Phishing With A Darknet: Imitation of Onion Services

Frederick Barr-Smith
Department of Computer Science
University of Oxford
Oxford, England
freddie.barr-smith@cs.ox.ac.uk

Joss Wright
Oxford Internet Institute
University of Oxford
Alan Turing Institute
Oxford, England
joss.wright@oii.ox.ac.uk

***Abstract*—In this work we analyse the use of malicious mimicry and cloning of darknet marketplaces and other 'onion services' as means for phishing, akin to traditional 'typosquatting' on the web. This phenomenon occurs due to the complex trust relationships in Tor's onion services, and particularly the complex webs of trust enabled by darknet markets and similar services. To do so, we built a modular scraper tool to identify networks of maliciously cloned darknet marketplaces; in addition to other characteristics of onion services, in aggregate. The networks of phishing sites identified by this scraper are then subject to clustering and analysis to identify the method of phishing and the networks of ownership across these sites. We present a novel discovery mechanism for sites, means for clustering and analysis of onion service phishing and clone sites, and an analysis of their spectrum of sophistication.**

I. Introduction

The *darknet* [33], [49] is commonly used to refer to a set of websites hosted in anonymous and untraceable overlay networks, the most common of which is Tor and, specifically, its *onion services*. Amongst many legitimate and illicit uses of darknet websites, the rise of marketplaces for restricted drugs [38] and other illegal products are perhaps the most well-known.

Darknet markets are sites that utilise the capabilities of onion services in order to host and facilitate illegal commerce. These take advantage of the Tor network's capability of providing anonymity and onion services' provision of relative anonymity for the operator of an onion service [60]. These markets are ephemeral and often rotate domains to maintain availability and anonymity, particularly following heavy DDoS attacks using the Stinger-Tor[1] tool during the period under analysis [11]. The currency used to undertake transactions on these marketplaces are typically cryptocurrencies [2], [24] such as Bitcoin, Dash, Monero and Litecoin.

Darknet markets trading in illicit goods [18] result in large volumes of hard-to-trace cryptocurrencies being transferred amongst mutually untrusting users, and is thus an attractive target for a variety of scams, theft, and fraudulent behaviour. Phishing taking place on such sites is unlikely to be reported to law enforcement, although this is not to say that anonymous reports on forums do not occur on occasion. Winter et al. [67], theorised about the potential existence of phishing onion domains based on data from their experiments.

Phishing on the clearnet is a well established threat [15], [29]. A particularly effective variant being that of imitating sites [52] to obtain credentials or financial details from victims. We discovered this to be the dominant technique utilised in onion services hosting phishing sites.

One complexity in assessing potential phishing in onion domains is that we must identify whether purposeful 'typosquatting' of domains is occurring in order to target people in an equivalent manner to similar behaviour on the normal web. For onion services, there is a key difference in these behaviours - the objective is to deceive potential victims into clicking, visiting, and potentially using a site; rather than capitalising on a victim's mistyping of a given domain. The use of precise and sophisticated mimicry was identified through the similarity of domains that were part of phishing campaigns to the domains that they were imitating.

Crucially, onion domains are not created manually and registered via a DNS system as in the normal web. Instead, onion domains [43] are generated from a public key, resulting in a string of characters that are hard for humans to validate [64]. The complexity of onion domain names, and lack of fully human-memorable addressing schema, enables easier mimicry of marketplaces and cloning of other sites for phishing and malicious purposes. Our research discovered that these phishing sites were prolific throughout the darknet and that many sites were imitations or clones.

As an illustration of this complexity even for experienced users, during "Operation Onymous", in which the US Federal Bureau of Investigation investigated and shut down

[1] https://github.com/whitepacket/Stinger-Tor via archive.org

several darknet markets. The FBI reportedly experienced difficulty discerning the genuine from the imitation with regards to shutting down darknet markets. The resulting shutdowns of various phishing-focused clones of major darknet forums, included possible mistakenly-identified sites [14]. Some of the key findings from this research were:

1) We established that over half the sites on the darknet are duplicates, many of them benign.
2) We discovered that a subset of onion domain names are imitated by phishing groups and that phishing is a widespread phenomenon on the darknet.
3) We clustered phishing campaigns on the darknet into individual campaigns, grouped by HTTP headers.
4) Our analysis showed that the vast majority of onion sites are used for criminal purposes.

The structure of this paper begins with describing the background and related work. This is followed by a discussion of the complex ethics involved. The main body of the paper then proceeds, which is a description of methods of analysis to identify, cluster and quantify phishing on the darknet; and the results of this analysis. A further analysis of the aggregated contents follows. The paper concludes with an assessment of limitations, areas for future work and a conclusion.

II. Background

Despite the existence of other forms of darknet, in this work we focused on hidden, or "onion" services hosted in the Tor network, and only accessible via users that have connected to that network [16]. Due to the anonymity and untraceability provided by onion services these attract, amongst others, those who have a vested interest in operating anonymously for illegal purposes. Long-term centralised directories of sites and chains of trust are largely unavailable for service authentication in the darknet; due to the fragmented nature of onion services and lack of reliable public key infrastructure. This situation has partially begun to resolve with the development of nascent search engines and listing sites such as Grams [25] and DeepDotWeb. Such services are themselves still partial, unreliable and prone to shutdown by law enforcement. The ephemerality of onion services, indexes and marketplaces [28] contributes to this fractured trust environment, creating fertile ground for malicious onion service operators to defraud users.

Onion services, considered in aggregate, experience a significant level of churn and flux of domains. As such, attempts to crawl and index services are far less effective than the more stable clear web; representing a snapshot of a moment in time rather than a stable directory.

A. Timing of Research

The period during which we conducted this research, was from May to July 2019. This was during and immediately following the public shutdown and arrest of the administrators of the Wall Street Market and Valhalla [23] and the murky closure of Dream Market in the wake of several vendor arrests [66]. During this period there were also reported DDoS attacks of unknown origin, causing downtime. This strongly incentivised administrators of illegal sites to close their operation before they were arrested, from compromised informant vendors and operators or through deanonymisation.

DeepDotWeb, a major centralised link repository and Tor news site, was also shut down by law enforcement during this period. As a result of this disruptive activity, there was significant potential for malicious operators to engage in phishing activity to capitalise on this uncertainty, and as such our results also represent a particular snapshot in time that may not fully generalise to future behaviours.

B. Contents of Darknet

Despite a range of stated and demonstrated legitimate purposes for onion services; we discovered that the most common uses for websites on Tor onion services was criminal activity [41]. Approaches that damage the trust between and within sites and their users with respect to the anonymity and trustworthiness of sites leaves fertile ground for fraudulent activity to occur; as demonstrated in a recent law enforcement operation [62].

It is hopefully uncontroversial that the use of privacy enhancing technologies, cryptocurrencies, Tor, and related tools are not, and should not be, criminal acts in themselves. The uses of these tools to avoid censorship and to engage in communication and access to information against the wishes of regulators with whom users may legitimately disagree, are sufficient justification for their development and adoption.

Further, despite the illicit activities occurring in darknet marketplaces, users of these technologies are entitled to be free from fraud. The lack of trust and inability to have confidence in transactions in the darknet has implications for legitimate users there regardless of the specific effects on those committing criminal acts: if security mechanisms do not work for drug dealers, they also do not work for whistleblowers, independent journalists, and others.

III. Related Work

Phishing on the clearnet is well studied in the academic literature. Of particular interest is Clayton and Moore's early work [42], which identified a 'rock-phish' campaign

based on an element of its URL, similar to our identification and clustering of distinct campaigns. Recent measurement papers by Oest et al. provide an example of the increase in the scale [47], sophistication and profitability of phishing; in addition to an exploration of the effectiveness of countermeasures [45], [46].

