Only the Good… Get Pirated:
Game Piracy Activity vs. MetaCritic Score

Anders Drachen	Kevin Bauer	Robert W. D. Veitch
Department of Informatics	Cheriton School of Computer Science	Department of Informatics
Copenhagen Business School	University of Waterloo	Copenhagen Business School
ad.inf@cbs.dk	k4bauer@cs.uwaterloo.ca	rv.inf@cbs.dk

ABSTRACT

The practice of illegally copying and distributing digital games is at the heart of one of the most heated and divisive debates in the international games environment, with stakeholders typically viewing it as a very positive (pirates) or very negative (the industry, policy makers). Despite the substantial interest in game piracy, there is very little objective information available about its magnitude or its distribution across game titles and game genres. This paper presents a large-scale analysis of the illegal distribution of digital game titles, which was conducted by monitoring the BitTorrent peer-to-peer (P2P) file-sharing protocol. The sample includes 173 games and a collection period of three months from late 2010 to early 2011. A total of 12.6 million unique peers were identified, making this the largest examination of game piracy via P2P networks to date. The ten most pirated titles encompass 5.27 million aggregated unique peers alone. In addition to genre, review scores were found to be positively correlated with the logarithm of the number of unique peers per game ($p < 0.05$).

Categories and Subject Descriptors

K.8.0 [Games]; K.7.m [The Computing Profession]: Miscellaneous – Ethics.

General Terms

Economics, Security, Human Factors, Legal Aspects.

Keywords

Digital games, game piracy, BitTorrent, economics.

1.   INTRODUCTION

Game piracy, which involves the illegal copying and distribution of digital games [16], is a complex phenomena that occurs across multiple channels and has a magnitude that is difficult to estimate [1,6,8,9,10,14]. It is the cause of heated debate, with pirates on one side and game developers, game publishers and legislators/policy makers on the other [10]. Despite the interest in game piracy and the controversy surrounding the activity, there is only limited information available on the subject that spans across game titles. The information that does exist often comes from industry organizations or operators of Peer-2-Peer (P2P) networks [e.g., 18], but lacks objectivity and a transparent methodology. The purpose of this paper is to address the need for objective information on digital game piracy. We report analyses on game piracy data obtained via tracking of the BitTorrent P2P protocol.

The BitTorrent network was chosen because it is regarded as one of the main channels for game piracy and the de facto standard for distribution of digital files via P2P networks [2,12,18]. Monitoring was carried out for 173 game titles over a three month period running from late 2010 to early 2011, spanning all genres and multiple hardware platforms (e.g., PC, X360, PS3, Wii, DS, PSP), and the recording of 12.7 million unique peers. The analysis of these data indicates that the major commercial, AAA-level, action-oriented titles account for the highest proportion of activity on the BitTorrent network. The distribution of the torrent activity across game titles was highly asymmetric. For example, the 10 most popular titles comprised 41.5% of the total number of unique peers in the dataset. Finally, aggregated review scores were found to be positively correlated with the logarithm of the number of unique peers per game ($p < 0.05$). This indicates review scores may be one determinant of BitTorrent activity.

2.   METHOD

In order to obtain the data necessary for the analysis of game piracy, two data streams are necessary: 1) BitTorrent data on the online distribution of illegal copies of digital games; 2) Information about the characteristics of these products. The steps towards obtaining these data are comprised of a series of steps, as follows (for a full description, please see Drachen et al. [7]):

2.1 Obtaining Unique Peers from BitTorrent

For the current study, a list of 173 game titles was compiled across genres and hardware platforms including Xbox360, PlayStation 3, Nintendo Wii, PC, Nintendo DS (DS) and PlayStation Portable (PSP). This list included a series of games released in the fall 2010, and every game released since November 17th 2010 until 29th January 2011, the end of the tracking period (note that games can appear on BitTorrent prior to the official launch date). The sample consisted of games from all genres (irrespective of the specific definition system), ranging from AAA-level major commercial titles (e.g., Bioshock 2) to casual and indie games (e.g., Majin and the Forsaken Kingdom). Of these 173 titles, within the period of tracking (November 17th 2010 to February 6th 2011), 127 were located on BitTorrent, indicating these games had been cracked of any copyright protection (DRM, Digital Rights Management), and released on BitTorrent. During the period of tracking, 12.6 million unique peers were identified for all these games, making this the to date largest study of BitTorrent-based game piracy, surpassing even the report of the Entertainment Software Association (ESA) from 2009, who reported 9.58 million downloads for about 200 unspecified titles produced by members of the ESA during one month in late 2009 (no detailed information has been revealed, including methodology and the specific titles involved). Forty of 127 titles made their first appearance on BitTorrent during the period of tracking (comprising 1.16 million unique peers).

