

P2P Survey 2006

ipoque has observed and analyzed the user behavior in the most popular P2P networks used for file sharing between March and October 2006. Depending on the time of day, P2P has a proportion of 30% (daytime) to 70% (nighttime) of the complete Internet traffic in Germany. The absolute value of P2P data volume has increased by 10% between June and October. BitTorrent has passed eDonkey in popularity. Together they are responsible for more than 95% of all P2P traffic and have nearly replaced older networks. The most popular files are still current movies, music and computer games, but to a growing extent also eBooks and audio books. A significant proportion still is pornographic content.

Overview

The prevalence of Internet file sharing based on peer-to-peer (P2P) networks sparks controversy not only among insiders. How much traffic does it generate? What content is shared among file sharers? This survey answers these questions based on a current empirical investigation.

During March and October 2006 ipoque has gathered aggregated statistics on the traffic generated by P2P networks. These data allow for conclusions on general trends in traffic development.

In addition, ipoque has collected about 30 terabytes of detailed traffic data during two consecutive weeks in September and October 2006, which provide the basis for a thorough content analysis. The amount and data volume of shared files have been extracted. Files have been classified according to their content type.

Data collections were conducted at different locations in Europe. The major part of these data was gathered in Germany representing about 100,000 households, as opposed to only 10,000 in Europe. Thus the results generally apply to Germany, but, with some limitations, can also be applied to Europe.

The statistics data have been kindly supplied by various partners running ipoque's traffic management solutions. All data is available only anonymously thus allowing no conclusions on the activities of individual users. Also, this survey is limited to unencrypted P2P protocols, which

significantly reduces the difficulties linked to the content classification. Note that as of the time of this analysis, the amount of file sharing traffic using encrypted P2P protocols is still insignificant.

Particularly interesting and a source of much controversy is just how much P2P networks are used to share copyright-protected material, which is illegal in most countries. This survey deliberately refrains from judging if shared content is legal or illegal. Instead, it provides an overview of all exchanged content and classifies it into defined categories based on content type. A proper legal assessment would require an authoritative and unambiguous list of titles, which would - with reasonable effort - always remain incomplete. This would leave a high proportion of encountered content unclassified or ambiguous. However, the herein presented statistics, in particular the ranking of the most popular titles, may provide some guidance to draw conclusions on the prevalence of copyright violations in P2P networks.

This survey carries on a previous one conducted in spring 2006.

Development of P2P Traffic

Germany

In comparison to ipoque's earlier survey in spring 2006, P2P traffic volume has seen an absolute increase of 8.9% over five months (see Figure 1).

With a proportion of 30 to 70% depending on time of day, P2P still contributes the lion's share to the overall Internet traffic volume. As the overall Internet traffic has seen an even larger increase, the relative P2P proportion has slightly decreased.

This is due to other protocols such as Voice over IP, streaming and encrypted P2P - which is not part of this survey - having strongly grown recently.

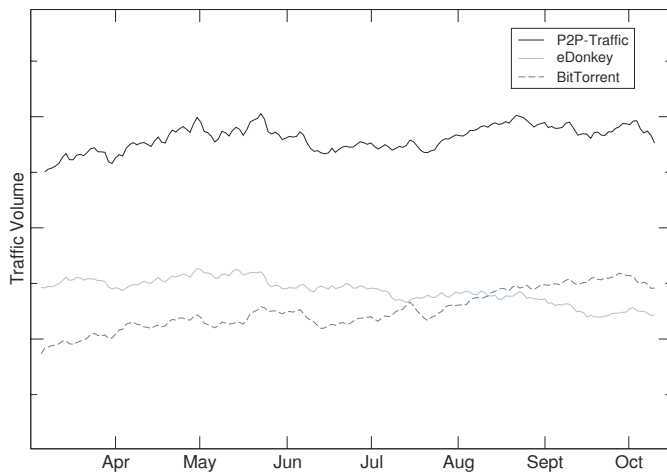


Figure 1: Development of P2P traffic in Germany between March and October 2006

There are many different P2P networks, but only the two most popular ones currently contribute a significant traffic volume (see Figure 2). BitTorrent - current share 53% - has surpassed eDonkey - current share 43% - in August. One reason is that many file sharers - mistakenly - assume that BitTorrent offers better anonymity than eDonkey. In addition, BitTorrent, with the new Azureus client, now provides similar usability as eMule, the most popular eDonkey client. Last but not least, BitTorrent suffers less from fakes than eDonkey does. Fakes are files with wrong and intentionally misleading names.

