

Movie piracy and (the lack of) Cinemas in Hungary

Working paper. Draft date: 3/27/2009

Do not cite or circulate without written permission from authors.

Comments welcome. Contact bodo@mokk.bme.hu

Bodó Balázs, Budapest University of Technology and Economics¹

Lakatos Zoltán, Eötvös Lóránd Univeristy of Arts and Sciences

Abstract

In this paper we analyze data collected from three of the biggest Hungarian bittorrent based file-sharing communities between 2008 April-June, and Hungarian cinema distribution data from the same period. We tracked the downloading activity of file-sharers with known location information and asked if the number of downloads for any given film correlates with ticket sales, revenues, the number of cinemas where the film was shown or any other statistical data on the traditional movie distribution infrastructure we had access to. Our results show that more than 9 out of 10 downloaded films were not available in any of the Hungarian cinemas during our timeframe and only 1 out of 20 downloaded films actually screened. This is in itself an indicator of the importance peer-to-peer networks play in the diffusion of cinematographic content. We have found that though most of the peer-to-peer traffic comes from off-screen content, the most sought after titles are those actually played in the cinemas. **The regression analysis shows that the number of movie theatres in which a film was shown has a strong impact on the number of downloads. However we have found neither causal nor correlation relation between how a film fares at the box offices and how often it is downloaded.**

¹ This work started in 2006 with the publication of (Bodó, Halácsy, Korsós, Prekopcsák, & Szalai, 2007). Our colleagues of that study, Prekopcsák Zoltán and Korsós Milán of Kitchen Budapest continued to work with us in this study and proved to be of invaluable help throughout the work. Ducsay Tamás and Halász Péter, students at the Budapest University of Technology and Economics (BUTE) were instrumental in designing, implementing and running the crawler. Without them it would have been much more difficult to accomplish this study. We are also indebted to Veszélovski Zsolt and his staff at the online program guide port.hu for letting us use their database. Vincze Gábor, a PhD candidate at BUTE has helped us with infrastructure and with his insights for which we are grateful. And last but not least we would like to thank those anonymous hundreds who helped us in pairing torrents with movie titles. Thank You.

Placed in a broader context, these findings suggest several important conclusions. First we could not confirm that the Hungarian box office revenues would suffer from movie file-sharing. Second, we have found that the marketing efforts of movie distributors generate demand that they cannot adequately serve: 33% of the downloaders with known location live in a settlement where there is no cinema, not even a temporary one. Third: we have identified something we call the media-effects of cinematographic distribution. Films highly sought after on peer-to-peer networks gain their popularity not because many people have seen it, but because of their presence in the (mass) media.

These conclusions reinforce our initial hypothesis that file-sharing plays a crucial role in the distribution of cinematic works in a situation where the traditional retailers leave much of the demand unserved. This also means that it is the official distributors themselves who can do the most to decrease the level of file-sharing (if that is the goal): they need to provide better service for their would be customers. For policymakers this should serve as a warning to resist the efforts of traditional market players to combat file-sharing with stronger copyright protection and better enforcement.

Introduction

Though the unauthorized reproduction of someone else's creative output (for fame or for profit) is as old as creativity itself (Alford, 1995; Lendvai, 2008. Június), it was the advent of the technologies of mass reproduction which has made it an everyday, mass phenomenon. Free riding on an already existing investment is always profitable. Even if there are firm institutions to curb this practice (such as copyright laws and treaties), the potential gains are so huge, that up until the era of file-sharing, we hardly see anyone resisting the temptation of reaping huge sums by copying without asking first. Even though the reasons that called pre-internet pirates into existence vary greatly (Bodó, in press) there is one thing that is common in all of them: their love of and quest for profit.

Online, peer-to-peer file swapping is unique in the sense that those who participate in it are not interested in the potential monetary gains. Of course there are many in the digital age who do unauthorized copying of CDs, DVDs, software, etc for profits, but this type of activity dwarfs in comparison with the number of those individuals who engage in the non-profit gift economy of online file-sharing.

If monetary incentives do not explain p2p file-sharing, we have to examine other factors that drive this activity. In this paper we hope to shed light on one factor among the many which may explain online file-sharing. That factor is the failure of traditional markets. By traditional markets we mean the well-established institutions that engage in the marketing, distribution, retailing, lending of cultural goods: libraries, cinemas, broadcasters, video rental outlets, etc. By the failure of these institutions we mean such a deficiency (lack of retail outlets, price, lack of variety, etc) that leaves a sizeable demand on the market unserved.

With the identification of such failures we hope to serve several aims. First we would like to explain why p2p file-sharing - something that many actors from the traditional markets see an undesirable, or even as criminal activity -, is so popular. Second, we hope that we can urge the traditional actors to do whatever they can to improve on those points in their businesses that contribute the most to these market failures. Third, we hope to help policy makers in devising a consumer and citizen

friendly policy environment in which citizens, consumers on the cultural markets do not get prosecuted and punished for acting up if the traditional actors don't, or - due to structural deficiencies -, simply cannot.

In this paper we describe our findings from measuring the traffic of movies on three Hungarian, bittorrent based file-sharing networks between April and June, 2008. During this period we have tracked the new titles that appeared on these networks, the location of individual users and the instances of users downloading, seeding, uploading these titles. We are therefore able to tell who, downloaded what from where for how long. We compare this dataset to another set of databases that track the performance of the traditional movie distribution system: cinemas. We have mapped the geographical distribution of the cinema network, and analyzed the distribution patterns that are defined by the producers and distributors of audiovisual works.

By matching these two datasets we were able to estimate to what extent file-sharing traffic can be explained by the way traditional markets operate.

Previous research on the effects of file-sharing on traditional markets

We see file sharing networks as markets, with their own logics of supply and demand. It is clear, that the workings of these peer-to-peer markets differ greatly from the rules of traditional markets.

The source of these differences may be:

- The price. The zero price at file-sharing network is much lower than the price set by the competition of street pirates and lower than the monopolist price set by the traditional producers/distributors. Though traditional actors are convinced that one cannot compete with something that is free, it is clear that price alone cannot explain the existence of p2p markets.
- The timing of putting goods on various markets. While traditional actors utilize release windows to discriminate among different markets to extract the highest profits from these markets, file sharing networks honor the quickest releasers the highest, fuelling a race that makes new titles available in the same time at all territories worldwide.
- The size of the catalogue, in other word the variety of supply. Digital platforms, even though they promise to break the bottlenecks of physical distribution are yet to solve the problem of titles being out of print. Legal hurdles, lack of resources, business considerations hinder the release of all back catalogs in digital format. On the other hand as file-sharers do the digitization, storage, transmission of titles they deem worthy, there is a good chance that any title that has at least one person to care for it will be available, bringing out of print works back into the market (Bodó, 2006; Freeman, November 2008)
- File sharing is as much cultural consumption as a social activity. (Becker & Clement, 2006; Condry, 2004; J. Cooper & Harrison, 2001; M. N. Cooper, March 2005; Giesler & Pohlmann, 2003; Huang, 2005; Hunter & Spitz, 2003; Keenan, November 2008; Manuel, 1993; Marshall,

2004; Rojek, 2005; Strahilevitz, 2003) In fact, the first file sharing service, Napster was created to solve the problem of fans chatting about music, but not able to show to each other what they are talking about. This horizontal, bottom-up self-organization of consumers adds an extra layer to the top-down structures created around, and created by broadcast media.

Since 1999, when Napster, the first p2p technology appeared on the Internet, a growing body of research has emerged on the impact of file-sharing on traditional markets of cultural goods. (Becker & Clement, 2003; Blomqvist, Eriksson, Findahl, Selg, & Wallis, ; Dejean, 2008; "Digital Music Report 2006," 2007; Domon & Nakamura, 2007; "The Economic Impact of Counterfeiting and Piracy," 2008; Givon, Mahajan, & Muller, 1995; Ram D. Gopal, Bhattacharjee, & Sanders, 2006; R. D. Gopal & Sanders, 1998; Gu & Mahajan, 2004; Huang, 2005; IFPI, 2001, 2006; Liebowitz, 2006; Oberholzer-Gee & Strumpf, 2007; Peitz & Waelbroeck, 2006a, 2006b; Rob & Waldfogel, 2006; Sheikh, Rashed, Qudah, & Peace, 2006; Zentner, 2006) The results are more than ambiguous: there are (mostly industry sponsored) studies that link file-sharing to massive economic losses in the cultural industries, while others find little or no correlation between file-sharing activity and sales data. Still others find positive effects of file-sharing in the case of certain groups of artists, and in relation to overall social welfare. This variety of often contradicting findings only demonstrates that it is impossible to take file-sharing out from those cultural, economic, legal, social contexts in which the users of these services are situated. File-sharing *per se* might be a truly networked, global phenomena, but its impact on the traditional markets of culture are as local as those markets and its customers are.

We would like to add to this body of research by digging into the workings of a segment of the cultural markets of a post-communists country 20 years after its re-integration into the global flows of capital and culture.

Why cinemas and file-sharing?

