) in two files.

Trading in both films reached a peak the weekend of May 11, when traffic spiked to unprecedented heights across all visible channels. The following chart illustrates the jump in traffic on some of the major Internet networks:



Note: IRCNet, UnderNet, Efnet, and DALNet are major IRC Networks. IMesh, Gnutella, and FastTrack are P2P Networks. KaZaA and Grokster operate on FastTrack, while Morpheus, Bearshare, and Limewire operate on Gnutella.

Source: Average periodic observation during peak hours, and Slyck.com

Send in the Clones

New network capacity spontaneously appeared to accommodate the vastly increased traffic. IRC, which normally consists of the four major networks shown above and about 50 ancillary networks, on May 11 added over 300 new ancillary networks in response to increased demand. During the same day, mirc.com, the distributor of one of the most popular IRC client applications for PCs, reported more than 2 million downloads of their software, far above the norm.

All of the new networks we sampled in the period from May 9 to May 12 appeared, based on queries of file server statistics, to be heavily engaged in the up- and downloading of the "Star Wars" and "Spider-Man" movie files.

Taking these new networks into account,the online trading population soared to over 9 million simultaneous users, of which more than 2.5 were observable on IRC -- about five times the norm.

We estimate that over the May 9-12 period more than 2 million new users signed on to IRC (potentially after their attempts at using the more popular P2P network services failed to yield positive results). That they were inexperienced was evident from their behavior in the chat rooms: much of the chat room message traffic expressed confusion and pleas for assistance in obtaining the movies that they'd heard were available for the taking. Some experienced users suspended their usual hostility toward outsiders and directed them to "helpful"resources.

Of the nearly 10 million Internet seekers of "Star Wars"and "Spider-Man," it's likely that only perhaps 2-3 million have, at the time of this writing, been successful in obtaining complete copies of either of these films. Many attempted the downloading feat with an analog modem, which is generally a losing proposition at any time, but particularly fruitless during a traffic jam. Others gave up in frustration after several aborted attempts. An average broadband home user probably had to struggle for more than six hours to obtain one of the file sets. The procedure could not be easily automated, as most peer servers only allow one or two queue spots per user (while obtaining a complete "Star Wars" required four file transfers), and overall connectivity tends to be spotty. So the process would require frequent user intervention. All in all, those who chose to brave the theater lines probably had a much more rewarding experience, despite having to pay for their tickets.

The most successful seekers are likely to be on college campuses, where a successful download to a server on a campus LAN could in turn serve hundreds or thousands of students who therefore would not need to brave the Internet queues to obtain their copies. These LAN-based trading activities generally take place behind a firewall and therefore cannot be externally monitored.

It's also highly likely that a great many corporate LANs have been put surreptitiously to the same purpose.

With Great Power Comes Great Responsibility

This phenomenon suggests a few conclusions:

- **Online piracy continues to rise.** Although the climate for peer-to-peer start-ups like Napster has been dramatically worsened both by forceful legal measures on the part of copyright interests as well as by a severe downturn in online advertising revenue, Internet piracy has apparently not diminished over the past year. On the contrary, spurred by this summer's blockbuster movies, we believe it may currently be undergoing a period of rapid expansion. If the current capacity and interest levels remain intact,we would estimate that this represents roughly a 20% increase over traffic levels observed a year ago, to the range of 400-600,000 films per day.
- **Eliminating P2P networks will not solve the problem.** Despite their superior ease-of-use and audience sizes, P2P networks continue to lag IRC in file transfer efficiency (and perhaps usenet as well, although this is more difficult to verify). To illustrate, although IRC accounts for only about 14% of the observable file-trading traffic online, we estimate it accounts for 25%-30% of observable file transfer traffic. Put another way, an IRC user is more than twice as likely to successfully complete a movie file download than a P2P network user, while the selection on IRC is both broader and more recent, and the files less likely to be corrupt or mislabeled. This is significant because compared with P2P networks, the IRC network (along with usenet newsgroups), has a strong legal claim to substantial non-infringing use, making providers of network clients and access more difficult legal targets. Moreover, much peer file trading is done directly over legitimate channels like instant-messaging services provided by AOL, Yahoo, and Microsoft, which all support direct peer-to-peer transfers that cannot easily be tracked.

- **Post-theatrical windows represent the greatest area of exposure for movie studios.** Although the question of file-trading's impact on movie box office receipts remains difficult to address, there is clearly a greater threat to subsequent post-theatrical release windows, most notably home video. It is no accident that movie assets are broken into file sizes that can fit on a standard CD-R disk, at the cost of considerable downloading inconvenience: most of these movies are burned to disk, collected,shared, traded and sold in physical form.
- **The opportunity to co-opt piracy with legitimate services is still quite strong.** Despite this recent data, the movie download experience remains intensely challenging and frustrating for most consumers. The fact that so many continue to attempt to obtain content is a powerful testament to the popularity of entertainment products and the demand for online access. Despite some significant investment in their own online services, movie and music companies continue to struggle with the online environment, with very limited online licensing in evidence.

