Clickonomics: Determining the Effect of Anti-Piracy Measures for One-Click Hosting

Tobias Lauinger,* Martin Szydlowski,† Kaan Onarlioglu,* Gilbert Wondracek,† Engin Kirda* and Christopher Kruegel‡

*Northeastern University, Boston, MA. Email: {toby, onarliog, ek}@ccs.neu.edu
†Vienna University of Technology, Austria. Email: {msz, gilbert}@seclab.tuwien.ac.at
‡UC Santa Barbara, CA. Email: chris@cs.ucsb.edu

Abstract—Piracy is a mass phenomenon on the Internet today. Various file sharing platforms offer free access to unauthorised copies of copyrighted works such as media content and software. Copyright holders are using a range of legal and technical methods to protect their rights, and they are lobbying for legislation that would give them additional ways of enforcing their copyright online. However, little is known about how effective current forms of copyright enforcement are and how enduring the effect of proposed new measures can be. In this paper, we report on the results of a large-scale measurement study of the piracy ecosystem that has emerged around One-Click Hosters or "cyberlockers" such as Rapidshare, Megaupload, Mediafire, and Hotfile. Our data shows that current anti-piracy efforts are visible, but their overall impact appears to be rather limited. Furthermore, our analysis of the file sharing ecosystem suggests that future anti-piracy measures that are currently under discussion may not be as successful as their proponents might expect. Ongoing legal proceedings and efforts by payment processors, however, may force hosters to increase their own anti-piracy efforts.

I. INTRODUCTION

The culture of sharing files regardless of potential copyright infringement is older than the Internet. It began in the 80s when underground Warez groups traded their *releases* on bulletin board systems. The recent uproar caused by the proposed U.S. Stop Online Piracy Act (SOPA) and the Anti-Counterfeiting Trade Agreement (ACTA) as well as the spectacular seizure of file hosting site Megaupload's assets on 19 January 2012, once ranked among the 100 largest web sites worldwide, have made it clear that copyright infringement or *piracy* continues to be a current and controversial topic.

The first One-Click Hosters (OCHs) Rapidshare and Megaupload started operating in 2004 and 2005, respectively. They simplified the exchange of large files between users by allowing them to upload files through a simple web interface. The files could then be shared by sending the corresponding download links to the intended recipients of the files. Pirates soon discovered OCHs as a platform for their activities: They started uploading popular movies, TV shows, music, ebooks, games and applications to OCHs and posted the corresponding download links on external *indexing sites*, so-called direct download or streaming sites. As of today, there are more than 300 OCHs and tens of thousands of indexing sites. Millions of pirated files are uploaded each day.

The U.S. Digital Millennium Copyright Act (DMCA) grants OCHs immunity from liability for copyright infringements committed by their users if they fulfil certain conditions. One of its provisions allows copyright holders to request that infringing links be taken down. According to the criminal indictment[1] against Megaupload, in September 2009, Warner Bros. had 2,500 infringing links removed from Megaupload on a daily basis. Furthermore, trade associations such as the Motion Picture Association of America (MPAA), the Recording Industry Association of America (RIAA) and their equivalents in other countries conduct or aid in investigations that lead to the seizure of file sharing sites' assets such as server hardware and domain names.

Many file sharing sites are difficult to reach for law enforcement because the servers are located in foreign jurisdictions and the sites use addresses under foreign top-level domain names that cannot easily be seized. Therefore, recent lobbying efforts such as SOPA have aimed at giving copyright enforcement actors new tools to fight piracy. One of the main proposals included in SOPA was a takedown notice system similar to the DMCA, but targeting entire foreign sites instead of single hosted files. While some argued that this notice-based regime would make it possible to prevent copyright infringement in a timely manner, others were concerned that this system would be ripe for abuse of all sorts and they saw it as a threat to freedom of speech.

Currently, the public debate mainly relies on partisan arguments and opinions in favour of or against these new anti-piracy measures. There is little empirical data that would make it possible to assess the effectiveness of anti-piracy measures in an objective manner. We believe that in addition to its desirableness, a proposed law should be judged based on its expected effectiveness—especially when the law is criticised for having far-reaching negative side effects.

In this paper, similar to the recent line of work that has measured and analysed spam [1], [2], [3], counterfeit pharmaceutical sites [4], [5], [6] and other aspects of underground economies [7], [8], [9], we undertake an empirical effort to quantify the supply side of One-Click Hosting-based piracy

---

[1]Superseding indictment, U.S. v. Kim Dotcom et al., 1:12-cr-00003-LO (E.D. Va., Feb. 16, 2012) at ¶ 73 zzz.

with respect to current and proposed anti-piracy measures. We characterise the dynamics of the ecosystem based on our extensive data sets and observe the impact of current anti-piracy measures. Furthermore, we analyse how vulnerable the ecosystem might be to a range of proposed anti-piracy measures by investigating the relationships between a range of OCHs, indexing sites, and payment processors.

Due to the size of the ecosystem and its decentralised structure, it is not practically feasible to characterise certain of its aspects in a globally representative way. The approach that we follow in this paper is to select a few large indexing sites and to measure metrics such as the lifetime of download links and the number of mirror copies per content object. In a strict sense, these results hold only for the few observed (albeit large and popular) sites. Yet, they can be seen as a *lower bound* for the availability of content in the overall ecosystem since there are many more independent sites where additional copies of the same content can be found.

In detail, we crawled the indexing sites `rlslog.net` and `scnsrc.me` to discover links to pirated files. Our crawler recognised 500 types of links pointing to 300 different OCHs and extracted around 6 million links overall. Over a duration of 12 months in total, we periodically checked the availability of 1.4 million of these files on a selection of 50 OCHs in order to quantify how long infringing files could survive. Furthermore, over a duration of 19 months, we uploaded hourly test files to up to six OCHs and computed the OCH-wide file upload rate. To estimate the size of the ecosystem, we analysed the results of a popular specialised search engine.

Our data suggests that current anti-piracy efforts are visible, but their effect is limited when compared to the overall availability of pirated content. Furthermore, we find some parts of the OCH piracy ecosystem to be rather resilient to anti-piracy approaches targeting their sources of revenue, while other actors appear to be more exposed. Other proposed measures can be circumvented with technology that is already in widespread use today. These results imply that different anti-piracy strategies may be required; from a purely technical point of view, targeting the demand instead of the supply of pirated content would be the most promising direction.

With this paper, we make the following contributions:

- We conduct the most in-depth investigation of the OCH-based piracy ecosystem to date and quantify the dynamics of OCH-based file sharing.
- We show that current anti-piracy efforts do not succeed in making pirated content unavailable.
- We discuss how effective several proposed future anti-piracy strategies can be expected to be, given the current state of the OCH file sharing ecosystem.

II. BACKGROUND & RELATED WORK

With a simple web interface, One-Click Hosters (OCHs) such as Rapidshare, Hotfile or Mediafire allow users to upload potentially large files that can then be reached under unique download links. Uploaded files are private as long as the corresponding download links remain secret; this is why OCHs are sometimes also referred to as *cyberlockers*.

As with every other file distribution channel of the Internet, OCHs are commonly used to share copyrighted works without the rights holders' consent. Uploaders seeking to make files publicly available post their download links on external *indexing sites*, also called direct download or streaming sites. The function of these sites is very similar to torrent sites such as The Pirate Bay in the BitTorrent (BT) ecosystem: They offer a searchable catalogue of popular (and mostly commercially exploited) content such as movies, TV shows, music, ebooks and software. Users visit these indexing sites or use search engines to find the content they are looking for and follow a link to an OCH to download the file.

One-Click Hosting-based piracy has reached a significance at least similar to BT. In October 2011, Alexa[2] ranked between 8 and 17 OCH indexing sites among the 100 most visited sites in France, Germany and Spain but only 1 to 3 BT indexing sites. (In contrast, BT was more popular in the U.S. with 5 BT indexing sites and 2 OCH indexing sites.) Despite these numbers, we stress that OCHs have various legitimate use cases. Whether legitimate or illegitimate uses of OCHs prevail is a contentious topic and beyond the scope of this paper. In the following, we exclusively focus on abusive uses of OCHs and leave out legitimate parts of the ecosystem.

