2019 11th International Conference on Cyber Conflict:
Silent Battle
T. Minárik, S. Alatalu, S. Biondi,
M. Signoretti, I. Tolga, G. Visky (Eds.)
2019 © NATO CCD COE Publications, Tallinn

Permission to make digital or hard copies of this publication for internal use within NATO and for personal or educational use when for non-profit or non-commercial purposes is granted providing that copies bear this notice and a full citation on the first page. Any other reproduction or transmission requires prior written permission by NATO CCD COE.

Hidden in the Shadow: The Dark Web – A Growing Risk for Military Operations?

Robert Koch
Fraunhofer FKIE
Bonn, Germany
robert.koch@fkie.fraunhofer.de

Abstract: A multitude of leaked data can be purchased through the Dark Web nowadays. Recent reports highlight that the largest footprints of leaked data, which range from employee passwords to intellectual property, are linked to governmental institutions. According to OWL Cybersecurity, the US Navy is most affected. Thinking of leaked data like personal files, this can have a severe impact. For example, it can be the cornerstone for the start of sophisticated social engineering attacks, for getting credentials for illegal system access or installing malicious code in the target network. If personally identifiable information or sensitive data, access plans, strategies or intellectual property are traded on the Dark Web, this could pose a threat to the armed forces.

The actual impact, role, and dimension of information treated in the Dark Web are rarely analysed. Is the available data authentic and useful? Can it endanger the capabilities of armed forces? These questions are even more challenging, as several well-known cases of deanonymization have been published over recent years, raising the question whether somebody really would use the Dark Web to sell highly sensitive information. In contrast, fake offers from scammers can be found regularly, only set up to cheat possible buyers. A victim of illegal offers on the Dark Web will typically not go to the police.

The paper analyses the technical base of the Dark Web and examines possibilities of deanonymization. After an analysis of Dark Web marketplaces and the articles traded there, a discussion of the potential risks to military operations will be used to identify recommendations on how to minimize the risk. The analysis concludes that surveillance of the Dark Web is necessary to increase the chance of identifying sensitive information early; but actually the ‘open’ internet, the surface web and the Deep Web, poses the more important risk factor, as it is – in practice – more difficult to surveil than the Dark Web, and only a small share of breached information is traded on the latter.

Keywords: *Dark Web, military operations, data breaches, data leaks, data sale, marketplaces, anonymity, Tor, deanonymization, operational security, OPSEC, PII*

1. INTRODUCTION

The so-called Dark Web has been in the focus of the media in recent years, regularly in a negative context. With the takedown of the ‘Silk Road’ website in October 2013 by the FBI, the Dark Web entered the awareness of large parts of the population. In February 2015, the FBI took the infamous Dark Web site ‘Playpen’ offline, which hosted more than 23,000 child pornographic images and videos and had more than 215,000 users. As part of the preparation for the terrorist attacks in Paris in November 2015, the communication was anonymized by using the software Tor; while the weapon used in the shooting rampage in Munich in July 2016 was also acquired over the Dark Web. Beside drugs, weapons, and child pornography, every kind of information is sold via marketplaces on the Dark Web: from credit cards to sensitive information captured during data leaks or hacking attacks. The latter can pose new challenges for the armed forces.

Since sensitive data is repeatedly looted (see the overview of the world’s biggest data breaches (McCandless 2018)), the possibilities of the Dark Web can increase the motivation of attackers even further: based on the anonymity of the users, as well as the easy to use but (in the sense of the user, not fully traceable transactions) hard to track digital currencies like Bitcoin, illegal activities can be executed with apparently low risk for criminals.

To analyse the possible influence of the Dark Web on military operations, an overview is provided in Section 2, including an analysis of the technical background. Based on that, possibilities of deanonymization attacks are discussed; the security and reliability of the Dark Web may have an influence on the offered content. Next, an analysis of Dark Web marketplaces and the goods traded there is provided in Section 3, followed by a discussion of the resulting potential risks for military operations in Section 4. Finally, the main arguments of the paper are summarized in Section 5.