Though technically it would have been possible to measure the flow of any cultural goods swapped online, we have settled on focusing on films for several reasons.

Conducting such a study requires access to a wide variety of data beyond file-sharing traffic alone. In order to be able to assess the impact of file-sharing on traditional distribution channels, one naturally needs an exact picture of those markets: what is being sold on the market, at what price, where, for how long, with what success. In other words one needs detailed statistics on the production/distribution/consumption patterns on the traditional markets. The studies mentioned before rely heavily on publicly or commercially available data on the workings of the given traditional market they are investigating, which is possible only because such data is collected and published by either governmental organizations, trade groups or third party organizations.

In Hungary there are very few such data sources. The market research databases are completely missing or are still underdeveloped, and even if individual members of the trade do have data on the shelf life of their wares, they are reluctant to share it with anyone. The problem of data scarcity on

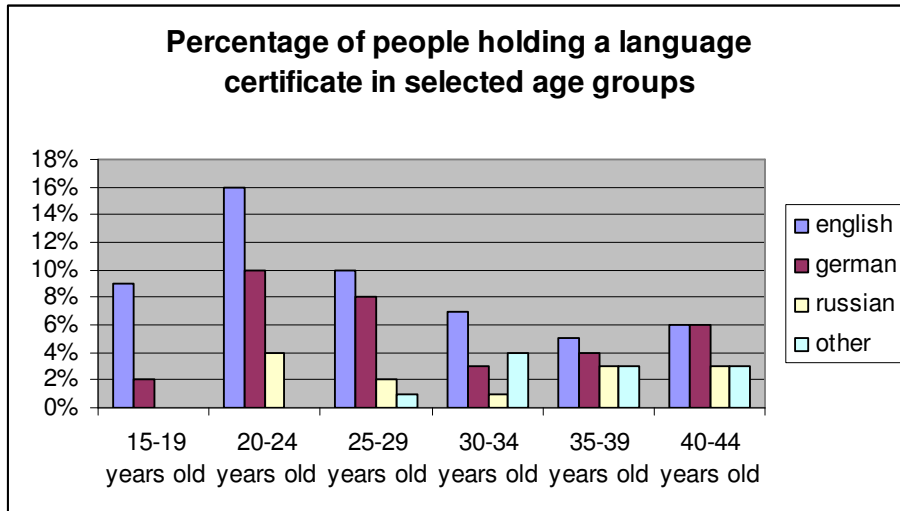
the working of traditional markets was the foremost factor determining which aspect of file-sharing we should deal with. We could not acquire access to any meaningful data on the music market: music publishers, industry members do not release useful data, and there is no third party data available either. On the other hand we were able to get access to the database of port.hu, an online program guide which contains screening information on each and every cinema in the country from 2000 onwards.

The second factor that determined our decision to measure film traffic was methodological. We had to decide between analyzing a sample of all file-sharing traffic, and trying to analyze the whole population of a selected field. The decision between focusing on films or focusing on music was determined by several factors: the number of titles to track, certain characteristics of the users and the characteristics of file-sharing hubs we could track.

In itself alone, the sheer number of titles could have been a determining factor. According to the Internet Movie Database there are around 1 million movie, TV and entertainment titles globally, while at cddb.com which collects titles that were published in CD format, there are more than 6 million albums and over 80 million tracks. This nearly two orders of magnitude of difference (if we take the track as the unit of consumption) is further aggravated by the specificities of the markets of these two different cultural products. In the local movie markets the mainstream US producers share the market with a few European and Hungarian films. The overwhelming majority of these titles are from the last few years (Kanzler, 2009). The number of titles on the movie market at any given time is therefore limited to a small fraction of all possible titles. According to the port.hu database there were only 11805 different titles shown in Hungarian cinemas between 2000 and 2008, and 2008 has seen only 202 new releases.

Tracking music titles would have been a much more difficult task. To put the above number in perspective, on a popular Hungarian file-sharing service examining one randomly chosen user we have found more than 53.000 tracks in her classical and jazz collection alone. There is much more room for diversity in musical tastes, due to the lower cost of music production, free access to music through various online and offline channels, etc. This means that supply and demand on the traditional and on the file-sharing markets are made up of a narrower selection of fewer titles in the case of movies, and a wider selection of a larger number of titles in the music domain. This creates a significant difference in the number of movie and music titles which can be observed at any given time. As a result we decided on tracking movies as the number of movies to track was more manageable.

Certain demographic variables have also contributed to this decision. Hungarian users have a relatively low command of foreign languages: according to a 2005 research (Szénay, 2005) only 9 percent of the 15-44 age group holds a certificate for English as a foreign language.



Source: (Szénay, 2005)

This fact alone limits the demand for foreign language films, but certainly does not affect demand for foreign music, further aggravating the difference described above. The importance the language dimension is stressed by the early appearance (~2001) and the huge popularity of Hungarian online fansubbing communities that translate films, Tv series and produce freely downloadable subtitles.

Finally there are certain technical aspects that make tracking movies so much easier than tracking music. The prime vehicle for movie-sharing is the bittorrent protocol, due to its comparative advantage in terms of speed compared to other protocols, such as DC++, or direct downloading from a web server. Music sharing is partly track based, where bittorrent speed and efficiency factors are lost in comparison to other protocols. Music also has other important vehicles of digital transmission such as file-hosting services (like rapidshare). These (non-p2p) alternatives are inconvenient for sharing large movie files, therefore movie sharers tend to concentrate around a few popular file sharing hubs, among which bittorrent hubs are clearly preferred due to the speed of the network. In turn music sharers are scattered among a number of protocols and services making them much harder to track.

In conclusion: by deciding to track movies on Hungarian file-sharing networks and in cinemas we were able to gather all the film sharing data from the 3 biggest Hungarian torrent networks and compare this data to a detailed program database which tracks film distribution in cinemas.

Changes in the Hungarian movie distribution infrastructure

To put the current state of the Hungarian movie distribution infrastructure into perspective, one needs to go back to the decade before 1989. Due to the seemingly limitless state sponsorship and a strong cultural drive of the ruling party elite (György, 2005) Hungary enjoyed a dense network of libraries, cinemas, and other cultural institutions. Most villages had some kind of a multi-functional institution, a small cultural center, that served as a concert or meeting hall, but could be converted into a screening hall as well. The high number of cinema screens (and libraries) during the eighties reflects this situation.

<i>Year</i>	<i>Number of libraries</i>	<i>Cinema screens</i>
1980	10498	3624
1981	10490	3552
1982	10272	3556
1983	10010	3700
1984	9580	3794
1985	9647	3745
1986	9320	3600
1987	9049	3279
1988	8731	2943
1989	8215	2608
1990	7350	1960
1991	6585	1025
1992	5848	697
1993	5264	638
1994	4727	595
1995	4468	597
1996	4248	558
1997	4092	594
1998	3908	628
1999	3786	604
2000	3585	564
2002		498
2004		464
2007		369

Source: Central Bureau of Statistics

The collapse of the planned economy in 1989 put an end to the financial background of this network. The sudden collapse of the infrastructure was of course not limited of cinemas. It encompassed each and every field in the formerly state sponsored cultural industries including production and distribution of films, theatre, libraries, etc.(Cserta, 2002)

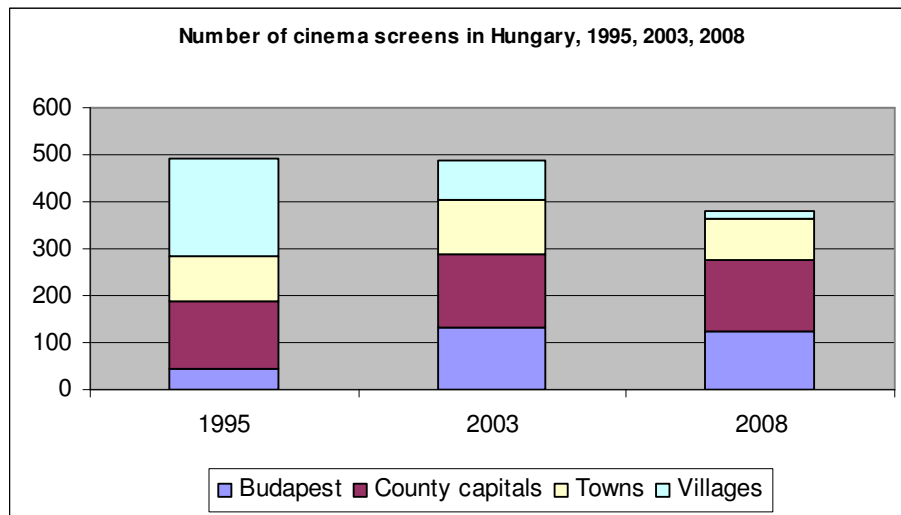
Beyond the changes in the basic political and economic governing principles, several other factors also contributed to the post-1989 transformation of the movie distribution infrastructure.

1. With the disappearance of public funding, ticket prices rose rapidly which occurred alongside the sudden and dramatic drop in per capita GDP and therefore a drop in disposable income.

