

TorrentGuard: stopping scam and malware distribution in the BitTorrent ecosystem

Michał Kryczka^{*†}, Ruben Cuevas[†], Roberto Gonzalez[†] Angel Cuevas[‡] and Arturo Azcorra[†]

^{*}Institute IMDEA Networks

[†]University Carlos III Madrid

[‡]Telecom SudParis

Abstract—In this paper we conduct a large scale measurement study in order to analyse the fake content publishing phenomenon in the BitTorrent Ecosystem. Our results reveal that fake content represents an important portion (35%) of those files shared in BitTorrent and just a few tens of users are responsible for 90% of this content. Furthermore, more than 99% of the analysed fake files are linked to either malware or scam websites. This creates a serious threat for the BitTorrent ecosystem. To address this issue, we present a new detection tool named TorrentGuard for the early detection of fake content. Based on our evaluation this tool may prevent the download of more than 35 millions of fake files per year. This could help to reduce the number of computer infections and scams suffered by BitTorrent users. TorrentGuard is already available and it can be accessed through both a webpage or a Vuze plugin.

I. INTRODUCTION

BitTorrent is one of the most popular applications in the current Internet. It is daily utilised by millions of users and is responsible for a major portion of the Internet traffic [21]. This success motivated the research community to investigate different aspects of BitTorrent covering performance [14], [19], economics [4], [6], [25] and incentives [11], [22] issues. However, to the best of the author knowledge, the research community has dedicated less attention to BitTorrent security aspects. Some previous works have analysed the vulnerabilities of BitTorrent protocol to free-riders [15], [16], [24] whereas some others [2] address the lack of privacy offered by BitTorrent. More recently, in a previous work [5] we demonstrated that the BitTorrent ecosystem is suffering from a continuous *poisoning index* attack by which 30% of the published content is associated to fake content. Furthermore, this fake content produces 25% of the download events it means every 4th download which happens in BitTorrent is associated with fake content. These initial results highlight a serious issue that, to the best of the authors knowledge, has not been still covered by the research community.

In this paper, we thoroughly analyse the *fake publishing* phenomenon in BitTorrent in order to understand its real impact on the system performance as well as the potential risks of fake content for BitTorrent users. Furthermore, we propose a practical solution to mitigate this problem. We base our study on data collected from torrents published in The Pirate Bay portal during a period of 14 days from 30-04-2011 to 13-05-2011. The 35% of almost 30K analysed torrents are

associated to fake content. This depicts a 5% increment in the presence of fake content within the BitTorrent ecosystem in a period of one year between our two measurement studies. This justifies (even more) the necessity of the research conducted in this paper.

In order to fight the fake publishing phenomenon, the first step is to properly characterise the fake publishers and their behaviour. The current BitTorrent portals solutions identify fake publishers through the user account that they use to upload fake torrents to the portal. We show in the paper that this technique is inefficient since the fake publisher can generate as many user accounts as needed in those portals. Instead, the parameter that uniquely identifies the fake publisher is the IP address it uses to perform its activity. Surprisingly, our data reveals that just 20 fake publishers (whose IP we identify) are responsible for injecting 90% of fake content in the BitTorrent ecosystem. Moreover, most of these IP addresses belong to Hosting Providers where the fake publishers rent dedicated high-resource servers to perform their activity.

Therefore, the fake publishing activity is time consuming since a fake publisher needs to manually create the user accounts used in the different portals (in some cases up to 4 accounts per day). Furthermore, this activity requires dedicated resources (e.g. rented servers). This investment in time and resources can be only justified by a strong motivation behind the distribution of fake content. We have downloaded and manually inspected a large number of fake content published during our measurement period and found 3 different profiles among the fake publishers: (i) a first group of fake publishers aims to spread malware using the popular BitTorrent system; (ii) a second set of users tries to attract BitTorrent users to scam websites in order to get economical benefit from the victims by using different scam techniques; (iii) the last group is formed by antipiracy agencies that upload fake versions of those content that they want to protect.

Our data shows that more than 99% of the published fake content is associated with the two first profiles. This supposes a very serious threat for the BitTorrent ecosystem since the activity of these publishers may lead to thousands of undesirable episodes of scammed users and computer infections. These findings suggest that new solutions need to be proposed in order to eliminate or at least reduce the number of fake content available in the BitTorrent ecosystem. Towards this

end, we have designed and implemented TorrentGuard. This is a novel detection tool that allows to identify the IP address of the fake publisher, thus being able to report as fake each content published from this IP address at the moment of its publication. Based on the performed evaluation, TorrentGuard would be able to avoid more than 35 millions fake content downloads every year. This means, preventing hundreds of thousands of users to suffer from computer infections or scam incidents every year. TorrentGuard can be currently used through a publicly available website and a Vuze plugin.

The rest of the paper is structured as follows. Section II presents the background information. In Section III we describe our measurement methodology and present our dataset. Next, Section IV characterises fake publishers, while Section V classifies them depending on the goal they pursue with their activity. Section VI shortly characterises the downloaders of the fake content. In Section VII we describe and evaluate our solution to improve the detection of fake content. Section VIII describes relevant works to this paper. Finally, Section IX concludes the paper.

II. BACKGROUND

In this Section we briefly describe the main aspects of the BitTorrent ecosystem making a special emphasis on the procedure of publishing content on The Pirate Bay (and by extension on other BitTorrent portals) and specifically, how fake publishers do it. This is summarised in Figure 1. For a full description of the BitTorrent ecosystem we refer the reader to [13] and [26].

A. Main elements of BitTorrent ecosystem

- *BitTorrent Portals*: these are webpages which index .torrent files, classify them into different categories and provide basic information for each file. These portals serve as rendez-vous points between content publishers and BitTorrent downloaders. The publishers upload their .torrent files to BitTorrent portal and the clients download them.