2. THE ONION ROUTER AND ANONYMITY

To understand the opportunities and weaknesses when using the Dark Web, some knowledge of how anonymization networks work is required. Therefore, terms with respect to the Dark Web are explained. These are often mixed, but must be clearly separated. This is followed by an investigation into the security levels of the Dark Web, since this is fundamental for an evaluation of the transactions to be expected there.

A. Terminology

Quite often, the terms *Darknet, Deep Web* and *Dark Web* are improperly mixed or used interchangeably. Due to insufficient separation and misuse of terms, data and evaluations can be incorrectly assigned and falsify the actual situation.

Deep Web. The Deep Web "refers to any Internet information or data that is inaccessible by a search engine and includes all websites, intranets, networks and online communities that are intentionally and/or unintentionally hidden, invisible or unreachable to search engine crawlers" (Janssen 2018). The term, Deep Web, "relates to deep sea/ocean environments that are virtually invisible and inaccessible" (Janssen 2018). Therefore, the Deep Web "contains data that is dynamically produced by an application, unlinked or standalone Web pages/websites, non-HTML content and data that is privately held and classified as confidential. Some estimate the size of the Deep Web as many times greater than the visible or Surface Web" (Janssen 2018).

Darknet. From a technical and historical point of view, the term 'Darknet' is used to describe the part of the IP address space which is *routable, but not in use*. This must be differentiated from addresses, which should not be routed by definition. In the still predominantly used internet addressing architecture, Internet Protocol version 4 (IPv4), specific addresses are defined as private.[1] By using them, a router can provide connectivity to numerous attached devices by using its own public address, translating the traffic between the private network and the internet. The respective private addresses are not visible on the internet; therefore, they should *not* be routed, and only routable addresses can be seen. By monitoring these unused but routable addresses, a lot of observations with respect to security can be made: normally, nobody should interact with them. So if some interaction can be seen, the underlying behaviour is typically malicious, e.g., an automated worm run looking for target addresses to infect. This security-relevant part of the address space is called the *Darknet*.

One of the early uses of the term with regard to digital content can be found in an article about content protection. It described Darknets as a 'collection of networks and technologies used to share digital content' (Biddle 2002). Nowadays, the term

[1] Subnetworks 10.0.0.0/8, 172.16.0.0/16 and 192.168.16.0/24, RFC1918.

is mainly used for o*verlay networks* providing anonymous network connectivity and services. An overlay network is a layer of virtual network topology on top of the physical layer, which directly interfaces with users (Zhang 2003). Tor is an example of an overlay network, and the biggest and most widely used anonymisation network; but there are numerous others, such as I2P, Freenet or ZeroNet.

It is important to recognize that the term *Darknet* originally refers to the network itself, and therefore the *technical base* like the protocol and devices; but not the content which may be transported through the network, or can be found on its respective servers.

Dark Web. The Dark Web refers to the websites which are hosted within overlay networks, and are *normally*[2] not accessible without special software like the Tor Browser. Nowadays, usage of the Tor network is easy and straightforward: the Tor Browser is a complete bundle ready to use without installation by providing a fully configured Firefox Browser. As in the case of the Deep Web, search engine crawlers are not able to index the websites of the Dark Web. But in contrast to it, its most important feature is that the users of a service stay anonymous - neither a provider of a website can identify the visitors, nor can a visitor identify the service provider. Given this, the respective services are also called ‘hidden services’; more recently, ‘onion services’.

B. Anonymity on the Internet

The history of privacy-enhancing technologies dates back to 1981, with a technique to hide the communicating participants of an electronic mail system and their messages (Chaum 1981). Since then, much work has been done in the area of anonymization techniques, with the Tor Project one of the most well-known. The acronym Tor stands for ‘The Onion Router’, based on the underlying principle of onion routing (Reed 1998). It was developed as a research project of the Naval Research Laboratory in the 1990s, with the purpose of protecting the online communication of US intelligence agents. The first pre-alpha of Tor was published in 2002 (Dingledine 2002). In 2004, the second generation of the system was published (Dingledine 2004), and the code released under a free licence.