Phishing in onion domains was mentioned briefly by [7]. Research related to the initial Onymous campaign that indicated that a large number of the over 410 or 600 domains [20] in the Operation Onymous takedown conducted by the FBI and the Department of Justice were discovered to be clones of real markets [14], [53] .

A recent paper by Yoon et al. [68] analysed data from 2017 related to darknet clones, focusing on the now-defunct AlphaBay and Empire Markets. The work presented here focuses on a different and later time period, 2019 as opposed to early 2017. More significantly, Yoon et al.'s analysis relies on a relatively simple extraction of page titles to identify similarity between pages, as opposed to the more complex and detailed header-based analysis presented here. Beyond this, in this work we extend the analysis of darknet clones significantly by identifying and clustering phishing domains apparently operated by the particular groups across multiple sites and targets, and identify both their complexity and mode of operation.

A. Historical Data

Phishing as a general technique has been widely exploited by a number of criminals. Due to the fleeting nature of these criminal operations and the lack of archiving of onion sites, historical evidence of the practice is somewhat scarce. Despite this, the Web Archive at *archive.org* does index a number of onion sites that have an equivalent clearnet presence [10].

A 2015 post on the `tor-talk` mailing list linked to incidents in which cloned sites for the onion service search engine *ahmia.fi* were discovered, as well as for the onion service iteration of the search engine `DuckDuckGo`. As `ahmia`, `DuckDuckGo` and the other imitated sites did not process Bitcoin or other cryptocurrencies the motives for this remain unclear.

From the darknet operator community itself, a 2017 interview with the onion service listing site operator 'Abrupt-dismissal' of 'FreshOnions' [8] stated that he detected a number of darknet market clones using these phishing methods, specifically 350 clones of AlphaBay and over 4000 in general.

B. Previous Arrests and Campaigns

The ephemerality of markets and domains in the darknet mean that several of the domains studied here and elsewhere, both legitimate and imitated, are no longer active. Despite this, the phenomenon itself is of long standing.

In 2017, a man was sentenced for performing this method of fraud upon customers of darknet markets. Within this court deposition there is evidence of the method of this attack technique. The discovered attack technique was *"fake links on forums that when clicked would port forward the users through RICHO's computer server to the actual marketplace site*" [61]. This correlates with the described method of exploitation identified by Van Riper of using a *"fully self written proxy software*" [54]; which would intercept and steal credentials, to facilitate the theft of cryptocurrencies.

C. Typosquatting Onion Domains

As noted above, the random alphanumeric string assigned as an onion address, resulting from the generation of the public/private keypair is not generally usable or memorable. Given this, users' capacity to recognise the domain name string relevant to the market they are trying to reach is extremely limited. As such, potential victims may need to visit link repositories or forums in order to find the address; exposing them to be targeted by phishers seeding fake links.

While onion URLs are problematic for users in general; there are tools such as `Shallot` [2], that allow a degree of control over the onion URL strings. In particular, it is technically feasible to specify a desired string prefix to be produced. Short prefixes of up to eight characters can feasibly be generated on consumer hardware within a number of days or weeks, whilst shorter prefixes may require only hours or seconds. As such, replicating the initial few characters or an otherwise random string, combined with replication of the user experience of a site, is a potent means to deceive unwary users.

Comparing typosquatting domains against genuine domains is typically achieved via calculation of the Levenshtein distance [36], or one of its variants. Typosquatted domains can be identified by measurement of the *edit distance* between the intended URL and the typosquatted imitation [1], [57].

Whilst there are key differences between onion domains and traditional chosen web domains, for vanity onion domains the same technique can be applied to discern the distance between the intended destination and the cloned site. The default Levenshtein distance within typosquatted domains is likely to be higher for those domains in the Tor network due to the peculiar nature of Tor domains.

We identified that onion sites typically only attempt to mimic the first few characters of a given onion service

[2] https://github.com/katmagic/Shallot

domain. We provide evidence of some of these mimicry attempts in Section VI-F.

IV. Ethics

In a general sense, this research does not aim directly to assist criminals in conducting their activity. We instead study activity that is already taking place and illuminate means by which users' security and privacy is being violated against their expectations.

This research was subject to prior ethical review by our host institution[3]. We adhered to principles as laid out by the Menlo report on Ethical Principles for Information and Communication Technology Research [17], prioritising beneficence, avoiding harms to the subjects of study, and maximising societal benefit.

To avoid accidental access to particular significant classes of illegal imagery, most notably child abuse imagery, only the textual content and headers of onion service sites were stored and processed in this research. [51] propose a robust pipeline for research involving child abuse imagery with relative safety, but specific analysis of such content was outside of the scope of this research and as such those guidelines were used here for general advice on handling sensitive content. Redaction of any imagery at the ingestion stage also prevented any inadvertent exposure of researchers to illegal or traumatic images.

Respect for individuals whose data may be collected, and the recognition of users of darknet forums as having both an expectation of privacy and, potentially, as vulnerable users, was core to the ethical conduct of this research. From a legal perspective, Recital 26 of the European General Data Protection Regulation [13] states that the GDPR does not apply to anonymised information which, due to the nature of the technologies involved, covers the vast majority of the content accessed during this research [32]. Despite this, to avoid any possible violations of privacy or data protection concerns, our analysis and reporting procedures focused on metadata and aggregate data. Further, the ethical approval for this research explicitly prohibited the dataset from external release due to the sensitivity of the topic of analysis. This concurs with published research on ethics of cryptomarket research [39], where *partially obfuscated data* has been suggested as a potential solution for sharing darknet market data.

V. Methods

A. Onion Service Spider

Core to this work was our creation of a onion service spider. This tool collected data including the textual content of the site, the site's URL, the title, and the HTTP headers for each page. We conducted this by visiting the HTTP service of each server using the Python requests library and then processing the returned response.

The spider we developed in this work uses a regular expression taken from the alleged `xkeyscore` source code leaked by Edward Snowden; which included regular expressions to parse .onion addresses [26]. The crawler, given an onion service website, works simply by parsing the query response to identify the domains on the page. We then added these domains to the database of onion sites. The onion service spider would then recursively visit this list of onion sites, adding additional domains that were discovered.

1) Seed Data: The seed data for this was 'The Hidden Wiki' and various online repositories of .onion links that were found. This replicates the information source for many potential victims of this attack. We can assume that the domains within our dataset are as complete a representation of Tor hidden services as is possible, as our 11,533 sites are similar in number to 7,257 that Sanchez-Rola et al. [55] discovered in their research or 13,326 by Yoon et al. [68]. Due to the fractured trust environment and complex domain names; visitors to darknet markets may visit forums, wikis or link repositories that could be poisoned by phishing site operators.

Given a list of onion services scraped data, we manually combine a number of trusted sources; such as the addresses listed on *https://dark.fail* to ascertain which were genuine sites. We then used these sites as the baseline from which to compare mimicked phishing sites. We identified the malicious imitations of a site by virtue of the fact that they were not included within the list of known reliable Tor sites, but masqueraded as them in terms of appearance.

B. Site Comparison

We used a number of methods to cluster and compare phishing sites, as well as to approximate of the scale of the phishing sites. Not all phishing sites in use on onion services could be identified by crawling alone; leaving a small amount of these sites not included in our sites.

In analysing the data retrieved by the spider, we aimed to identify several characteristics. These were the number and type of phishing domains, the nature of the sites and the techniques that they use to defraud users, including any information that determined links between multiple fraudulent sites. To determine the nature of sites, and the links between them in terms of shared operators, we relied on a number of techniques. Our use of these techniques identified shared resources, underlying server software and shared phishing techniques across phishing imitations of multiple markets.

[3]Specific ethical review information and approval code elided for anonymous peer review.