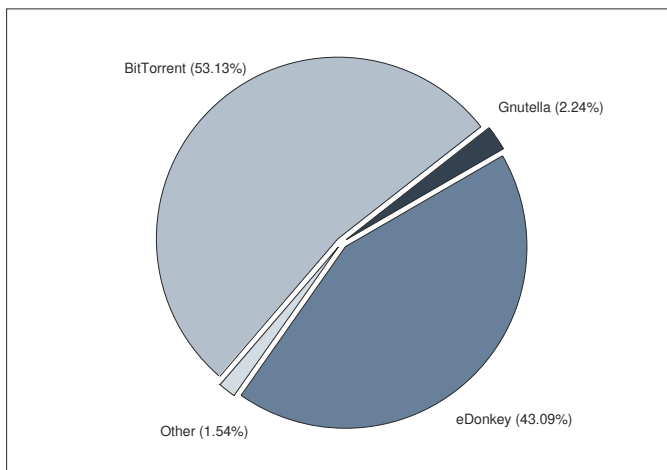


Figure 2: Distribution of P2P protocols in Germany in October 2006

Rank 3 of the most popular P2P networks takes Gnutella, although far behind with only about 2% share of the P2P data volume. Other networks with a combined share of only 2% are of no practical importance. Particularly FastTrack, the most popular P2P network in 2003 thanks to Kazaa, has all but ceased to exist with a minute remaining share of just 0.06%.

Europe

No comparative data from spring 2006 are available, so no trend can be inferred. Currently, eDonkey is the most popular P2P network in Europe with a share of 50%, ahead of BitTorrent with 45% (see Figure 3 and 4). Just as in Germany, Gnutella is at position 3, but with a slightly higher share of 3%, and all remaining protocols with a combined proportion of only 2% are insignificant.

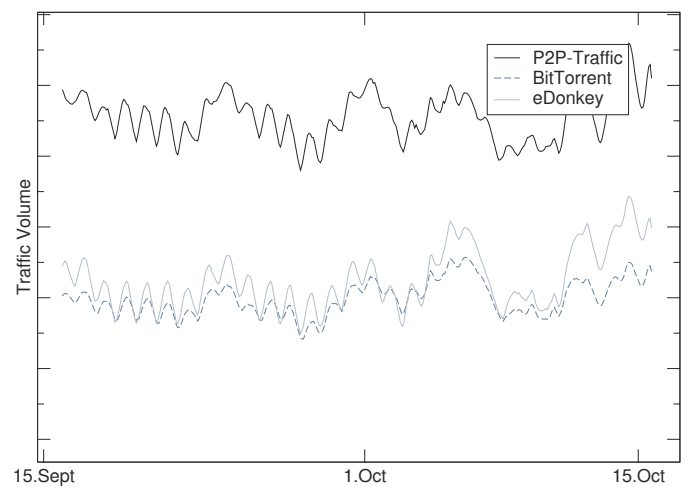


Figure 3: Development of P2P traffic in Europe between September and October 2006

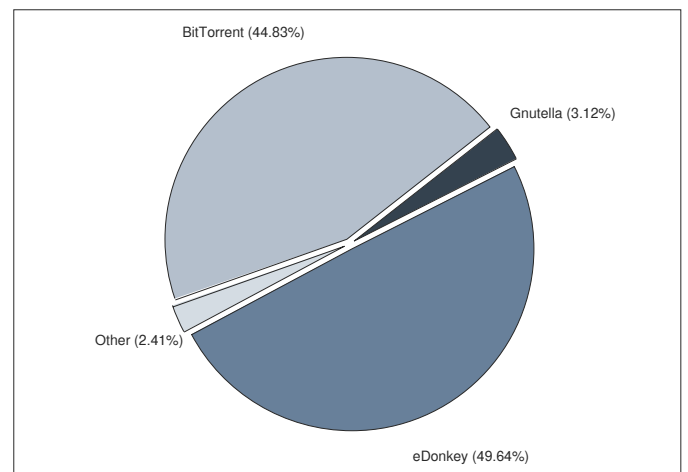


Figure 4: Distribution of P2P protocols in Europe in October 2006

Classification Scheme

Files shared with the by far most popular protocols BitTorrent and eDonkey have been classified based on the two-tier scheme with categories and subcategories listed in Table 1.