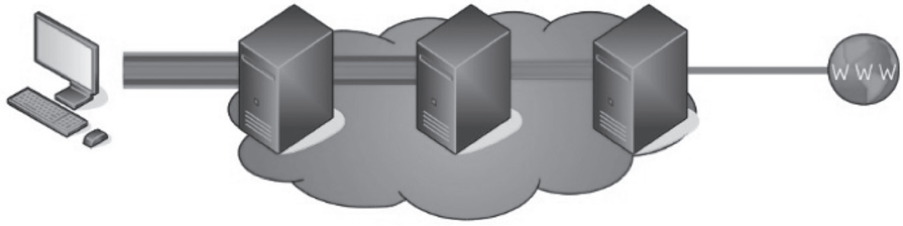
Becoming Anonymous. Two basic modes of application are offered by Tor: anonymous access to the internet, and onion services. In the first case, the traffic is routed through the Tor network and returns to the internet via so-called Exit relays. When accessing a website on the internet, it does not see the real IP address of the user, but that of the Exit relay; the IP of the user is not traceable. In the second case, the traffic stays *within* the Tor network: users can offer services like websites or instant messaging servers,

[2] Nowadays, there are also ways to access the content of the Dark Web without the use of special software. For example, the website tor2web.org enables browsing and accessing content on the Dark Web without the use of Tor software; though one must be aware when using this service that only the provider of the content stays anonymous, not the requesting user.

while others can access them via so-called ‘rendezvous points’. Both sides, the visitor as well as the service provider, stay anonymous.

To get a better idea of how Tor works, anonymous access to the internet is briefly described. Tor generates an overlay network in which each relay maintains a Transport Layer Security (TLS) connection to every other relay. Based on that, Tor establishes a circuit - a random pathway through the network - by selecting an Entry, Middle, and Exit relay.[3] The Exit relay is chosen based on a weighted random selection and changes regularly.[4] When sending data through Tor, the client encrypts it multiple times with the relays’ keys, including the predecessor’s and successor’s addresses for their respective relays. Each relay has the key for only one layer, uses the key to remove that layer, then forwards the data. In this way, it sees only the IP address of where the packet came from and where it must go. The Exit relay sends the packet to its final destination, which sees only the exit relay’s IP address. When the answer returns, each relay adds its encryption layer only the sender can finally remove them all and thus read the answer. Figure 1 visualizes the routing and anonymizing process of Tor.

FIGURE 1. FUNCTIONAL PRINCIPLE OF ONION ROUTING. EVERY RELAY ADDS RESPECTIVELY REMOVES ONE LAYER OF ENCRYPTION, AND ONLY KNOWS ITS IMMEDIATE PREDECESSOR AND SUCCESSOR.



Becoming Deanonymized. Due to the broad application possibilities of the Tor network, positive as well as negative/illegal ones, there is a strong interest in deanonymizing providers as well as users of onion services. For example, repressive regimes can try to locate those who use Tor for freedom of expression; while government agencies can try to fight illegal drug trafficking or child pornography. Therefore, many efforts to deanonymize users have been made and three basic categories can be identified, which will be explained briefly:

CAT 1 The first category includes attacks at the technical level. This is the most dangerous, but in practice also the rarest type of deanonymization attack. These can be

[3] Tor can extend the circuit by adding relays; but a circuit typically has only one Middle relay, so that communication latency remains at an acceptable level.

[4] By default, the circuit for a new TCP stream is rotated all 10 minutes to avoid profiling attacks; long-lasting single TCP streams (e.g., an IRC connection) are not rotated and will stay on the same circuit (Tor 2015).

directed against implementation flaws of the Tor software, but also attack weaknesses in the design of the network protocol of Tor. Attacks based on actual technical shortcomings of Tor are rare, but can have severe impact. An important example is the 'relay early' traffic confirmation attack, which was identified and executed between January 30, 2014 and July 4, 2014 by the Software Engineering Institute of Carnegie Mellon University (Dingledine 2014). The identified IP addresses were subpoenaed by the FBI and used in the trial against Brian Farrell:

> The record demonstrates that the defendant's IP address was identified by the Software Engineering Institute ("SEI") of Carnegie Mellon University (CMU") [sic.] when SEI was conducting research on the Tor network which was funded by the Department of Defense ("DOD") […] Farrell is charged with conspiracy to distribute cocaine, heroin, and methamphetamine due to his alleged role as a staff member of the Silk Road 2.0 dark web marketplace (Cox 2016).