2. Profit oriented private companies replaced the access-conscious state distributors, releasing fewer copies to fewer cinemas, focusing their distribution efforts on high density markets only, denying local cinemas the possibility of displaying the latest releases.
3. Municipal owners of the cinemas had little funds to maintain, modernize the buildings and the equipment of movie theatres (Borsos, 2007. november), as a result these institutions quickly became run down and/or were privatized and put to other use.
4. In the second half of the 1990's , the rapid proliferation of cable TV, VHS and DVD actors, later the cheap far-eastern home theatre sets posed a serious competition as well. At the end of 2008, 52% of the Hungarian adult population owned a PC/laptop, 59% reported owning a VHS player, 64% owned a DVD player, while 68% had a cable tv subscription. (Source: Szonda Ipsos, National Media Analysis, 2008 december)

These external factors resulted in rapid changes in the structure of the movie distribution infrastructure of the country:

- The number of screens literally decimated compared to the 1980s
- The remaining screens recessed to bigger urban centers leaving (in 2006) as much as 99% of villages, and 70% of towns without a cinema screen. On another level: in 57% of the Hungarian statistical regions there are no settlements with cinemas.(Borsos, 2007)



Source: Central Bureau of Statistics, National Film Bureau

- The screens in urban centers are more and more located in shopping malls, and operated by a handful of US owned companies. Such multiplexes controlled 49% of all screens, 50% of all seats, sold 76% of all tickets, and controlled 84% of all revenues in 2008. It goes without saying that multiplexes have a fundamental effect on what is being shown in cinemas, skewing movie supply towards popular US titles.
- Public subsidies aimed at reconstructing smaller "arthouse" cinemas that show movies outside of the mainstream culture did nothing to change the uneven distribution of cinemas

and resulted in upgrading already established institutions without founding new ones. (Borsos, 2007. november)

In conclusion: in the last two decades movie theatres, along with other cultural retailers have receded to where effective, solvent demand was to be found: into urban centers. "The quick change in the economic and legal environment erodes the basic cultural supply. This is true in qualitative, content-wise terms, in terms of the physical state of infrastructure, costs of operation and in human resources, which is an especially serious problem because due to their cheap accessibility these institutions were mostly used by lower income social groups in need for an access to cultural goods."(Bárdosi, Lakatos, & Varga, 2004) This process of regression proved to be a fatal one: the lack of solvent demand and adequate funding ruined the distribution infrastructure, and the collapse of the distribution infrastructure left those unserved who had been able to pay for these services, but who weren't numerous enough to be served economically.

The shift from independent cinemas with one or two screens to multiplexes in shopping malls also transformed the content that was shown in movie theatres. Multiplexes focus on the few most profitable titles, while those institutions that could serve midlist titles (to borrow a term from the publishing industry) have all but vanished. The lack of cinemas is a problem in itself, but it also generates another one: the lack of diversity in titles.

The structure of p2p file-sharing markets

The structure of illegal online content (films, music, TV programs, e-books, software, etc) markets is a complex one, where p2p communities sharing with and downloading from each other represent only the last step in an intricate and mostly hidden pyramid of middlemen, who participate in the process of acquiring, digitizing and distributing cultural items intended for official release. Based on Howe (January 2005) before a release hits the file-sharing networks, there are several groups whose participation is needed to make something widely accessible.

There are the insiders, „[i]ndustry and theater employees [who] run their own straight-to-video operations. Hackers looking for prerelease videogames target company servers. And before that long-awaited CD hits Amazon.com, moles inside disc-stamping plants have already got a copy."(Howe, January 2005) Then, release groups digitally repackage multi-gigabyte movie files for easy online distribution, rip CDs into mp3, or create cracks that bypass DRM. Many release groups have exclusive relationships with sites on the top of the distribution hierarchy. When a file appears on a so-called top-site, the distribution chain-reaction begins. Couriers step in to copy and transfer files from the top-sites to lower-level dump sites, and then from there to P2P networks and hubs. The couriers are working for such rewards as fame and respect, or for props from their peers and credits redeemable for goods on upper levels of the pyramid. (b-bstf, Summer 2004) The p2p using public mostly trades what is made available for them through these distribution channels. However, local (in many instances semi-amateur) release groups and individuals also participate in the digitization and publication processes, releasing mostly locally relevant titles to local hubs.

This structure of the underground cultural markets suggests two different factors that shape the supply of pirated online goods. On the one hand there is a steady stream of the global supply of the most recent titles. Weeks or months before the official release dates music, film, software is made available through the shadow distribution pyramid. At the same time, local releasers and individuals continuously release titles the local community deems important. The balance between the two for a given hub is defined by the demographics and by the interests of the community that gathers around a specific hub.

File-sharing is undoubtedly a mass phenomenon and this massive demand for such services has called to life a wide variety of file-sharing protocols and each of these protocols support a number of different business models. Bittorrent for example is a file-sharing protocol, the method which describes how peers can connect and communicate with each other. Likewise, Gnutella, Freenet, etc. are similar protocols with different technical characteristics. Apart from a small number of these protocols, which are proprietary, software developers are able to develop with a variety of client software to connect to the network.

Open protocols also enable the proliferation of services that coordinate the users using any given protocol. Sometimes called hubs, at other networks trackers, these services serve as meeting points for users who wish to share and download or engage with each other in any other fashion. It is up to the service providers to decide what kind of business model they want to adopt. Some hubs operate in a truly communitarian fashion: the system administrators finance the operating costs of an open, ad-free service, as they believe they are engaged in a cultural / political mission. Others follow a closed, ad-supported model. The notorious Pirate Bay is a global, open torrent tracker, which exposing its users to advertising content, it is however unclear whether ad revenues cover the operating costs of the service. On the other end of the scale we find open, for-profit distributors: illegal warez servers which sell a flat rate access for a relatively high monthly fee as well as authorized distributors who operate their services with the approval of rights-holders.

Services/clients	Open service (anyone can become a member)	Closed (invitation only) service
Non-profit (with no ad revenue / without membership fee / donations based)	Elite DC Hub service (DC protocol) Soulseek client (soulseek protocol)	Karagarga tracker (bittorrent protocol)
Ad supported	Piratebay tracker (bittorrent protocol) Mininova tracker (bittorrent protocol)	Bithumen tracker (bittorrent protocol) Ncore tracker (bittorrent protocol)
For profit (adware / spyware / membership fee)	Kazaa client (fasttrack protocol) Bearshare client (Gnutella protocol)	X

It should be clear though that - apart from the warez FTP servers, which not being a p2p service should not be included in this table anyway -, all of these services are free for the users in the sense

that they cannot request and they don't receive any compensation neither from the tracker service nor from their fellow users for uploading or downloading content through these services. There are however several third party services which build on these networks, offering, for example server space for seeding files, enabling individual users to achieve a higher upload ratio thus more downloads for them. We do not take such services into account in this report, as their use is no way necessary to participate in file-sharing.

Apart from their business models we can differentiate file-sharing services according to their membership policies. An open service means that there are no registration requirements, or that gaining membership is easy. Closed services accept a limited number of users only, usually through invitations by existing members. Openness has obvious advantages: the more connected people there are, the wider the catalogue and the faster the downloads are. On the other hand openness raises several issues: that of free-riding and that of the risk of being caught and litigated if rights-holders think the services breach - or help users breach - copyright.

Both of these factors have played a role in the proliferation of closed / secretive services in the last few years. Invitation only services offer a good solution to the problem of anonymous users free-riding on others and/or polluting the catalog with garbage. At the same time it offers some level of protection against rights-holders who hope to solve what they perceive as a threat to their business by taking legal actions against individual users as well as service providers. Exclusivity has always served a third, non-related function in the file-sharing scene as well: the social hierarchy of the scene is created, maintained and measured by having access to certain sites: members of the most exclusive services are the highest ranking in the unofficial sub-cultural hierarchy.

Apart from these third type of sites that wish to maintain their exclusivity at all costs, all other torrent tracker services need to balance their interests between having a relatively large userbase, and limiting the access to their services, , therefore they all allow new users to join their services from time to time.