A. Related Work

Antoniades et al. conducted the first study of One-Click Hosting [10], arguing that it offered a better file sharing experience than BT (e.g., better anonymity, higher availability and throughput). Data extracted from indexing sites revealed that most content on the sites was published by only a small number of users. While the authors showed the availability of pirated content on Rapidshare, the effect of anti-piracy measures was not a topic of their study.

The results of later work by Mahanti et al. [11] and Sanjuàs-Cuxart et al. [12] reflected the presence of pirated content in the workload of OCHs, but the studies did not distinguish legitimate and illegitimate traffic or further highlight the effects of copyright enforcement. In [13], Mahanti et al. investigated how users discovered and shared content. They also analysed Web analytics data by `Compete.com` to characterise the popularity evolution of OCHs and indexing sites, reporting similar downloader-driven trends as we found in our uploader-driven data in Section IV-B.

Nikiforakis et al. [14] found that the download links used by many OCHs were vulnerable to guessing attacks. They showed that attackers were already exploiting this weakness to compromise the privacy of uploaders.

In a previous study [15], we analysed the demand for pirated content posted on three indexing sites and estimated

---

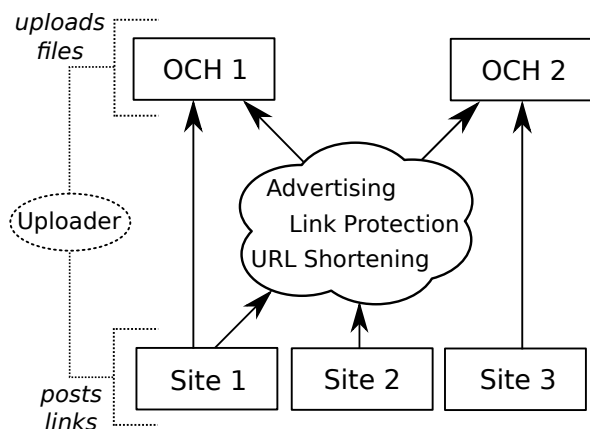[2]http://www.alexa.com/topsites

Figure 1. Overview of the OCH ecosystem: An uploader uploads a file to one or more OCHs, optionally wraps the download links using redirection services such as link shorteners, and posts the links on indexing sites.

the corresponding income that content uploaders could derive from OCHs' affiliate programmes. In these three particular communities, we found that only very few uploaders were *potentially* earning significant amounts of money. Furthermore, these high-income uploaders were not essential for the supply of pirated content because other uploaders (with lower income) provided alternative copies of most content. Zubin Jelveh and Keith Ross [16] used payment screenshots posted in a webmaster forum to deduce the range of *actual* uploader income through Filesonic's affiliate programme. Based on 151 earnings screenshots, they reported a daily uploader income ranging from $\$\,0$ to $\$\,226.27$ with an average of $\$\,33.69$.

B. The OCH Ecosystem

Beyond the indexing sites described in prior work, a whole ecosystem of specialised services has emerged around OCHs. Various *search engines* specialise in OCH download links extracted from indexing sites; Filestube is the most popular representative of them. So-called *multi-upload services* spread files over several different OCHs on the users' behalf and *multi-download services* provide their paying subscribers with premium download access to several OCHs at once. Because large files are sometimes split into smaller parts (these can be hundreds of parts for very large files), some sites offer so-called container files or use external *container services* that provide all the download links for the individual parts grouped together under one single link. These container files can be opened with specialised *download managers* such as JDownloader. Specialised *advertising services* such as adf.ly or linkbucks.com work similar to URL shorteners, but display interstitial or pop-up advertisements and allow uploaders to monetise their links.

One-Click Hosters typically disable infringing download links when they receive DMCA complaints from the respective copyright holders. In order to prevent automated link extraction from indexing sites (that might correspond to

takedown efforts by copyright owners), many uploaders hide their download links behind a layer of redirection, such as general-purpose URL shorteners or specialised *link protection services* such as linksafe.me. In addition to hiding the original download link from the indexing site, uploaders can optionally request that downloaders solve a Captcha before they are redirected to the OCH. Furthermore, many sites and services automatically check if download links are still valid and mark, remove or replace them if they are not. Undeadlink was one such service that automatically reuploaded a file when the original copy was deleted from the OCH. Figure 1 illustrates the optional use of various kinds of services in a redirection chain between a link posted on an indexing site and the actual link target hosted on an OCH.

C. Funding in the OCH Ecosystem

Most OCHs sell premium services to users who wish to enjoy higher download throughput and less advertising. According to the criminal indictment,[3] Megaupload received at least 150 million dollars in subscription fees and 25 million dollars for advertising between September 2005 and 5 January 2012. In order to attract popular content (and new paying members) to their service, most OCHs used to operate (and several continue to operate) affiliate programmes for uploaders and indexing sites. These affiliate programmes financially reward uploaders and indexing sites based on the number of downloads and member subscriptions that they generate. Megaupload, for instance, used to reward one million downloads with $\$\,1,500$ and Wupload used to pay up to $\$\,40$ per one thousand downloads.

These affiliate programmes are controversial for allegedly encouraging users to upload infringing content and thereby funding piracy. Megaupload's former affiliate programme and their alleged knowledge that affiliates uploaded pirated content were a central element of the criminal indictment[4] that lead to the seizure of Megaupload's assets, the detention of its operators, and the shutdown of the site on 19 January 2012. In the aftermath, most OCHs modified or terminated their affiliate programmes. Several OCHs (including Filesonic, Wupload and X7) quit the file sharing business altogether.

In addition to the OCHs' affiliate programmes, uploaders, indexing sites and service providers such as link protectors can generate revenue through advertisements. Some indexing sites collect donations, and a few require members to suscribe in order to access "premium" content. In contrast, especially smaller sites do not have any apparent source of income and may be run as hobby projects. Despite the end of most OCH affiliate programmes and additional evidence that it has become harder to earn money in the ecosystem (as discussed

---

[3]Superseding indictment, U.S. v. Kim Dotcom et al., 1:12-cr-00003-LO (E.D. Va., Feb. 16, 2012) at ¶ 4.

[4]Superseding indictment, U.S. v. Kim Dotcom et al., 1:12-cr-00003-LO (E.D. Va., Feb. 16, 2012) at ¶ 58; ¶ 73 g−j, v, y, bb, jj, pp, qq, uu, ppp, qqq, www, xxx; and ¶ 102.

in Section V-A), many indexing sites and other services continue to exist, which implies that they are still profitable in either monetary or altruistic terms.

D. Current Anti-Piracy Measures

Anti-piracy measures for OCHs are dominated by the difficulty for copyright holders to identify uploaders or downloaders of pirated content. P2P networks, for instance, reveal the IP addresses of peers to other participants, while communication in OCH-based file sharing is direct between the uploader or downloader and the indexing site or OCH. Furthermore, deterrence in the case of OCH may be less effective because the fines for downloading infringing content are lower than for uploading, which is usually automatic for P2P downloaders but not for OCH downloaders.

Section 512(c) of the U.S. Digital Millennium Copyright Act (DMCA) provides copyright holders with a *takedown* mechanism to disable access to infringing files. The DMCA "safe harbour" provision grants compliant hosting providers such as OCHs (but also YouTube, for instance) immunity against copyright infringements committed by their users. Most OCHs, even if located outside of the U.S., claim to honour takedown requests. Similar DMCA rules apply to other online service providers such as search engines.

Copyright owners use this mechanism to have infringing files removed from OCHs. They either manually search the Internet for infringing files, or they utilise the services of specialised agencies that automatically crawl the web and extract download links from indexing sites and search engines. Corresponding takedown notices are then sent to OCHs and search engines, respectively. Chilling Effects[5] maintains a database of the takedown notices sent to services such as Google and Twitter, and Google themselves provide statistics about copyright-related takedown requests in their Transparency Report.[6] Urban and Quilter [17] performed an analysis of 876 such notices, including all 734 notices submitted to Google from March 2002 until August 2005, and found that at least one third of the notices were flawed.

Several lawsuits against OCHs with allegations of copyright infringement are currently pending. As of the time of writing, none of them has been decided by a court of final appeal. Indexing sites have been targeted by arresting the operators and seizing the server hardware, such as in the case against Ninjavideo[7] in the U.S. If an indexing site was located outside of the U.S., Immigration and Customs Enforcement have in the past seized domain names that were under U.S. jurisdiction (`.com` and `.org` in the case of the Spanish site Rojadirecta[8]). Courts in various European countries have

ordered ISPs to block access to The Pirate Bay based on either their domain name or IP addresses; similar measures are conceivable for OCH-based sites.

