Film Factory Losses: is BitTorrent a Major Responsible?

3 authors:

Praboda Rajapaksha
Institut Mines-Télécom
10 PUBLICATIONS **21** CITATIONS

SEE PROFILE

Reza Farahbakhsh
Institut Mines-Télécom
71 PUBLICATIONS **379** CITATIONS

SEE PROFILE

Noel Crespi
Institut Mines-Télécom
402 PUBLICATIONS **2,861** CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:

Project   FUSE-IT View project

Project   Fake News and Bot Detection View project

Film Factory Losses: is BitTorrent a Major Responsible?

Praboda Rajapaksha*†, Reza Farahbakhsh*, Noël Crespi*

*Institut Mines-Télécom, Télécom SudParis, CNRS Lab UMR5157 Evry, France
†Uva Wellassa University, 90000, Badulla, Sri Lanka
{praboda.rajapaksha, reza.farahbakhsh, noel.crespi}@telecom-sudparis.eu

Abstract—Online piracy is an important challenge in the motion-pictures industry. Several studies claimed that unauthorized content in online venues are reducing substantially the BoxOffice revenues while few other studies were not in favor of this claim. To understand better the impact of this phenomena, a study based on a large dataset is needed to analyze how different portals (e.g. BitTorrent) permitted movie downloads that help to increase revenue movie industry and it is also important to identify what type of movies were mainly affected. This paper aims to answer this question based on a dataset containing almost 15 million records from around 3.25 million torrents' data collected from the BitTorrent portal and their detailed movie related records extracted from IMDB. In this study, we observe (i) the impact of online movie downloads on movie revenues, which predominantly affect on low budget and independent movies, (ii) the correlation between screen period of movies in cinema to the availability of torrents, (iii) the fake torrents that are injected to the portals before and during the screen period of a movie. Apart from that this work analyses the movie viewer's feedbacks gathered from a questionnaire survey on user opinion and experiences with online downloads. We explored that, people used to be aware more about the online downloads and their related portals after introducing anti-piracy laws and starting to shut down different portals. We also identified several other ways that can help to reduce online download rates of movies.

Index Terms—BitTorrent, motion pictures industry, movie's BoxOffice, online piracy, anti-piracy law.

I. INTRODUCTION

There exists a number of online portals such as BitTorrent allowing people to share and download media content, for instance video (movies, music, TV shows etc.) and audio files over the Internet. Previous studies have shown that a large portion of content available in these portals are videos [1] and unauthorized [2]. The main after-effects of sharing media files via these portals is streaming or downloading without any charge or at a minimal cost. As a result, movie industry players arguing that their revenues are declining with the distribution of pirated copies in an online venues. As a consequence, the impact of unauthorized multimedia content shared in online platforms is a hot and debatable topic. In the past decades, number of researchers have analyzed the impact of online media content downloads to the movie revenue. Few of them reported that this impact was very low [3] while others proved considerable impacts on the movies' boxoffice revenue. In addition, few other studies were argued on the benefits of online file sharing (e.g, distribution of the pre-release hype of a movie) and they said this has increased the attraction of the movie viewers to visit the movie in cinema [4].

In this paper, we investigate the impact of keeping movies in online platforms and allowing to download them and, the effect on their budget and revenues. We model movies revenue losts with respect to the non-copyrighted compliant downloads. However, trying to study all non-copyrighted content is too ambiguous as there exists many different online file sharing platforms, such as Cyberlockers (e.g. Uploaded, WUpload, Letitbit, etc) and streaming services (e.g VideoZZ). Hence in this study, we only focus on BitTorrent, one of the most successful file-sharing technologies during last decade.

We implemented an advanced data crawler in-order to collect very large dataset from BitTorrent, which monitors available torrents in 18 major trackers for a period of 47 days. There are 15 million(M) torrent records collected and among them, 3.25M of them are unique torrents. We filtered and identified 926K qualified torrents that included 241K torrent files associated to the movies category. Their infohashes are used to crawl TorrentZ Web portal to collect more information about each movie. BoxOffice information (production budget, BoxOffice revenue, etc.) of each identified movie is also collected from IMDB portal and other available sources to study correlation between movies' downloads with their budget and worldwide gross. To sum up, in this study we implemented four different crawlers to obtain required information. In addition, we conducted one questionnaire survey among 500 participants to understand the their awareness about online piracy and what can be a motivation for them to have legal access to the movies instead of downloading them unlawfully.In summary, the results of this study shows the impact of movie downloads vs movie revenues. Additionally, we identified that very low budget movies were more affected due to online downloads.
The main contributions of this work are:

(i) we propose a methodology to identify the impact of downloads on the film industry and their revenue. This methodology can be generalized across other domain such as music and software market.