CAT 1b Another attack on a technical basis is much more common – but not directed against the Tor software or the protocol itself, but against *the used browser*. While Tor can be used with any browser, this must be configured accordingly. The Tor Browser, which is based on a Mozilla Firefox browser, makes this much easier, as it just needs to be downloaded and started; it is preconfigured and no installation is required, which should make it particularly attractive to many users. Therefore, *vulnerabilities of the browser* can present an interesting target and be exploited to deanonymize the users. A famous example is the shutdown of the 'Playpen' Dark Web child pornography website by the FBI in February 2015. The FBI used a so-called 'Network Investigative Technique' (NIT), which was exploiting a non-publicly-known vulnerability of the Mozilla browser to break into suspected visitors' computers and identify their real IP addresses (Cox 2016). Instead of shutting down the website, the FBI continued to run it from a government server for 13 days to collect the IP addresses of potential visitors. In further action, the FBI broke into more than 8700 computers in 120 countries due to a court decision of a single judge. The procedure was heavily criticized. Of the 100,000 people worldwide who visited the site, 8700 were hacked but only 214 were arrested.

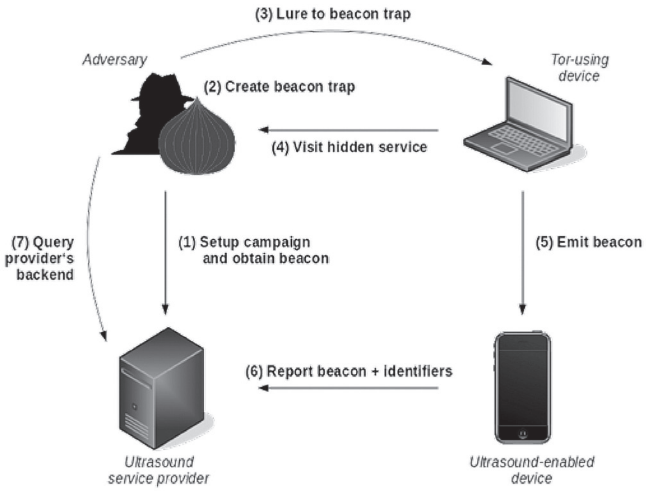
Because of deanonymization attacks like that one, the Tor Project provided a hardened version of the Tor Browser, beginning from November 2015 (Tor Browser 5.5a4-hardened), providing additional hardening against the exploitation of memory corruption bugs and adding debugging features. Anyway, in part because of, *inter alia*, the confusion among users caused by the two series, regular and hardened, the second one was discontinued in April 2017.

CAT 2 The second attack category is not based on technical characteristics of the Tor Browser or the respective protocol, but exploiting *indirect shortcomings, which are not based on technical vulnerabilities*. A prominent example is the use of default configurations: on most distributions, the Apache server ships with a feature called `mod_status` enabled, which provides a website at `/server-status`, containing statistics like resource usage and virtual hosts. For security reasons, this page is by default only reachable from localhost. Yet the Tor demon for onion services *is* running on localhost, which allows connections to the status page from external clients if the configuration is unchanged. Due to this, sensitive information can be leaked; even a .onion search engine was identified as having the module enabled, exposing *all* search queries sent to the page.

Another example highlights the endangerment of the indirect attack vectors included in this category even better: back in 2014, a new advertising technique called 'ultrasound cross-device tracking' (uXDT) was deployed. The idea behind uXDT is embedding unique sound codes, inaudible to humans, into advertisements. The inaudible sounds are replayed when the ad is presented to a user. Unknown and unrecognizable to the user, the sound pattern may be noticed by another device nearby. Software supporting uXDT is listening for such patterns; if it recognizes one, it sends it back to a central server - together with information about the device. The central server knows the pattern as it was created in a unique way, and therefore knows the targets to which it was sent. In this way, it is possible to identify and merge multiple devices owned by a user, optimizing ad campaigns to all their devices, even if they were never involved in an action like searching for a specific product, resulting in a purposive ad.