1) HTTP Headers: The key method to identify cloned websites in this work was by comparison of HTTP headers between sites. In general, whilst particular care was taken by malicious site operators to replicate the *appearance* of sites; our experiments showed that operators were far less diligent in replicating the header information that accompanies the web request. As such, many onion services inadvertently revealed information that they did not intend to via HTTP headers. Specifically, this information is that when a number of sites in the same or disparate campaigns have identical HTTP headers, it is possible to cluster them based on these headers as the same actor. We performed this clustering by querying the database in which these headers were locally stored, to identify exact matches of header elements between sites. This allowed us to identify individual campaigns. This source of information leakage can be used to cluster servers and sites that are part of a given phishing campaign, even when the cloned sites themselves differ.

An example of this is the following HTTP headers for the legitimate Empire Market, that was hosted at *http://mttzugnjjcmwe6pp.onion*:

```
[...]
`Date': `Wed, 12 Jun 2019 12:58:06 GMT',
`Server': `Apache/2.4.7 (CentOS)' ,
`Set-Cookie':
↪  `shop=b9l36ilbg25ag5dv9g5bv74ap8ioq8dq;
Expires=Wed, 12-Jun-2019 14:58:06 GMT;
↪  Max-Age=7200 path=/;  HttpOnly',

[...]
```

This can be programatically or manually compared to the HTTP headers for an Empire Market phishing site at *http://hjr4un7jexg-5o5r6.onion*:

```
[...]
`Date': `Wed, 12 Jun 2019 23:30:08 GMT',
`Server': `Apache',
`Set-Cookie':
↪  `shop=r8827fj468jfn3p0320vs8pp8u1lslp7;
Max-Age= 7200; path=/; HttpOnly',
`Expires': `Thu, 19 Nov 1981 08:52:00 GMT',
[...]
```

In the headers for this imitation site we observed that the phishing imitation had a cookie expiry date of 1981, as opposed to the legitimate market, whose cookie expired within twenty-four hours of data collection. Alternatively, whilst both servers were *Apache*, the legitimate version specified the version number of '2.4.7' and specified that it was running on a CentOS system. Other minor differences also existed in the headers, despite attempts to replicate elements of the legitimate site such as the 'shop' cookie.

There were subtle differences between the HTTP headers of these sites, highlighted in bold, that provided us a reliable means of differentiating services. Such as some headers being formatted differently or having different cookies, fields and default values. There was a syntactical and structural similarity between the HTTP headers of sites that were hosted by the same actor, or as virtual hosts on the same server. This is how we identified and clustered individual operators. We did not discover specific server or header types that were found to be symptomatic of phishing, indicating a diversity of reverse proxies in use.

2) HTML Comparison: We used the `html-similarity`[4] library to quantify the similarity between an original site and any apparent clones. This tool compares the respective HTML content of pages to determine a value from 0 to 1 assessing how similar 2 pages are; however it was not effective in differentiating many of the more sophisticated phishing sites. Most notably, as will be discussed, those using *reverse proxies* that made minor modifications to Bitcoin addresses. Despite this, the tool was largely effective at extraction and discovery of a range of lower tier, less sophisticated, phishing efforts.

Additionally, we manually determined the difference between legitimate versions of sites and their illegitimate clones, through analysing the differences between the respective pages. This was from a subset of pages that had an identical page title to the legitimate site. For instance the following two snippets of code may generate and render an image that looks precisely the same; however it was clear that two different methods and site architectures were used to generate these images. These were from a legitimate mirror of the Empire market and a phishing link, respectively.

```
<img src='images/1558002190.8484.jpg' class='ui
medium image captcha' width='250' height='70'
/>
```

```
 <img src='http://zjtixcqfxugkqu77.onion
/public/captcha img/1559670801.875.jpg'
style='width: 250; height: 70; border: 0;' alt='
' />
```

There were also discrepancies between different mirrors of legitimate sites; however, these differences were less pronounced than those between legitimate sites and phishing sites.

VI. Results

A. Scale of Imitation

The results of our analysis, visible within Table I show that a significant number of the hidden services discovered are

[4]https://github.com/matiskay/html-similarity

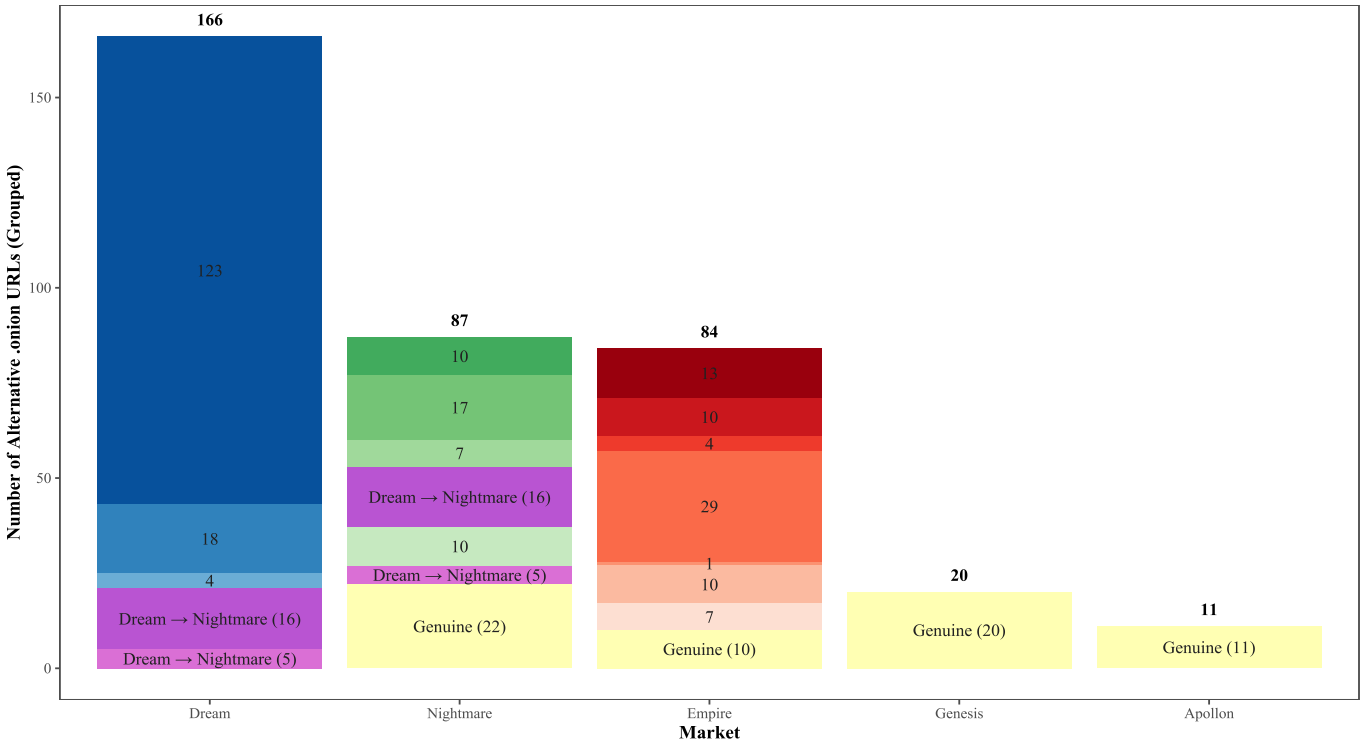

Figure 1: Darknet Markets and Imitations. Genuine Markets in cream. Variations in colour, with each shade as a different cluster of phishing sites, grouped by headers.

Site Category	Amount of Sites	Percentage
Unique	3,162	27.42
Imitation	5,922	51.35
Default	2,449	21.67

Table I: Sites By Category

imitations, a subset of 3 separate types *default*, *imitation* and *unique* domains. The method by which sites are automatically ascertained to be *unique* is if no other site has the same content and title. Other sites, for which there is an identical page title are clustered as imitations . Default sites are those with default page titles such as "*Index of /*" "*Apache2 Ubuntu Default Page*" .