(ii) a large dataset of torrents associated to movies, music, and software, is collected (available for further research[1]).

---

[1]The dataset will be available to the community in the camera ready version.

(iii) a set of observations that helps to mitigate movies company lost by the solutions extracted from the feedbacks received from people to a large questionnaire.

II. RELATED WORK

A large and growing body of literature have investigated the impact of online piracy on different sectors including movie and music markets. However, most of these studies were based on small datasets collected from questionnaires or/and surveys and only a few of them have used a real dataset collected from movies portal. Based on our knowledge, limited number of studies have considered very large amount of information to analyze the music industry, DVD trades, and box-office gross. This section focuses on exploring previous studies on the impact of pirated content on revenues and solutions to reduce this impact.

A. Analysis on the impact of piracy

Online piracy is associated with sharing copyrighted content without authors consent in the Internet. Researchers use various approaches to find the relationship between legal and piracy content. In a recent work [5], authors have measured the relative impact of movie piracy on motion picture industry and shown that, if piracy could be eliminated then the BoxOffice revenue would increases by 15%. In another study [6], authors identified that unauthorized CAM copies downloads reduces their box-office sales 19.1% reduction in the legitimate sales. Hollywood movies are released first in US and subsequently returned for further screening in other countries. BitTorrent and other similar portals distribute those movies illegally to the other countries before their official release. An analysis to this phenomena is the study carried out by [7], across 17 countries, found that longer international release lags decline box-office returns by at least 7% in the presence of pre-release of movies. In a very recent work [8], authors studied emerging streaming cyberlocker ecosystems and they discovered these environments actively involved in copyright infringement with an aggressive injection of recent releases.

The study done by Marc R. [9] used multiple data sources to collect more information about movies BoxOffice related information and unauthorized CAM copy downloads from BitTorrent. Their results proved that CAM downloads have a low probability of equating to a lost sale. Another study [10], looked at the influence of online file sharing on the movie theaters targeting on the Wolverine movie which was available in the online platforms 4 weeks before its theatrical releases. Their results indicated that downloads of this movie prior to its release shows positive impact on the box-office sales.

B. Taken actions to reduce the impact

Several major actions have been taken by governments and major content providers and companies to reduce the availability of the illegally published copyrighted content in online platforms. As an example, Danaher et al., [11] investigated the impact of ceasing the major piracy cites such as Megaupload and associated Web portals that are used for

movie sales. Their study shown that, because of these close down events, revenues of three major motion picture company increased by 6.5-8.5%. Apart from that, the study [2] identified how antipiracy actions effect on online piracy by considering closure of Megaupload and the implementation of the French antipiracy law (Hadopi). They have shown that Megaupload closure diminished the activities of professional BitTorrent publishers who are running their own BitTorrent portals.

Considering the stated literature, this paper aims to look in to this problem from a deeper perspective by using a large scale dataset collected from several relevant sources including BitTorrent and IMDB. We use a novel methodology to measure this impact and provide some guidelines on how this impact can be mitigated based on the survey results.

III. DATA COLLECTION METHODOLOGY

This section describes our data collection methodology and provide details about different steps of preparing the final dataset including filtering the raw data, classification of movies, and collection of movies meta data.

A. Raw data collection

In order to collect movies information from BitTorrent, we have implemented an advanced crawling tool which connects to different trackers[2] using trackers' scrape mode. Two following filtering steps are executed to select more suitable tracker list from more than 100 unique trackers.

Firstly, we investigated that a group of URLs are referring to the same tracker (analyzed using dig utility), for instance *publicBT* tracker refers to a wide number of URLs. Therefore, in this study, we first filtered those individual trackers. Secondly, we excluded a group of trackers those who do not permits to use scrape mode as our crawler is failed to connect with them. Also, we excluded some of the trackers who provide their scrape file in a very strange codification as these files could not be analyzed in this study. Finally, we queried filtered trackers and received in a file all information about its torrents list. In particular, for each torrent in this list, we can obtain: i) infohash (a unique ID of a content file); ii) number of downloads; iii) number of seeders at present (i.e. those who are completed its download and in the swarm to send pieces of the content to others); and iv) number of leechers at present (i.e. those who are not completed its download).