Category	Subcategory
Video	Movie
	TV
	Anime
	Porn
Audio	Music
	Audio book
Software	Applications
	Games
eBooks	eBooks
Pictures	Porn
	Other

Table 1: Classification scheme for shared files

For eDonkey we have defined two additional categories: Archive for files combined and compressed in an archive file, and Image for CD and DVD images. Files from both types do not allow to infer the content type. While file names are available as part of the transmission in the eDonkey network, they are missing for BitTorrent and need to be investigated by other means.

Three different classification methods have been used:

Automatic classification (only eDonkey)

- category based on file name extension
- subcategory based on file name patterns (e.g. XXX hints at pornography, or the pattern SiiEjj with ii and jj being numerals for „season“ and „episode“ hints at TV series).

Querying of Internet release repositories (particularly BitTorrent)

- time consuming process; some sites limit the query rate; others have changed the page format during the investigation
- main advantage: categories and subcategories are provided by the repository

Manual classification (particularly eDonkey)

- difficulties arose from such file names that did not allow for an either automatic or quick manual classification (e.g. names coded with Cyrillic or Asian character sets)

Goal of this survey is to provide an as comprehensive as possible overview of shared content in P2P networks. More than half of all encountered files could be classified (56% of BitTorrent and 50% of eDonkey files). Thus, the observed sample clearly exceeds the significance threshold. Nearly all of the top-5000 files have been classified.

Content Shared with BitTorrent

More than 56,000 BitTorrent files have been observed. Video dominates the data volume statistic with a share of 70%. Software, with a high proportion of games, is at position 2. See Table 2 and Figure 5 for details. Based on the number of different files, the ranking slightly changes. Music assumes the top position ahead of video. This results from the different average file size. Software files are usually larger than music files.

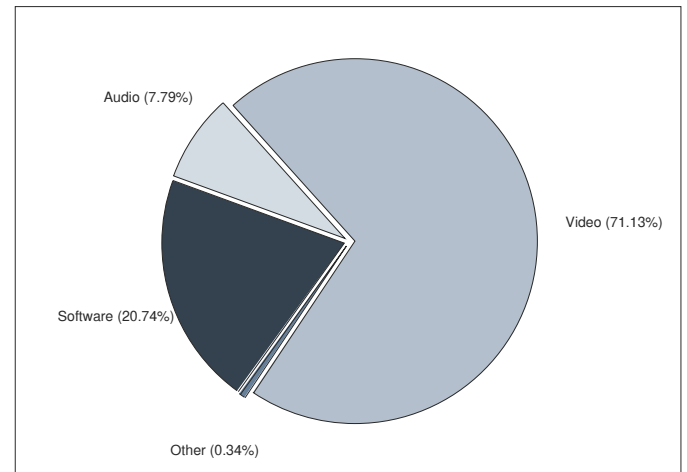


Figure 5: Relative data volume of file types for BitTorrent

Note that the number of different files is not equal to the amount of exchanged titles, as one title often occurs in various versions and sometimes is distributed over a number of files. In addition, many downloads are interrupted early after their beginning making these files appear in the statistics with no contribution to the overall download volume. It is, however, possible to infer rough estimates from the number of shared files on the exchange volume of a certain title.

Category	Subcategory	Vol.	Files	Σ Vol.	Σ Files
Video	Movie	37.42%	19.66%	71.13%	56.64%
	TV	16.12%	17.55%		
	Anime	2.72%	3.00%		
	Porn	14.30%	15.57%		
	unknown	0.57%	0.34%		
Audio	Music	6.76%	22.25%	7.79%	24.93%
	Audio book	1.03%	2.68%		
Software	Applications	6.78%	6.41%	20.74%	14.08%
	Games	13.51%	6.57%		
	unknown	0.45%	1.10%		
eBooks	eBooks	0.13%	2.83%	0.13%	2.83%
Pictures	Porn	0.20%	1.32%	0.21%	1.52%
	Other	0.01%	0.20%		

Table 2: Data volume and number of shared files percategory for BitTorrent

For an estimation of the relative amount of actually downloaded files, their average size has been determined based on

the collected data. The overall transfer volume divided by this average yields an approximation of the amount of actually transferred files (see Figure 6 and table 3).