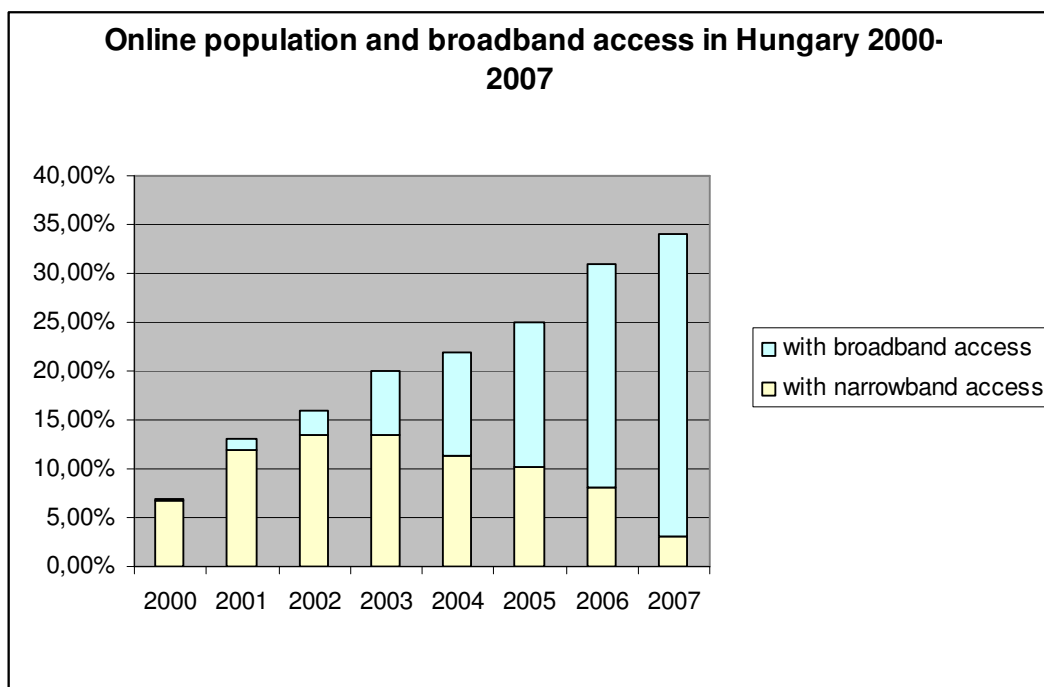
The Hungarian file-sharing scene

Although sharing computer files is as old as computers themselves, the first mainstream file sharing applications emerged at the turn of the new millennium.

First released	P2P Protocol
July 1999	Freenet
September 1999	Napster
November 1999	Direct Connect
March 2000	Gnutella
September 2000	eDonkey2000
April 2001	BitTorrent

By 1999 all the necessary preconditions for wide-scale file-sharing were already in place: reasonable size individual digital libraries, an increasing level of PC penetration and bearable download times

even with a modem. Nevertheless, it was at the colleges, equipped with broadband connections, where file-sharing first took off.



Source: National Communications Authority

College networks also hosted the few first Hungarian file-sharing hubs. The first Hungarian hub using the DC protocol started in 2001 and by 2002 as many as 6 hubs were running. All of these were on university networks, and only started to move to dedicated servers at commercial hosting services around 2004, when dealing with university network administrators became more difficult. ("Az Elite Hub történelme,")In the meantime residential broadband access started to gain momentum.

The first Hungarian bittorrent tracker, bitHUMen started in July, 2004 with a few hundred users. **(sct, 2009 február 01. 15:20)** Soon others followed. With the rapid growth of residential broadband access the number of trackers and the number of users increased rapidly. At the end of 2008 the top 10 Hungarian torrent trackers had the following registered user-base and peer numbers (blue represents trackers participating in our study):

Tracker	registered users	peers (downloaders + uplodaers) on 2008. december 28.
nCore	78612	308330
Independence	68315	90327
Moobs	59989	125913
bitHUMen	51318	218731
Malcka	45067	24054
PREtorians	39979	39250

1st Torrent	37692	46811
GigaTorrents	35001	95085
Spiryt	33349	17898
Blue-Dragon	31826	66153

Source: <http://asva.info/2008-magyar-bittorrent-trackerei-2008-12-29.html>

This of course does not provide us with any real estimate on how many file-sharers there actually are in Hungary. Simply totaling the number of users for each tracker would result in taking into account those users more than once who are registered at multiple trackers. Others, who are not members of these sites would go unaccounted for. According to one estimate there were around 300.000 file sharers in Hungary in early 2008. (Turcsán, 2008. február 7.) Based data from countries with similar population sizes (Huygen et al., 18 February 2009) we need to conclude that this figure of 300.000 downloaders must be a very conservative estimate.

Trackers we tracked

We have decided to track 3 of the most popular Hungarian bittorrent trackers: bitHUMen, nCore and Independence. The choices were made based on reputation, stability, number of users/peers, number and type of titles and finally access. Information on these dimensions were based on interviews of and recommendations by community members.

BitHUMen service is the oldest Hungarian tracker with a solid reputation and a committed community. BitHUMen is also the world's 26th most sought after tracker in terms of the number of invitation requests found on the Internet (sharky, December 17, 08). It is relatively ad-free suggesting a non-profit operation.

NCore, another closed Hungarian hub we have tracked, ranks #40 in the same list. The specialty of this tracker is that it accepts releases from outside of "the scene", the unofficial circle of trusted release groups. Anyone can release on nCore, which means a wider variety of titles, but sometimes also lower quality and lower download speeds. nCore is also an ad based service.

Independence is a relative newcomer compared to the other two. Any user can register to the site, however registration is not free. Independence therefore reaches out to those users who cannot get into the other, more reclusive trackers and offers them a chance to buy themselves into a world they cannot otherwise have access to. This, and the site's strong emphasis on monetizing its user-base seemed to create a bad reputation for the site and its owner among Hungarian file-sharers, who seem to deem such an unabashedly commercial approach objectionable. Nevertheless exactly because of its relative openness we included it in our study.

Even if there is apparent (social) value in exclusiveness, all the torrent tracker services need to balance their interests between exclusivity and the advantages of a wide user-base, therefore they all let new users join the service from time to time. It is possible to join these services, even if it takes some time and effort. In the case of bitHUMen and nCore, we decided to track these closed

/invitation only services, and we interpreted their entry barrier as a variable that separates casual file-sharers from those who engage in file-sharing in a more systematic fashion.

Methodology of torrent traffic tracking

There are several approaches to measuring peer-to-peer file-sharing traffic (Chu, Labonte, & Levine, 2002; Gummadi et al., ; Guo et al., ; Pouwelse, Garbacki, Epema, & Sips, 2005; Saroiu, Gummadi, & Gribble, 2002; Schulze & Mochalski, 2008; Sen & Wang, 2002) utilizing deep packet inspection techniques, protocol level sampling and other approaches. Our decision to develop a new method was the result of a simple factor: as we do not have access to any data source that would let us connect IP addresses to settlement level geographical data. As geographic analysis is crucial in our research, we needed to come up with a different approach that enabled us to acquire user location data.

Luckily the most popular/influential torrent tracker services have all enabled their users to communicate on their user profile pages the settlement where they live. Not everyone has filled out this field in his/her profile, and there are significant differences in the list of settlements these services offered their users to choose from. Nevertheless in 40% of the cases we had proper location information which gave us enough torrent traffic data for the research to be feasible.

Our approach therefore focuses on the hubs that serve the Hungarian file-sharing community instead of monitoring the actual data flows over the network.

Besides serving as community hubs, providing users with self-identification and communication (forums, ratings, polls, etc) services, torrent trackers coordinate the p2p flock. They maintain the information about which user has which part of which file in the network. The users need to communicate with the tracker if they want to download or share something from the others, as it is the server which knows which user has the necessary piece of the file in question. Therefore the server knows, and publishes this type of information which is then relatively easy to gather.

Open-access hubs are easy to monitor, as they do not make an attempt to hide their activities. Closed hubs require more precaution, so the monitoring activity does not get detected by the administrators of the site. Such a monitoring effort raises several ethical issues. We have addressed these issues on several levels. First we gather only such data that is available for each and every ordinary member of the torrent tracker. We respected the privacy decisions of the site administrators, and did not try to gather more information that was intended by them to be public. Also, we did everything we could to respect and protect the privacy of the individuals who use these services. We did not collect any information that could be used either by us or by other parties to connect the online user profiles with real life identities. On the other hand we did engage in a monitoring effort without the knowledge and consent on either the site administrators or the users. This was necessary as acquiring the same amount and depth of information from the administrators of these sites would have been impossible: either because they don't archive such information, or because they do everything they can to protect their communities. Before, during and after the data gathering period we have communicated clearly to the Hungarian file sharing scene that we are

doing research on the effect of file-sharing on traditional markets. We were also trying to be present in the online discussion boards so we could personally answer any questions about the research.

In order to achieve a non-intrusive, difficult-to-detect monitoring of closed hubs we have developed the appropriate monitoring technology. The software has three main functionalities. (1) Its crawler collects data from the torrent hub, (2) the parser extracts relevant information, stores in a (3) database and instructs the crawler on which page to crawl next.

The crawler

The crawler collects the traces of the data flow generated by the torrent communities. As the torrent-tracker is the information hub to which all users go for new downloads, it contains relevant information on what is accessible at any given moment through the given hub. The tracker also contains information on the health and status of the torrent files, it contains the profiles of the users, offering us a chance to extract some information on the users themselves. The crawler crawls these pages to extract every possible information available on the hub through the web interface.

The task of ensuring data quality requires that we crawl the hub (or parts of the hub) frequently. The status of freshly released popular titles (like a Hollywood blockbuster) changes quickly: dozens of users can appear and disappear from the downloading flock within minutes. This requires an intensive presence of the crawler on a site that tries to do everything it can to protect its users from scrutiny and possible prosecution. By dispersing the queries among a number of different proxies we were able to conduct an aggressive but non-intrusive, non-detectable monitoring.