Our crawling tool runs 6 times per day in every 4 hours to collect scrape file information (around 15 million values for 3.25M infohashes) and store the values of downloads, seeders, and leechers. As a summary, table I lists the names of 18 trackers, number of torrents, and number of movies available for each tracker (identified in Section III-C). Next step is to collect information of each torrents that shows name of the file and type of content (movie, music, etc.) of each torrent.

---

[2]"Tracker" keeps track of where file copies reside on peer machines and helps coordinate efficient transmission and reassembly of the copied file.

Tracker-Name	#Torrents	%Movies
blazing	9,783	31
podtropolis	49,698	40
eztv	580,555	25
firebit	8,614	41
publicbt	579,581	25
vtv	579,354	25
harrry	568,022	25
torrentbay	238,897	32
fr33dom	184,499	32
xxx	34,537	15
exodus	279,092	29
opensharing	24,623	62
yify	11,000	47
hotplug	31,577	38
9you	6,498	34
elitezones	2,218	17
mytracker	12,741	37
anime	51,058	21
Total	**3,252,347**	**Avg=32%**

Category	#Torrents	%Torrents
Movie	241,129	26.03
Video	240,956	26.01
None	240,652	25.98
Audio	92,321	9.97
Software	43,369	4.68
Games	30,238	3.26
Ebooks	27,747	3
Porn	1,078	0.12
Pictures	8,931	0.96
Total	**926,421**	**100**

video TV, Game, audio, software, picture, porn) etc. As an example, queried details belongs to the infohash: *0000bf2359476d55166d891c22d61cc48af6df1a* are; infohash name: *"The Girl With The Dragon Tattoo" (2011)*; torrent category: *movie*; added date in the portal: *2012-01-18*.

Table II summarizes the distribution of torrents based on the obtained data from TorrentZ portal. It shows a large number of torrents (more than 50%) are from movies and videos categories. The analysis of our study is based only on the movie category which includes $241,129$ movie torrents. As a potential future work, the research question of this paper can be studied in other categories such as audio and Software.

B. Data filtering

We have collected 3.25M unique torrent information and among them, large number of torrents attracted very few users (in many cases number of downloads, seeders and leechers are 0). Therefore, to analyze torrents with reasonable amount of metadata, following policies are used to filter torrents that are having considerable amount of downloads, seeders and/or leechers. i) Infohashes>500 downloads, ii) Torrents>10 seeders/leechers if #downloads is not available. These steps filtered only 926K torrents.

C. Modeling number of Downloads

As was mentioned earlier, tracker's response includes #downloads, #seeders, and #leechers. However, some of the trackers (4 out of 18 in our study) do not provide real #downloads as they reset #downloads frequently, but gives correct #seeders and #leechers. However our inspections on the snapshots with valid #downloads shows that #leechers in each timestamp is almost equal to #downloads. Therefore we considered #leechers as #downloads for torrents those who not provided #downloads. We also use following steps to calculate #downloads when a snapshot missed the value of #downloads but, have valid #downloads counts previously. i) If #downloads is zero for the present crawling period, then the previous value (snapshot of 4 hours ago) is considered. ii) If #downloads is smaller than previous value (when the tracker reset #downloads), sum of those two values are considered.

D. Content classification

In order to differentiate the type of the content (e.g. Film, TV Show, Game, etc) of torrents another crawling application is implemented to query Torrentz portal indexing the infohashes. This tool collects different information such as name, date when it was first available in one of the torrent portals, type (e.g. movie, video,

E. Collecting movies information

We collect movie budget and revenue information by using available APIs of IMDB portal (filter by film name or by name or imdb-ID)) for each of the identified movie. Further, we queried the details of each movie including IMDB unique ID, movie title, production year, released date, country, director, actors, etc. In addition, another crawler executed to collect box-office information from: (i) "IMDB business" portal, (ii) "The-Numbers" portal, and (iii) "Worldwide box-office"portal. All box-office information acquired from above web portals are amalgamated to gather accurate information. This approach allows to collect complete box-office information for each IMDB ID such as movie's budget, domestic and overseas sales, worldwide gross revenues, domestic DVD sale and released date of the movie.

F. Final Datasets

Table III summarizes our dataset collected during the crawling period of 47 days starting from 2012-02-17 and we categorize movies into the 3 groups: during, after and before.