Even worse, this technique can be used for deanonymization attacks on Tor users as well (Mavroudis 2017). If someone enters the Dark Web, they will quickly recognize there are a lot of ads, for example embedded in well-known search pages and even in popular marketplaces. Using the default configuration of the Tor Browser, these ads are presented to the user. Therefore, if someone opens a web page which presents an ad with an embedded uXDT sound, there is the risk that a device nearby, maybe a smartphone, another computer or even one of the numerous IoT gadgets which are now so popular, is listening. By applying the same technique, sending back such a unique beacon trap to a central server, the attacker can directly merge the anonymized access to the regular, public connection, and easily deanonymize the user. Figure 2 illustrates the attack scheme.

FIGURE 2. ULTRASOUND TRACKING BASED ATTACK SCHEME TO DEANONYMIZE TOR USERS. VISUALIZATION BASED ON (MAVROUDIS 2017).


(3) Lure to beacon trap
Adversary
Tor-using device
(2) Create beacon trap
(4) Visit hidden service
(7) Query provider's backend
(1) Setup campaign and obtain beacon
(5) Emit beacon
(6) Report beacon + identifiers
Ultrasound service provider
Ultrasound-enabled device

These two examples highlight the wide range of opportunities through which Tor users can be deanonymized if they are not extremely careful when using the network. CAT 3 In fact, user mistakes and human behaviour are the most common reason for deanonymization. A prominent example is the shutdown of the Dark Web marketplace, "Silk Road", which specialized in drug trafficking and was one of the first of its kind on the Dark Web. The creator, Ross Ulbricht, who used the pseudonym "Dread Pirate Roberts", revealed himself by several momentous mistakes. First, he used the pseudonym "altoid" to announce and promote his marketplace in early January 2011. In October of the same year, the same pseudonym was used for a post on a Bitcoin talk, and his email address was included as a contact opportunity for interested users: rossulbricht@gmail.com. This was discovered by the authorities, enabling them to trace Ulbrich back, eventually resulting in his imprisonment.[5] Blake Benthall failed to heed this; he was arrested in November 2014 for establishing and running the Silk Road 2 marketplace, after the first one was closed. Benthall could be identified because he registered the server where the anonymous website was running with his email address, blake@benthall.net; the same category of mistake as that of Ulbricht.

Another example was an online drug dealer, caught in 2017 because he was conspicuous at the post office. To avoid fingerprints, he always delivered the postal packages wearing latex gloves at the counter. However, this eventually caught the attention of the postal employees, so they informed the police. When the dealer was

[5] Also, there was a report that Ulbricht ordered several fake IDs to rent the required servers for the Silk Road website. The fake IDs were sent from Canada to the US, and found at the border as part of a routine mail search. The packet contained nine fake IDs - each with a different name, but all of them with the same photo: a *real* photo of Ulbricht. As the packet was even addressed *directly* to Ulbricht, that was another low-hanging fruit for the officers. However, the careless handling of the pseudonym 'altoid' seems to have been the root cause of the identification.

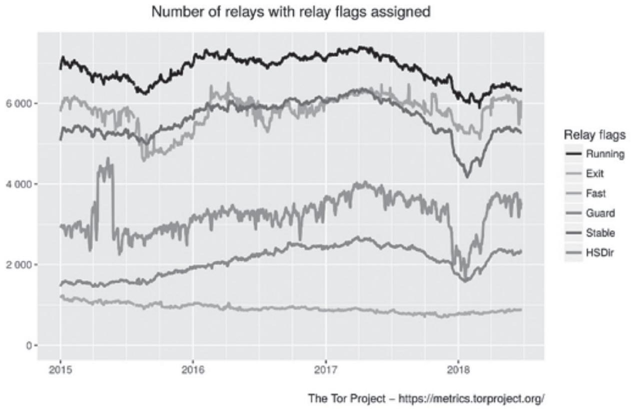
arrested, further traces on his mobile phone linked him to an entry on a Reddit website about drug dealers on the Dark Web. He had not deleted the history.