- *.torrent file*: this is a meta-information file including relevant information for the BitTorrent protocol such as: (i) the content infohash, this is a unique identifier of the content in the BitTorrent ecosystem; (ii) the IP address of the BitTorrent Tracker managing the content distribution process; (iii) the size of the content and the number of pieces forming the file.

- *BitTorrent Trackers*: these are servers which manage the BitTorrent download process of a given content. The set of all peers downloading a given file is named *swarm*. The tracker maintains a list with the IP addresses and the download progress of all the peers forming the swarm associated to a specific content. Furthermore, when a new peer joins the swarm, it contacts the tracker in order to obtain a list of IP addresses of other peers participating in the swarm. By doing so, the new incomer is able to retrieve pieces of the content from these peers.

- *BitTorrent downloaders (peers)*: these are clients forming the swarm that download and/or upload pieces of the content. We distinguish two types of peers. A *seeder* is a peer that

possess a complete copy of the content, thus only uploads pieces whereas a *leecher* does not have the complete file so that it uploads and downloads pieces.

- *BitTorrent publishers*: these are the clients that make available the first copy of the content in the BitTorrent ecosystem.

B. Publishing a content in BitTorrent

When a publisher wants to publish a content in the BitTorrent ecosystem, it firstly creates a .torrent file. After creating the .torrent file, the publisher uploads it to one or more BitTorrent portals. For this purpose, it uses a user account (with a specific username) created in these portals. Furthermore, the publisher distributes the first copy of the content by acting as the initial seeder in the associated swarm. Therefore, *the content publisher can be identified by the IP address of the initial seeder distributing the content and by the username utilised to upload the content to a BitTorrent Portal.*

In this paper we specifically address the fake content publishing phenomenon in BitTorrent. A fake publisher is a user that exploits the BitTorrent ecosystem to publish fake content, this is, content that is different than what is expected from the content name. A fake publisher makes available the fake content from a single IP address (or limited number of IP addresses) that corresponds to the initial seeder of all its published content. Furthermore, a fake publisher typically creates a user account in a BitTorrent portal from which it uploads .torrent files associated with its fake content. Some portals, such as The Pirate Bay, removes this user account after some client reports that it is being used to publish fake content. Then, the fake publisher reacts by creating a new account to publish new .torrent files and this loop keeps repeating. Hence, contrary to the case of regular publishers (that can be identified by its associated username in the BitTorrent portal), fake publishers can exclusively be identified by its IP address. Finally, it must be noted that, to the best of the authors knowledge, the previously described technique based on users' reports is the only one used nowadays for detecting and deleting fake content.

C. Downloading a content in BitTorrent

When a user wishes to download a content, it first downloads the .torrent file associated to the content from a BitTorrent portal such as The Pirate Bay. Then, the user retrieves the IP address of the Tracker managing the swarm from the .torrent file and connects to it. The Tracker provides the user with a list (50 to 200) of IP addresses participating in the swarm along with the number of seeders and leechers forming the swarm. Finally, the user starts downloading the pieces of the content from the obtained IP addresses.

D. BitTorrent Portals, the case of The Pirate Bay

We use The Pirate Bay as the reference BitTorrent Portal for our study. Previous works [26] have demonstrated that the Pirate Bay is a key element and the most important portal in the BitTorrent ecosystem. A publisher needs to create a user account in order to upload .torrent files to The Pirate

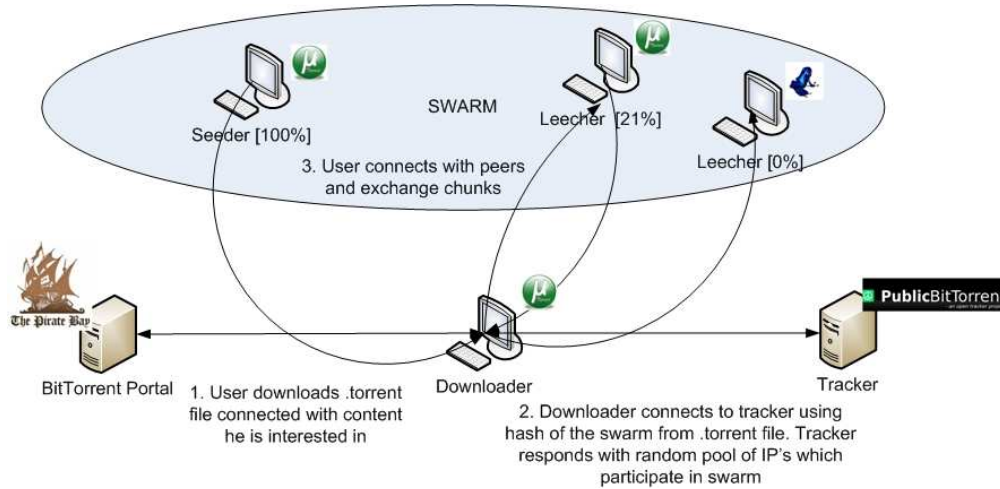


Fig. 1. BitTorrent ecosystem basic functionality

Bay whereas other portals such as IsoHunt [10] use crawling techniques to obtain the offered content from third portals such as The Pirate Bay. Hence, The Pirate Bay is the most interesting portal to be considered in order to understand the content publishing phenomenon in BitTorrent. Specifically, The Pirate Bay offer the following relevant services to our study: (i) an RSS feed system in which each new published content is announced along with the username that uploaded the .torrent file to the portal; (ii) each user registered within The Pirate Bay portal has an individual webpage in which its published torrents are listed and (iii) The Pirate Bay removes the accounts, webpages and .torrent files of those users whose content is detected as fake. Typically, this happens after a client, who downloaded the content, reports its falseness to The Pirate Bay administrators.