(1) *During*: All movies that are released during 47 days of our crawling period (2012-02-17 to 2012-04-04) and all associated torrents to these movies that are uploaded after movie's released date (1,240 torrents, 81 movies).
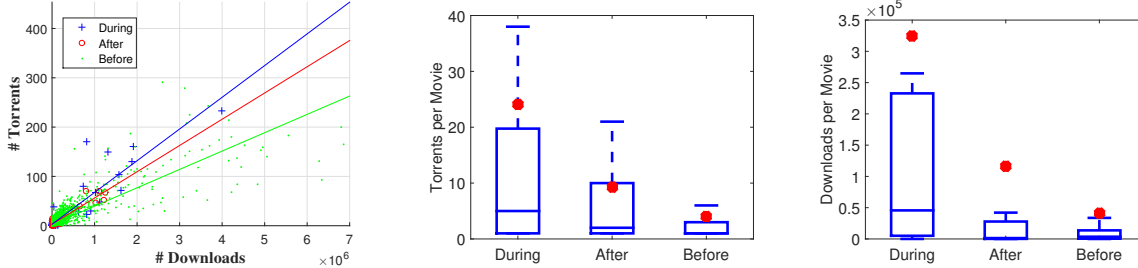
(2) *After*: All movies that their released date is after stop crawling (682 torrents for 124 movies). This category includes all torrents uploaded before the movie is released and our investigation shows a large portion of this torrents.

(3) *Before*: All movies that their released date is before starting the crawling period (57,578 torrents, 14,618 movies).

There are 2 types of torrents each movie can belongs to and the main focus of this study is based on the 2nd category of

TABLE III
MAIN FIGURES OF THE DATASET AFTER FILTERING AND CONTENT IDENTIFYING STEPS

Facts	Figures
Start date of data collection campaign	17/02/2012
Crawling period (#days)	47
#snapshot collected per day	6 (Q4h)
#trackers	18
#snapshots (total collected infohashes)	15,079,905
#Unique torrents (unique infohashes)	3,252,347
#torrents (after filtering step)	926,421
#torrents corresponding to movies (TorrentZ step)	241,129
#torrents corresponding to movies (IMDB step)	59,500
#movies (unique movies identified from IMDB step)	14,823
#movies (& #torrents) in *During* category -"released during the crawling period"	81 (1,240)
#movies (& #torrents) in *After* category -"released after the crawling period"	124 (682)
#movies (& #torrents) in *Before* category -"released before the crawling period"	14,618 (57,578)

(a) Linear distribution of #torrent vs #download   (b) Distribution of #torrents per Movie   (c) Distribution of #download per Movie

Fig. 1. Availability of torrents (#torrents) vs. popularity of torrents (#downloads) across movies (the red dots show the average value)

torrents mentioned below. (i) Torrents that are uploaded in the BitTorrent portal before releasing the movie. (ii) Torrents that are uploaded after movie's released date.

IV. ANALYSIS OF DOWNLOADS VS. MOVIES' BOX-OFFICE

In this section we investigate how movie downloads affect on movie's worldwide gross and production budget.

A. Availability vs. Popularity

First, we look to the relationship between availability of the torrents in BitTorrent and their popularity (#downloads) for *During*, *After*, and *Before* categories as depicted in Figure 1. Figure 1(a) shows the linear relationship between total number of movie downloads and torrent uploads. The first observation is that, many movies with large #Downloads (>4M) are from the torrents belongs to *Before* category (r=0.84). Based on the linear regression analysis, r^2 for *Before* category is 0.706, which indicates that 70.6% of the total variation in #torrents can be interpreted by the linear relationship between #downloads and #torrents. This is reasonable because when a movie is released higher chance of increasing the number of torrents uploaded to the BitTorrent . Figures 1(b) and 1(c) depicted distributions of #torrents per movie and #downloads per movie. In *During* category there were more uploaded torrents and more movie downloads in compared to other 2 categories. This indicate that recently released movies are more popular among content publishers and consumers. On average, 24 torrent files are uploaded to BitTorrent in *During* category which is almost 3 times more than the *After* category (9 Torrents in Average) and 6 times more than the *Before* category (4 Torrents in average). However, we also found that

the top 3 movies with highest number of torrents are from the *Before* category (*"Journey 2: The Mysterious Island"* - 291 *"Ghost Rider: Spirit of Vengeance"* - 279 *"Mission: Impossible - Ghost Protocol"* - 243). Another interesting observation in 1(a) is that many movies having at least 200 torrents per movie, shows less than 2M downloads and few other movies downloads count goes up to 7M. As shown in Figure 1(c) most movie downloads occurred in *During* category (avg - 0.35M downloads). Which means that, many movie are downloaded just after the release. As indicated in Table III, #movies and #torrents in *Before* category is the highest compared with other categories but, average movie download count in this category is 411k, which is ~10× and ~3× lower than *During* and *After* categories, respectively. Maximum #downloads per movie is from *Before* category around 6.9M, and these are the widely popular and already released movies.