Out of 11,533 onion domains we analysed in this work, 5,922 were *imitations* - sites discovered to be duplicates or phishing clones of others. This number excludes those *default* domains with titles such as '404 Not Found', 'Apache2 Ubuntu Default Page: It works','Dir' and 'Index of /'. Whilst these were not active phishing attempts, if these default pages were included, cloned sites would have comprised over 70% of onion services. In total, these 5,922 sites were each clones from an original set of 710 distinct sites. There were also *unique* domains, for which there is only 1 instance of the domain.

Site Type	Number
100x Your Coins in 24 Hours	319
Thank You Guys	182
Dream Market Login	148
The Open Road - Marketplace	81
Clone CC : No.1 Trusted onion site	77
Nightmare Market	51
Empire Market	48
Alabama Market	46
Grow Your Bitcoin	42
Agartha: Underground Anonymous Darknet Market	30

Table II: Most Cloned Sites - Title and Frequency

1) Taxonomy of Imitation Types: The first distinction made between different onion services was that between phishing sites, clones and mirrors. These sites are inserted into wiki sites, malicious directory mirrors or search engines by malicious actors. This partially explains the presence of dozens of hidden wikis and their respective different links, leading to different subsets of sites.

2) Profitability: On a subset of these sites there were Bitcoin addresses that were discovered. These addresses were noted and run through sites that facilitate basic cryptocurrency tracing, such as `Blockchain Explorer`, to identify

any activity on the wallet addresses referenced. We could observe profit from looking at the addresses involved. In some cases up to 1.2 BTC ($8,750 on 18/07/2020) in singular instances for the operators of the criminal enterprise. This was due to the addresses listed on some phishing sites showing inflows of Bitcoin.

We also discovered inflows of cryptocurrency to the respective associated addresses for some of the less sophisticated scams. These less sophisticated scams were observed to use the limited anonymity provision mechanism of Bitcoin address rotation; however, it was still possible to see that there was profit achieved.

B. Sophistication of Imitation

There was a spectrum of sophistication seen amongst cloned sites.

1) Official Mirrors: Mirrors are distinct from phishing sites. Although they are imitations of the original service, they are set up by the site in order to add resilience to an onion service and are also typically set up by the operators of that service. Authorised and legitimate mirrors exist for users to access sites that may want to remain hidden or more resilient to distributed denial of service attacks. One example of this was Dark0de, a prolific cybercrime forum, following its shutdown by Europol [12], [21].

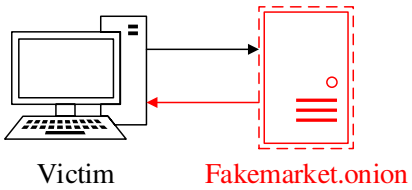

Figure 2: Unsophisticated Imitation

2) Unsophisticated: Many sites, such as many clones of the Wall Street Market at time of writing; made little attempt to mimic the original site faithfully and instead simply ask for a Bitcoin 'registration fee' using similar interface design and fonts. Another example of a particularly unsophisticated phishing attempt was the 'Dream Market Bitcoin Mixer'; which capitalises on the Dream Market brand reputation and attempts to offer the service of Bitcoin laundry to facilitate darknet trade. Figure 2 shows the architecture of unsophisticated clones.

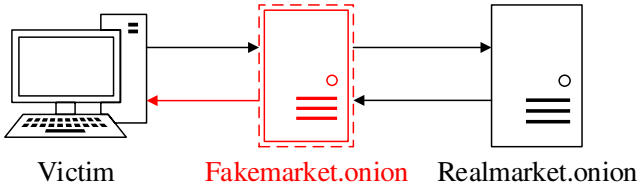

Figure 3: Sophisticated Imitation

3) Sophisticated: At the other end of the spectrum are sites such as the `Onion Cloner` and `Rotten Onions` scam [40]. These sites operated reverse proxies, and similar approaches, to replicate the user experience of the site they are imitating dynamically; whilst capturing full user credentials and withdrawal codes. The operators of these more sophisticated imitations may also substitute cryptocurrency deposit addresses with illegitimate addresses to defraud users. Figure 3 shows the conceptual architecture used by the sophisticated phishing sites we observed, using reverse proxies.

C. Clones By Site

The sites which we found to have the most dedicated pages, as visible in Table II were mostly markets; these likely being clones and phishing attempts. Of specific interest here was the presence of 'Dream Market Login', a site for the defunct Dream Market. This appears to have been an attempt to capitalise on uncertainty in the darknet market ecosystem. In general, some of these pages may have been legacy from previous phishing campaigns that were left in order to entrap people unaware of the market shutdown.

By contrast, the scale of the illicit drugs trade on darknet markets is worth fraudulently imitating, explaining the greater prevalence of phishing sites for darknet markets. As shown by a RAND study from 2016 [35] to be on the order of between 14.2 and 25.0 million dollars monthly, when prescription drugs and tobacco are not included. In the period from 2016 to 2019, the percentage of drug users in the UK utilising darknet markets more than doubled from 12.4% to 28.6% [65].

We also observed a menagerie of smaller sites, motivations, and behaviours. It is also worth mentioning the 182 pages, stating 'Thank you guys', the retirement confession of a hacker who profited by 200 BTC, by their own admission. SecureDrop, which had 108 instances, and Red Room, which is an imitation murder site had 63 instances. The presence of an imitation murder site to capitalise on customers' morbid fascination [25] does not preclude the existence of sites that host similar content. The sites charging Bitcoin to access violent abuse imagery under the guide of this particular 'Red Room' site and its predecessors are deceptive in nature. As mentioned earlier in the paper, this particular scheme has achieved some measure of success; with Bitcoin addresses linked to these sites showing inflows.

D. Clustering of Clones

Clustering based on HTTP headers was possible, as seen in Figure 1. The shared infrastructure that was repurposed from an initial Dream Market phish such as that visible in Figure 4 to a Nightmare Market phish like that visible

in Figure 5 during the course of the research is shown in purple and lilac. For the other colours, they all represent separate clusters of imitation, so individual threat campaigns.

Figure 4: Dream Market Phish

Figure 5: Nightmare Phish

We could detect the same server being reused via HTTP headers and presumably other indicators available via active fingerprinting of the servers. This clustering and separation was achieved via linking the metadata specific to given servers.

One interesting finding from our clustering based on HTTP headers was the detection of the same header signature being used across Dream and Nightmare markets, for 2 clusters of 5 and 16 phishing domains respectively as visible in Figure 1. This demonstrates shared tools, techniques, and procedures across several markets, in combination with the more obvious domain reuse.

1) Reuse of Phishing Domains: Over the course of the study we observed that phishing sites that had previously been used to host Dream Market phishing sites, as cached by *ahmia.fi*, had been repurposed by their operators to host phishing domains for Nightmare market. This provides further evidence of reuse of tools, techniques, and procedures between market imitations. We identified 21 iterations of this repurposing between the Dream and Nightmare markets; demonstrating that not only did these phishing campaigns span months, they also targeted multiple markets. Unsurprisingly, the most popular markets at the time given the churn and high mortality rate of darknet markets. This is due to simple market economics of targeting the biggest potential victim base and to take advantage of uncertainty.

E. Comparison of HTML

As enumerated within the methods section, we used comparison tools for the HTML code contained within these sites extracted by the scraper, to establish similarity. A notable result of this is that the child abuse imagery sites were similar at a structural level. This would appear to have indicated these sites were direct clones or 'Official Mirrors' of the actual site; according to the taxonomy in Section VI-B. This was further re-enforced by the fact that the observed headers for the given sites were precisely the same; indicating that they were operated off the same server infrastructure, or group of servers.