III. MEASUREMENT METHODOLOGY

Our analysis in this paper is based on a range of data sets extracted from different actors of the OCH file sharing ecosystem: The INDEXING data set contains download links found on several indexing sites, along with metadata about when and by whom the links were published. The LIFETIME data set contains time-based availability data for a subset of these links. In IDLOGGER, we measure the global file upload rate of several OCHs, and in FILESTUBE, we crawl a specialised search engine for file-sharing related links in order to estimate a lower bound of the OCH file sharing ecosystem's size. Table I summarises the time of our measurements and the key characteristics of these data sets.

A. Analysis of Indexing Site Data Sources

In order to obtain structured information about files being shared, we crawled several indexing sites to extract public download links together with metadata about their contents and the users who posted the links. There is a great variety of different types of indexing sites, ranging from smaller blogs with occasional download links to specialised sites using sophisticated web applications with social networking features. We extracted links from 14 different indexing sites in total, but for the sake of a clearer presentation, in this paper we use only data gathered from `rlslog.net` and `scnsrc.me`, two so-called *release blogs*. These sites are specialised in the timely dissemination of fresh releases originating from the P2P or Warez Scene, that is, content with an often high commercial value, such as movies, TV shows, music, ebooks and software shortly after their official release date in stores, and sometimes even before. (For an introduction to the Warez Scene, refer to [18], [19], [20].)

The sites are organised like blogs: For each content object, staff members post a main entry with a review of the content and a few download links. Most download links, however, are added by independent, unregistered uploaders who post them in (anonymous) comments to the blog entries made by the site staff. There are no rules that would restrict the choice of the OCH, for instance. This characteristic allows us to better observe how changes in the OCH ecosystem affect the behaviour of uploaders independent of policy constraints. The amount and speed of posts as well as the use of particular OCHs and advertising services suggest a *higher* degree of profit-oriented uploader behaviour on these sites than what we observed on a different set of sites in prior work [15].

According to our observations, the posts published on `rlslog.net` and `scnsrc.me` contain almost no legitimate content. We manually analysed two random samples of 1,000 content objects posted on each of the two sites. Only

[5]http://www.chillingeffects.org/dmca512/notice.cgi

[6]http://www.google.com/transparencyreport/removals/copyright/

[7]enigmax, "NinjaVideo Founder Sentenced To 22 Months in Prison," http://torrentfreak.com/ninjavideo-founder-sentenced-to-22-months-in-prison-120106/, 6 Jan. 2012.

[8]Verified complaint, U.S. v. Rojadirecta.org et al., 1:11-cv-04139-PAC (S.D.N.Y., June 17, 2011).

(a) INDEXING

Site	rlslog.net	scnsrc.me
Alexa Rank	1,557	14,099
First Post	01/2006	03/2008
Last Post	04/2012	04/2012
# Content	70,423	25,880
# Links	4,997,040	2,200,814

(b) LIFETIME

Site	rlslog.net	scnsrc.me
Links Extracted I	$12/2010 - 07/2011$	$04/2011 - 07/2011$
Links Extracted II	$03/2012 - 04/2012$	$03/2012 - 04/2012$
Links Checked I	$01/2011 - 11/2011$	$01/2011 - 11/2011$
Links Checked II	$03/2012 - 05/2012$	$03/2012 - 05/2012$
# Content Obj.	13,133	2,974
# Links Found	1,376,334	350,136
# Links Checked	1,294,739	322,611

(c) IDLOGGER

OCH	Rapidshare	Hotfile	Duckload	Filesonic	Wupload	Easy-share
Time Span I	$06/2010-05/2011$	$12/2010-11/2011$	$05-06/2011$	$06-10/2011$	$06-11/2011$	$07-10/2011$
Time Span II	$-$	$02/2012-05/2012$	$-$	$02-03/2012$	$02-04/2012$	$02-02/2012$
# Test Uploads	16,150	20,741	1,041	8,328	10,022	6,904

two freeware tools and some news reports (without download links) appeared to be legitimate. All remaining posts corresponded to material that was commercially exploited, and therefore most likely not meant to be freely available.

The proportion of polluted content, that is, files that were corrupted or did not contain the advertised content, appeared to be significantly lower on rlslog.net and scnsrc.me than what had been observed in the past on P2P networks [21], [22]. We downloaded and manually inspected the links posted for ten content objects representing one item of each category present on the indexing sites, that is, TV shows, movies, ebooks, music, games, and software. Our sample included an episode of the TV shows *Gossip Girl* and *Fringe*, an early copy of the James Bond movie *Skyfall* (which had just been released to movie theatres), a prerelease copy of the latest album by Robbie Williams, *Take the Crown*, as well as a DVD rip of *The Expendables 2* and several game and software titles. Overall, we downloaded 194 files. Four of them were linked in the main blog post written by site staff; all others were extracted from mostly anonymous comments.

At least 93 % of the downloaded files appeared authentic. The remaining files were either authentic content posted in the wrong category (e.g., a wrong episode or wrong TV show), or they were password-protected and we could not open and verify them as authentic because we did not have access to the password. Consequently, we can assume that the vast majority of download links posted on rlslog.net and scnsrc.me were leading to "authentic," copyright-infringing content. One might speculate that the difference to P2P systems is due to the possibility of sending DMCA takedown notices to OCHs. It may be more effective for copyright owners to have infringing files *removed* instead of *adding* fake files. Legal issues might be another reason.

While we verified the authenticity of the files by inspecting them in a virtual Linux machine, we did not carry out any detailed analysis regarding potential maliciousness. Kammerstetter et al. [23] found that the majority of cracks and key generators downloaded from the Internet, including OCHs, were malicious and attempted to infect the victim's machine with some kind of malware. However, the authors did not further analyse the *provenance* of their download links—that is, whether they had initially been posted on scam sites or on more reputable indexing sites.

In our sample, once all archives had been extracted, we detected only few different file versions for the same content, which corresponded to different scene releases. In the case of executable files, all of them had identical hash values, which made them rather unlikely to be malicious. We cannot fully exclude the presence of malware on the two indexing sites, and we observed a low number of user complaints to that end. Yet, software accounted for only 8.94 % of the download links on scnsrc.me and 15.11 % on rlslog.net, which limits the maximum impact of malware on our overall results.

B. Full Crawls of Indexing Sites

For the INDEXING data set, we performed a *full crawl* of rlslog.net and scnsrc.me during which we extracted the entire archive of links posted on the two sites. The full crawls permit us to analyse how the behaviour of link posters evolved over time. Our crawler contained regular expressions to recognise 500 different types of links, corresponding to 300 different OCHs and several further "service providers" such as link protection and advertising services. We also automatically extracted download links hidden behind URL shorteners and advertising or link protection services if this was possible without solving a Captcha, for instance.

C. Lifetime of Links on OCHs

In order to establish the LIFETIME data set of infringing files on OCHs, we carried out additional *real-time* crawls of `rlslog.net` and `scnsrc.me`. Our goal was to detect new content and links as soon as they were published on the indexing sites. For the real-time crawls, we retrieved each site's RSS feed once an hour to discover new blog posts. We then extracted the metadata and links contained in the page similar to the full crawls. Additionally, we continued to fetch each blog post repeatedly to detect new links in comments that were added after publication of the original post.

To check the validity of the links discovered in the real-time crawls, we periodically requested availability information from 50 OCHs. Our link checker supported the 20 most frequently used OCHs on the indexing sites and a selection of less popular OCHs. To check the status of a link, we used the OCHs' link checking APIs where available. In many cases, such an API permitted to check 100 links at a time and returned for each link the file name, the file size, and whether the file was still available or deleted. In the case that the OCH did not offer an API, we resorted to requesting the link's download page and inferred the status of the file from the server's HTML response. We internally stored the status of a file as *online*, *offline*, or *unknown*,[9] and we kept track of the last time that we had seen a file online. We checked each *online* or *unknown* file's status at least once every 48 hours (and more frequently during the first weeks of its discovery); *offline* files were checked weekly. All in all, more than 1.4 million links published on `rlslog.net` and `scnsrc.me` were periodically checked in this way.