To summarize, although newly released movies were popular (#downloads) and available (#torrents) in BitTorrent some other movies are very popular yet not released.

B. Worldwide Gross vs. Budget and Downloads

This section focuses on the impact of movie downloads to the worldwide gross, considering their initial production budget. Figure 2 shows a 3D plot of the movies worldwide gross in compared to #downloads and their budget. This graph clearly indicates the effect on worldwide gross in different movies with respect to their #downloads and budget. We noticed that movies with high worldwide gross were the once with very large budget and few #downloads. This indicates
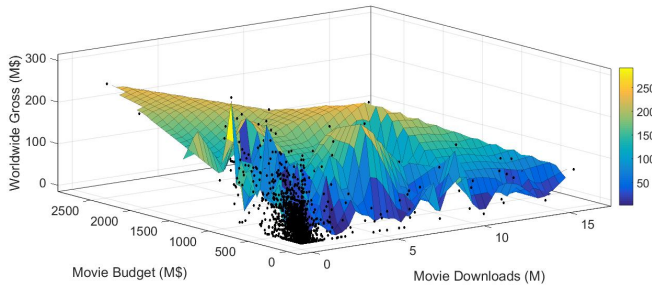
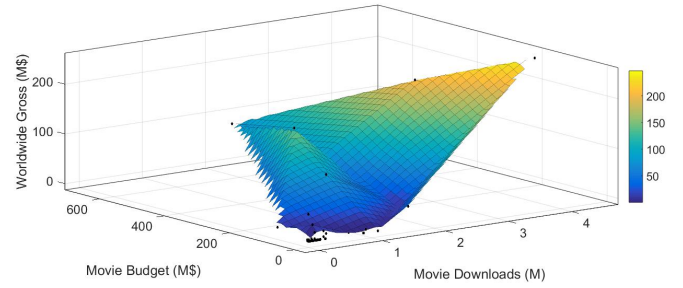Fig. 2. WW Gross vs. budget and #downloads for all categories

Fig. 3. WW Gross vs. budget and #downloads for *During* category

that movies with low budget and/or less #downloads have much lower worldwide gross and many downloaded movies were less costly (independent movies). Whereas, regression coefficient between movie budget and worldwide gross is shown to be, r^2=0.435, where only 43.52% of the movie's worldwide gross can be predicted using its budget. One of the reasons behind this pattern is the copyright control of movie production companies on popular movies with the goal to reduce the availability of the movies in online platforms. Besides these facts, regression between movie's worldwide gross and movie downloads appeared to be 4.7% and there is no linear relationship between these two parameters.

In summary, low budget movies were downloaded more than others and had very low worldwide gross, and their revenues impacted a lot due to online downloads.

On average, many movie downloads and torrent uploads are from the *During* category. Therefore, in-order to have a better understanding of the above analysis, we look more in detail to the *During* category and pictorially presented in Figure 3. In contrast with the analysis performed on all categories (Figure 2), results from *During* category presents that (Figure 3), higher the number of downloads per movie higher the worldwide gross (r^2=0.735), but still a many low budget movies were also shown a large #downloads.

V. POTENTIAL LOSS ANALYSIS

This section analyses movies losses (absolute and relative losses) based on BitTorrent downloads in the *During* category.

A. Absolute Losses

In our hypothesis to study the absolute loss we assume one BitTorrent download as a potential audience who could pay for a movie in cinema. Therefore BitTorrent downloads are proportional to $\sum_{j=1}^{n} X_i^j$; where cinema viewer i visit theater for j^{th} time (up to n times) and this is equivalent to a user who may view the same movie multiple times from online platforms. The analyses are focused on the *During* category as many downloads during the period of screening the movie and few month later. Despite that, some movies in the *During* category continue their screening after our crawling period, but 62% of the identified movies in this category were released before the first half of the crawling period. As a result, the probability of identifying movie download patterns is very high in the first few days just after the release date.