III. MEASUREMENT METHODOLOGY

This Section describes our measurement methodology to identify and characterise the main properties of the fake publishers (i.e. users publishing fake content). For this purpose we crawl The Pirate Bay, the most popular BitTorrent portal (as reported by previous works [26] and by Alexa Ranking [1]).

The main objective of our measurement study is to identify fake publishers. Towards this end, our measurement tool has three independent modules. The first one is responsible for finding the IP address and username of the publisher associated with each announced content in The Pirate Bay. For this purpose, the module is subscribed to the RSS feed of The Pirate Bay in order to learn each torrent just after its birth. After getting a new .torrent file the tool obtains the username that uploaded the .torrent file to The Pirate Bay and also connects to the associated Tracker to obtain the IP addresses of the peers forming the swarm in its very initial stage. Then, it is very likely that we can find the IP address of the content publisher (initial seeder). Specifically, we face three different situations: (i) The tracker only reports the IP address of the initial seeder. This is likely to happen since we connect to the

swarms just after the torrent birth. (ii) The tracker announces the presence of one seeder and few leechers in the swarm. Then, by connecting to all these peers and obtaining their bitfields (vector that shows the number of pieces that a peer possesses) we are able to identify which one is the initial seeder, and thus the content publisher. (iii) In some cases, the Tracker announces the presence of quite a few seeders in the swarm thus we cannot identify the initial seeder. This happens because the swarm has been formed before the torrent is announced in the RSS feed of The Pirate Bay portal. Therefore, using the described methodology we are able to characterise the content publisher by both its username and IP address in many cases.

The second module of our tool is responsible for identifying those publishers that are in fact fake ones. For this purpose our tool connects periodically (every 5 minutes) to the Pirate Bay webpage of each known publisher. If at some point the Pirate Bay webpage has been removed we consider that the IP address associated with the removed account belongs to a fake publisher. Furthermore, we also collect the time that The Pirate Bay requires to detect and eliminate each fake publisher account.

Finally, our tool has a third module that counts the number of peers that connect to the swarm of each fake content in order to download it. Specifically, our tool systematically queries the Tracker managing the download of each fake content to obtain those IP addresses participating in the swarm. In order to accelerate this process we perform this task from four independent machines.

A. Dataset description

We have applied the described methodology between 30-04-2011 and 13-05-2011, in addition to 5 days of warm-up phase dedicated to identify the initial fake publishers' IP addresses. During the measurement period we have collected 29330 torrents, from which 10206 (35%) were identified as fake ones. Furthermore, we have collected the IP addresses of those peer participating in swarms associated with fake content

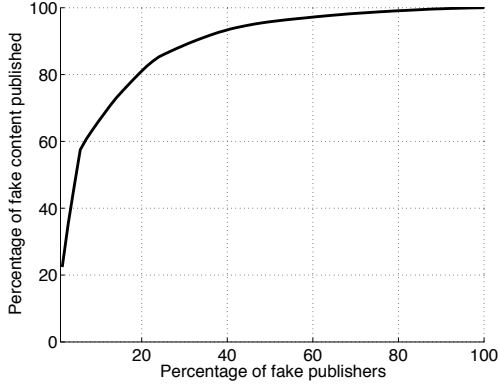


Fig. 2. Percentage of fake content published by the top x% fake publishers

until two instants: (i) the moment the content is removed from The Pirate Bay and (ii) the end of our measurement study.

IV. FAKE PUBLISHERS CHARACTERIZATION

Our results reveal that more than 1/3 of the content published in the Pirate Bay is fake. This shows an increasing trend in the number of fake content regarding our previous study done one year ago when the fake content represented 30%. Therefore, it is critical to eliminate or at least reduce this huge number of fake content in BitTorrent ecosystem. The first step towards this end is identifying who is responsible for publishing this fake content and characterising their behaviour. In this Section, we address this issue using the collected data. More specifically, we aim to answer questions such as: *How many fake publishers (i.e. IP addresses) are uploading fake content to the BitTorrent Ecosystem?*, *From where (i.e. which ISP) they perform their activity?* or *How frequently they upload fake content?*.

A. Number and Contribution of Fake Publishers

Unexpectedly, we observe that only 71 IP addresses are responsible for those 4779 fake content for which we identified the initial seeder. This implies almost 70 fake content published from each of these IPs in average. However, it is interesting to investigate the level of the contribution of each one of these fake publishers. Towards this end, Figure 2 depicts the percentage of fake content published by the top x% of these fake publishers. The graph shows a skewed distribution where 10 IPs (14%) are responsible for publishing almost 75% of all the fake contents. Moreover, this number increases to 90% if we consider the top 20 IP addresses. Therefore, we can conclude that a reduced number of just 20 fake publishers are responsible for poisoning the BitTorrent ecosystem. In the rest of the paper we focus on thoroughly studying this group of 20 fake publishers that we refer to as *Top Fake Publishers*.

B. Location of fake publishers

We have mapped the IP address of each one of the Top Fake Publishers to its correspondent ISP using the MaxMind database [17]. Surprisingly, 17 out of the Top 20 fake publishers operate from Hosting Providers. These are companies dedicated to rent high-resources (cpu, memory and bandwidth)

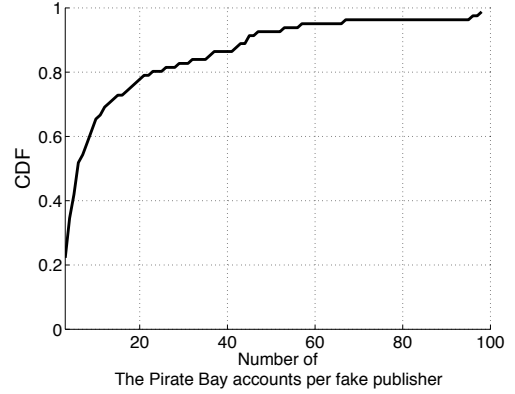


Fig. 3. CDF of the number of The Pirate Bay accounts per fake publisher

provisioned servers. Moreover, 70% of the fake content is seeded from just two Hosting Providers named *OVH Systems* and *Obtrix* located at France and New Zealand respectively.