In order to obtain the BitTorrent metadata files [4] for the 173 game titles, a custom web crawler was developed. The web crawler periodically issues queries to a popular BitTorrent search engine website for each title, extracting the metadata files. Having located the metadata files, the web crawler obtains the tracker server URIs. Having obtained the URIs, an HTTP GET request is issued to each URI over periodic intervals (the tracker servers are queried every few minutes) to obtain a list of IP addresses for the peers who are currently participating in sharing the specific file.

When searching for game torrents, a number of false positives can occur as torrents related to a game or with a similar name may not contain the full game. In order to eliminate such torrents, all torrent lists for each game title were manually inspected and filtered, leaving only the torrents that contain the full game. It is a common practice for copyright investigators to spread false information on BitTorrent in an attempt to make it difficult for peers to access copyrighted material [3,5,12]. In order to filter out the false IP addresses, standard publicly available blacklists were applied to filter out IP address blocks controlled by well-known copyright investigators, thus ensuring that the final set of unique peers consists only of real peers who are actively engaged in file sharing (Available from: http://www.iblocklist.com/list.php?list=bt_level1).

2.2 Obtaining Game Feature Data

A substantial challenge for the monitoring of newly released digital games on P2P networks and other piracy channels is that the game developers and publishers often do not adhere to announced release dates. This practice makes it very difficult to determine when tracking of a specific game title should begin. Given the propensity for digital games to be available on Bit-Torrent before the official release date, the best approach is to start tracking the game as soon as a title is reported nearing completion (e.g., beta-testing stage). Unlike other media such as movies and music, there are no central repositories for information on digital games, covering release dates, sales figures, product features etc. Instead, a wide variety of websites attempt to provide parts of this information with greater or lesser degrees of accuracy (e.g., metacritic.com, gamerankings.com, gamestats.com, vgnchartz.com, gamespy.com, ign.com, gamespot.com). Aggregating the information derived from these sites forms the best current approach towards obtaining the most reliable information possible.

2.2.1 Genre

Another challenge towards defining product features is the variety of digital games and the many different systems for categorizing games into genres. Game genre systems are nebulous at best, and therefore a similar aggregation approach was adopted here to build a genre system based on majority consensus. In practice, a variety of websites (e.g., mobygames.com, ign.com, gamespy.com, metacritic.com) were mined and genre definitions for the individual titles based on majority definitions.

2.2.2 Aggregated Review Score

In order to obtain a measure of the quality of a game title, aggregated review scores were obtained from several recognized metacritic sites (metacritic.com, gamerankings.com, gamestats.com). Not all the games in the sample were available on all three of these sites. Aggregated scores could not be found for 12 of the 127 torrented games. Ten of these were commercially small titles (e.g., Stardrone and Brain Puzzles 2). For one of these titles, aggregated review scores could be built manually by recording review scores from game sites such as gamespy.com and ign.com.

The remainder was eliminated from any analysis involving review scores. Average review scores on a scale from 0-100 range from 26 (Deca Sports Freedom) to 94.67 (Mass Effect 2), with a mean score of 70.44 and Std. Dev. = 15.83 (n=113).

3. ANALYSIS AND RESULTS

Following data collection and pre-processing, the final dataset with BitTorrent activity and game-feature information comprised 127 games. The BitTorrent activity data contained some noticeable characteristics. The frequency distribution of unique peers per game was highly asymmetrical; the majority of the game titles had relatively limited activity on BitTorrent (i.e., less than 50,000 unique peers observed) (Figure 1). The 10 most popular games in the sample accounted for 5.37 million unique peers alone (Table 1), all of which were major commercial titles.