According to the statistics of the The Hollywood Reporter, average cinema ticket price during the second quarter of 2012 (our crawling period) was 8.12$. Hence, we consider 8.12$ as the ticket price of movies in-order to calculate total amount of losses using above hypothesis as follows:

$$Absolute\ Loss = \sum_{j=1}^{n} X_i^j \times Ticket\ Price$$

Figure 4 elucidates the distribution of the absolute losses for the *During* category. We assume that, each user stream/download any video only once (j=1). The average movies potential loss is around 4.81M$ in the *during* category. However, several movies appeared to have a huge loss; e.g. *"John Carter"* movie had the largest absolute loss, around 32M$ (Maximum value), and 4M online downloads, where the loss of this movie can be considered as approximately 1/8 of its worldwide gross. This signifies that, even though a movie had high revenue, still there is a possibility to increase its income by reducing online downloads.

In summary, movies can increase their income if people visit cinema instead of downloading them from online portals and online portals can charge from the users per stream/download by keeping them as piracy that permits to increase income.

B. Relative Losses

Absolute losses analysis was not focused on how harmful BitTorrent is and it is important measure the relative impact of BitTorrent on movies. We define *HIR* (Harmful Income Ratio) in-order to understand the relative impact of the BitTorrent downloads on movie revenues.

$$Harmful\ Income\ Ratio\ (HIR) = \frac{AbsoluteLoss}{ActualGross}$$

Actual Gross = Worldwide Gross - Movie Budget

Negative values of *HIR* metric represent movies with very large revenue losses and their worldwide gross is less than its budget. On the other hand, positive value of *HIR* represent movies of the reasonable or high worldwide gross and may have higher absolute loss or lower actual gross compared to other movies. If the *HIR* value is approximately equal to 0, then the movies have considerably large worldwide gross regardless of the #downloads. Also, if the *HIR* is 1, then one movie has a similar actual gross and absolute loss values indicating that the movie has a lost equal to its revenues. We
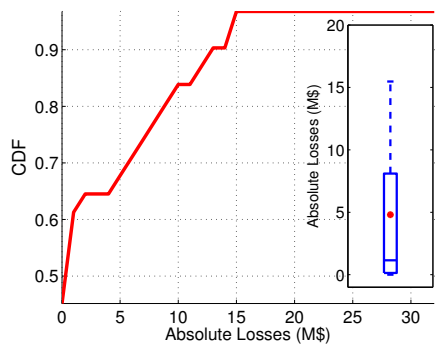
Fig. 4. Distribution of Movie's Viewer Losses

Fig. 5. Harmful Income Ratio (HIR) vs (a)#downloads, (b)Movie budget (M$), (c)Worldwide gross (M$)

calculated *HIR* for the most affected movies in the *During* category identified from their absolute losses and analyzed the correlation between *HIR* with #downloads, worldwide gross and movies' budget. The results are shown in Figure 4.

Figure 5 illustrates the relationship between movie's budget, #downloads, worldwide gross, and *HIR* in a 4D graph. The *HIR* values are represented by different colors in the graph and shows the variability of *HIR* with other variables. There are few movies with negative *HIR* values and positive *HIR* values. However on average, *HIR* value in many movies in the *During* category is almost zero (mean: 0.24). We also can identify the maximum *HIR* value (4.8) having a huge absolute loss and very low actual gross belongs to *"London Paris New York"* movie. This movie had 1.7M$ worldwide gross during the crawling period with 1.3M$ budget and having only 0.25M #downloads. Even though #downloads of this movie is less than the average #downloads in *During* category (0.32M) absolute loss is higher than the actual gross. The next largest positive *HIR* value is belongs to *"Posti In Piedi In Paradiso"* movie and followed by *"London Paris New York"* movie.

On the other hand, the most negative *HIR* value (-1.43, the darkest blue marker in Figure 5) represents *"Chaar Din Ki Chandni"* movie. This can be considered as an unpopular movie. The worldwide gross of this movie is less than the budget and during the crawling period #downloads was 68,142. The *"John Carter"* movie is having its *HIR* value almost equal to 1. The highest number of downloads (4M), the largest budget (250M$) and around 283M$ worldwide gross is from this movie. As its *HIR* is equal to 1 its absolute loss and actual gross are almost equal.

In summary, the mean *HIR* for many movies from the *During* category is 0.24 indicating that they have very high actual gross compared to absolute loss and movies that were released before crawling period obtained considerable amount of revenues regardless of the #downloads.