As an interim conclusion, it can be stated that the protection afforded by the Dark Web for criminal activities can be quickly lost through numerous possibilities of deanonymization. This can involve particularly careless behaviour by users, but can also be originated by attacks on the software or the protocols.

Since the need for secure anonymization can be anticipated when dealing with information relevant to military operations on the Dark Web, a closer look should be taken at the functional principles and their weaknesses. In particular, the question arises whether the Dark Web offers sufficient protection when used cautiously.

Traffic Analysis and its Relevance. To answer that question, a closer look at the working scheme of the overlay network, and the resulting possibilities of deanonymization without an exploitation of protocol and programming vulnerabilities, should be taken. As such an analysis would go beyond the scope and technical depth of this article, only a few key findings are outlined as follows. Tor is the largest, most widely used anonymization network; yet it has the problem that the number of relays in the network is relatively limited and barely growing. In some cases, it can even be observed that the number of relays involved is decreasing – which may also be due to legal reasons.[6] However, there are also relatively spontaneous, very large changes in the number of specific relays – often a sign that an attack on the Tor network is being attempted again, or that research institutions or other bodies are trying new analyses. Figure 3 shows the development of the number of relays since January 2015, as provided by The Tor Project (Tor 2018). In particular, the Exit relays stagnated for years and only increased again recently.

FIGURE 3. DEVELOPMENT OF THE NUMBER OF RELAYS SINCE JANUARY 2015 (TOR 2018).

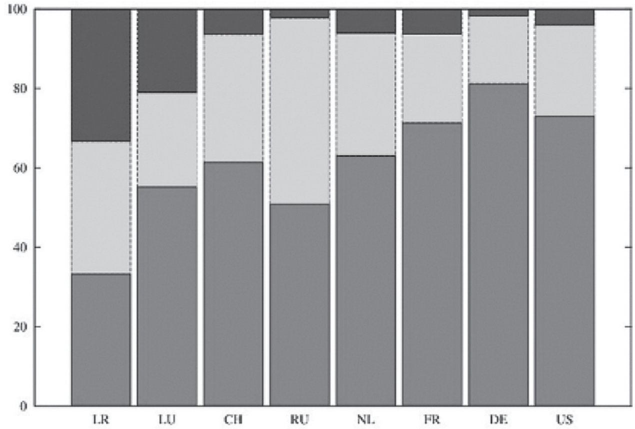


[6] For example, because of violations of copyright infringement when the Exit Nodes are misused (Ferner 2017).

Some areas of the curves are striking: a sudden, rapid increase in the number of Hidden Service Directory (HSDir) relays can be observed from mid-April 2015 until the end of May 2015. On the other hand, a sudden drop of HSDir and Stable relays can be identified in December 2017: this was affected by a DDoS attack on the Tor network. Multiple servers went down because of the attack; the HSDir relays were badly affected, because if such a system goes down, it does not get back the HSDir server flag immediately after rebooting, but takes 96 hours. The loss of HSDir relays also affected the reachability of onion services (Goulet 2017). Other strong jumps in the number of relays may also be related to, e.g., C&C infrastructure ran over the Tor network or bots.

As we can see, only a small number of relays are providing the core functionality of the Tor network, and the chances are high that they include quite a number of malicious ones. Moreover, not only is the number of Exit relays already quite low, but the way they are selected by the underlying algorithms reduces the actually used relays significantly. Figure 4 provides an example of the actual Exit relay use per country relative to available Exit relays based on a three-week observation (Koch 2016). Each bar shows the ratio of available Tor relays (red) to relays configured as Exit relays (green) to selected Exit relays (blue). Nearly a quarter of all nodes were located in the US, but Tor selected only 5.53 per cent of these (blue section of US bar). Likewise, 8.53 per cent of all exit nodes were located in Germany (green section of DE bar), but Tor selected only 2.22 per cent of these (blue section of DE bar).

FIGURE 4. RATIO OF AVAILABLE TOR RELAYS TO EXIT RELAYS TO SELECTED/USED EXIT RELAYS. THE SMALL SHARE OF ACTUALLY USED EXIT RELAYS SIMPLIFIES TRAFFIC ANALYSIS ATTACKS (KOCH 2016).