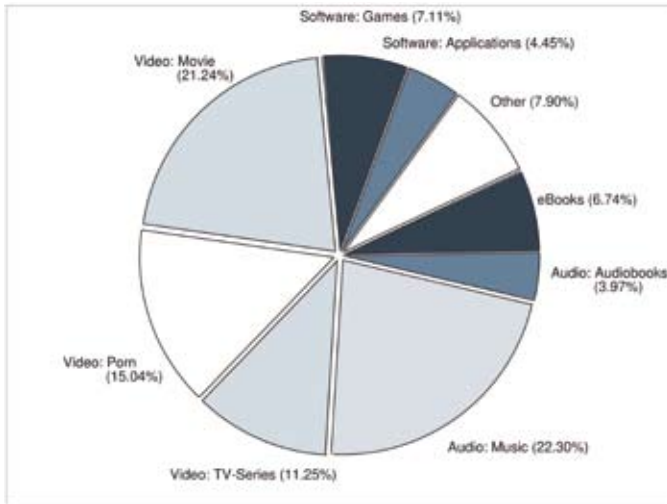


Figure 6: Relative number of files shared with BitTorrent

Category	Subcategory	Ø File Size	Shared Files	Σ Shared Files
Video	Movie	1,390 MB	21.24%	50.37%
	TV	1,130 MB	11.25%	
	Anime	945 MB	2.27%	
	Porn	750 MB	15.04%	
	unknown	500 MB	0.57%	
Audio	Music	239 MB	22.30%	26.27%
	Audio book	200 MB	3.97%	
Software	Applications	1,200 MB	4.45%	12.29%
	Games	1,500 MB	7.11%	
	unknown	422 MB	0.73%	
eBooks	eBooks	15 MB	6.74%	6.74%
Pictures	Porn	40 MB	3.91%	4.33%
	Other	40 MB	0.42%	

Table 3: Estimation of actually transferred files - based on average file size - per category for BitTorrent

Distribution of BitTorrent trackers

BitTorrent requires no central servers for coordination of file exchange. Instead, it uses so called trackers. These trackers mediate between seeders (i.e. clients offering a file) and leechers (i.e. clients downloading a file) of a file without transmitting any content themselves, even without any knowledge about the type of the content. All information necessary to download a file are stored as meta information in a so called torrent file. These information comprise, among others, the tracker that stores contact data of seeders. While a tracker could theoretically run by anyone, our investigation have shown that more than half of all BitTorrent traffic is coordinated by only 20 trackers. The top-5 trackers are listed in Table 4.

Rank	Subcategory	Volume	Cumul. Vol.
1	tracker.bitreactor.to:2710	10.83%	10.83%
2	tracker.zerotracker.com:2710	8.55%	19.38%
3	tracker.torrent.to:2710	8.54%	27.92%
4	www.torrent-downloads.to:2710	5.28%	33.20%
5	www.ultimate-bit-board.to:6969	4.10%	37.31%

Table 4: The five most popular BitTorrent trackers

The first trackers with evidently legal content are far behind. At rank 33, eu.tracker.worldofwarcraft.com generates only 0.34% of all BitTorrent traffic. This tracker is used to distribute updates for the popular online game „World of Warcraft“.

The next one at rank 77 with a mere 0.14% traffic volume is tracker.opensuse.org, which is used for the updates of the Suse Linux distribution.

Content Shared with eDonkey

The statistics data of this section is based on 250,000 different files shared with eDonkey. As with BitTorrent, video is the dominating content type with a share of nearly 70% of all eDonkey traffic (see Table 5 and Figure 8). Half of all videos are movies. TV shows and pornography each account for about one quarter. This distribution is very similar to BitTorrent.

Accordingly, software is at rank 2 of the data volume statistic. However, based on the number of different files, music takes rank 2. The user behavior in the two most popular P2P networks eDonkey and BitTorrent is very similar.