VI. QUESTIONNAIRE AND CONCLUDING REMARKS

To better understand user opinions and experiences about online downloads, we conducted a survey (https://goo.gl/Mk6iiZ) and asked people about their download activities. This section briefly presents some of these findings and results.
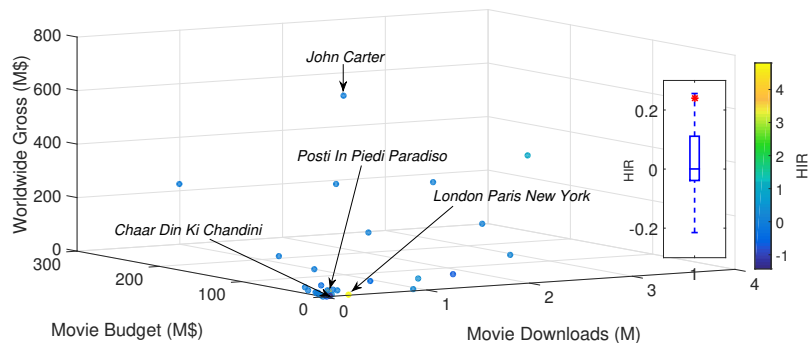
A. Findings from the people feedbacks to the Questionnaire

The survey received around 500 replies from people across different; countries, age groups, and internet expertise levels. The main objective of this survey was to: (i) understand user download patterns from online sources (ii) understand user interests on visiting to cinema or willingness to pay for a video content instead of downloading illegally, (iii) and finally have a vision on the awareness of people about anti-piracy events related to illegal content availability and sharing.

The main findings from this survey are as follow. First of all, we observed very high awareness among users about antipiracy activities conducted to reduce the amount of available unauthorized content. This observation comes from people awareness to the Megaupload closure (on 19 January 2012). Megaupload was one of the largest online file sharing application which caused to reduce revenue of the entertainment industries about US$500 millions. According to the statistics, the shutdown of Megaupload and its associated sites caused digital revenues for major motion picture studios to increase by 6.5-8.5% [11]. Survey results shown that most of the users (around 68%) had basic knowledge about the Internet and was aware about Megaupload closure and more than 84% of them are from age<50 years. Among all the people who attended our survey in the age group <20 years were aware about the Megaupload closure. Moreover, 97% of the surveyed people are sensible on this closure and they have used Megaupload to download multimedia content. Also, all French nationalities take part in this study had the knowledge on the anti-piracy law introduced by the French Parliament namely HADOPI. In summary, these findings indicate that enforcing this type of law/rules for pirated contents is positive effect to limit non-paid consumption.

Based on the results obtained from this survey, we found many people refused to download media content illegally from the Internet due to different laws established in their countries. More than 53% of the advanced and expert Internet users were not downloaded illegitimate content due to the punishments undertaken (e.g. Internet disconnection, high fines, sending to the prisons, etc.).This shows introducing punishment rules in country level can increase people concern and help to reduce #downloads as detailed in [2]. Almost half of the participants

Price ($)	0	1	2	3	4	5	6	7
#viewers	35	22	39	67	69	116	32	76
Probability	1	0.92	0.88	0.79	0.64	0.49	0.24	0.17

Fig. 6. Overall movie analysis of the *During* category for different cinema ticket prices

was not considered that the Internet disconnection was as a major reason for not downloading from online portals and 25% of them claimed that, nothing can prevent their downloads.

In summary, result shown that people are willing to download online multimedia content but they refuse to undertake any punishment. One approach is to allow consumer to download/view multimedia content paying less amount of money than the actual prices i.e. Netflix. This benefit for both content provider and consumer and content provider can deliver services to the registered users (who paid membership fee) and they can also limit download capacity of the files or number of downloads per some period (one month) for each user. Moreover, ad-supported streaming and downloads are able to monetize free consumption behaviors.

B. How to increase cinema viewers?

One possible solution to attract cinema viewers is by observing their preferred ticket prices. We have collected these information from our survey and summarized results are shown in Table IV including #movie viewers for different ticket prices. Among them, 25% of the users were preferred to visit cinema if the ticket price is 5$ (ticket price range: 0-7) without downloading the movie. We use following formula to analyze how to increase movie revenues by increasing #cinema viewers and reducing online downloads in *During* category.