It can be seen that only a small fraction of the available Exit relays is selected and used. This simplifies attacks that analyse traffic flows through the Tor network, as the number of relays to be monitored drops sharply. But not only Exit relays are endangered. With respect to onion services, malicious HSDir relays can be used to identify new onion services on the Dark Web. For example, more than 100 snooping HSDir relays were identified on the Tor network (Noubir 2016) – a technique typically used by companies providing Dark Web intelligence, or by federal agencies.

These intense activities of various actors, which aim at the analysis of actions up to the deanonymization of Dark Web users, should be kept in mind ahead of the further discussion.

3. DARK WEB MARKETS AND DATA

Based on the knowledge of the function, opportunities and weaknesses of anonymising networks, an analysis of Dark Web marketplaces and their trading is performed, before specifically looking into the trading of sensitive information.

A. Data Economy and Marketplaces

Of course, a central aspect of the question whether the Dark Web is a growing risk for military operations involves the nature, extent and quality of information which can be found there. While crawling the Dark Web can be challenging, e.g., finding new websites or entering closed marketplaces, DARPA's Memex program sought to develop software to advance search capabilities, especially with regard to the *Deep* Web, and a series of tools was made public (DARPA 2014). Some studies tried to shed some light by analysing onion services in the Dark Web provided by Tor. e.g., 39,824 hidden service descriptors were analysed on 4 February 2013 (Biryukov 2014). After scanning the hosts, 3,050 HTTP services were identified, and the content classified. Only hidden services offered in the English language had been analysed: 2,618 services in total. From these pages, 805 showed a default page and no actual content; 44 per cent of the identified topics were devoted to drugs, adult content, counterfeit, and weapons, while 56 per cent were devoted to topics like politics.

Another study identified a share of 57 per cent in services with illicit content (Moore 2016). The used categories are shown in Table 1.

TABLE 1. CATEGORIES AND ACCESS NUMBERS OF CONTENT IN THE DARK WEB (MOORE ET AL. 2016).

Category	Category
None	2,482
Other	1,021
Drugs	423
Finance	327
Other illicit	198
Unknown	155
Extremism	140
Illegitimate pornography	122
Nexus	118
Hacking	96
Social	64
Arms	42
Violence	17
Total	5,205
Total active	2,723
Total illicit	1,547

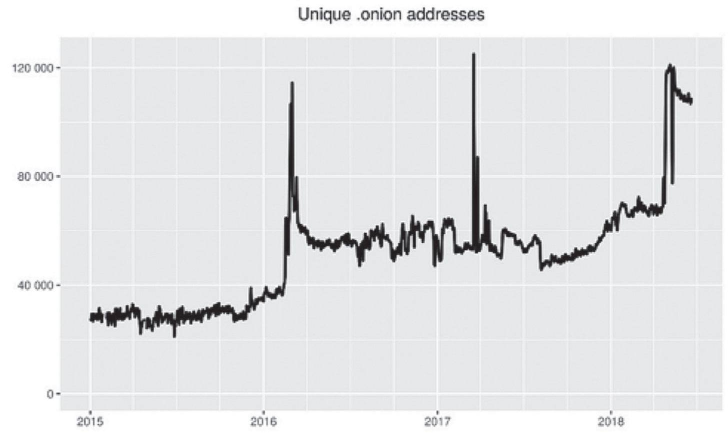
Repeatedly, it is argued that most parts of Tor traffic are illicit; the rough numbers seem to confirm this. A study presented by the University of Portsmouth even highlighted that 80 per cent of traffic to Tor hidden services is related to child pornography. While these are shocking results at first glance, a closer look at the underlying data reveals that the corresponding values are highly uncertain and only marginally justify such statements: based on the nature of the Dark Web, respective measurements can typically only be made indirectly. Regularly, requests to (malicious, therefore especially set up for the measurement task) hidden service directories will be counted. The respective numbers are often used to derive relative numbers of users; but they say more about the behavioural differences of different types of users (Mathewson 2014). Another important and often unnoticed aspect is that child protection agencies also regularly crawl the Dark Web for websites containing illicit pornography. Law enforcement agencies do so too. Therefore, it is interesting to look at the evaluations of these agencies to get a better idea of the actual situation. The results presented in the recent reports of the Internet Watch Foundation (IWF) are highlighted in Table 2.