F. Manual Analysis of Domain Imitation

Traditional typosquatting research, such as that in related work makes use of the Levenshtein distance between domain names; however, this is less useful in onion domains due to the way that names are generated. Rather than applying the Levenshtein distance directly, therefore, we analysed the dataset for mimicking of *prefixes*.

We discovered that there were a small number of attempts to mimic prefixes of some domains. Given the churn in onion service domains and poor usability of these domains in general; it appears that phishing service operators have deemed this approach not worth the effort.

One interesting exception to this, was the imitation of prefixes for which the original site had standardised its own domains. As an example, the first 7 characters of domains operated by the Empire Market were themselves *empire*: *http://empiremktxgjovhm.onion*. For phishing domains targeting Empire Market, operators had taken the trouble to generate appropriate vanity domains:*http://empirembpcpuelxd.onion*, *http://empire2uiax76ofj.onion*, and *http://empir7gxe2th2bu6.onion*. This trend was also seen in other markets, such as the recognisable legitimate prefix of *nightmare*: *http://nightmareocykhgs.onion*, imitated by the phishing domains contained at *http://nightmarepwitgei.onion*, *http://nightmareiflmewa.onion* and *http://nightmarefkzlzxk.onion*. This pattern is similar to clearnet phishing, wherein domains are also imitated.

VII. Classification of Site Types

We evaluated the structure of the darknet in aggregate to ascertain the type of material that was in the Tor network. We discovered that these onion services were primarily criminal in nature. However this criminal activity was split into a number of different criminal sub-groups. There was also other non-criminal activity that occurs on onion services.

We used string matching as the method to classify site type, the results of which are visible in Table III. The approach taken was to identify the strings that would be present within certain site types and titles. After reviewing more complicated methods such as Support Vector Machines and other machine learning based approaches we chose to use a simple keyword based grouping. We performed this in an evolutionary way wherein initial search terms were entered and used to classify sites, then those that did not match were used. This method bears a resemblance to Moore and Rid's methodology

for 'Cryptopolitik and the Darknet' [41], except our site classification was accomplished via string matching instead of Support Vector Machines. There have been other papers that have taken similar approaches to trawling and collecting .onion domains en masse for the purpose of collecting [6], [55] and classifying [48] with differing conclusions [7].

Site Type	No. of Sites
Drugs	7,011
Fraud	6,289
Child Abuse Imagery	3,816
Other	3,164
Hacking	1,461
Forums	1,454
Indexes	1,100
Murder	1,068
Bestiality	918
Human Rights	201
Seized	64

Table III: Number of Unique Domains By Category

We achieved this through identifying a number of terms in the titles and text of websites that could be extrapolated to be relevant to a particular site type, crime or sub-class of crime. For instance we used the strings 'lolita', 'PTHC' (Pre-Teen HardCore) and 'Infant' to determine a site as belonging to child abuse imagery distributors. In contrast, we entered the strings 'CVV' (card number verification value), 'Dumps', 'Fullz' and 'Mixer', amongst others, into a financial crime wordlist. We then programatically compared these wordlists against the text and titles for all of the sites and pages discovered by the onion service crawler and weighted according to the amount of matches. The output of this classification was a weighting applied to each of the sites indexed. Only a very small amount of the sites in our dataset could not be categorised.

Sites could match multiple different categories. One site could be in both in the Drugs and Fraud market categories, or simultaneously the Fraud and Hacking categories. This would be indicative of a crossover of certain elements of criminal activity; for instance a market may have been seen to sell both drugs and data for the commission of fraud, simultaneously.

A. Child Abuse Imagery

We discovered a significant amount of the sites on onion services to be hosting child abuse imagery. Analysis of the keywords and textual content of the site revealed that sites were hosting forums, image hosting, video hosting and other such services. We also found a subset of sites were hosting 'Hurtcore' content [44]; content wherein blackmail and extortion is used to satisfy the sadistic impulses of the criminals involved. This is the creation of content that humiliates the victims, in a manner similar to sextortion; but with the aim being non-consensual pain of the victim rather than financial profit.

B. Weaponry Sites

A report by the United Nations states that the arms trade on the Tor network was small in scale in comparison to traditional arms markets, traditional illicit arms markets and also darknet markets [50]. Additionally a further complication may be related to deterrence for consumers. There is a far larger penalty for possession of a weapon in most jurisdictions than there is for drug sales.

C. Drug Markets

Our analysis determined that the most frequent utilisation of the Tor network by density of sites, was for the ostensible distribution of controlled drugs. A wide variety of academic and government research has been conducted on this well-established use of onion services [19], [31]. These drug distribution networks are located typically in marketplaces [38], which may offer other goods, but for which the main business is typically drug distribution.

D. Hacking and Fraud Sites

Not necessarily distinct from the darknet markets that offer drugs for sale, we observed a number of sites offering the opportunity to rent services, software and data for the perpetration of fraud. These darknet markets frequently offer data such as credit card numbers or malware for sale. As well as being present on the established darknet markets, there are markets, forums and stores entirely for the provision of one type of criminal accessory, for instance sites dedicated exclusively to credit card fraud or malware sales. There tends to be a colocation of financial fraud services offered in some of these sites [30].

E. Murder Sites

We observed there to be a subsection of sites that were similar in structure. These sites appeared to offer contract killers to dupe potential buyers; with millions in illicit profit achieved according to prior research [56]. These sites were seen to offer contract murder services with surrounding violence and threatening behaviour for investigators and detractors. These sites had similar HTML code, page layout and similar operating models that appeared to be related to taking upfront payment in Bitcoin for services rendered.

F. Journalists, Law Enforcement and Freedom of Speech

We identified onion services that are run by corporate entities or for the purpose of freedom of speech. Corporations that have utilised .onion services are those that have a high level of technical sophistication such as Facebook. The United States Central Intelligence Agency recently set up a onion service [9] as a proxy to their website on the clearnet; ostensibly to enable more functional anonymity for those submitting information in addition to the public relations boost this provided.

We discovered a multi-agency task-force, the *Northern California Illicit Digital Economy Taskforce* (NCIDE), with a website at `http://ncidetf3j26mdtvf.onion`; established for the purpose of providing information to the task force and also to advertise which vendors in Northern California have been apprehended as a result of their investigation efforts on Tor sites [63].

Journalism sites such as `ProPublica` [59] and `The Intercept` [58] have set up public-facing .onion sites for confidential communication with sources and also to show that they take the anonymity and privacy of their journalistic sources seriously.

We discovered that the journalism security software `SecureDrop`, formerly known as `DeadDrop` and `Strongbox` hosted a number of relevant sites hosted in the Tor network. These sites are part of a decentralised network through which journalist's sources are able to submit confidential documents to capitalise on the anonymity providing properties of the Tor network.

VIII. Countermeasures

The way in which it is possible for users of onion services to ensure that they are not defrauded when trying to access a given domain for a given onion service is the same as in forensic science; that being confirmation and validation of findings by cross-referencing with multiple trusted sources. It is also worthwhile questioning whether this is a practice that needs to explicitly be reduced via countermeasures, despite the setting up of these honeypot websites being fraudulent in nature. The existence of these sites may prevent some money from going into the online cybercriminal marketplaces. The public health and ethical implications of the existence of darknet markets and imitations are beyond the scope of this analysis but covered in other papers; some of which suggest there may may in fact be benefits to darknet markets from a harm reduction perspective [3]–[5].