Files on OCHs can be deleted for several reasons, such as copyright infringement, expiry after a fixed amount of time or a fixed amount of downloads, system failures, or when users manually delete files or close their accounts. Note that most OCHs do not publicly disclose why a specific file has been removed. As we have not seen any notable number of links to legitimate content on the indexing sites, we assume that all files in the data set infringe copyright and could be rightfully removed by the respective copyright owners.

We define the *lifetime* of a link (or file) as the time span between the time when the link was published, that is, the first time that a copyright holder could become aware of the existence of the file, and the last time that we saw the file online. Our lifetime estimates are conservative in that they underestimate the file lifetime in case of errors or delays.

Depending on the indexing site, the experiment lasted for up to 10 months (see Table I(b)). We stopped adding new links 45 days before we halted our link checker. That is, the lifetime distribution of all links is accurate for a duration of up to 45 days; data censoring affects only longer time spans.

[9] A file's status became *unknown* in the case of network errors, or if the server's response could not be understood, such as when the API or HTML template of download pages had changed.

D. Upload Rate of Files to OCHs

The download links of several OCHs contained file identifiers that were sequential [14]. We leveraged this fact to estimate the total number of files uploaded each day to Rapidshare, Hotfile, Duckload, Filesonic, Wupload, and Easyshare. We gathered the IDLOGGER data set by uploading a test file to each OCH every 30 minutes and by extracting the file identifier reported in the corresponding download link. The number of third-party uploads between two test uploads is the numerical difference between the two file identifiers. We found that all OCHs except for Filesonic effectively assigned every possible identifier; Filesonic assigned roughly 11 % of the identifiers at the time of the experiment (only identifiers ending with 1 or 4 but not always both of them). We continued the experiment until the OCHs stopped using sequential identifiers or quit the file sharing business.

E. Domains and Addresses of Indexing Sites

In order to estimate a lower bound for the number of OCH indexing sites, we queried `filestube.com`, a search engine specialised in OCH download links, for 43 piracy-related search terms. From the search results, we extracted the domain names of the indexing sites where the download links had been found. Because Filestube provided only the first 100 result pages with 10 hits each, we increased the number of results by additionally following up to 50 suggested related search terms. These were typically refinements of the original search term, such as "season 2," "dvdrip" or "french." The FILESTUBE data set contains all the domain names found in the search results. Additionally, we looked up the domain names on the same day in April 2012 from a computer located in California and recorded the primary and alias domain names as well as the IP addresses returned in response to the lookups. The latter data provides us with an approximation of the jurisdictional diversity of the OCH indexing site ecosystem.

F. Limitations

The design choices that we have made in our study imply certain limitations as to the conclusions that we can draw from our data. Firstly, our methodology of detecting links based on regular expressions means that we might not detect links of OCHs that are used very infrequently. Certain types of protected links that we did not extract, such as those protected by a Captcha, might have a longer lifetime that what we have measured. However, our resulting findings are conservative in that they can only underestimate the size of the OCH ecosystem and the availability of pirated content.

Secondly, while the chosen release blogs are among the largest of this category of indexing sites, they are not representative for the whole OCH-based piracy ecosystem. Yet, they mirror many general trends that we have also informally observed on other types of indexing sites, including the lifetime of links, the diversity of OCHs, and the number of

available mirror copies. Furthermore, since we argue in this paper that anti-piracy measures for OCH are failing, using just one popular and generally accessible site to measure a lower bound on link lifetime and ecosystem diversity is sufficient to show that pirated content remains available.

Note that the indexing sites crawled for this paper reference only a tiny fraction of the files uploaded to OCHs. Sources of uploads not covered by our work include files posted on indexing sites that we did not crawl and files uploaded for private use with download links that were never published. Therefore, our study does not permit to draw any conclusions about the fraction of piracy-related content stored on OCHs.

G. Ethical Considerations

All the data sets were gathered from public data sources only. Even though the INDEXING and LIFETIME data sets occasionally contain the user names of uploaders, this information is publicly reported by indexing sites and OCHs and freely chosen by the uploaders of the files. We have no means to link user names to real-world identities or IP addresses. Therefore, we do not consider the data sets to jeopardise any person's privacy. Furthermore, we implemented best practices such as rate limitation and exponentially increasing retry intervals in order not to overload the queried sites.

IV. IMPACT OF ANTI-PIRACY MEASURES

Based on our measurement data, we assess the real-world effect of a range of measures that aim to reduce the availability of pirated content in the OCH ecosystem: DMCA takedown notices in Section IV-A, anti-piracy measures implemented by OCHs in Section IV-B, and hardware seizure to take down OCHs or indexing sites in Section IV-C.

A. DMCA Takedown Notices

Copyright owners can ask OCHs to take down infringing files. OCHs typically provide an email address that accepts DMCA takedown notices. Some OCHs even give access to an internal link removal tool that allows copyright holders to directly remove infringing links. In order to assess how effective this mechanism is for copyright enforcement, we observe the lifetime of infringing links in the LIFETIME data.

Figure 2 shows the lifetime CDF of links posted on `rlslog.net` in June 2011 and from 8 to 19 March 2012 for a selection of OCHs. We observe two general trends. Firstly, for most OCHs, more than half of the links survive for at least 30 days, even though they are all infringing. Secondly, the shape of the curves is rather steep in the first three to five days and flattens out afterwards. For some OCHs, link deletions increase again after several weeks.

We do not have any ground truth data that would allow us to explain with certainty the shape of the link lifetime curves. However, we make two observations: Users who deliberately publish download links to pirated content appear to have few reasons to delete their files themselves just a few hours or
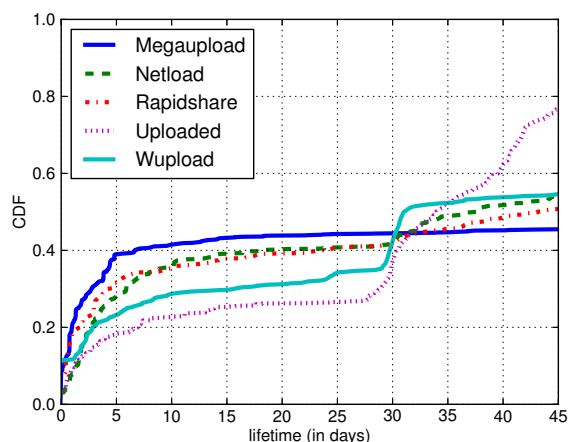
days after the upload—storage capacity is either unlimited or so large that it can hardly be exhausted within a few days. On the contrary, uploaders interested in profit have incentives to keep the files available for longer time periods in order to increase their income. Consequently, link deletions within the first few days are most likely due to copyright holders' takedown efforts with the goal of removing links as quickly as possible. Under this interpretation, infringing files on some OCHs such as Megaupload or Rapidshare appear to be taken down much faster than on others such as Bitshare.

Our second observation is that many OCHs enforce expiry times for files that were uploaded by non-premium users. Wupload, for instance, deleted files uploaded by non-premium users 30 days after the upload. (Note that the statements made on several OCHs' web sites are inconsistent with their actual expiry policies observed in our data and in ad hoc experiments.) A clear increase in the link lifetime CDF curve after a period of stability, such as around 30 days for Wupload and Uploaded in Figure 2(a), makes it plausible that these links were most likely automatically deleted due to expiry. If we admit this explanation, it is striking to observe that on some OCHs, more pirated files appear to be deleted due to expiry rather than because of takedown requests.
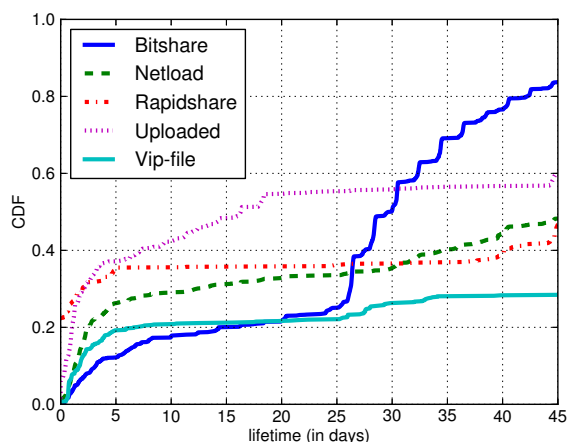
When comparing the situation before and after the shutdown of Megaupload in Figures 2(a) and 2(b), respectively, the curves for Rapidshare and Netload exhibit the same general trend. Uploaded, on the other hand, nearly doubled its deletion rate within the first five days. Furthermore, the linear increase in its curve between day 7 and 18 in Figure 2(b) corresponds to a single mass deletion event: On 26 and 27 March 2012, Uploaded deleted more than 2,800 links from our data sets. The daily average for the other days of the month was only around 400. One might speculate whether the shutdown of Megaupload may have encouraged other OCHs to increase their own anti-piracy diligence.