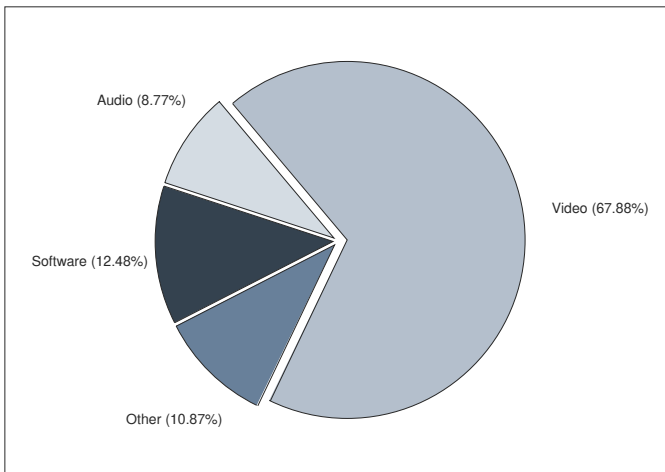


Figure 7: Relative data volume of file types for eDonkey

Category	Subcategory	Vol.	Files	Σ Vol.	Σ Files
Video	Movie	24.18%	10.81%	67.87%	62.45%
	TV	13.02%	11.72%		
	Anime	0.30%	0.34%		
	Porn	13.42%	12.45%		
	unknown	16.95%	27.13%		
Audio	Music	8.33%	17.41%	8.77%	17.84%
	Audio book	0.44%	0.43%		
Software	Applications	2.55%	1.33%	12.48%	4.51%
	Games	8.60%	2.68%		
	unknown	1.33%	0.50%		
eBooks	eBooks	0.56%	0.35%	0.56%	0.35%
Pictures	Porn	0.02%	1.25%	0.05%	2.90%
	Other	0.03%	1.65%		
Archive		9.35%	11.21%	9.35%	11.21%
Image		0.92%	0.74%	0.92%	0.74%

Table 5: Data volume and number of shared files per category for eDonkey

Archive and Image are responsible for about 10% of P2P traffic. As discussed above, these file types cannot be associated with any of the other categories. Fortunately, their share is so small that they resulting distortion is only minute.

Looking again at the average file size for the various file types, the picture is somewhat different. While, for instance, music

files for eDonkey have an average size of 59 MB, it is 239 MB for BitTorrent. The reason is that BitTorrent is used to exchange only complete albums, while eDonkey users also share individual songs. For details see Figure 8 and Table 6.

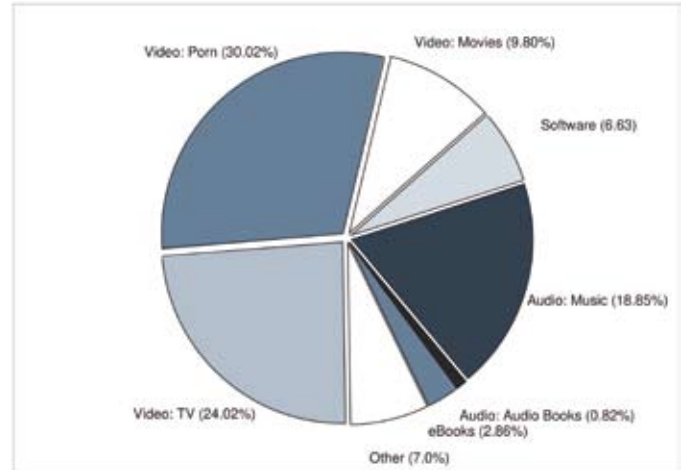


Figure 8: Relative number of files exchanged with eDonkey

This distortion in the rankings is more pronounced for eDonkey than for BitTorrent. So, pornography takes rank 1 based on the number of shared files, followed by TV shows and music. Movies only account for 10%, but generate 25% of the transferred data volume.