Fake publishers need on the one side resources in order to sustain the distribution of a large number of fake files [5] and on the other side anonymity due to the *illegitimate* activity being performed. The usage of rented servers in Hosting Providers covers both requirements.

Hence, the use of dedicated servers in Hosting Providers reveals that most of the fake publishers perform their activity from a stable IP since those servers typically have a static IP address configured. This makes them easily identifiable. In this sense, the usage of anonymity services such as TOR [20] or proxy services seems to be useful for fake publishers in order to difficult their identification. However, we have not found that the fake publishers identified in our dataset use such services. This suggests that the severe performance degradation associated to these anonymity services prevent the fake publishers from using them.

C. Pirate Bay accounts utilisation

The Pirate Bay solicits to solve a CAPTCHA [3] in order to create an account to avoid the automatic generation of accounts. Hence, fake publishers are obliged to create their accounts manually. Figure 3 shows the CDF of the number of The Pirate Bay accounts used by each one of the 71 identified fake publishers. A fake publisher use (in median) 6 accounts in a period of 14 days. However, a 5% of the fake publishers injects content using more than 58 different accounts in the same period. This represents an average number of 4 accounts per day. This result suggests that fake publishers need to dedicate time to track the availability of their accounts in order to manually generate new ones if needed.

Interestingly, we also observe a second strategy that although marginal is worth to report. In these cases, fake publishers hijack the accounts with a legitimate publishing history. This provides a trust reputation among the downloaders. Therefore, this could extend the time that fake user could be injecting fake torrents before being reported. However, due to the required technical skills for applying this technique, this case represents less than 1% of all fake accounts.

D. Publishing Strategies

Fake users follow two different strategies to upload fake contents into The Pirate Bay portal. On the one hand, we found users that publish a large number of fake content in a row (typically around 10) in just few seconds after creating a user account. Once the account is deleted, they repeat the same process from a new account. Around 70% of Top Fake Publishers use this technique. On the other hand, 30% of the Top Fake Publishers upload just one or two fake contents with a username. This is a more conservative technique that extends the time that those fake accounts are active before being eliminated when compared to the previous case. Specifically, the accounts of those publishers using the first strategy are detected and then deleted in 92 minutes (in average) whereas the accounts of those using the second strategy are deleted in 253 minutes, thus being their content available 2.75 times more time in The Pirate Bay. Unexpectedly, although the second strategy offers longer accounts' lifetime, it attracts only 47 downloaders per torrent (in average) in front of the 84 attracted by fake publishers using the first strategy. This happens because the fake publishers using the first strategy typically use popular names associated to their content whereas publishers using the second more conservative strategy do not use such popular names.

E. Strategies to attract downloaders

The main goal of fake publishers in BitTorrent is to produce as many downloads of their content as possible. Therefore, they need to offer torrents that sound very attractive for the downloaders. Towards this end, we have observed that fake publishers use three different strategies: (i) they assign to the content a very popular name such as the title of the last released Hollywood movies; (ii) creating the false impression that the content has been published by a well-known and trusted user. For this purpose, the fake publisher names its content in the same way as the trusted one. For instance, eztv one of the most popular publisher in The Pirate Bay adds the signature [eztv] at the end of the title of its published files. Then, some fake publishers also add this signature to the title of their fake content; (iii) presenting attractive performance statistics (i.e. a high number of seeders and leechers) for the fake torrent. In this way, the fake torrent is perceived as a very popular torrent by the downloaders, that assume they will obtain a high download rate in case of selecting that torrent. In order to generate these fake statistics the publisher connects to the Tracker many times using a single IP but different ports. Then the tracker considers each one of these IP+port pairs as a single peer and reports a high number of seeders and leechers. The Pirate Bay retrieves and presents these statistics from the Tracker.

In summary, the fake content publishing activity is performed from Hosting Providers facilities by just few dozens of users. Furthermore, fake publishers are aware of how the BitTorrent ecosystem works, thus they use sophisticated strategies in order to improve the success of their activity.

V. FAKE PUBLISHERS PROFILES

After characterising the Fake Publishers behaviour, we still need to answer an important question: *What incentives a user has to publish fake content?*. To answer this question we have downloaded up to 10 files published by each fake publishers in our dataset and manually inspected them. Our analysis reveals the presence of three different profiles: malware propagators, scammers and antipiracy agencies. Next, we describe in detail each one of these profiles.

A. Malware propagators

These users exploit the popularity of BitTorrent system in order to rapidly propagate malware among thousands of users. On the one hand, for some of the users in this group the published content is the malware itself. In this case, the content including the malware pretends to be typically a patch for a popular game, a key generator, etc. On the other hand, a second set of users use a more sophisticated technique. They publish a movie with a catchy title. The content has the standard size of a DivX movie (i.e. between 700MB and 1GB), and even sometimes includes a second small file with a real sample of the movie. Hence, the file has the appearance of a non-fake legitimate content. However, when a user downloads the content and tries to play the movie, it is requested to reproduce it using Windows Media Player (WMP) in case a different player is run instead. When the movie is finally reproduced with the WMP a pop-up window appears requesting to install new codecs along with an url link from where these codecs can be downloaded. Of course, the file including those pretended codecs is reported as a malware by security and anti-virus software.