Determining the exact reasons for link removal may be an interesting topic for future work. However, from a downloader's perspective, the reason why a link becomes unavailable has little importance; what really matters is that most of them remain valid for rather long periods of time.

The next natural question is how the availability of individual links translates to the availability of content. To this end, Figure 3 ranks the content objects added in March of the years 2006 to 2012 on `rlslog.net` by the number of alternative download links. Recall that the indexing site groups download links by content object. Uploaders may add arbitrary amounts of (alternative) download links for each content object. Consequently, the index page for each content object typically contains links to copies uploaded to different OCHs, and frequently also several versions of the same content on the same OCH. At the inception of `rlslog.net` in 2006, the site was mainly posting links to other BitTorrent sites, resulting in few OCH download links. However, from 2009 to 2011, the median number of

(a) Links posted in June 2011.

(b) Links posted from 8 to 19 March 2012.

Figure 2. Link lifetime CDFs on `rlslog.net`. For each OCH, between 742 and 42,092 links were checked for at least 45 days after they had been posted. Link deletions during the first five days appear to be most likely due to copyright violations. Later deletions, such as after 30 days, can additionally be due to link expiry on the OCH.
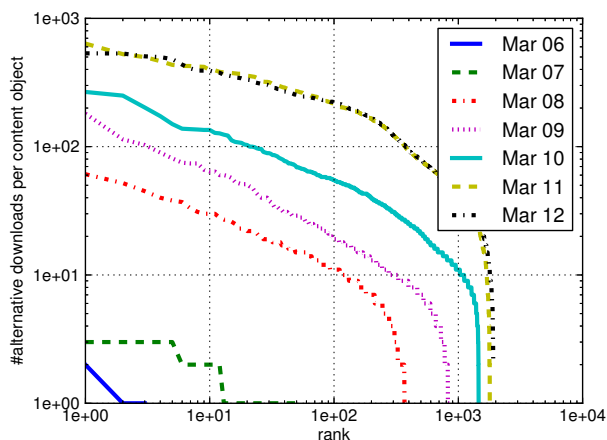

Figure 3. Alternative links per content object on `rlslog.net`. In March 2012, each content object had a median of 43 mirrors. Large numbers of copies would need to be taken down in order to impact content availability.

alternative downloads per content object roughly doubled each year, reaching 48 in 2011, and decreasing slightly to 43 in 2012. We ascribe the latter observation to the aftermath of the Megaupload shutdown, which saw several other OCHs discontinue their affiliate programmes or quit the file sharing business altogether. However, the current situation illustrates that it still takes a lot of effort to effectively take down all copies of pirated content. The high number of copies makes it reasonable to expect the availability of content to be significantly better than the availability of individual links.

In conclusion, the effects of current DMCA takedown efforts as an anti-piracy tool are visible and may cause a nuisance to downloaders who encounter occasional broken links. Yet, our data suggests that they remain limited as to the total extent of copyright infringements. In particular, the availability results presented in this section refer to only one specific indexing site; content availability is strictly larger when considering the whole OCH ecosystem. In the future, the content industry might be able to increase their takedown efforts, but indexing sites could equally aim to better "protect" their links. It is not clear how an increased use of Captchas, for instance, would affect the efficacy of takedown efforts.

B. Effects on Cooperative OCHs

Lawsuits and voluntary action to combat piracy on their services have lead OCHs to implement a range of anti-piracy measures, including file name or hash blacklists, termination of upload incentive programmes, and removal of repeat infringer accounts. Using the global upload rate data provided by IDLOGGER and OCH popularity on indexing sites, we look into how uploaders have reacted to the implementation of such measures. The goal is to assess whether these *local* measures were *globally* effective in reducing piracy.

Figure 4 shows the total number of daily uploads to a range of OCHs according to the IDLOGGER data set. Rapidshare's uploader popularity sharply decreased in the end of June 2010; this coincided with the termination of their rewards programme.[10] Hotfile experienced a similar decrease in upload activity in the middle of February 2011 when they

[10]Ernesto, "RapidShare Kills Reward Program Over Piracy Concerns," http://torrentfreak.com/rapidshare-kills-reward-program-over-piracy-concerns-100620/, 20 Jun. 2010.
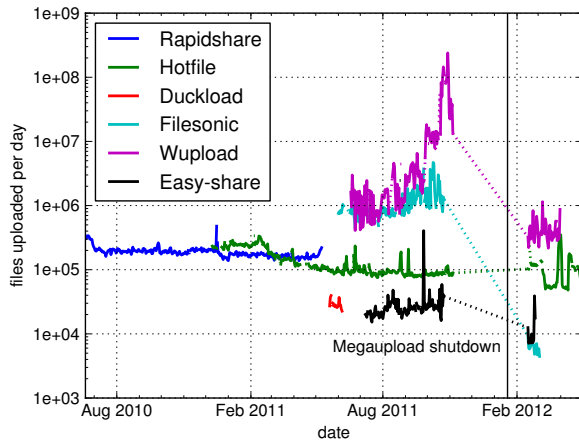
Figure 4.   The global number of files uploaded daily to several OCHs. The dotted parts correspond to offline periods of the IDlogger.
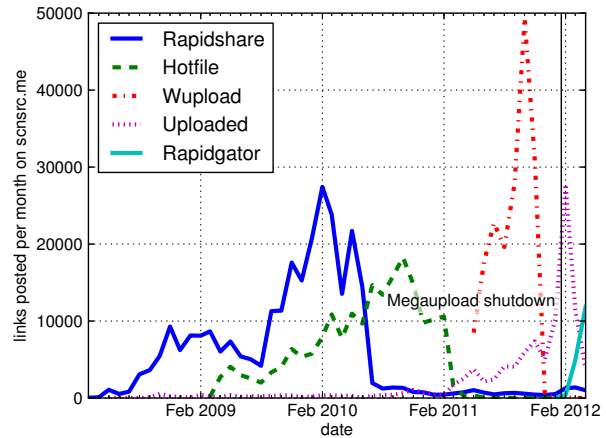

Figure 5.   The number of links posted monthly on `scnsrc.me` for a selection of OCHs. Uploaders quickly switched to other OCHs in the event of unfavourable policy changes.

started to delete the accounts of repeat copyright infringers.[11] On the other hand, other OCHs were able to exponentially increase the number of uploaded files. Wupload, for instance, started operating around May 2011 and initially offered a competitive affiliate programme. However, their upload traffic nosedived when they announced in November 2011 that they were going to put their affiliate programme to an end.[12]

Even though the upload rates reported by IDLOGGER include legitimate files, changes in the upload rates closely coincide with policy changes that mainly affected infringing users of OCHs. This suggests that a non-negligible fraction of some OCHs' upload traffic may have been related to piracy activities, and that the policy changes have led to a local reduction of infringing uploads.

To address this issue from a more global point of view, we look at the number of links per OCH posted monthly on the sites in the INDEXING data set. Figure 5 plots the evolution of OCH popularity for a selection of OCHs on `scnsrc.me`. It makes clear that the fading popularity of one OCH was compensated by the rising popularity of another OCH. Consequently, anti-piracy policies of single OCHs appear to have displaced piracy to other OCHs rather than reducing it on a global scale. As we will show in the next section, the OCH ecosystem is large and diverse, and there are new OCHs entering the market who can potentially take over infringing traffic. After the shutdown of Megaupload and the subsequent voluntary shutdowns of other OCHs, even initially abandoned OCHs such as Rapidshare gained new

popularity with pirates and had to take drastic measures[13] to curb the surging piracy. Interestingly, we observed that the new use of Rapidshare and other "traditional" OCHs was often associated with the use of advertising services such as `adf.ly` to compensate profit-oriented uploaders for the absence of an OCH affiliate programme.