The parser

The parser has three tasks. First it extracts the relevant data from the files sent by the crawler. Second, it anonymizes user related data and dumps all the data into the database. It also adjusts the frequency by which the crawler needs to request a specific page. To avoid data loss we crawl popular, thus quickly changing pages more often than pages of less popular or dead torrents. The parser adjusts the time of the next crawl of a page based on the amount of changes in the flock around the torrent since the last crawl.

Data cleaning, title identification

The torrent files form the basis of the analysis, representing movie titles. This data needs further cleaning as there are several issues to be solved. Even in the case of global releases of global titles there might be several competing versions of the same title in various formats (VCD, DVDrip, CAM), released at different times, by different release groups.

For example the fourth installment in the Die Hard movie franchise is available under the following names: [07.11.17.Live.Free.Or.Die.Hard.Blu.Ray.All.Disk@Ht](#), *Die Hard - Quadrilogy. Untouched Box Set.Nordic*, *Die.Hard.SE.Trilogy.BOXSET.PAL.6DISC.DVDR-SPLiNTER*, *Die Hard Series*, *Die.Hard.4.0.Yippee.Ki.Yay.Edition.2DiSC.NORDiC.PAL.DVDR-ViSiON*, *DIE_HARD_4_0. PAL. R2. SUBS DK,NO,SE.FI. DVD9*, *Live Free Or Die Hard (La Jungla De Cristal 4)*, *Die.Hard.4 x264.720p*, *Die Hard 4. 720p BluRay AC3-5.1 x264*, *Die Hard 4*, *Die Hard 4.0*, *Die Hard 4.0 (Live Free or Die Hard) 2007 DUTCH!*, *Die.Hard.4[2007]MultiSub.DvDR-Gothicmaster*.

It is clear that all of these versions contain the Die Hard 4 movie. The consolidation of these versions and establishing the connection with the titles stored in other datasets was a difficult task. We needed to consolidate more than 7000 movie torrent files that were uploaded within the examined timeframe with nearly 12.000 movie titles that track traditional markets. Proxies such as IMDB ID proved to be massively unreliable, and automatic pairing algorithms provided noisy results. Therefore we decided to crowdsource the task of pairing and asked the file-sharers themselves to participate in connecting torrent files with titles. The results were beyond every expectation: several hundred anonymous users finished this task in less than a week, with very few (less than 1%) errors.

Similar, but much smaller task was to standardize user location info.

The results

Our focus being the relationship between peer-to-peer trading and movie distribution of films, we had to conduct our data mining and analysis in a way that accounts for the multiple connections that exist between these two domains. As explained above, torrent files were identified making it possible to analyze not simple torrent but content-related trading patterns. Without meticulously linking torrent files to specific film titles, we could not have undertaken a content-oriented analysis of the peer-to-peer networks. The following data sources were used for the purposes of this analysis:

1. A database of the connections between users and torrent files. Each connection accounts for a different case, whereby users downloading several different torrent files appear in separate lines in the form of individual “transactions”. Both the users and the torrent files are identified. This is a comprehensive database as it includes all the transactions that occurred on the three selected Hungarian trackers between April 1st and June 30th 2008.
2. A database with information on torrents: size, upload and creation date, tracker, etc. using unique identification.
3. A table linking torrent files to film titles. (Films, just like torrent files, have a unique identification code.)
4. A film database with general information on the film including IMDB, title in Hungarian and/or English (if any), etc. For Hungarian premiers between February 2007 and December 2008, detailed box office data are also available (premier dates, tickets sold, revenue, copies, etc.).
5. A table linking film identification codes to cinema distribution data. For each film, screening dates and the cinemas where the film screened are available.
6. A cinema database with the geographical parameters of the cinemas including the name of its location, etc.
7. A user database with all available information on the users registered at the three torrent trackers in our data collection. User location is of particular importance as it allows for examining the geographical aspect (availability of films downloaded within users’ catchment area).
8. A geographical database with elementary information on settlements, including distances between pairs of them. This is especially useful when exploring the relative shortage of films within one’s catchment area.

Using the above sources, we are able to analyze in detail the relationship between a film’s peer-to-peer and cinema distribution within a three month timeframe that is between April 1st and June 30th 2008. Below are presented the major findings of this relationship. (Though our analysis is based on transactions related to three Hungarian specific torrent trackers within a three-month time frame, the results can be extrapolated as representative of mechanisms on a larger scale).

The following table shows how films downloaded using the three selected trackers break down to categories according to its cinema presence and peer-to-peer transactions.

	Frequency	%
In theaters and downloaded	191	5,1
Not in theaters but downloaded	3427	92
In theaters but not downloaded	107	2,9
Total	3725	100

The very high portion of films not screened but downloaded² during the three-month timeframe is in itself an indicator of the importance peer-to-peer networks play in the diffusion of cinematographic content: more than 9 out of 10 downloaded films were not available in any of the Hungarian cinemas during our timeframe³ and only 1 out of 20 downloaded films actually screened. On the other hand, what's in cinemas tends to get downloaded: only a smaller portion of screened films are not downloaded.

This disproportion is also reflected in the global torrent traffic, as evidenced from the below table. The total number of downloads is app. 3 times higher in the "not-in-theater but downloaded" than in the "in theaters and downloaded" category. However, this preponderance of "not screened content for download" hides a structural imbalance that is characteristic of the peer-to-peer traffic (and which, as we will see later, will be an important factor in determining what will be a popular download): the average number of downloads is significantly higher for on-screen films than in the case of off-screen films, the difference being more than sixfold (887 vs. 141). Off-screen stuff may well account for most of the downloads, but real popular stuff is what's actually being screened. **Any on-screen film attracts more downloaders than any off-screen film.**

Macro-statistics of theatrical distribution and peer-to-peer traffic (base=all transactions)

		Film category			Total
		In theaters and downloaded	Not in theaters but downloaded	In theaters but not downloaded	
Torrent life-span (days)	Mean	73	28	0	29
	Maximum	617	440	0	617
	Sum	14 014	94 675	0	108 689
	<i>Std Deviation</i>	93	36	0	42
Number of downloads	Mean	887	141	0	175
	Maximum	9 108	6 736	0	9 108
	Sum	169 335	481 700	0	651 035
	<i>Std Deviation</i>	1 709	316	0	519
Film life-span (days)	Mean	35	0	34	3
	Maximum	89	0	89	89
	Sum	6 689	0	3 601	10 290
	<i>Std Deviation</i>	35	0	35	14
Screenings	Mean	415	0	209	27
	Maximum	5 831	0	4 273	5 831
	Sum	79 290	0	22 314	101 604
	<i>Std Deviation</i>	1 076	0	599	280
No. of theaters where film screened	Mean	12	0	10	1
	Maximum	105	0	88	105
	Sum	2 376	0	1 116	3 492
	<i>Std Deviation</i>	21	0	19	7

² Here we consider total (100%) downloads by excluding downloads that were started but never finished.

³ For simplicity reasons, the cinema timeframe is cast in perfect symmetry with its the peer-to-peer equivalent. Availability in the cinemas means at least one screening of the film in any Hungarian cinema between April 1st and June 30th 2008. If the film screened only before or after this period, it is classified as unavailable.

On the other hand, **downloaded films get more screenings than not downloaded ones** (in average 415 vs. 209, or a two-to-one ratio). This difference is further accentuated in the total number of screenings. Downloaded on-screen content is also available in more cinemas than not downloaded films. In sum, though most of the peer-to-peer traffic comes from off-screen content, the most sought after titles are those actually played in the cinemas.

Other aspects of film distribution, especially box office statistics are also important. As these are not available for the totality of the films downloaded during the three-month time frame, our detailed analysis is restricted to a category of films for films with a Hungarian premier between February 2007 and June 2008 where these figures are available.⁴

The table on the next page is similar to the previous table but it also includes box office statistics. These corroborate our first findings with regard to the imbalance between the two categories of on-screen content: download vs. not downloaded. **Downloaded on-screen films are not only distributed to more cinemas but attract a much larger audience and generate higher revenues than not downloaded ones.** But this is only one part of the story. A closer look at these data reveals that a film's ability to attract larger audiences does not trigger a higher number of downloads: there is no correlation between the number of downloads for a film and its box office figures (number of tickets and revenues). (See correlation table.) In contrast, a film's popularity for download is highly correlated with the number of cinemas where it is screened and, to a lesser extent with its number of screenings and copies. Cinema distribution appears to play a complex role:

- on one hand, films with a heavier cinema promotion are more popular with downloaders;
- on the other hand, the number of **cinema visitors for a film and its number of downloads are not correlated.**

Placed in a broader context, these findings reveal the media-effects of cinematographic distribution. **Films highly sought after on peer-to-peer networks gain their popularity not because many people have seen it, but because of their presence in the (mass) media.** It is not necessary to conduct an additional media check to infer that films with the highest frequency of screenings are also the most marketed films with heavy presence at all levels of both commercial promotion and press commentary. In this sense, a film is more visible from the downloader's point of view when it is broadly distributed and mediatized, but it can fail to attract attention when it is only a "good film" with many visitors but a narrower distribution and not much media noise. Widely distributed but box office-wise poorly performing films will be popular stuff for downloading because of their sheer visibility, not because of their quality.