A. Site Verification via PGP Keys

One method to create webs of trust is through the use of saved and verified PGP public keys. These PGP keys need to be saved on an initial visit to the genuine market to allow for potential future verification of mirrors, else they are not useful as verification mechanisms. We observed that one particular cluster of these mirror sites had an identical reverse proxied version of Empire Market; however when this site was loaded we discovered that the user cannot pass the CAPTCHA to access PGP keys and as such was attempted to be softly dissuaded from pursuing further. This occurred because PGP keys cannot be falsified to prove authenticity.

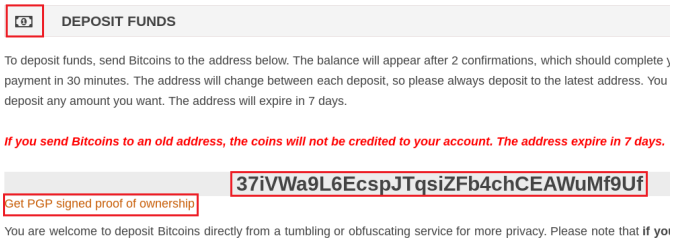

Figure 6: Empire Market - PGP Verification for Deposits

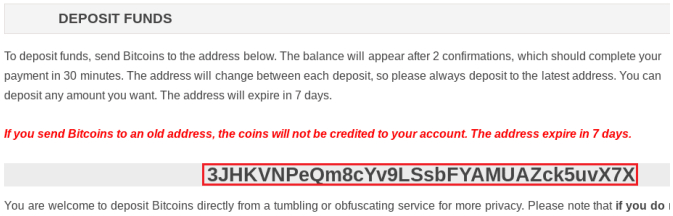

Figure 7: Empire Market Phish - No PGP Verification

As we show in Figure 6 and Figure 7, PGP verification is a method used to ascertain whether or not the market to which the user is depositing their Bitcoins is legitimate. The first, a real market, had PGP verification in the deposit section. The second, a phishing site, subtly removed it. This proves the value and unfalsifiability of PGP and its requisite infrastructure for verification to provide trust. The most reliable indicator of legitimacy for a consumer is the verification of PGP key on a site. However this requires forethought, for the consumer to have saved the correct PGP verification key on their initial visit to the domain.

B. Site Verification via Central Listings

Central guide sites such as previously Deep Dot Web, may enable people to prevent themselves from being led to incorrect sites whilst attempting to conduct commerce via onion services.

Deep Dot Web, despite its harm reduction approach [27] and enabling of information exchange; was shut down by law enforcement and its owners were arrested, for money laundering. This was due to the administrator's practice of taking commission from the markets to which it linked.

The Hidden Wiki, another comparable service, the longest-running, having been established in 2007 [37] was also shut down by law enforcement in an earlier iteration of itself. There are search engines that enable the conduction and identification of relevant onion services. One of these `Ahmia`, was resultant from a Google Summer of Code project.

To an extent these search engines can help remove phishing links, but .onion domains switch so fast that it may indeed be impossible to realistically vet the links. It is also necessary to note that many of these sites have differing attitudes towards certain crime classes. Many of these sites will remove phishing and cloned links from them, in addition to child abuse imagery [34], showing a tendency to view child abuse imagery distributors as especially morally reprehensible in comparison to other criminals.

There are indexing sites and Wikis which are used by users of onion services to determine the correct URLs to visit. These sites themselves were seen to be subject to imitation. An example of this is The Hidden Wiki, once a central indexing site that was present on a onion service but that now has been succeeded by a number of sites that contain various links. An example of a current site that was subject to imitation is *https://dark.fail/*, that has phishing imitations, *https://darkfail.com/*, *https://darkfail.org/* and *https://därk.fail/*, that have imitated its aesthetic, but do not contain links to legitimate darknet markets.

The reputation of index sites and link sites is a contentious issue, particularly following the seizure and arrests of DeepDotWeb's administrators and site for money laundering and assisting criminal enterprises [22]. As this site was well regarded in terms of providing reliable information the void left in the wake of its shutdown has been occupied by a number of sites purporting to be indexing sites.

IX. Limitations and Future Work

There are several limitations that can be noted to be present in the experimental methodology we adopted and used for this study. We collected data from a finite time period, 2019. As such, there is a lack of longitudinal relevance. Though prior campaigns and historical data have illustrated that this is a long-term phenomenon they have not been subject to a standardised and rigorous analysis, especially to identify and cluster given campaigns.

Due to the particular nature of the crawler, we were only able to collect sites that were linked to from other onion services or that had been aggregated by onion address listing sites. This may have left some of the more obscure sites, or sites that are operating an HTTP service on obscure ports out of the collected data. Also, services such as `Cloudflare` or other crawler detection tools would have prevented complete data collection.

The estimation of the size of the profitability for these sites was based on the transactions that were visible for the sites in question. These Bitcoin addresses were also not subject to aggregated analysis which would have given a more precise indication of the scale of the profitability of this phenomenon.

An important area for future work in this field is to propose methods that wed the increased anonymity and security offered by the Tor network with some methods for validating the providers of certain onion services.

There may also be avenues for research following on from our discoveries illuminated, that involve the disruption of the trust architecture of online criminal networks. This may involve concerted attacks on the centre of gravity for trust in anonymity networks. It may also involve analysis of phishing domains as effective honeypots or information harvesting portals. Investigation also needs to be made of the effects of allowing indexing sites and search engines to list and dictate the overarching trust model available for onion services. Indexing sites are perceived and treated as the central authority that is available with regards to sites hosted on onion services. To this end it is possible to view the indexing sites, particularly the more well regarded, as the closest thing to a reliability indicator for onion services.

X. Conclusions

This paper makes several key contributions, towards a better understanding of the phishing sites on the darknet.

We provide practical evidence that over half the sites on the darknet were imitations or benign duplicates. Whilst we observed some of these sites used for the purpose of phishing, a larger quantity of them were clones; to add redundancy for legitimate sites, or to increase the amount of visitors for a particular fraudulent or non-fraudulent venture.

We also identified that imitation of darknet markets is a phenomenon that has occurred through a number of sites, historically and recently. There were also provable victims and income generated from these phishing sites.

We have also shown that it is possible to use HTTP headers and other methods to fingerprint and cluster these sites into identifiable, concerted campaigns that span across different marketplaces.

We observed a wide variety of uses for onion services. Most sites that we found and classified in terms of volume were hosted for criminal purposes. However there are also onion services for legal purposes and the use of the Tor network is not restricted to criminals. Political dissidents, governments, the privacy conscious and other citizens use these services.

Onion services frequently experience cloning and imitation via phishing sites due to the nature of the Tor onion service environment and easily exploited trust hierarchies. Tor-based phishing is a criminal activity that principally occurs to facilitate defrauding customers using darknet markets. Tor onion services are uniquely vulnerable to cloning due to the nature of trust and fluctuating reputation of onion services.

References

[1] P. Agten, W. Joosen, F. Piessens, and N. Nikiforakis, "Seven Months ' Worth of Mistakes : A Longitudinal Study of Typosquatting Abuse," *Network and Distributed System Security Symposium*, pp. 8–11, 2017. [Online]. Available: https://securitee.org/files/typosquatting{_}ndss2015.pdf

[2] R. Anderson, I. Shumailov, A. Rietmann, and M. Ahmed, "Bitcoin Redux," *Workshop on the Economics of Information Security*, pp. 1–33, 2018. [Online]. Available: https://www.cl.cam.ac.uk/{~}rja14/Papers/bitcoin-redux.pdf

[3] A. Bancroft, "Responsible use to responsible harm: illicit drug use and peer harm reduction in a darknet cryptomarket," *Health, Risk and Society*, vol. 19, no. 7-8, pp. 336–350, 2017. [Online]. Available: https://doi.org/10.1080/13698575.2017.1415304https://www.tandfonline.com/doi/pdf/10.1080/13698575.2017.1415304?needAccess=true

[4] A. Bancroft and P. Reid, "Concepts of illicit drug quality among darknet market users: Purity, embodied experience, craft and chemical knowledge," *International Journal of Drug Policy*, vol. 35, pp. 42–49, 2016. [Online]. Available: http://dx.doi.org/10.1016/j.drugpo.2015.11.008

[5] M. J. Barratt, S. Lenton, A. Maddox, and M. Allen, "âĂŸWhat if you live on top of a bakery and you like cakes?'âĂŤDrug use and harm trajectories before, during and after the emergence of Silk Road," *International Journal of Drug Policy*, vol. 35, pp. 50–57, 2016. [Online]. Available: http://dx.doi.org/10.1016/j.drugpo.2016.04.006

[6] M. Bernaschi, A. Celestini, S. Guarino, and E. Mastrostefano, "Spiders like Onions : on the Network of Tor Hidden Services," *World Wide Web Conference*, 2019.