In conclusion, convincing or coercing OCHs to implement more advanced anti-piracy measures can drive away pirates. However, instead of ceasing to pirate, uploaders appear to move to less cooperative OCHs. Lower levels of piracy subsist even on the more diligent OCHs.

C. Hardware Seizure

When the degree of cooperation of OCH file sharing actors is deemed insufficient, it may be an option to locate and seize the server hardware to shut down the site. The impact of such an event on the ecosystem is different according to whether it is an OCH or an indexing site.

1) One-Click Hosters: In order to assess the impact of an OCH shutdown on an indexing site, we analyse the monthly and cumulative use of OCHs on `rlslog.net` as shown in Figure 6. As of April 2012, the site contained links to almost 300 different OCHs, and since June 2009, the site has consistently been using at least 50 different OCHs each month. These numbers illustrate that uploaders have a large choice of OCHs available, and that the OCH ecosystem has grown so diverse that there is no single indispensable actor.

After the shutdown of Megaupload, the slope of the cumulative curve increased slightly, which means that even more new OCHs have entered the ecosystem, despite the now apparent risks. Many of these new OCHs are based on

---

[11]enigmax, "Hotfile Goes To War Against Copyright Infringers," http://torrentfreak.com/hotfile-goes-to-war-against-copyright-infringers-110219/, 19 Feb. 2011.

[12]enigmax, "FileServe and Wupload Exit The File-Sharing Business," http://torrentfreak.com/fileserve-and-wupload-exit-the-file-sharing-business-120403/, 3 Apr. 2012.

[13]enigmax, "RapidShare Slows Download Speeds To Drive Away Pirates," http://torrentfreak.com/rapidshare-slows-download-speeds-to-drive-away-pirates-120224/, 24 Feb. 2012.
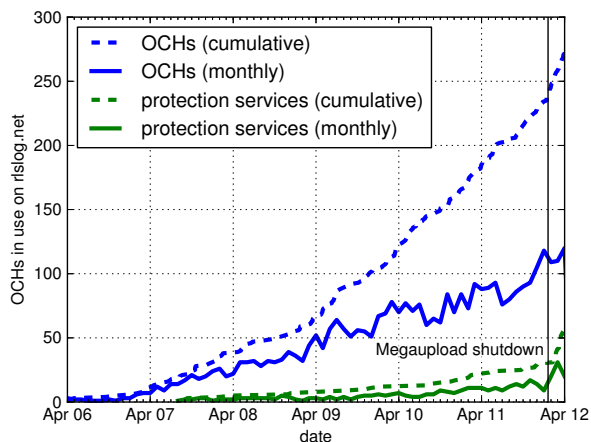
Figure 6. The number of different OCHs used over time, per month and cumulative, and the number of link protection services used on `rlslog.net`. The slope of the cumulative curves increased in February 2012 after the shutdown of Megaupload and the affiliate programmes of several other large OCHs.
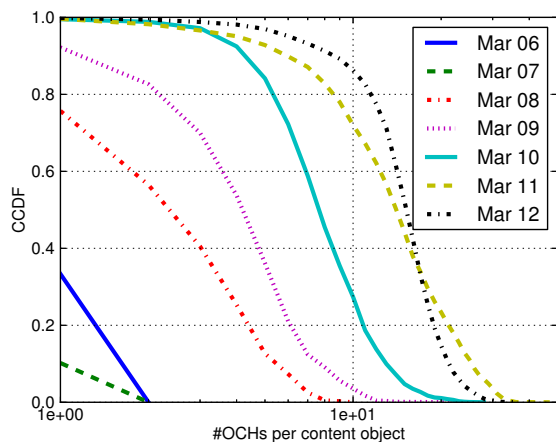

Figure 7. The number of different OCHs used in the download links of content objects on `rlslog.net`.

XFileSharing Pro, a commercial script package that allows webmasters to operate their own OCH for a one-off software licence starting at $ 200 for the basic script and a payment gateway[14] and requiring only moderate programming skills. Overall, we are aware of more than 125 current and former OCHs using this script; the technical barrier of entry to the OCH market appears to be low. Figure 6 also illustrates that uploaders are increasingly using protection services, which include advertising services that help to compensate for the many affiliate programmes that were discontinued in the aftermath of the Megaupload shutdown.

Figure 7 shows that the large OCH diversity also translates to single content objects: In March 2011, more than 80 % of all content objects on `rlslog.net` had links to at least 8 different OCHs; in March 2012, this value increased to over 90 % and the median number of OCHs per content object was 16. The large number of OCHs per content object means that most objects will remain available on other OCHs if any single OCH is shut down. However, other types of indexing sites and especially smaller sites may rely on fewer OCHs and may be more vulnerable to OCH takedowns.

Taking down an OCH is a long and costly legal process, while setting up a new OCH is not. This imbalance results in a large and diverse OCH ecosystem that is difficult to physically take down in its entirety. Uploaders and indexing sites can leverage this advantage to become more resilient to OCH takedown efforts.

2) Indexing Sites: In contrast to OCHs, indexing sites have moderate storage and traffic requirements, which means that indexing sites can be taken "offshore" more easily. It is not unusual at least for larger indexing sites to have servers located in different countries, with the true locations often being hidden behind a load balancing or anti-DDoS service. Investigations against such sites can be time-consuming. The most popular German-language streaming site `kino.to`, for instance, was closed down during a police raid in several European countries on 8 June 2011. According to the rights holder organisation GVU, their preliminary investigation against the site had lasted "for years."[15] However, few weeks later a clone of the site resurfaced at `kinox.to`. It is still online as of December 2012, albeit not as popular.

The FILESTUBE data set allows us to estimate a lower bound on the number of indexing sites on the Internet. As shown in Table II, links pointing to OCHs and associated with piracy-related search terms were found on 9,825 different domains. These domains range from otherwise legitimate services such as Facebook or Blogger to pure "stand-alone" indexing sites. Out of the domains, 814 could not be resolved any more, indicating that the corresponding sites have ceased to exist or moved to a different location. The true number of indexing sites is likely to be much higher because we used only 43 search terms and because Filestube indexes only unprotected links of a limited number of OCHs. Yet, these numbers already illustrate that physically taking down all indexing sites would be a very resource-consuming option.

Table III gives an overview of the 10 most frequent top level domains of the resolved primary and alias domain names as well as the countries of the associated IP addresses looked up in the Maxmind GeoLite database. Some of the disproportionately popular top-level domain names such as `cc` (Cocos Islands) and `ws` (Samoa) have a reputation for allowing anonymous registration and for not being subject to

---

[14]Sibsoft XFileSharing Pro, available at http://sibsoft.net/xfilesharing.html, accessed 6 May 2012.

[15]GVU, "Internationale Durchsuchungsaktion gegen das System kino.to," http://www.gvu.de/media/pdf/768.pdf, press release pub. on 8 Jun. 2011.

Table II

DISTINCT DOMAIN NAMES AND IP ADDRESSES FOUND IN THE FIRST 1,000 RESULTS FOR PIRACY-RELATED SEARCH TERMS ON FILESTUBE.COM (APRIL 2012). INCLUDES 50 SUGGESTED RELATED TERMS PER SEARCH TERM.

	Search Term	Domain Names	Unresolved	Primary	Alias	IP Addresses
specific	californication	257	15	210	43	251
	for dummies	506	31	367	115	388
	lady gaga	721	34	446	264	485
	need for speed	570	36	392	157	444
	photoshop	1,531	147	974	445	951
	other (27 terms)	4,828	325	3,159	1,434	2,766
	total (distinct)	7,045	506	4,352	2,316	3,564
generic	crack	916	100	659	181	629
	dvdrip	1,019	101	747	189	702
	other (9 terms)	3,217	314	2,128	827	1,748
	total (distinct)	4,225	420	2,784	1,085	2,224
	total (distinct)	9,825	814	5,971	3,198	4,478

Table III

TOP 10 TOP-LEVEL DOMAIN NAMES AND IP ADDRESS COUNTRIES FOUND ON FILESTUBE.COM (RESOLVED DOMAINS FROM TAB. II).