Movie Revenue = Gross + X% × #downloads × Ticket price

Where X represents cumulative probability of #cinema viewers for different ticket prices where we assumed that if a user is willing to pay 7$ then she visit cinema for lower prices as well. With this assumption, we calculated total movie revenues ranging the ticket price from 0$ to 7$ and results are shown in Figure 6. It demonstrates the fluctuation of the revenue for all movies in *during* category with reference to different ticket prices. As shown, the largest mean value and maximum revenue can be observed when the ticket price equals to 4$ and due to this average movies income is increased by 13.56%. Similarly, a small variation to this result can be observed for 3$ (by 12.48% increment) and 5$ (by 12.94% increment). Thus, we can conclude that movie income decreases if the cinema ticket price is very low and very high and keeping a middle range value (5$) can increase the income of movies by 13.56%.

In summary, there are different ways to increase the revenue of movies boxoffices and one possible way is to reduc ticket price and keep in an average value. This will encourage people to visit cinema and helps to reduce the #download.

VII. CONCLUSION

This paper study the impact of online movie downloads to the motion picture sales using a collected dataset from
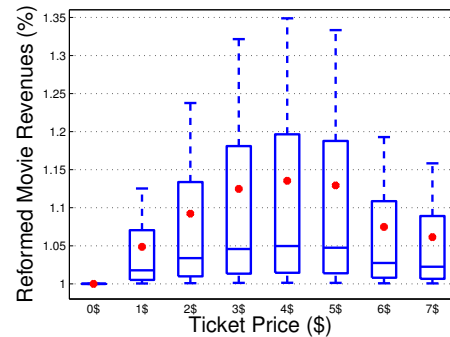
BitTorrent and from other portals (e.g. TorrentZ, IMDB portal, and Worldwide BoxOffice portal). The analyses indicated that soon after releasing a movie there is a linear relationship between total #torrent uploads and movie downloads, indicating that recently released movies are popular among movie viewers and publishers. We also explored that low budget movies were downloaded more than others and shown a very low worldwide gross. In addition, based on the people feedback to our questionnaire, we concluded that although there are solutions such as shutting down a portal or enforcing anti-piracy laws that can reduce #downloads, but people are willing to support other solutions such as low price online possibility to watch a movie instead of downloading it and also willing to pay for promotional tickets.

REFERENCES

[1] R. Farahbakhsh, N. Crespi, Á. Cuevas, R. Cuevas, R. González, Understanding the evolution of multimedia content in the internet through bittorrent glasses, IEEE Network 27 (6) (2013).

[2] R. Farahbakhsh, A. Cuevas, R. Cuevas, R. Rejaie, M. Kryczka, R. Gonzalez, N. Crespi, Investigating the reaction of bittorrent content publishers to antipiracy actions, in: Peer-to-Peer Computing (P2P), 2013 IEEE Thirteenth International Conference on, IEEE, 2013 (2013).

[3] M. D. Smith, R. Telang, Assessing the academic literature regarding the impact of media piracy on sales (2012).

[4] B. Cammaerts, B. Meng, R. Mansell, Copyright and creation: a case for promoting inclusive online sharing (2013).

[5] L. Ma, A. L. Montgomery, P. V. Singh, M. D. Smith, The dual impact of movie piracy on box-office revenue: Cannibalization and promotion, Available at SSRN 2736946 (2016).

[6] L. Ma, A. L. Montgomery, P. V. Singh, M. D. Smith, An empirical analysis of the impact of pre-release movie piracy on box office revenue, Information Systems Research 25 (3) (2014) 590–603 (2014).

[7] B. Danaher, J. Waldfogel, Reel piracy: The effect of online film piracy on international box office sales (2012).

[8] D. Ibosiola, B. Steer, A. Garcia-Recuero, G. Stringhini, S. Uhlig, G. Tyson, Movie pirates of the caribbean: Exploring illegal streaming cyberlockers, in: Twelfth International AAAI Conference on Web and Social Media, 2018 (2018).

[9] M. R. Milot, Testing the lost sale concept in the context of unauthorized bittorrent downloads of cam copies of theatrical releases, Browser Download This Paper (2014).

[10] K. Strumpf, F. Oberholzer-Gee, The impact of file sharing on movies, Tech. rep., Working Paper, Northwestern (2010).

[11] B. Danaher, M. D. Smith, Gone in 60 seconds: The impact of the megaupload shutdown on movie sales, International Journal of Industrial Organization 33 (2014) 1–8 (2014).