[7] A. Biryukov, I. Pustogarov, and R.-p. Weinmann, "Trawling for Tor Hidden Services : Detection , Measurement , Deanonymization," *IEEE Symposium on Security and Privacy*, 2013. [Online]. Available: https://www.ieee-security.org/TC/SP2013/papers/4977a080.pdf

[8] Caleb, "An Interview with the Creator of the Darknet Search Engine âĂIJFresh OnionsâĂİ," 2017. [Online]. Available: https://medium.com/@c5/always-use-a-trusted-source-for-darknet-market-links-189de792bb39

[9] Central Intelligence Agency, "CIA's Latest Layer: An Onion Site," 2019. [Online]. Available: https://www.cia.gov/news-information/featured-story-archive/2019-featured-story-archive/latest-layer-an-onion-site.html

[10] N. Christin, "Traveling the Silk Road : A Measurement Analysis of a Large Anonymous Online Marketplace," *World Wide Web Conference*, 2013.

[11] C. Cimpanu, "Tor Project to fix bug used for DDoS attacks on Onion sites for years," 2019. [Online]. Available: https://www.zdnet.com/article/tor-project-to-fix-bug-used-for-ddos-attacks-on-onion-sites-for-years/

[12] L. Clark, "Hacker Forum Darkode is Back and More Secure Than Ever," 2015. [Online]. Available: https://www.wired.co.uk/article/darkode-back-and-more-secure

[13] Council of European Union, "Regulation (eu) 2016/679 of the european parliament and of the council of 27 april 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing directive 95/46/ec (general data protection regulation)," 2016. [Online]. Available: https://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1528874672298&uri=CELEX:02016R0679-20160504

[14] N. Cubrilovic, "Large Number of Tor Hidden Sites Seized by the FBI in Operation Onymous were Clone or Scam Sites - Archive.org," 2014. [Online]. Available: https://web.archive.org/web/20150414144737/https://www.nikcub.com/posts/onymous-part1/

[15] R. Dhamija and J. D. Tygar, "Why Phishing Works," *CHI 2006 Proceedings*, no. November 2005, pp. 581–590, 2006.

[16] R. Dingledine, N. Mathewson, and P. Syverson, "Tor : The Second-Generation Onion Router," *ACM Special Interest Group on Data Communication*, 2018. [Online]. Available: https://svn.torproject.org/svn/projects/design-paper/tor-design.pdf

[17] D. Dittrich and E. Kenneally, "The Menlo Report," Department of Homeland Security: Science and Technology, Tech. Rep. August, 2012. [Online]. Available: https://www.caida.org/publications/papers/2012/menlo{_}report{_}actual{_}formatted/menlo{_}report{_}actual{_}formatted.pdf

[18] M. Dittus, J. Wright, and M. Graham, "Platform Criminalism," *World Wide Web Conference*, 2018.

[19] D. S. Dolliver and J. L. Kenney, "Characteristics of Drug Vendors on the Tor Network : A Cryptomarket Comparison," *Victims & Offenders: An International Journal of Evidence-based Research, Policy, and Practice*, vol. 4886, no. May, p. 8, 2016.

[20] Europol, "Global Action Against Dark Markets on Tor Network: Press Release," 2014. [Online]. Available: https://www.europol.europa.eu/newsroom/news/global-action-against-dark-markets-tor-network

[21] ——, "Cybercriminal Darkode Forum Taken Down Through Global Action," 2015. [Online]. Available: https://www.europol.europa.eu/newsroom/news/cybercriminal-darkode-forum-taken-down-through-global-action

[22] ——, "DeepDotWeb Shut Down," 2019. [Online]. Available: https://www.europol.europa.eu/newsroom/news/deepdotweb-shut-down-administrators-suspected-of-receiving-millions-of-kickbacks-illegal-dark-web-proceeds

[23] ——, "Double Blow to Dark Web Marketplaces," 2019. [Online]. Available: https://www.europol.europa.eu/newsroom/news/double-blow-to-dark-web-marketplaces

[24] S. Foley, J. R. Karlsen, and T. PutniĹĘš, "Sex, Drugs and Bitcoin: How Much Illegal Activity Is Financed through Cryptocurrencies ?" *Review of Financial Studies*, 2018.

[25] L. Gayard, *Darknet: Geopolitics and Uses*. John Wiley & Sons, 2018.

[26] R. Graham, "XKeyScore: regex foo," 2014. [Online]. Available: https://blog.erratasec.com/2014/07/xkeyscore-regex-foo.html

[27] A. Greenberg, "Feds Dismantled The Dark-Web Drug Trade - But It's Already Rebuilding," 2019. [Online]. Available: https://www.wired.com/story/dark-web-drug-takedowns-deepdotweb-rebound/

[28] Gwern, "Darknet Market Mortality Risks," 2019. [Online]. Available: https://www.gwern.net/DNM-survival

[29] C. Hadnagy and M. Fincher, *Phishing Dark Waters*. Wiley, 2015.

[30] A. Haslebacher, J. Onaolapo, and G. Stringhini, "All Your Cards Are Belong To Us : Understanding Online Carding Forums," *APWG Symposium on Electronic Crime Research (eCrime)*, 2017. [Online]. Available: https://seclab.bu.edu/people/gianluca/papers/haslebacher{_}forums.pdf

[31] HM Government, "2017 Drug Strategy," no. July, p. 20, 2017. [Online]. Available: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment{_}data/file/628148/Drug{_}strategy{_}2017.PDF

[32] Information Commissioner's Office, "Anonymisation: managing data protection risk code of practice," 2012. [Online]. Available: https://ico.org.uk/media/for-organisations/documents/1061/anonymisation-code.pdf

[33] E. Jardine, "The Dark Web Dilemma: Tor, Anonymity and Online Policing," *Global Commission on Internet Governance Paper Series*, no. 21, 2015. [Online]. Available: https://www.cigionline.org/sites/default/files/no.21.pdf

[34] Juha, "Ahmia Search After GSoC Development," 2014. [Online]. Available: https://blog.torproject.org/ahmia-search-after-gsoc-development

[35] K. Kruithof, J. Aldridge, D. Décary-hétu, M. Sim, E. Dujso, and S. Hoorens, "Internet-Facilitated Drugs Trade An Analysis of The Size , Scope and The Role of the Netherlands," *RAND Europe*, p. 24, 2016.

[36] V. Levenshtein, "Binary Codes Capable of Correcting Deletions Insertions and Reversals," Tech. Rep., 1965. [Online]. Available: https://nymity.ch/sybilhunting/pdf/Levenshtein1966a.pdf

[37] K. Loesing, "Length of New Onion Addresses," 2007. [Online]. Available: https://lists.torproject.org/pipermail/tor-dev/2007-June/001442.html

[38] J. Martin, "Lost on the Silk Road : Online drug distribution and the âĂŸ cryptomarket '," *Criminology & Criminal Justice*, p. 2, 2014.