Table 1: The 10 most torrented game titles encompass 5.37 million of the unique peers in the dataset, averaging 536,727 peers per game and an average review score of 74.5 (on a scale from 0-100). Genre definitions here from www.metacritic.com.

Title	Genre	Unique Peers	Avg. Review Score	Developer
Fallout: New Vegas	Role-Playing	962,793	83.7	Obsidian Entertainment
Darksiders	Action Adventure	656,296	82.7	Vigil Games
Need for Speed: Hot Pursuit	Racing	656,243	88	Criterion Games
NBA 2k11	Basketball	545,559	86.7	Visual Concepts
TRON Evolution	Action Adventure	496,349	59.5	Propaganda Games
Call of Duty: Black Ops	First-Person Shooter	469,864	83.8	Treyarch
Starcraft 2	Real-Time Strategy	420,138	89.5	Blizzard Entertainment
Star Wars the Force Unleashed 2	Action	415,021	61	Lucas Arts
Two Worlds II	Role-Playing	388,236	73.3	Reality Pump
The Sims 3: Late Night	Virtual Life Games	356,771	77.5	The Sims Studio

In general, casual games and indie games were much less frequently pirated, with a few exceptions, notably Bejeweled 3 with over 250,000 unique peers recorded.

The aggregated review score of a digital game is generally expected to be linked with its financial success, although this is not always the case [11,13]. Similarly, it is possible that review scores are also related to how much a game is distributed on BitTorrent – which is indicated by the high proportion of major commercial titles in the 127 game sample, as well as the observation that 7 of the 10 most torrented games in the sample have aggregated review scores over 75 (on a 0-100 scale). In order to explore a possible relationship between torrent activity and aggregated review scores, Pearson's Product-Moment Correlation Coefficient for Metacritic Scores (mean = 70.44, Std. Dev. = 15.83) and number of unique peers per game (mean = 99894.43; Std. Dev. = 156028.6) was calculated (r= 0.28; p<0.05 significance (two-tailed) given df = 125 (n-2)) [15]. Please note that log(unique

peers) was used due to the non-normal distribution of the peers data. The result indicates a statistically significant positive relationship between the number of unique peers and aggregated review scores. Put differently, Metacritic Scores explain 14% of the variance in the unique peers per game on BitTorrent.
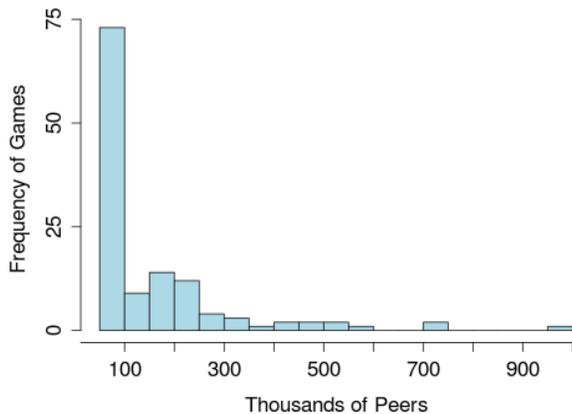

Figure 1: Frequency distribution of the number of unique peers associated with the games in the dataset.

4. DISCUSSION AND CONCLUSION

The illegal copying and distribution of digital games stands at the heart of one of the central controversies in the international interactive entertainment environment. Despite the substantial interest in the problem [e.g., 6,8,9,10,12,14], the wealth of industry-based reports of piracy rates, and the size of the industry, there is minimal objective information available about the magnitude of game piracy and its distribution across game titles or genres. In this paper, a first step has been taken towards addressing this knowledge gap, via the analysis of a 12.6 million unique peer dataset obtained from BitTorrent; the major channel for game piracy and the standard for P2P distribution of [2,18]. The first and perhaps most important contribution of this paper is to provide objective documentation of the magnitude of distribution of digital game files via BitTorrent. Out of 173 game titles in the study, released during the fall 2010 or early 2011, 127 were found on BitTorrent networks. Approximately 12.6 million unique peers accessed these files, indicating the prevalence of game piracy via BitTorrent-based distribution. Unlike previous work, the data reported here are objective, quantitative and developed using state-of-the-art techniques and with a public and open methodology. Analysis of the data indicate that it is the major commercial titles that are the most heavily distributed games on BitTorrent, with some exceptions from the more casual form. The ten most pirated titles encompass 5.27 million aggregated unique peers. Finally, the evidence presented indicate that review scores and torrent activity for game titles correlate positively (p<0.05). Additional analyses are available in Drachen et al. [7].