⁴ Premiers between February 2007 and June 2008 are considered for this analysis. Though torrents for some of the films with premiers past June 30 may be available in advance (before July 1), expanding the premier timeframe beyond the upper limit of the transaction timeframe would lead to misleading results as end-of-the-year blockbusters might be classified as failing to attract downloader interest, which is not true.

**Macro-statistics of theatrical distribution and peer-to-peer traffic
(base=2007 and '08 Hungarian premiers)**

		Film category		Total
		In theaters and downloaded	In theaters but not downloaded	
Torrent life-span (days)	Mean	99	0	65
	Maximum	617	0	617
	Sum	9 138	0	9 138
	<i>Std Deviation</i>	<i>106</i>	<i>0</i>	<i>98</i>
Number of downloads	Mean	1 471	0	967
	Maximum	9 108	0	9 108
	Sum	135 314	0	135 314
	<i>Std Deviation</i>	<i>2 068</i>	<i>0</i>	<i>1 814</i>
Film life-span (days)	Mean	53	62	56
	Maximum	89	89	89
	Sum	4 877	2 979	7 856
	<i>Std Deviation</i>	<i>33</i>	<i>29</i>	<i>32</i>
Screenings	Mean	813	462	693
	Maximum	5 831	4 273	5 831
	Sum	74 822	22 160	96 982
	<i>Std Deviation</i>	<i>1 429</i>	<i>831</i>	<i>1 264</i>
No. of theaters where film screened	Mean	22	22	22
	Maximum	105	88	105
	Sum	2 027	1 039	3 066
	<i>Std Deviation</i>	<i>24</i>	<i>24</i>	<i>24</i>
Revenue (million HUF)	Mean	83	24	63
	Maximum	675	198	675
	Sum	7 657	1 172	8 829
	<i>Std Deviation</i>	<i>110</i>	<i>39</i>	<i>96</i>
Tickets sold	Mean	84 622	29 785	65 821
	Maximum	739 672	193 883	739 672
	Sum	7 785 259	1 429 697	9 214 956
	<i>Std Deviation</i>	<i>118 404</i>	<i>41 398</i>	<i>102 177</i>
Tickets per screening	Mean	20	15	18
	Maximum	59	42	59
	Sum	1 817	715	2 533
	<i>Std Deviation</i>	<i>9</i>	<i>7</i>	<i>9</i>
Copies	Mean	19	12	17
	Maximum	54	99	99
	Sum	1 775	589	2 364
	<i>Std Deviation</i>	<i>13</i>	<i>16</i>	<i>14</i>

**Correlations of theatrical distribution and peer-to-peer traffic
(base=2007 and '08 Hungarian premiers)**

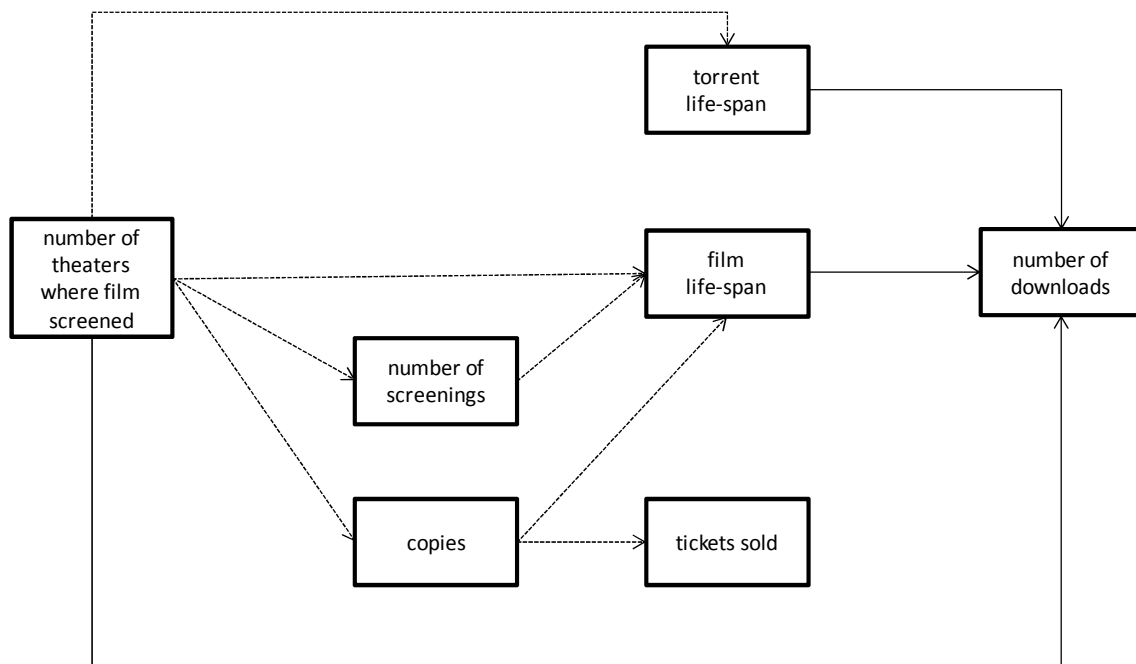
		Torrent life-span (days)	Number of downloads	Film life-span (days)	Screenings	No. of theaters where film screened	Revenue (million HUF)	Tickets sold	Tickets per screening	Copies
Torrent life-span (days)	Pearson Corr. Sig. (2-tailed) N	1 140								
Number of downloads	Pearson Corr. Sig. (2-tailed) N	,718(**) 0,000 140	1 140							
Film life-span (days)	Pearson Corr. Sig. (2-tailed) N	0,06 0,479 140	0,011 0,897 140	1 140						
Screenings	Pearson Corr. Sig. (2-tailed) N	0,115 0,176 140	,229(**) 0,006 140	,188(*) 0,026 140	1 140					
No. of theaters where film screened	Pearson Corr. Sig. (2-tailed) N	,184(*) 0,029 140	,282(**) 0,001 140	,511(**) 0,000 140	,715(**) 0,000 140	1 140				
Revenue (million HUF)	Pearson Corr. Sig. (2-tailed) N	,228(**) 0,007 140	0,161 0,057 140	-0,002 0,983 140	,182(*) 0,031 140	0,124 0,144 140	1 140			
Tickets sold	Pearson Corr. Sig. (2-tailed) N	,208(*) 0,014 140	0,13 0,125 140	0,009 0,912 140	0,128 0,131 140	0,068 0,424 140	,977(**) 0,000 140	1 140		
Tickets per screening	Pearson Corr. Sig. (2-tailed) N	,176(*) 0,038 140	0,143 0,091 140	-0,052 0,544 140	-0,032 0,705 140	-0,089 0,297 140	,814(**) 0,000 140	,837(**) 0,000 140	1 140	
Copies	Pearson Corr. Sig. (2-tailed) N	,189(*) 0,025 140	,180(*) 0,033 140	-0,005 0,951 140	,304(**) 0,000 140	,299(**) 0 140	,610(**) 0,000 140	,601(**) 0,000 140	,410(**) 0,000 140	1 140

**Regression coefficients for the explanatory model of a film's popularity among downloaders
(base=2007 and '08 Hungarian premiers)
 $R^2=0,536$**

	Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95% Confidence Interval for B	
	B	Std. Error				Beta	Lower Bound
(Constant)	503,56	144,63		3,48	0,00	217,49	789,62
No. of theaters where film screened	21,02	4,44	0,28	4,73	0,00	12,24	29,81
Copies (residual)	182,80	107,80	0,10	1,70	0,09	-30,43	396,02
Screenings (residual)	45,06	108,19	0,03	0,42	0,68	-168,94	259,06
Tickets sold (residual)	113,47	108,59	0,06	1,05	0,30	-101,31	328,26
Film life-span (residual)	-266,82	108,99	-0,15	-2,45	0,02	-482,40	-51,24
Torrent life-span (residual)	1193,79	109,40	0,65	10,91	0,00	977,41	1410,17

Dependent Variable: Number of downloads

**Path model of a film's popularity among downloaders
(base=2007 and '08 Hungarian premiers)
 $R^2=0,536$**



The above results from a regression equation and the path model present the causal relationship between downloading intensity (dependent variable) and five explanatory variables of cinematographic distribution. The model also includes a second peer-to-peer traffic variable, torrent life-span. Variables are entered into the final regression equation in their residual form (i.e., that part of their variance that is not explained by the other explanatory variables) where necessary.