B. Scammers

In this case, the fake publisher uses a similar technique to the sophisticated one described above. However, when the user plays the movie with WMP, it is automatically redirected to a website in the Internet. A second variant used by scammers is to provide a file protected with a password (typically .rar), and offer the user a website in which the password can be obtained. Once the user gets into one of these websites, a credit card payment is requested in order to obtain some privilege to watch the downloaded movie (e.g. the password of the .rar file). In some other situations the user is informed that in order to check he is not a bot, a survey must be filled previously to watch the movie. This survey results to be a contest in which you are obeyed to subscribe to a paid premium SMS service. These websites are often reported as scam on different forums, one example of them is <http://movieyt.com>.

We have performed a more detailed analysis of these websites. On the one hand, when a user wants to abandon the webpage several pop-up windows appear trying to change user mind and making leaving the webpage at least bothersome. On the other hand, when a user enters some of these webpages, a pop-up window advertising a Facebook group of the webpage shows up. This pop-up does not react to the explorer close button, rather, just by clicking on the "I like it" Facebook

Country	Percentage of BitTorrent users downloading fake content	Percentage of BitTorrent users	The ratio
United States	12.40%	10.42%	1.19
China	6.27%	4.20%	1.49
Great Britain	4.60%	6.26%	0.73
Brazil	4.26%	2.68%	1.59
Italy	3.88%	4.13%	0.94
India	3.78%	5.71%	0.66
Canada	3.29%	3.85%	0.85
Spain	2.79%	5.95%	0.47
Austria %	2.73%	2.83%	0.96
Poland	2.66%	2.86%	0.93

TABLE I
DEMOGRAPHICS OF BITTORRENT USERS VS FAKE CONTENT
DOWNLOADERS PER COUNTRY (THE THIRD COLUMN REPRESENT THE
RATIO COLUMN 1/COLUMN 2)

button the window closes. This method aims to increase the trust of the webpage so that users interpret it is a legitimate website. More importantly, this finding suggests that these scammers do not limit their activity to BitTorrent but they also try to capture victims from other popular applications such as online social networks.

C. Antipiracy Agencies

The two previous profiles have dishonest purposes. Antipiracy agencies instead, publish fake version of the copyrighted content that they want to protect. This content however, is not what downloader is expecting from the title (e.g. copyrighted movie). Sometimes this content includes antipiracy adverts. The action performed by antipiracy agencies is limited in the number of contents (under request from a company) and time (in the weeks before and after the content, e.g. movie, is released).

In summary, we distinguish three different profiles among fake publishers that motivates them to perform their activity. On the one hand, 65% of the Top Fake Publishers in our dataset are malware propagators and are responsible for around a 30% of the published fake content. On the other hand, a 35% of the Top Publishers are scammers and they published a 70% of the fake content during our measurement period. Finally, antipiracy agencies represent a very small fraction of the fake content published due to the specificity of their actions.

In conclusion, it is worth to mention that the content published by malware propagators and scammers is potentially harmful, specially for not technically skilled downloaders. Hence, they represent a serious risk for the BitTorrent ecosystem that should be erased or at least mitigated. We address this issue in Section VII.

VI. CHARACTERIZING THE DOWNLOADERS OF FAKE CONTENT

In this Section we look at the studied phenomenon from the victims side. First we analyse the demographics of the victims and group them per country in order to understand which countries suffer more from the reported problem. In order to provide full meaningful results we have compared

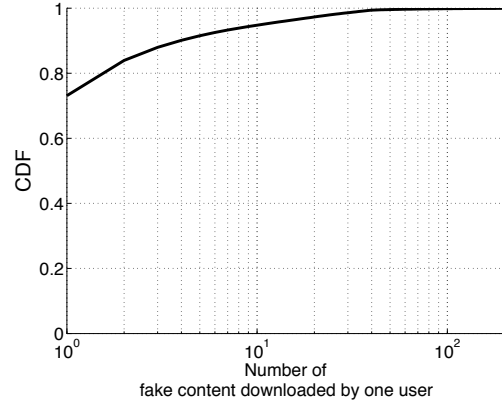


Fig. 4. CDF of the number of fake content downloaded by one user

the demographic distribution of the victims of fake content and the demographic of distribution of BitTorrent clients obtained from the dataset used in our previous work [5] that includes the IP address of 27M clients associated to around 40k torrents.

Table I offers the obtained results. It shows the percentage of victim downloaders of fake content, the percentage of BitTorrent users and the ratio between these two percentages for the 10 countries with a larger number of victims. If the victims were randomly selected, this ratio would be close to 1. However, this is not the case. On the one hand, we observe that some countries such as US, China and Brazil shows a ratio > 1 . For instance, Brazil has a ratio equal to 1.59. This means that Brazil has 59% more victims than expected from a random process. On the other hand, countries such as UK, India or Spain shows a value < 1 . For instance Spain has a ratio equal to 0.47. This means, Spain only has 47% of the victims it should have from a random process.

Next, we study the number of fake content downloads performed by a single user. This help to understand whether there are users that are highly vulnerable to the described threats. Figure 4 shows the CDF of the number of fake content downloaded by a each victim. We can see that 70% of the victims downloaded just 1 fake content. However, it is worth to note the presence of users who downloaded multiple fake torrents during the measurement period.

The obtained results suggest that users from some specific countries (those having a ratio less than 1) are more skilled to identify fake content so being more protected against possible infections and/or scam episodes.