5. ACKNOWLEDGMENTS

6. REFERENCES

[1] Banerjee, D., Khalid, A. M. and Sturm, J.-E. 2005. Socio-economic development and software piracy. An empirical assessment. *Applied Economics* 37, 2091 - 2097.

[2] Bauer, K., Grundwald, D. and Sicker, D. 2009. The Arms Race in P2P. In *Proceedings of the 37th Research Conference on Communication, Information and Internet Policy*, TPRC, 1-18.

[3] Bauer K., McCoy, D., Grundwald, D. and Sicker, D. 2009. Bitstalker: Accurately and Efficiently Monitoring BitTorrent Traffic. In *Proceedings of the First IEEE Workshop on Information Forensics and Security,* IEEE Publishers, 181-185.

[4] *BitTorrent Protocl Specification*. 2010. Available from: http://wiki.theory.org/BitTorrentSpecification.

[5] *BTGuard: Anonymous BitTorrent Services*. 2011. Available from: http://btguard.com/.

[6] Business Software Alliance. 2010. *Seventh Annual BSA and IDC Global Software Piracy Study*. Available from: http://www.bsa.org/country/Research%20and%20Statistics.aspx

[7] Drachen, A., Bauer, K., Veitch, R. 2011. *Piracy Activity vs. Product Features in Digital Games*. Technical Report, Copenhagen Business School, 6th May 2011. 8 pp. Available from: http://openarchive.cbs.dk/handle/10398/8299

[8] Goldacre, B. 2009. Illegal downloads and dodgy figures. The *Guardian*, June 5 2009. Available from: http://www.guardian.co.uk/commentisfree/2009/jun/05/ben-goldacre-bad-science-music-downloads.

[9] International Intellectual Property Alliance (IIPA). 2010. *2010 Special 301 Report on Copyright Protection and Enforcement*. Available from: www.iipa.com.

[10] Kirk, R. 2010. *Special 301 Report 2010.* Published by the Office of the United States Trade Representative, United States of America. Available from: http://bangkok.usembassy.gov/root/pdfs/2010_special_301_report.pdf.

[11] Klepek, P. 2008. Bethesda Is Tired Of Spending Money Supporting Software Pirates. *MTV Multiplayer*, October 13 2008. Available from: http://multiplayerblog.mtv.com/2008/10/13/bethesda-deals-with-pirates/.

[12] Piatek, M., Kohno, T. and Krishnamurty, A. 2008. Challenges and directions for monitoring P2P file sharing networks-or: why my printer received a DMCA takedown notice. In *Proceedings of the Proceedings of the 3rd conference on Hot topics in security,* USENIX Association, 1-7.

[13] Rabin, S. 2010. *Introduction to Game Development*. Charles River Media.

[14] Sanchez, J. 2008. 750,000 lost jobs? The dodgy digita behind the war on piracy. *Ars Technica*, October 7th, 2008. Available from: http://arstechnica.com/tech-policy/news/2008/10/dodgy-digits-behind-the-war-on-piracy.ars

[15] Sigle, D. 2010 *Critical Values of the Pearson Product-Moment Correlation Coefficient*. Neagle School of Education - University of Connecticut. Available from: http://www.gifted.uconn.edu/siegle/research/correlation/corrchrt.htm

[16] Stryszkowski, P. and Scorpecci, D. 2009. *Piracy of Digital Content*. OECD Publishing.

[17] *The Tor Project: Anonymity Online*. 2010. Available from: http://www.torproject.org/

[18] Wolchok, S., and Halderman, J. A. 2010. Crawling BitTorrent DHTs for Fun and Profit. In *Proc. of the 4th USENIX Workshop on Offensive Technologies*, USENIX Association, 112-113.