Category	Subcategory	Ø File Size	Shared Files	Σ Shared Files
Video	Movie	690 MB	9.80%	64.55%
	TV	92 MB	24.03%	
	Anime	71 MB	0.70%	
	Porn	72 MB	30.02%	
Audio	Music	59 MB	18.85%	19.67%
	Audio book	71 MB	0.82%	
Software	Applications	196 MB	1.74%	6.64%
	Games	668 MB	1.72%	
	unknown	56 MB	3.18%	
eBooks	eBooks	26 MB	2.86%	2.86%
Pictures	Porn	7 MB	0.30%	1.03%
	Other	5 MB	0.73%	
Archive		250 MB	5.00%	5.00%
Image		500 MB	0.25%	0.25%

Table 6: Estimation of actually transferred files, based on average file size, per category for eDonkey

Ranking of the Most Popular Files in Both Networks

Looking closer at the exchanged files reveals, that mainly very current content is exchanged. Both BitTorrent and eDonkey exhibit a very similar „timeliness“ of their content. It is obvious that most of the exchanged content clearly infringes copyright protection. So are, for instance, all movies at the top ranks currently showing in movie theaters. Similarly, most audio files are no free, legally downloadable copies. The same is true for software.

Besides the „classic“ file type movies, music and software, audio books have appeared among the top ranks recently. So are, for instance, „The Three Investigators“ and „Harry Potter“ among the BitTorrent top-5. eBooks are dominated by computer-related text books.

The following tables list the top-5 rankings for the most important categories video, audio and software.

Rank	BitTorrent	eDonkey
1	Klick	World Trade Center
2	World Trade Center	Parfum
3	Der tierisch verrückte Bauernhof	Der tierisch verrückte Bauernhof
4	Der Teufel trägt Prada	Cars
5	Crank	Miami Vice

Table 7: Top-5 video

Rank	BitTorrent	eDonkey
1	German TOP-100 Single Charts	Rosenstolz - Das große Leben
2	Kuschelrock Vol 20	Bon Garçon - Freak U
3	Die drei Fragezeichen Folgen 1-120	David Bisbal - Buleria
4	Fleedwood Mac - Discography	Depeche Mode - 23 Full Albums
5	Harry Potter und der Orden des Phönix	Mozart - 46 Symphonien

Table 8: Top-5 audio

Rank	BitTorrent	eDonkey
1	Fifa 07	Fifa 07
2	Company of Heros	Gilde 2
3	Just Cause	Para World
4	Para World	Joint Task Force
5	Scarface	Age of Pirates - Caribbean Tales

Table 9: Top-5 software

Conclusions

Peer-to-Peer traffic in Germany has risen by nearly 10% from March to October 2006. Depending on time of day, P2P generates between 30 and 70% of the overall Internet traffic. BitTorrent and eDonkey are still the networks that generate by far the most traffic - together about 95% of all P2P traffic. In Germany, BitTorrent - with a share of 53% - has surpassed eDonkey (43%). In Europe, eDonkey is still leading by a small margin.

The shared content of these two networks has been analyzed in detail. While video is responsible for about 70% of the data volume for both, looking at the number of exchanged files reveals a different pictures. Half of all BitTorrent files belong to the category video, most of which are movies followed by pornography and TV shows. Nearly one third are music files, 12% software and 10% audio books and eBooks. For eDonkey, in contrast, nearly one third of all exchanged files are pornography, one quarter TV shows and one fifth music. Movies only account for 10%, even though they are responsible for about one quarter of the eDonkey traffic volume due to their larger average file size. Software, audio books and eBooks are less common in the eDonkey than in the BitTorrent network.

The top positions in the ranking of the most popular files are taken by current titles such as „Click“, „World Trade Center“, „German Top-100 Single Charts“ and the computer game „Fifa 07“. The ranking is dominated by titles that are not freely available. Copyright is massively violated.

The two large P2P networks differ in some aspects due to their different approaches to file sharing. So it is extremely easy to offer a file with eDonkey. Thus, the overall amount of available files is much higher than in the BitTorrent network. In our sample we have encountered 250,000 different eDonkey files and only 56,000 different BitTorrent files. The average file size is smaller for eDonkey, as - in contrast to BitTorrent - pictures, pdf documents and songs are offered individually.

The main problem from the eDonkey users' perspective is the high number of fakes. This is particularly common for current titles such as movies or audio CDs. The search for titles is not an integral part of BitTorrent. Instead normal Web sites are used to offer and categorize titles. These repository sites often provide a rating scheme for the offered files, which significantly reduces the fake problem. It appears that the average quality of offered files is higher for BitTorrent than for eDonkey.