VII. TORRENTGUARD

In the previous Sections we have demonstrated that a large number of fake content (35%) is currently being published in the BitTorrent ecosystem, and what is worse, most of these fake contents are potentially harmful for those users that download them. We have also seen that the techniques used to remove these contents are inefficient and requires human intervention to: first, detect and report the falseness of a given content, and second, remove it from the BitTorrent portals (this is done by the portal administrator). Furthermore, the scope of the user reports is limited to a single BitTorrent Portal, thus

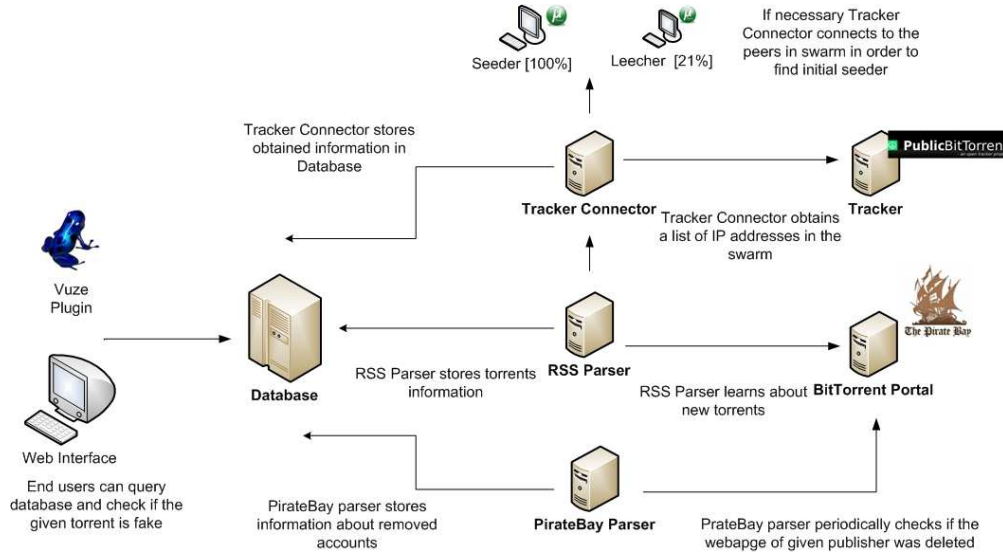


Fig. 5. The schema of TorrentGuard

the content is removed exclusively from this portal instead of the whole BitTorrent ecosystem.

In this Section we present our tool, named TorrentGuard, that aims to automatise and accelerate the process of detecting fake publishers. For this purpose, TorrentGuard identifies a fake publisher by its IP address instead of its username as it is done by BitTorrent portals such The Pirate Bay nowadays. By doing so, a fake content can be identified just after its birth since we can identify that the IP address of the initial seeder belongs to a fake publisher. This allows to accelerate the detection process.

Furthermore, contrary to current techniques used by BitTorrent portals, TorrentGuard removes the fake content from the whole BitTorrent ecosystem because it reports the content infohash. Since the infohash uniquely identifies a content in the BitTorrent ecosystem, a user of TorrentGuard can identify the content as fake independently of the portal from which the content was retrieved (or even if it comes from the BitTorrent DHT service).

In the rest of the Section we present the details of the TorrentGuard implementation as well as the performance results obtained over a testing period of 14 days.

A. TorrentGuard Implementation

Figure 5 depicts a complete schema of TorrentGuard. It is composed by the following modules:

- **RSS Parser:** this module continuously monitors the RSS feed of The Pirate Bay portal. For each new published torrent the RSS Parser gathers the publisher's username and the content infohash. Furthermore, the RSS Parser sends requests to the Tracker Connector.
- **Tracker Connector:** this module is responsible for connecting to the tracker for every torrent obtained by the RSS Parser. The main objective of the Tracker Connector is to obtain the IP address of the initial seeder. In those swarms where the list of IP addresses returned by the

tracker contains more peers than just one seeder, this module connects to all the peers and retrieves their bitfield in order to identify which one is the initial seeder. If the IP address of the initial seeder matches with one of those included in the blacklist of fake IP addresses, this torrent is marked as fake.

- **The PirateBay Parser:** this module periodically connects to the Pirate Bay webpage associated to the different discovered publishers. Eventually, when a publisher's webpage (i.e. account) is removed from The Pirate Bay, the Pirate Bay Parser marks this username as fake.
- **Database:** It stores all the relevant information for the detection and evaluation of TorrentGuard. For each inspected torrent it stores detailed information such as the publisher's username and the initial seeder IP address (in case this is possible to obtain). More importantly, it includes two blacklists. The first one contains the infohashes of all the discovered fake torrents whereas the second one includes the IP addresses of fake publisher's found so far.
- **Website Interface and Vuze plugin:** The TorrentGuard functionality is publicly available throughout two different interfaces: a website¹ and a Vuze plugging. These interfaces provide access to the blacklist of fake torrents allowing a user to verify if a torrent file is associated to a fake content before starting the download process.

Next, we describe the functionality of TorrentGuard. It uses The Pirate Bay portal in order to identify new fake publishers and the IP addresses from where they operate. Towards this end, the RSS Parser continuously monitors the RSS feed of The Pirate Bay portal to learn about new torrents and identify for each torrent the publisher's username. Furthermore, it sends a query to the Tracker Connector that retrieves the IP address of the initial seeder (if it is possible). Both, the

¹This application is available at http://torrentguard.netcom.it.uc3m.es/Fake_torrent/

publisher's username and IP address (i.e. IP address of the initial seeder) are stored in database. In parallel, the Pirate Bay Parser periodically connects to the webpage of the different discovered publishers within The Pirate Bay. If we find that a publisher's account is removed, this user and all its torrents are marked as fake. In addition, we annotate this publisher's IP address as *potential fake IP address*. If three different accounts associated to a given publisher's IP address are removed from The Pirate Bay, we consider that IP as a *fake IP address*. From this moment on, any content published from that IP address is identified just after its birth and reported as fake. Therefore, in the worst case, i.e. for new fake publishers, TorrentGuard employs the same time as The Pirate Bay to identify fake contents. However, once the fake publisher's IP address has been identified, TorrentGuard is able to report fake content immediately after its publication what provides a significant improvement compared to standard detection mechanisms. Moreover, with TorrentGuard it is not necessary to manually report each fake content. Besides, three reports can be enough to mark the malicious user as a fake and in consequence all its future torrents will be automatically classified as fake.