Domain Names			IP Addresses		
TLD	#	%	Country	#	%
com	6,197	67.59	U.S.	1,782	39.79
net	1,013	11.05	Germany	439	9.80
org	471	5.14	Netherlands	329	7.35
ru	391	4.26	Russia	307	6.86
info	161	1.76	France	223	4.98
pl	88	0.96	Ukraine	160	3.57
cc	61	0.67	Vietnam	145	3.24
vn	59	0.64	Turkey	120	2.68
ws	52	0.57	U.K.	89	1.99
in	49	0.53	Canada	69	1.54

domain name seizures, but most of the domain names on the list could still be seized by U.S. authorities. The countries of the IP addresses associated with the domain names are more widely spread, illustrating the potential jurisdictional complications in investigations against indexing sites. If an IP address corresponds to a load balancing service, the true server location of the indexing site can be different again.

Overall, the large size and diversity of the OCH-based file sharing ecosystem mean that there is no dependency on any single actor. Physical takedown efforts cannot realistically scale to the extent and speed that would be required to eradicate the entire ecosystem. Complicated and lengthy cross-border investigations against indexing sites and the relative ease with which a new site can be opened rather suggest that the physical takedown of an indexing site is viable only for high-profile cases.

D. Summary

In this section, we analysed our long-term measurement data to characterise the dynamics of the OCH file sharing ecosystem. The following key findings summarise the limitations of current anti-piracy measures:

- On most OCHs, more than 50 % of the files remain online for more than 30 days. Takedown notices might be responsible for only a fraction of pirated files becoming unavailable. Most content objects have so many mirrors that current levels of takedown requests fail to render content unavailable.
- Anti-piracy measures implemented by OCHs appear to successfully reduce piracy locally, but not globally as long as the measures are not globally applied. Because there are hundreds of OCHs and many of them are located abroad, this approach has natural limits.
- The high diversity of the OCH file sharing ecosystem means that shutting down a single actor, even if as large as Megaupload, has little immediate effect on file availability. Yet, such a precedent may cause an increased perceived risk in the ecosystem.

V. DISCUSSION: SOPA

Given the limited effects of current anti-piracy efforts, copyright owners have lobbied for the introduction of more legal tools to combat piracy. The Stop Online Piracy Act (SOPA) was proposed in the U.S. in October 2011. It contained a notice-based takedown system against "foreign U.S.-directed sites dedicated to the theft of U.S. property." The nature of these notices was similar to the DMCA, but targeting entire sites instead of single infringing files.

The SOPA proposal consisted of two main provisions: Interrupting the flow of money to infringing sites, and blocking user access to such sites. In detail, upon receipt of a notice from a copyright holder, payment processors and advertising services would have been required to end their business relationship with the infringing site, and ISPs would have been required to block their customers from accessing the site. These measures were aimed at sites out of reach

of the U.S. justiciary and would have prevented them from receiving visitors or funding coming from the U.S.

After protests, SOPA was put on hold in January 2012. It was controversial for many reasons, including for giving powers perceived as equivalent to direct executive action to private rights holders, and for the risk of intentional or unintentional misuse. By many, SOPA was also viewed as a threat to freedom of speech in general. The purpose of this paper is not to discuss these political views. Rather, our goal is to discuss the pertinence of the proposed technical measures with respect to the current state of the OCH ecosystem. In fact, similar measures are currently being discussed in other countries (e.g., in France [24]), and some anti-piracy strategies are already being implemented by advertising and payment services. Consequently, our discussion of these approaches applies even beyond the scope of SOPA.

A. Disruption of Means of Financing

One-Click Hosters and indexing sites that are run for profit, but also many more altruistic sites, rely on external sources of revenue to support their operation. A recent strategy is to cut off these sites' ability to receive funding, with the goal of making them economically unviable so that the sites will eventually cease operation. This strategy is also known to be used against the online counterfeit pharmaceutical sector [6].

The two main sources of income for sites in the OCH ecosystem are advertising and user subscriptions. Especially indexing sites often rely exclusively on advertising, and sometimes also on voluntary contributions by users and the operators themselves. Major advertising providers such as Google AdSense are known to ban piracy sites. Alternative "underground" advertising networks are available and widely used on file sharing sites, but it has been reported that they may generate lower revenue than mainstream advertising providers.[16] In our experience, some of these underground advertising networks can be difficult to identify. Furthermore, we have observed widespread advertisement leading to fake anti-virus software [25]. These characteristics raise the question whether such services would react to SOPA-style takedown notices. In fact, it seems probable that the status quo would persist in terms of advertising on indexing sites.

For online payments, however, the situation is different. OCHs are most dependent on this means of financing as only few indexing sites collect payments from their users. Recently, PayPal modified their terms of service concerning hosting providers, which effectively banned them from accepting PayPal payments.[17] While OCHs continue to accept credit cards, they now process these payments through services

[16]Steve Bass, "Eleven questions for a warez site owner," http://www.computerworld.com/s/article/9050203/Q_A_Eleven_questions_for_a_warez_site_owner, 30 Nov. 2007.

[17]Ernesto, "PayPal Bans Major File-Hosting Services Over Piracy Concerns," http://torrentfreak.com/paypal-bans-major-file-hosting-services-over-piracy-concerns-120710/, 10 Jul. 2012.

Table IV
AVAILABLE PAYMENT OPTIONS (PMOs) ON OCHS AND THEIR RESELLERS, ORDERED BY THEIR ALEXA RANK.

OCH	Rank	#PMOs OCH	#PMOs Resellers
Mediafire	62	2	> 9
Depositfiles	172	5	> 100
Uploaded	244	7	99
Putlocker	333	2	> 5
Hotfile	393	5	> 45
Turbobit	499	3	> 50
Rapidgator	908	9	> 9
Filefactory	997	4	> 11
Filegag	95,070	2	-

other than PayPal, which accounted for 90 % of user payments on some OCHs, according to the article. These developments might eventually lead to a situation similar to underground pharmacies. According to McCoy et al. [6], [5], only very few banks process credit card payments for underground pharmacies; these sites constantly run the risk of losing their capability of accepting credit card payments and of forfeiting funds held back by their payment processors.

A notable difference between OCHs and underground pharmacies is that many OCHs allow payments through *resellers*. These resellers are often listed on the web site of the OCH and accept a very wide range of payment options, as shown in Table IV. Payment methods include national and international wire transfers, international credit cards and other national payment cards, gift cards, and various online currencies. Users make their payments directly to the resellers and obtain a coupon code that they can redeem on the OCH's web site. While this system makes OCHs less reliable on their own payment processing capability, in our own experience it can be a tedious process for users. Resellers are not always trustworthy, they sometimes charge higher prices than the OCHs themselves, and some means of online payment may incur additional fees. Consequently, one would expect OCHs that lose the capability of accepting direct payments to experience a noticeable drop in revenue. The Usenet site Newzbin2, for instance, recently cited[18] their inability to accept payments as a reason for shutting down, stating that the alternative means of payment "Bitcoin [was not] credible as it [was] just too hard for 90 % of people." In other words, cutting off direct and convenient means of payment appears to be a promising anti-piracy strategy.

In conclusion, OCHs are probably most vulnerable to anti-piracy measures targeted at removing external sources of revenue. Indexing sites may be less affected, especially those that are less driven by (and reliant on) monetary gain.

[18]enigmax, "Newzbin2, the MPAA's Usenet Enemy #1, Calls it Quits," http://torrentfreak.com/newzbin2-the-mpaas-usenet-enemy-1-calls-it-quits-121129/, 29 Nov. 2012.

B. Access Blocking

If physically taking down a site is not feasible or too time-consuming, blocking access to the site may appear to be an attractive option. The two most common techniques for blocking access to a destination are to either prevent the DNS name from resolving to the associated IP address, or to drop network traffic based on the destination's IP address.

Both techniques can be circumvented: Operators of blocked sites can switch to another DNS name or IP address. Visitors of blocked sites can either directly access the site's IP address, use a different DNS server, or use existing tools such as VPNs, proxy servers, or anonymisation services such as Tor.

Such a measure will probably affect OCHs most heavily because of their potentially large number of servers and their higher bandwidth requirements. Indexing sites, on the other hand, are more flexible in changing their DNS name or IP address. Their bandwidth requirements are rather modest so that free services such as Tor or open proxies could be an adequate means of accessing the sites.