This survey is based on data collected with ipoque's PRX traffic managers. These data were supplied by various partners participating in this survey. The statistical data comprise general traffic statistics, which are automatically generated by the traffic managers. They provide a good overview of the development of P2P traffic and the P2P protocol distribution. In addition, all participating partners have installed a custom module that logs the file hashes and the transferred data volume per hash for BitTorrent and eDonkey, and for the latter also the file name. No information that could potentially infringe the privacy of P2P users, such as IP addresses, have been collected. All traffic data have been aggregated and stored in a central database.

Appendix

The following appendix lists longer versions of all rankings described in the previous sections.

Top-20 Video

Rank	BitTorrent	eDonkey
1	Klick	World Trade Center
2	World Trade Center	Das Parfum
3	Der tierisch verrückte Bauernhof	Der tierisch verrückte Bauernhof
4	Der Teufel trägt Prada	Cars
5	Crank	Fluch der Karibik 2
6	Deutschland, ein Sommermärchen	Disney Collection
7	Das kleine Arschloch	Miami Vice
8	Dead or Alive	Die Super-Ex
9	Jede Sekunde zählt - The Guardian	Mission Impossible 3
10	Das Parfum	Ich, du und der andere
11	Bierfest	Garfield 2
12	King of Queens	The Hills have Eyes
13	Kiriku und die wilden Tiere	Bloodrayne
14	The Black Dhalia	Eine total versaute Krankenschwester
15	DragonBall Z	Deutschland, ein Sommermärchen
16	Gilmore Girls	Bezaubernde Tyra
17	Monster SMX	Schwanzfieber
18	Die Olsenbande	Der Pate
19	Silent Hill	Vivian Schmitt - Feuchte Träume
20	Rambo 1-3	Chaos

Top-20 Audio

Rank	BitTorrent	eDonkey
1	German TOP 100 Single Charts	Rosenstolz - Das große Leben
2	Kuschelrock Vol 20	Bon Garcon - Freak U
3	Depeche Mode - Collection of 11 albums + CD with Remixes	David Bisbal - Buleria
4	70-s & 80-s Dance & Disco	Depeche Mode - 23 Full Albums
5	Die Ärzte	Mozart - 46 Symphonien
6	Santana - Seriously Santana	Elvis Presley - Original Elvis Collection 50 CD
7	The Rolling Stone Magazines 500 Greatest Songs Of All Time	Johnny Cash - Discography
8	Jahrescharts	German TOP 100 Single Charts
9	VA - Tunnel Trance Force Vol 38	Die Ärzte - Discography
10	TOP 40 ODC	Böhse Onkelz - 52 Alben
11	Top 500 Rock And Roll Songs Music	Led Zeppelin - Discography
12	Schranz - Big Pack Schranz	Depeche Mode - 2005
13	Westernhagen - 24 Alben inkl Cover	Metallica - The Unforgiven
14	Top 1000 Pop Hits of the 80s	Die Toten Hosen 1982-2004
15	Johnny Cash - American V A Hundred Highways	Hitman - Blood Money
16	Die drei Fragezeichen Folgen 1-120	Bach (2-5)
17	ABBA - The Complete Studio Recordings-Box	Hands on Ableton - Live
18	Sandra	Hüttenkracher Vol. 2
19	Fleetwood Mac - Discography	Paul Oakenfold Discography Trance
20	Harry Potter und der Orden des Pheonix	Pink Floyd 1967-2001 - 31 CD

Top-20 Software

Rank	BitTorrent	eDonkey
1	Fifa 07	Fifa 07
2	Company of Heroes	Gilde 2
3	NBA 07	Para World
4	Para World	Joint Task Force
5	Caesar 4	Die Siedler 2
6	Scarface	Age of Pirates
7	Geheimakte Tunguska	Just Cause
8	Microsoft Flight Simulator X Deluxe	FlatOut 2
9	Magix Music Maker 2006 Deluxe	The Sims 2
10	Need for Speed	Microsoft Encarta Enzyklopädie
11	Titan Quest	GTR 2 - FIA GT Racing
12	Microsoft Encarta 2007 Lernen und Wissen	Super Bike Riding Challenge
13	Devil May Cry 3	Tom Tom Navigator 6 - Western Europe
14	Heroes of Annihilated Empires	Need for Speed Most Wanted
15	Mage Knight: Apocalypse	Grand Theft Auto San Andreas
16	Magix Video Deluxe 2006/2007 Plus	E Matador
17	Magix Digital Foto Maker 2007	Adobe Photoshop Cs2
18	The Sims 2	Anno 1503 - Königsedition
19	Der Pate	Company of Heroes
20	Battlefield 2	Titan Quest