The regression output confirms what we could only infer from the correlation matrix: **there is no causal relation between how a film fares at the box offices and how often it is downloaded.** Ticket sales and download intensity of a film do depend from a number of common factors among which the number of cinemas where the film screened is the single most important. But they can as well go in opposite directions. Further, the impact of a film's lifespan (the number of days of its being on cinemas' programs) on its downloads is negative: the longer a film is screened, the less it is downloaded. **Most popular for download are heavily marketed, widely distributed, short live movies that are not necessarily popular with cinema going audiences.**

The modeling of the causal relationship between a film's download traffic and its cinematographic distribution include two subversions. These regressions have been performed using catchment area as filter. Catchment area is defined as the 30 km radius circle around a user's location. A film is in the user's catchment area when it is on screen during the three-month timeframe. We have split the 2007-/2008 premier film population into two subsets: one including films screened inside, the other outside users' catchment area.⁵ The idea was to look for differences in the mechanisms affecting download intensity depending on whether downloaded films screen inside or outside the catchment area. (Regression coefficients and the path analyses are presented on the following two pages.)

The explanatory power of the model rises significantly for those downloaded films that are on screen inside users' catchment area (R^2 is 0,650 vs. 0,536 for the original model). **Cinematographic distribution explains more of the download traffic within the catchment area.** There is also a new causal relationship: the number of tickets sold (or the size of the cinema audience) does impact, albeit indirectly (via torrent files life-span) and to a very small extent on the number of downloads.

Outside the catchment area, the explanatory power of the cinematographic network decreases (R^2 is 0,512). Moreover, there is no relationship between torrents' lifespan and any of the cinematographic variables.

In conclusion, the catchment perspective corroborates previous findings on the double faceted relationship between a film's downloading trajectory and cinema life: **while most downloaded films are off-screen, on-screen content, and especially films within users' catchment area are intensifying downloads for a number of films with heavy media promotion. Cinematographic offer has that dual role in conditioning downloading activity in that while**

- its shortages expand its horizon
- its most marketed output define its focus.

Box office numbers however fail to explain peer-to-peer demand for movies—except within users' catchment area to a very limited extent. Peer-to-peer demand is triggered more by media presence than by actual audience size.

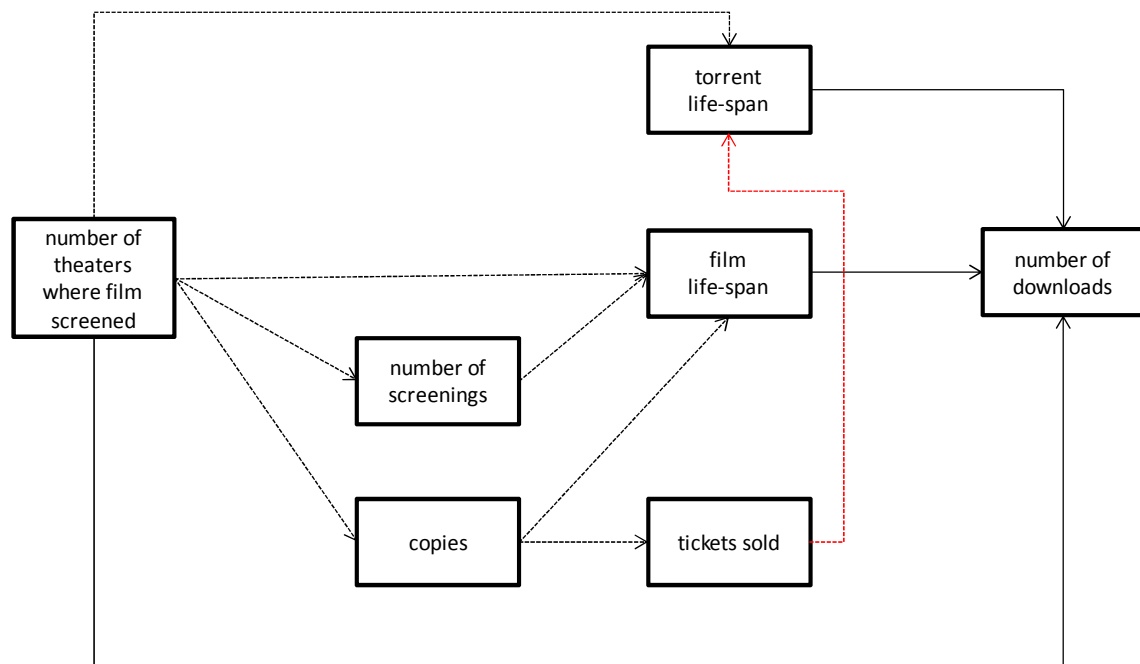
⁵ A third subset includes those users for whom no catchment area could be computed as we had no data on their location.

**Regression coefficients for the explanatory model of a film's popularity among downloaders
(base=2007 and '08 Hungarian premiers, downloaded films inside users' catchment area [30 km])
 $R^2=0,650$**

	Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95% Confidence Interval for B	
	B	Std. Error				Beta	Lower Bound
(Constant)	416,44	123,00		3,39	0,00	173,16	659,72
No. of theaters where film screened	22,82	3,78	0,30	6,04	0,00	15,35	30,29
Copies (residual)	151,84	91,68	0,08	1,66	0,10	-29,50	333,17
Screenings (residual)	49,76	92,01	0,03	0,54	0,59	-132,23	231,75
Tickets sold (residual)	125,51	92,35	0,07	1,36	0,18	-57,15	308,18
Film life-span (residual)	-207,30	92,69	-0,11	-2,24	0,03	-390,64	-23,97
Torrent life-span (residual)	1373,71	93,04	0,74	14,77	0,00	1189,69	1557,73

Dependent Variable: Number of downloads

**Path model of a film's popularity among downloaders
(base=2007 and '08 Hungarian premiers, downloaded films inside users' catchment area [30 km])
 $R^2=0,650$**

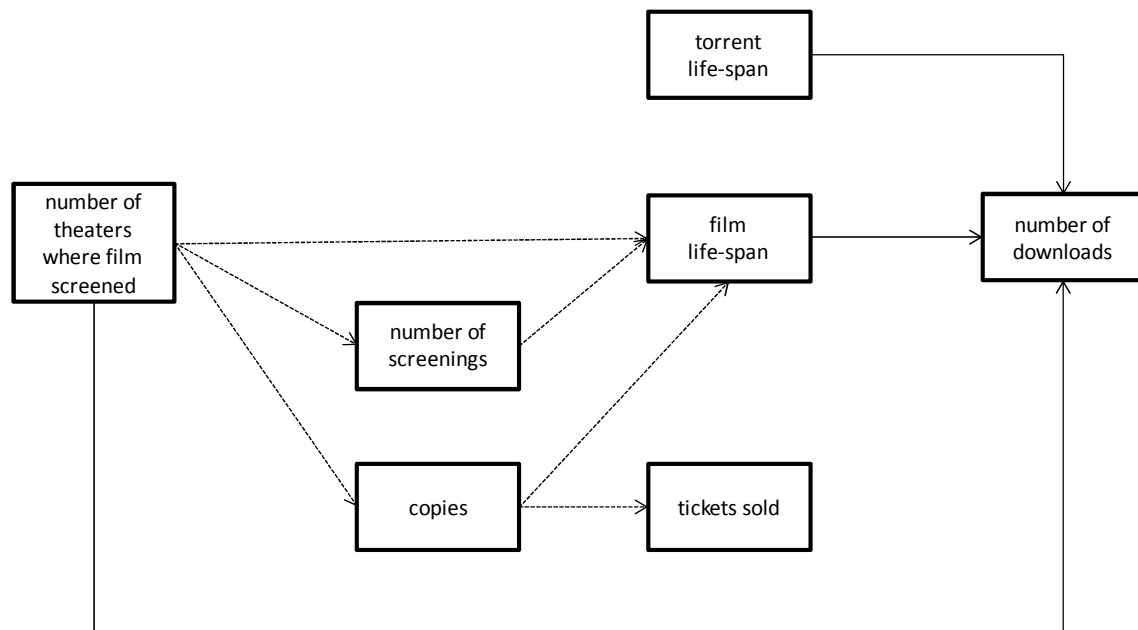


**Regression coefficients for the explanatory model of a film's popularity among downloaders
(base=2007 and '08 Hungarian premiers, downloaded films outside users' catchment area [30 km])
 $R^2=0,512$**

	Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95% Confidence Interval for B	
	B	Std. Error				Beta	Lower Bound
(Constant)	503,56	144,63		3,48	0,00	217,49	789,62
No. of theaters where film screened	21,02	4,44	0,28	4,73	0,00	12,24	29,81
Copies (residual)	182,80	107,80	0,10	1,70	0,09	-30,43	396,02
Screenings (residual)	45,06	108,19	0,03	0,42	0,68	-168,94	259,06
Tickets sold (residual)	113,47	108,59	0,06	1,05	0,30	-101,31	328,26
Film life-span (residual)	-266,82	108,99	-0,15	-2,45	0,02	-482,40	-51,24
Torrent life-span (residual)	1193,79	109,40	0,65	10,91	0,00	977,41	1410,17

Dependent Variable: Number of downloads

**Path model of a film's popularity among downloaders
(base=2007 and '08 Hungarian premiers, downloaded films outside users' catchment area [30 km])
 $R^2=0,512$**