While circumvention technology is commonly and freely available today, it is an open question how many users would install such software, and how much the traffic to the blocked sites would decline. As a first result in the area of BitTorrent, research from The Netherlands[19] has shown that when access to The Pirate Bay was blocked, there was no significant effect on BitTorrent piracy, suggesting that users have found and are applying means to circumvent that anti-piracy measure.

C. Summary

According to our findings, OCHs are probably more vulnerable to the anti-piracy strategies discussed in the context of SOPA than are indexing sites. It is not clear, however, whether OCHs would be the right target for SOPA-style takedown notices. According to our law laymen's intuition, it appears to be more challenging to declare OCHs as overall infringing or "dedicated to the theft of U.S. property" than indexing sites. The latter typically openly display the infringing content that they offer, while OCHs tend to be more opaque, honour DMCA takedown notices, and are also frequently used for entirely benign purposes. Indexing sites, on the other hand, appear to be more resilient against SOPA-style takedown notices. Hence, one may criticise the proposed SOPA takedown regime for potentially introducing an increased risk of abuse without being highly effective against the truly infringing sites. Furthermore, some of the anti-piracy strategies underlying SOPA are already being deployed in a less far-reaching and less controversial way than SOPA takedown notices.

Overall, the financial and legal pressure on OCHs to step up their anti-piracy efforts will probably increase. However, less cooperative OCHs may drift into underground situations similar to the providers of malware and spam-advertised goods. Hence, it is highly likely that illegal file sharing based on OCHs will not entirely go away, and that the arms race between pirates and copyright enforcement will continue.

Given our findings that highlight the difficulties of reducing the supply of pirated content, it appears to be promising to follow a complementary strategy of reducing the *demand* for pirated content, e.g., by providing legitimate offers that are more attractive to consumers than pirating content.

VI. CONCLUSION

There is a cat-and-mouse game between uploaders and copyright owners, where pirated content is being uploaded by the former and deleted by the latter, and where new One-Click Hosters and direct download sites are appearing while others are being shut down. Currently, this game seems to be in favour of the many pirates who provide far more content than what the copyright owners are taking down. This imbalance of powers, together with the size and diversity of the One-Click Hosting ecosystem, suggest that anti-piracy laws to reduce the availability and reachability of pirated content may be less successful than what their proponents might expect. On the other hand, strategies against economically motivated actors may effectively target certain parts of the ecosystem. Together with increased legal proceedings, we expect them to render many OCHs more diligent, and maybe even proactive, in their own anti-piracy efforts.

REFERENCES

[1] C. Kanich, C. Kreibich, K. Levchenko, B. Enright, G. M. Voelker, V. Paxon, and S. Savage, "Spamalytics: An empirical analysis of spam marketing conversion," in *CCS '08*. ACM, Oct. 2008.

[2] K. Levchenko, A. Pitsillidis, N. Chachra, B. Enright, M. Félegyházi, C. Grier, T. Halvorson, C. Kanich, C. Kreibich, H. Liu, D. McCoy, N. Weaver, V. Paxson, G. M. Voelker, and S. Savage, "Click trajectories: End-to-end analysis of the spam value chain," in *Symposium on Security & Privacy 2011*. IEEE, Mar. 2011.

[3] B. Stone-Gross, T. Holz, G. Stringhini, and G. Vigna, "The underground economy of spam: A botmaster's perspective of coordinating large-scale spam campaigns," in *LEET '11*. Usenix, Mar. 2011.

---

[19] Ernesto, "Censoring The Pirate Bay is Useless, Research Shows," http://torrentfreak.com/censoring-the-pirate-bay-is-useless-research-shows-120413/, 13 Apr. 2012.

[4] C. Kanich, N. Weaver, D. McCoy, T. Halvorson, C. Kreibich, K. Levchenko, V. Paxson, G. M. Voelker, and S. Savage, "Show me the money: Characterizing spam-advertised revenue," in *USENIX Security 2011*. Usenix, Aug. 2011.

[5] D. McCoy, A. Pitsillidis, G. Jordan, N. Weaver, C. Kreibich, B. Krebs, G. M. Voelker, S. Savage, and K. Levchenko, "PharmaLeaks: Understanding the business of online pharmaceutical affiliate programs," in *USENIX Security 2012*. Usenix, Aug. 2012.

[6] D. McCoy, H. Dharmdasani, C. Kreibich, G. M. Voelker, and S. Savage, "Priceless: The role of payments in abuse-advertised goods," in *CCS '12*. ACM, Oct. 2012.

[7] H. Fallmann, G. Wondracek, and C. Platzer, "Covertly probing underground economy marketplaces," in *DIMVA '10*. Berlin, Heidelberg: Springer-Verlag, 2010, pp. 101–110.

[8] G. Wondracek, T. Holz, C. Platzer, E. Kirda, and C. Kruegel, "Is the internet for porn? An insight into the online adult industry," in *WEIS '10*, Jun. 2010.

[9] J. Caballero, C. Grier, C. Kreibich, and V. Paxson, "Measuring pay-per-install: The commoditization of malware distribution," in *USENIX Security 2011*. Usenix, Aug. 2011.

[10] D. Antoniades, E. Markatos, and C. Dovrolis, "One-click hosting services: A file-sharing hideout," in *IMC '09*. ACM, Nov. 2009.

[11] A. Mahanti, C. Williamson, N. Carlsson, M. Arlitt, and A. Mahanti, "Characterizing the file hosting ecosystem: A view from the edge," *Performance Evaluation*, vol. 68, no. 11, pp. 1085–1102, Nov. 2011.

[12] J. Sanjuàs-Cuxart, P. Barlet-Ros, and J. Solé-Pareta, "Measurement based analysis of one-click file hosting services," *Journal of Network and Systems Management*, May 2011.

[13] A. Mahanti, N. Carlsson, and C. Williamson, "Content sharing dynamics in the global file hosting landscape," in *MASCOTS '12*. IEEE, Aug. 2012, pp. 219–228.

[14] N. Nikiforakis, M. Balduzzi, S. V. Acker, W. Joosen, and D. Balzarotti, "Exposing the lack of privacy in file hosting services," in *LEET '11*. Usenix, Mar. 2011.

[15] T. Lauinger, E. Kirda, and P. Michiardi, "Paying for piracy? An analysis of one-click hosters' controversial reward schemes," in *RAID '12*. Berlin, Heidelberg: Springer-Verlag, Sep. 2012, pp. 169–189.

[16] Z. Jelveh and K. Ross, "Profiting from filesharing: A measurement study of economic incentives in cyberlockers," in *P2P '12*. IEEE, Sep. 2012.

[17] J. M. Urban and L. Quilter, "Efficient process or "chilling effects"? Takedown notices under section 512 of the Digital Millennium Copyright Act," *Santa Clara Computer & High Technology Law Journal*, vol. 22, no. 4, pp. 621–693, Jun. 2006.

[18] A. Rehn, "The politics of contraband: The honor economies of the warez scene," *Journal of Socio-Economics*, vol. 33, no. 3, pp. 359–374, 2004.

[19] D. McCandless, "Warez world," *Telepolis*, Jul. 2001.

[20] J. Howe, "The shadow Internet," *Wired*, vol. 13, no. 01, Jan. 2005.

[21] J. Liang, R. Kumar, Y. Xi, and K. Ross, "Pollution in P2P file sharing systems," in *INFOCOM '05*. IEEE, Mar. 2005.

[22] R. Cuevas, M. Kryczka, A. Cuevas, S. Kaune, C. Guerrero, and R. Rejaie, "Is content publishing in BitTorrent altruistic or profit-driven?" in *Co-NEXT '10*. ACM, Nov. 2010.

[23] M. Kammerstetter, C. Platzer, and G. Wondracek, "Vanity, cracks and malware: Insights into the anti-copy protection ecosystem," in *CCS '12*. ACM, Oct. 2012.

[24] V. Bonneau and G. Fontaine, "Etude du modèle économique de sites ou services de streaming et de téléchargement direct de contenus illicites," http://www.hadopi.fr/sites/default/files/page/pdf/Rapport_IDATE.pdf, Hadopi, France, Tech. Rep., Mar. 2012.

[25] B. Stone-Gross, R. Abman, R. Kemmerer, C. Kruegel, D. G. Steigerwald, and G. Vigna, "The underground economy of fake antivirus software," in *WEIS '11*, Jun. 2011.