Top-20 eBooks

Rank	BitTorrent	eDonkey
1	Almost 500 Computer eBooks	Digital Publishing - Komplett
2	XXX Erotic Porn Comic	First Class Sprachkurs English v3.0
3	The Complete O'Reilly Series	C.H.Beck Anwalt Direkt 2006
4	Misc Programming Books	Fansadox Collection 73
5	135 For Dummies eBooks	Black & White Photography - A Basic
6	C.S.Lewis - Die Chroniken von Narnia 1	Das Telefonbuch Deutschland
7	Project Gutenberg Jul.06	National Geographic - Die Erde
8	Elektrotechnik Komplettstudium	Ken Follet - Die Säulen der Erde
9	XXX Magazine - Schulmädchen 24	Perry Rhodan - Silber Edition 10
10	Playboys 2006	Joomla verstehen und anwenden
11	1646 Sex-Geschichten	Duden - Die Deutsche Rechtschreibung
12	Cuvie Collection	John Irving - Witwe für ein Jahr
13	Gamestar 11-2006	Tattoo Vorlagen CD 2
14	Penthouse and Playboy pdf files	Jonathan Stroud - Bartimaeus Das Amulett
15	Battlefield 2 - das offizielle Lösungsbuch	Langenscheidt - Spanisch Kurs
16	Sprüche	Duden Korrektor Jura v3.51
17	Vampire - The Ark Ages und Vampire - Victorian Age	Collection English Grammar Oxford
18	Warhammer 40K Tyranids 4th Ed 2005	Note on Oxford Practice Grammar
19	Gamestar 10-2006	Sobotta Atlas der Anatomie 21. Auflage
20	Star Wars rpg Collection	LTB Sammlung Band 1-36

Top-20 BitTorrent Trackers

Rank	Subcategory	Volume
1	tracker.bitreactor.to:2710	10.83%
2	tracker.zerotracker.com:2710	8.55%
3	tracker.torrent.to:2710	8.54%
4	www.torrent-downloads.to:2710	5.28%
5	www.ultimate-bit-board.to:6969	4.10%
6	tracker.prq.to	1.86%
7	tracker.emporium.us:6969	1.55%
8	torrents-and-more.to:2710	1.53%
9	tracker.pornostube.to:6969	1.33%
10	tracker.deluxebits.to	1.17%
11	btfans.3322.org:8000	1.12%
12	tracker.torrentbox.com:2710	1.03%
13	btfans.3322.org:6969	0.67%
14	boyz.selfip.com	0.66%
15	inferno.demonoid.com:3393	0.59%
16	tpb.tracker.thepiratebay.org	0.56%
17	bt.torrents.ru	0.56%
18	inferno.demonoid.com:3391	0.55%
19	inferno.demonoid.com:3390	0.48%
20	20 btfans.3322.org:8080	0.44%

Contact

ipoque GmbH
Mozartstr. 3
D-04107 Leipzig
Germany

Tel.: +49 (341) 2 41 96 09
Fax: +49 (341) 2 41 96 08

E-Mail: pr@ipoque.com
URL: www.ipoque.com

About ipoque

ipoque has been founded in Leipzig, Germany, in the beginning of 2005.

The company specializes in professional solutions for application- and useraware traffic management and analysis. Its line of hardware traffic managers for Gigabit and Fast Ethernet allows an effective control of undesired network applications. These include file sharing in peer-to-peer networks (P2P), instant messaging (IM) and Voice over IP (VoIP) including the popular Skype. In addition to the bandwidth management, all network traffic can be captured and analyzed in real-time. ipoque offers its customers consultancy services and the development of individual traffic management and analysis solutions based on its standard products.

ipoque's customers comprise enterprises of all sizes, public and educational institutions and Internet service providers.

More information at www.ipoque.com.