[39] J. Martin and N. Christin, "Ethics in cryptomarket research," *International Journal of Drug Policy*, vol. 35, pp. 84–91, 2016. [Online]. Available: http://dx.doi.org/10.1016/j.drugpo.2016.05.006

[40] C. Monteiro, "Tor Wars: The Onion Cloner's legacy," 2016. [Online]. Available: https://pirate.london/tor-wars-the-onion-cloners-legacy-e28f35a42665

[41] D. Moore and T. Rid, "Cryptopolitik and the Darknet," *Journal of Strategic Studies*, 2015. [Online]. Available: https://www.tandfonline.com/doi/pdf/10.1080/00396338.2016.1142085?needAccess=true

[42] T. Moore and R. Clayton, "An empirical analysis of the current state of phishing attack and defence," *Workshop on the Economics of Information Security*, pp. 1–20, 2007. [Online]. Available: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.98.6098{&}rep=rep1{&}type=pdf

[43] A. Muffett and J. Appelbaum, "The ".onion" Special-Use Domain Name," *Internet Engineering Task Force*, 2015. [Online]. Available: https://tools.ietf.org/html/rfc7686{#}ref-Dingledine2004

[44] National Crime Agency, *National Crime Agency Annual Report and Accounts 2017âĂŞ18*. National Crime Agency, 2017. [Online]. Available: https://nationalcrimeagency.gov.uk/who-we-are/publications/177-nca-annual-report-accounts-2017-18/file

[45] A. Oest, Y. Safaei, A. Doupe, G. J. Ahn, B. Wardman, and K. Tyers, "PhishFarm: A scalable framework for measuring the effectiveness of evasion techniques against browser phishing blacklists," *Proceedings - IEEE Symposium on Security and Privacy*, vol. 2019-May, pp. 1344–1361, 2019. [Online]. Available: https://adamdoupe.com/publications/phishfarm-oakland2019.pdf

[46] A. Oest, Y. Safaei, P. Zhang, B. Wardman, K. Tyers, Y. Shoshitaishvili, A. Doupé, and G.-j. Ahn, "PhishTime : Continuous Longitudinal Measurement of the Effectiveness of Anti-phishing Blacklists," *29th USENIX Security Symposium (USENIX Security 20)*. [Online]. Available: https://www.adamoest.com/phishtime_usenix_2020_oest.pdf

[47] A. Oest, P. Zhang, B. Wardman, E. Nunes, J. Burgis, A. Zand, K. Thomas, A. Doupé, and G.-J. Ahn, "Sunrise to Sunset: Analyzing the End-to-end Life Cycle and Effectiveness of Phishing Attacks at Scale," *29th USENIX Security Symposium (USENIX Security 20)*, 2020. [Online]. Available: https://www.usenix.org/conference/usenixsecurity20/presentation/oest

[48] G. Owen and N. Savage, "Empirical analysis of Tor Hidden Services," *IET Information Security*, pp. 1–6, 2015.

[49] ——, "The Tor Dark Net," *Global Commission on Internet Governance Paper Series*, no. 20, p. 9, 2015. [Online]. Available: https://www.cigionline.org/sites/default/files/no20{_}0.pdf

[50] G. P. Paoli, "The Trade in Small Arms and Light Weapons on the Dark Web," United Nations Office for Disarmament, Tech. Rep. 32, 2018. [Online]. Available: https://s3.amazonaws.com/unoda-web/wp-content/uploads/2018/10/occasional-paper-32.pdf

[51] S. Pastrana, D. Thomas, A. Hutchings, and J. Tapiador, "Measuring ewhoring," *Proceedings of the ACM SIGCOMM Internet Measurement Conference, IMC*, pp. 463–477, 2019.

[52] R. S. Patel, *Kali Linux Social Engineering*. Packt Publishing, 2013.

[53] Phobos, "Thoughts and Concerns about Operation Onymous," 2014. [Online]. Available: https://blog.torproject.org/thoughts-and-concerns-about-operation-onymous

[54] H. V. Riper, "Dark Web Typosquatting: Scammers v. Tor," 2019. [Online]. Available: https://www.digitalshadows.com/blog-and-research/dark-web-typosquatting-scammers-v-tor/

[55] I. Sanchez-rola, D. Balzarotti, and I. Santos, "The Onions Have Eyes : A Comprehensive Structure and Privacy Analysis of Tor Hidden Services," *World Wide Web Conference*, 2017. [Online]. Available: http://s3.eurecom.fr/docs/www17{_}darktracing.pdf

[56] J. Simpson, "Hire-a-Hitman Website is a Scam and its Owner has Made a Killing," 2018. [Online]. Available: https://www.thetimes.co.uk/article/hire-a-hitman-website-is-a-scam-and-its-owner-has-made-a-killing-

[57] J. Szurdi, B. Kocso, G. Cseh, J. Spring, M. Felegyhazi, and C. Kanich, "The Long âĂIJTaileâĂİ of Typosquatting Domain Names," *USENIX Security Symposium*, 2014. [Online]. Available: https://www.usenix.org/system/files/conference/usenixsecurity14/sec14-paper-szurdi.pdf

[58] The Intercept, "The Intercept Welcomes Whistleblowers," 2019. [Online]. Available: https://theintercept.com/source/

[59] M. Tigas, "A More Secure and Anonymous ProPublica Using Tor Hidden Services," 2016. [Online]. Available: https://www.propublica.org/nerds/a-more-secure-and-anonymous-propublica-using-tor-hidden-services

[60] Tor Project, "Tor: Onion Service Protocol," 2019. [Online]. Available: https://2019.www.torproject.org/docs/onion-services.html.en

[61] United States District Court for the District of Conneticut, "United States of America vs. Michael Richo," Tech. Rep., 2016. [Online]. Available: https://assets.documentcloud.org/documents/3883353/Michael-Richo-Criminal-Complaint.pdf

[62] United States District Court for the Eastern District of California, "United States of America vs. Marcos Paulo De Oliveira-Annibale," vol. 91, 2019. [Online]. Available: https://www.justice.gov/opa/press-release/file/1159711/download

[63] U.S Department of Justice, "âĂIJDrfrostyâĂİ Indicted for Distribution of Methamphetamine from Modesto Using the Darknet," 2019. [Online]. Available: https://www.justice.gov/usao-edca/pr/drfrosty-indicted-distribution-methamphetamine-modesto-using-darknet

[64] M. Wachs, M. Schanzenbach, and C. Grothoff, "On the Feasibility of a Censorship Resistant Decentralized Name System," *Foundations and Practice of Security*, vol. 8352, p. 6, 2013. [Online]. Available: http://link.springer.com/10.1007/978-3-642-37119-6

[65] A. Winstock, "Global Drug Survey 2019Executive Summary," Global Drug Survey, Tech. Rep., 2019. [Online]. Available: https://www.globaldrugsurvey.com/wp-content/themes/globaldrugsurvey/results/GDS2019-Exec-Summary.pdf

[66] A. Winston, "How A Dark Web Drug Ring Was Uncovered After Suspicious A.T.M Withdrawals," Tech. Rep., 2019. [Online]. Available: https://www.nytimes.com/2019/04/16/nyregion/dark-web-drug-dealing.html

[67] P. Winter, A. Edmundson, L. M. Roberts, and N. Feamster, "How Do Tor Users Interact With Onion Services?" *USENIX Security Symposium*, no. June, 2018. [Online]. Available: https://arxiv.org/pdf/1806.11278.pdf

[68] C. Yoon, K. Kim, Y. Kim, S. Shin, and S. Son, "Doppelgängers on the dark web: A large-scale assessment on phishing hidden web services," *The Web Conference 2019 - Proceedings of the World Wide Web Conference, WWW 2019*, vol. 2, pp. 2225–2235, 2019.