References

- Alford, W. P. (1995). *To steal a book is an elegant offense : intellectual property law in Chinese civilization*. Stanford, Calif.: Stanford University Press.
- Az Elite Hub történelme. <http://www.4242.hu/hu/tortenelem>.
- b-bstf. (Summer 2004). A Guide To Internet Piracy, *2600 Hacker Quarterly*
- Bárdosi, M., Lakatos, G., & Varga, A. (2004). A kultúra helyzete Magyarországon. Budapest: Magyar Művelődési Intézet.
- Becker, J. U., & Clement, M. (2003). 'Generation Napster - Analysis of the economic rationale to share files in peer-to-peer-networks', *Wirtschaftsinformatik*, 45(3), 261-271.
- Becker, J. U., & Clement, M. (2006). 'Dynamics of illegal participation in peer-to-peer networks - Why do people illegally share media files?' *Journal of Media Economics*, 19(1), 7-32.
- Blomqvist, U., Eriksson, L.-E., Findahl, O., Selg, H., & Wallis, R. Trends in downloading and filesharing of music, *MusicLessons - Broadband technologies transforming business models and challenging regulatory frameworks – lessons from the music industry*.
- Bodó, B. (2006). '50bri J65k4 |)!9!t41', *Café Babel*(53).
- Bodó, B. (in press). *Necessity know no laws*. Budapest.
- Bodó, B., Halácsy, P., Korsós, M., Prekopcsák, Z., & Szalai, A. (2007). P2P hálózatok vizsgálata. Budapest: Kitchen Budapest Medialab.
- Borsos, Á. (2007). A mozi
- Borsos, Á. (2007. november). 'Tudatos fejlesztés vagy spontán folyamatok? A magyar mozihálózat az új évezredben'.
- Chu, J., Labonte, K., & Levine, B. N. (2002). 'Availability and locality measurements of peer-to-peer file systems', *In Proc. of ITCOM: Scalability and Traffic Control in IP Networks*.
- Condry, I. (2004). 'Cultures of music piracy: an ethnographic comparison of the US and Japan', *International Journal of Cultural Studies*, 7(3), 343-363.
- Cooper, J., & Harrison, D. M. (2001). 'The social organization of audio piracy on the Internet', *Media Culture Society*, 23(1), 71-89.
- Cooper, M. N. (March 2005). Time For The Recording Industry To Face The Music: The Political, Social And Economic Benefits Of Peer-To-Peer Communications Networks (Vol. 1, pp. 79): Consumer Federation Of America.
- Cserta, O. (2002). 'A kulturális piac (ki)alakulása magyarországon', *Statistikai Szemle*, 80(5-6).
- Dejean, S. (2008). *What Can We Learn from Empirical Studies About Piracy?: SSRN*.
- Digital Music Report 2006. (2007). IFPI.
- Domon, K., & Nakamura, K. (2007). 'Unauthorized Copying and Copyright Enforcement in Developing Countries: A Vietnam Case Study', *Review of Economic Research on Copyright Issues*, 5(1), 87-96.
- The Economic Impact of Counterfeiting and Piracy. (2008). OECD.
- Freeman, P. (November 2008). Getting it on the download, *The Wire*. London: The Wire Magazine Ltd.
- Giesler, M., & Pohlmann, M. (2003). 'The Social Form of Napster: Cultivating the Paradox of Consumer Emancipation', *ADVANCES IN CONSUMER RESEARCH*, 30, 94-100.
- Givon, M., Mahajan, V., & Muller, E. (1995). 'Software Piracy - Estimation of Lost Sales and the Impact on Software Diffusion', *Journal of Marketing*, 59(1), 29-37.
- Gopal, R. D., Bhattacharjee, S., & Sanders, G. L. (2006). 'Do artists benefit from online music sharing?' *Journal of Business*, 79(3), 1503-1533.
- Gopal, R. D., & Sanders, G. L. (1998). 'International software piracy: Analysis of key issues and impacts', *Information Systems Research*, 9(4), 380-397.

- Gu, B., & Mahajan, V. (2004). The Benefits of Piracy – A Competitive Perspective, *Sixteenth Workshop on Information Systems and Economics*.
- Gummadi, K. P., Dunn, R. J., Saroiu, S., Gribble, S. D., Levy, H. M., & Zahorjan, J. 'Measurement, modeling, and analysis of a peer-to-peer file-sharing workload'.
- Guo, L., Chen, S., Xiao, Z., Tan, E., Ding, X., & Zhang, X. 'Measurements, analysis, and modeling of bittorrent-like systems'.
- György, P. (2005). *Kádár köpönyege*. Budapest: Magvető.
- Howe, J. (January 2005). The Shadow Internet *WIRED* (Vol. 13): Condé Nast Publications.
- Huang, C. Y. (2005). 'File sharing as a form of music consumption', *International Journal of Electronic Commerce*, 9(4), 37-55.
- Hunter, S. D., & Spitz, D. (2003). The Social Construction of Napster: MIT Sloan Working Paper No. 4445-03.
- Huygen, A., Rutten, P., Huveneers, S., Limonard, S., Poort, J., Leenheer, J., et al. (18 February 2009). Ups and downs - Economic and cultural effects of file sharing on music, film and games, *TNO-rapport*. Delft: TNO Information and Communication Technology.
- IFPI. (2001). IFPI Music Piracy Report.
- IFPI. (2006). The Recording Industry 2006 Piracy Report: IFPI.
- Kanzler, M. (ed.) (2009). *focus 2008*: European Audiovisual Observatory.
- Keenan, D. (November 2008). Burning Chrome, *The Wire*. London: The Wire Magazine Ltd.
- Lendvai, Z. (2008. Június). 'Szerzői jog az ókorban', *Iparjogvédelmi és Szerzői Jogi Szemle*, 3 (113)(3).
- Liebowitz, S. J. (2006). 'File Sharing: Creative Destruction or Just Plain Destruction?' *Journal of Law & Economics*, 49(1), 1-28.
- Manuel, P. L. (1993). *Cassette culture : popular music and technology in north India*. Chicago: University of Chicago Press.
- Marshall, L. (2004). 'The effects of piracy upon the music industry: a case study of bootlegging', *Media Culture & Society*, 26(2), 163-+.
- Oberholzer-Gee, F., & Strumpf, K. (2007). 'The effect of file sharing on record sales: An empirical analysis', *Journal of Political Economy*, 115(1), 1-42.
- Peitz, M., & Waelbroeck, P. (2006a). 'Piracy of digital products: A critical review of the theoretical literature', *Information Economics and Policy*, 18(4), 449-476.
- Peitz, M., & Waelbroeck, P. (2006b). 'Why the music industry may gain from free downloading - The role of sampling', *International Journal of Industrial Organization*, 24(5), 907-913.
- Pouwelse, J. A., Garbacki, P., Epema, D. H. J., & Sips, H. J. (2005). The Bittorrent P2P File-sharing System: Measurements and Analysis, *4th Int'l Workshop on Peer-to-Peer Systems (IPTPS)* (Vol. 3640): LNCS.
- Rob, R., & Waldfogel, J. (2006). 'Piracy on the high C's: Music downloading, sales displacement, and social welfare in a sample of college students', *Journal of Law & Economics*, 49(1), 29-62.
- Rojek, C. (2005). 'P2P Leisure Exchange: Net Banditry and the Policing of Intellectual Property', *Leisure Studies*, 24(4), 357-369.
- Saroiu, S., Gummadi, P. K., & Gribble, S. D. (2002). A measurement study of peer-to-peer file sharing systems (Vol. 2002, pp. 152).
- Schulze, H., & Mochalski, K. (2008). Internet Study 2007: Ipoque.
- sct. (2009 február 01. 15:20). Boncasztal vol. 9 - bitHUMen
- Sen, S., & Wang, J. (2002). Analyzing peer-to-peer traffic across large networks (pp. 137-150): ACM New York, NY, USA.
- sharky. (December 17, 08). Invites: The Top 50 Most Requested Private Trackers <http://filesharefreak.com/2008/12/17/invites-the-top-50-most-requested-private-trackers/>.
- Sheikh, A. E., Rashed, A. A., Qudah, B. A., & Peace, A. G. (2006). 'An Exploratory Study of Software Piracy in Jordan', *The Electronic Journal of Information Systems in Developing Countries*, 23.
- Strahilevitz, L. J. (2003). 'Charismatic code, social norms, and the emergence of cooperation on the filesapping networks', *Virginia Law Review*, 89(3), 505-595.

- Szénay, M. (2005). Az idegennyelv-ismeret Jelentés az országos nyelvtudás-felmérés kvantitatív szakaszáról. In L. Lada (Ed.). Budapest: Nemzeti Felnőttképzési Intézet.
- Turcsán, T. (2008. február 7.). Cseresatározás Letöltők versus hatóságok, *Figyelő*.
- Zentner, A. (2006). 'Measuring the Effect of File Sharing on Music Purchases', *Journal of Law & Economics*, 49(1), 63-90.