Furthermore, the current existing solutions are limited to the portal where they operate. For instance, in the case of The Pirate Bay, once a content is identified as fake it is removed from the portal but not from the BitTorrent Ecosystem. Rather, TorrentGuard is a cross-portal solution, that is able to identify the infohash of the fake content preventing its download independently of the source from where the user obtained the .torrent file: any BitTorrent portal or the DHT service.

In short, TorrentGuard is a novel tool that: (i) reduces fake content detection time since it uses IP-based detection instead of username-based detection and (ii) allows to identify a fake content in the whole BitTorrent ecosystem rather than in a single portal because it identifies the fake content using the infohash (an unique identifier of the content in the whole BitTorrent ecosystem).

B. TorrentGuard Performance

We have evaluated the performance of TorrentGuard and compared it with the fake content detection mechanism used by The Pirate Bay during a testing period of 14 days. First, we count how many fake content published in The Pirate Bay are identified by the TorrentGuard just after its birth. Furthermore, we measure how long The Pirate Bay takes to identify these fake content. The obtained results show that TorrentGuard is able to early detect around 50% of the fake content uploaded to The Pirate Bay. Moreover, Figure 6 represents the CDF of the time difference between the detection instant of TorrentGuard and The Pirate Bay for these content. We observe, that TorrentGuard reduces the detection time 60 minutes in median. However, this reduction is higher than 2 hours for 20% of the fake contents, and for some cases it goes up to several days.

Although previous results already demonstrate the significant improvement provided by our tool compared to the state of the art solution, the final objective of TorrentGuard is

reducing the number of download events associated with fake content, thus preventing BitTorrent users facing malware and scam. Then, if TorrentGuard was widely used, it would have prevented almost 390K fake content downloads just during the 14 days of the evaluation period compared to The Pirate Bay. However, as described before the solution used by The Pirate Bay is specific for this portal whereas our solution works for the whole BitTorrent Ecosystem. Therefore, if we consider all the download events associated to fake content during the 14 days evaluation period, TorrentGuard could have avoided up to 1.35 millions fake content downloads. By extending this value to a complete year, we can state that TorrentGuard would be able to eliminate more than 10 millions fake content downloads per year compared to The Pirate Bay solution and more than 35 millions if we consider the whole BitTorrent ecosystem. This means, depending on the success of the fake publishers strategies, *preventing up to hundreds of thousands of malware infections and scam incidents per year*.

Hence, our initial evaluation show very promising results to incentive the BitTorrent community to use the TorrentGuard.

C. TorrentGuard Efficiency

A detection system is typically characterised by the number of false negative and false positive occurrences. On the one hand, the former is represented by those fake torrents escaping our detection tool. On the other hand, false positives refer to those content classified as fake, which actually are non fake ones. Unfortunately, it is not feasible to properly measure such parameter since it would require to manually inspect a huge amount (thousands) of contents classified as legacy (i.e. non fake) ones. Instead, we have performed an affordable evaluation by downloading few dozens of torrents classified as legacy by TorrentGuard and we did not find any fake torrent among them. We can state, however, that our tool discovers all fake contents which are also detected by The Pirate Bay.

In order to evaluate the false positives rate of TorrentGuard, we focus on those Pirate Bay usernames whose account has not been deleted from The Pirate Bay but their content has been classified by TorrentGuard as fake. The first intuition is that TorrentGuard may be mistaken for some of these usernames. We have downloaded content from each of referred Pirate Bay accounts and we did not find any non-fake content among them, thus these content belong to fake publishers that have still not been detected by The Pirate Bay. Hence, the performed evaluation suggests that TorrentGuard suffer from a low rate of both false positive and false negatives.

VIII. RELATED WORK

A. BitTorrent Measurement

Several authors have used real data collection in order to understand different aspects of BitTorrent [6], [8], [9]. Different methods of measuring the BitTorrent are described in [13]. However, only few works have looked at the content publishers [2], [26]. The most extensive study of characterisation of BitTorrent ecosystem is presented in [26]. This work includes discussion about BitTorrent publishers, defined by its

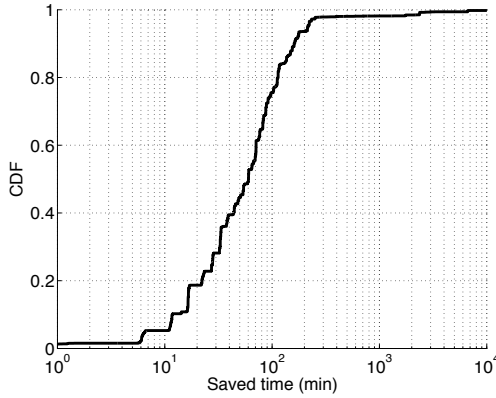


Fig. 6. CDF of the saved time in fake content detection when using TorrentGuard in front of The Pirate Bay

username. We demonstrate in this paper that fake publishers cannot be identified by its username, instead they are identified by its IP address. The presence of the fake publishers was firstly mentioned in our previous work [5]. Based on our initial observation, in this paper we perform a thorough analysis of fake publishers and their published content revealing their target, incentives and strategies and propose a novel solution to prevent users from downloading fake content.

B. Fake content

There are several studies presenting the possible threats in the Internet. In [28] authors state that 40% of all computers are infected by botnets and can be controlled by attackers. Another study [18] reports high presence of malware and spyware content in the Internet. Few previous works have studied the malware propagation through P2P systems [12], [23], [27]. Specifically, Kalafut et al. [12] analyse LimeWire whereas Shin et al. [23] analysed KaZaa. These authors look at the problem from the content perspective instead of the fake publisher perspective used in this paper. This avoids that they discover more sophisticated strategies as those reported in our study in which the content is not the malware itself but includes a link to the malware. Similar content-based approach is applied in FakeDetector program [7] that looks for fake hashes in DirectConnect hubs and reports found fake content to users and hub moderators. Finally, the authors of [12] propose to filter those content with a specific size since most of the malware content has specifically this size. Unfortunately, this solution is not valid for BitTorrent. Instead, we propose a more sophisticated solution (TorrentGuard) that provides early detection of fake content.