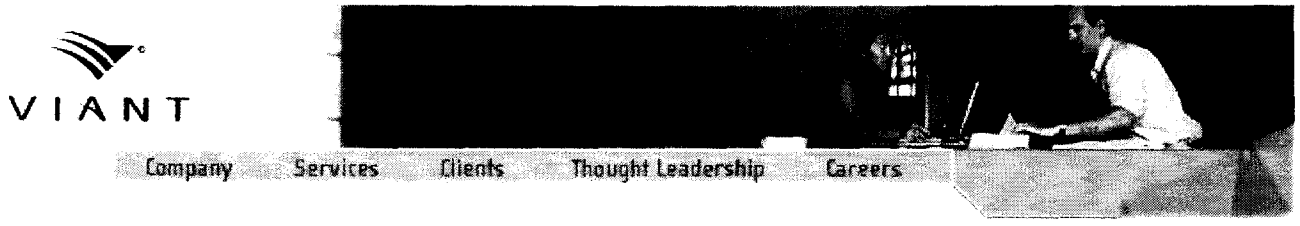
Additional Thoughts

To date, Hollywood has emphasized a strategy of applying legal and political pressure on technology companies to accelerate development and adoption of universal copyright protection technology into their digital media products. The Consumer Broadband and Digital Television Promotion Act (CBDTPA) introduced in Congress by Senator Hollings in March[1] represents a key thrust of this strategy.

The recent negative popular reaction to the bill, however, suggests that the movie industry must balance proposed mandated technology-based copyright restrictions with a promise of greater rewards for consumers. New services must preserve not only the time-shifting benefits of VCR, PVR and,ultimately, VOD, but also the space-shifting benefits of Internet-enabled portable laptop computers and wireless home networks that consumers will increasingly crave.

Consumer empowerment is not a myth. It is a reality to which content companies must aggressively adapt.

[1] S. 2048,introduced March 21,2002 by Senators Hollings, Inouye, Breaux,Nelson,and Feinstein,and referred to the Commerce Committee, http://frwebgate.access.gpo.gov/cgibin/getdoc.cgi?dbname=107_cong_bills&docid=f:s2048is.txt.pdf


VIANT
Company
Services
Clients
Thought Leadership
Careers

press
releases

Viant (ticker: VIAN, exchange: NASDAQ) News Release - 5/21/2002

press release
headlines

Viant Immediately Available to Comment on Any Online Film Piracy Stories Planned

Viant Investors:
Dwayne Nesmith
Direct: 617.531.3
dnesmith@viant.

Viant Media Inqui
Carol Trager
Direct: 617.531.3
ctrager@viant.co

BOSTON, May 21, 2002 /PRNewswire via COMTEX/ --

*****Updated 2002 Stats/Findings Are Also Now Available*****

In June 2001, Viant released The Copyright Crusade, a research report examining the phenomenon of peer-to-peer (P2P) file transfer of copyrighted works -- in particular music and feature-length films -- over the Internet.

Background:

In June 2001, Viant released The Copyright Crusade, a research report examining the phenomenon of peer-to-peer (P2P) file transfer of copyrighted works -- in particular music and feature-length films -- over the Internet (report available at www.viant.com). The report concluded that, on average, between 300,000 and 500,000 films per day were being transferred over Internet channels such as usenet, IRC, Gnutella, and FastTrack.

Viant has updated the Copyright Crusade Report's findings to reflect the unprecedented frenzy of recent illicit online trading activity that occurred around two of Summer 2002's highly-anticipated movie releases ON or PRIOR TO their public release dates.

Viant's View:

It's Viant's view that these latest numbers establish a new high-water mark for online film piracy activities, which may be marking the beginning of a new period of expansion. Viant's new estimate is that the films downloaded and pirated daily have increased to between 400,000 and 600,000.

Viant's Media & Entertainment Practice Technology Officer, Andrew Frank, is immediately available to comment on the current Internet piracy activity, including:

- Background, context
- Latest numbers, stats
- Significance, especially in light of Napster developments

The updated Viant report also is available for reference/ citation. Highlights include:

Over the weekend of May 11, after popular new movie features ("Star Wars" and "Spider-Man") began trading on movie piracy channels, the online trading population soared to over 9 million simultaneous users.

More than 2 million new users signed on to IRC (potentially after their attempts at using the more popular P2P network services failed to positive results).

During peak hours about 2.5 million simultaneous users were observable on IRC -- about five times the norm.

Of nearly 10 million Internet seekers of "Star Wars" and "Spider-Man," approximately 2-3 million are, at the time of this writing, likely to have been successful in obtaining complete copies of either of these films.

SOURCE Viant

CONTACT: Liana Miller or Karsha Chang, both for Viant, +1-949-595-4320