IX. CONCLUSIONS

This paper presents the first comprehensive study about fake content in the BitTorrent ecosystem. For this purpose we use real data collected during a large-scale measurement study. The obtained results demonstrate that 35% of all the content is fake. Moreover, just a few tens of users are responsible for most of the published fake content. Furthermore, more than 99% of the fake torrents are associated with either malware

or scam websites. This represents a serious threat for the BitTorrent ecosystem that must be eliminated or at least mitigated. Towards this end, we have implemented TorrentGuard, a novel tool for early detection of fake content. Based on our initial evaluation the widely usage of this tool may prevent the download of millions of fake content every year, thus contributing to reduce the number of computer infections and scam episodes faced by BitTorrent users.

REFERENCES

- [1] Alexa. <http://www.alexa.com/topsites/>.
- [2] S. Le Blond, A. Legout, F. Lefessant, W. Dabbous, and M. Ali Kaafar. Spying the world from your laptop. *LEET'10*, 2010.
- [3] CAPTCHA. <http://www.captcha.net/>.
- [4] D. R. Choffnes and F. E. Bustamante. Taming the torrent: a practical approach to reducing cross-isp traffic in peer-to-peer systems. *ACM SIGCOMM 2008*.
- [5] R. Cuevas, M. Kryczka, A. Cuevas, S. Kaune, C. Guerrero, and R. Rejaie. Is content publishing in bittorrent altruistic or profit-driven? *ACM CONEXT 2010, Philadelphia, USA*.
- [6] R. Cuevas, N. Laoutaris, X. Yang, G. Siganos, and P. Rodriguez. Deep diving into bittorrent locality. *IEEE INFOCOM 2011, Shanghai, China*.
- [7] FakeDetector. <http://sourceforge.net/projects/fakedetector/>.
- [8] L. Guo, S. Chen, Z. Xiao, E. Tan, X. Ding, and X. Zhang. Measurements, analysis, and modeling of bittorrent-like systems. In *ACM IMC'05*.
- [9] T. Isdal, M. Piatek, Krishnamurthy. A., and Anderson T. Leveraging bittorrent for end host measurements. In *PAM*, 2007.
- [10] IsoHunt. <http://www.isohunt.com>.
- [11] R. Izhak-Ratzin, H. Park, and M. van der Schaar. Reinforcement learning in bittorrent systems. In *In Proc. of INFOCOM 2011*.
- [12] A. Kalafut, A. Acharya, and M. Gupta. A study of malware in peer-to-peer networks. *6th ACM SIGCOMM conference on Internet measurement, IMC 2006*.
- [13] M. Kryczka, R. Cuevas, A. Cuevas, C. Guerrero, and A. Azcorra. Measuring bittorrent ecosystem: Techniques, tips and tricks. *IEEE Communications Magazine (accepted to publication)*.
- [14] N. Laoutaris, D. Carra, and P. Michardi. Uplink allocation beyond choke/unchoke or how to divide and conquer best. In *In Proc. of the CoNEXT 2008*.
- [15] N. Liogkas, R. Nelson, E. Kohler, and L. Zhang. Exploiting bittorrent for fun (but not profit). In *In IPTPS 2006*.
- [16] T. Locher, P. Moor, S. Schmid, and R. Wattenhofer. Free riding in bittorrent is cheap. In *In HotNets 2006*.
- [17] MaxMind. <http://www.maxmind.com/>.
- [18] A. Moshchuk, T. Bragin, S. Gribble, and H. Levy. A crawler-based study of spyware on the web. *Internet Society Network and Distributed System Security Symposium (NDSS)*, 2006.
- [19] M. Piatek, T. Isdal, T. Anderson, A. Krishnamurthy, and A. Venkataramani. Do incentives build robustness in bittorrent? In *4th USENIX Symposium NSDI 2007*.
- [20] TOR Anonymity Online Project. <https://www.torproject.org/>.
- [21] Sandvine. Fall 2010 Global Internet Phenomena Report. Available at: http://www.sandvine.com/news/global_broadband_trends.asp.
- [22] Alex Sherman, Jason Nieh, and Clifford Stein. Fairtorrent: Bringing fairness to peer-to-peer systems. In *In Proc. of the ACM CoNEXT 2009*.
- [23] S. Shin, J. Jung, and H. Balakrishnan. Malware prevalence in the kazaa file-sharing network. *6th ACM SIGCOMM conference on Internet measurement, IMC 2006*, 2006.
- [24] M. Sirivianos, J. H. Park, R. Chen, and X. Yang. Free-riding in bittorrent networks with the large view exploit. In *Intl. Workshop on Peer-to-peer Systems (IPTPS) 2007*.
- [25] H. Xie, Y. R. Yang, A. Krishnamurthy, Y. Liu, and A. Silberschatz. P4p: Provider portal for applications. *ACM SIGCOMM 2008*.
- [26] C. Zhang, P. Dhungel, D. Wu, and K.W. Ross. Unraveling the bittorrent ecosystem. *IEEE Transactions on Parallel and Distributed Systems*.
- [27] L. Zhou, L. Zhang, F. McSherry, N. Immorlica, M. Costa, and S. Chien. A first look at peer-to-peer worms: Threats and defenses. In *Proceedings of the IPTPS*, Feb. 2005.
- [28] Zhaosheng Zhu, Guohan Lu, Yan Chen, Z.J. Fu, P. Roberts, and Keesook Han. Computer software and applications. *COMPSAC '08*.