TABLE 2. URLS CONFIRMED CONTAINING CHILD SEXUAL ABUSE IMAGERY AS SEEN BY THE IWF (IWF 2015, IWF 2016, IWF 2017).

Year	URLs to Child Porn	Hidden Service Proportion	Proportion of the Dark Web
2015	68092	79	0.116
2016	57335	41	0.071
2017	78589	44	0.056

We can see that the number of identified hidden services related with child pornography is small in contrast to the actual identified links to websites with child pornography in the surface or Deep Web. The IWF highlights that 'hidden services commonly contain hundreds or even thousands of links to child sexual abuse imagery that is hosted on image hosts and cyberlockers on the open web' (IWF 2017). This must be combined with the fact that the Dark Web is very small and growing only very slowly. Figure 5 shows the number of unique .onion addresses between January 2015 and June 2018. There is only a slow increase in the number of onion services; and the numbers are often quite constant over longer periods of time, sometimes even declining. Very fast, large increases are typically indicative of an experiment or attack and do not represent a sudden increase in the number of available pages. It should also be noted that nowhere near all pages have content; many only present the default page of the web server, such as that already shown in the above-referenced analyses.

FIGURE 5. NUMBER OF UNIQUE .ONION ADDRESSES FOR SERVICE VERSION 2 FROM 1 JANUARY 2015 TO 23 JUNE 2018.



The Tor Project – https://metrics.torproject.org/

Independent from the number of onion services marketplaces, but very important, is the trading volume. Some calculations have been made of the sales volume of the ecosystem, including several famous and heavily used marketplaces like Silk Road, Black Market Reloaded and Silk Road 2.0 (Soska 2015). The trading volume was higher than previously thought, and is also subject to strong fluctuations. However, the total volume does not experience exponential growth. The study identified that "in the short four years since the development of the original Silk Road, total volumes have reached up to $650,000 daily (averaged over 30-day windows) and are generally stable around $300,000-$500,000 a day, far exceeding what had been previously reported" (Soska 2015).

It is important to keep these dependencies in mind, as it is the base from which to focus on the most significant aspects. Looking at leaked data, most occurrences are on the clear internet – and while there may be trades of the data on the Dark Web, the result normally provides a link to a page in the surface or Deep Web, where it can be found and downloaded; but normally, it is not hosted on the Dark Web. Paste services like pastebin are popular for that.

We can conclude that the growth and therefore, the evolution of the importance of and danger posed by, the Dark Web is often over-estimated. In particular, the sometimes assumed exponential growth of the Dark Web cannot be demonstrated by any measurable numbers: neither the number of onion services and Dark Web marketplaces, nor the traffic itself, nor the trading volume.

B. Trading Sensitive Data

Looking at the most important trading categories of the Dark Web marketplaces: drugs, counterfeit and adult, most of them are not really able to affect military operations.

Some companies are offering Dark Web intelligence, highlighting the footprints of companies on the Dark Web, based on data they find. For example, OWL Cybersecurity published a so-called 'Darknet [sic.] Index' which aims to measure how the availability of breached data affects the overall cybersecurity of a company (OWL 2017). For this purpose, OWL Cybersecurity has set up a database, which is "automatically and continuously updated with between 10 to 15 million pages per day, from more than 24,000 domains on the Tor network alone, as well as other darknet networks" (OWL 2017). It highlighted that every company in the 2017 Fortune 500 is exposed on the darknet [sic.]; the companies with the largest footprint are shown in Table 3.

TABLE 3. TOP 10 ENTRIES OF THE 'DARKNET [SIC.] INDEX' FOR THE FORTUNE 500 COMPANIES PRESENTED BY OWL CYBERSECURITY (OWL 2017).

DARKINT Rank	Company Name	Darknet [sic.] Index Score
1	Amazon.com	19.16
2	Alphabet (Google)	17.21
3	Apple	15.98
4	Facebook	14.99
5	eBay	14.55
6	American Express	13.33
7	Frontier Communications	13.29
8	Netflix	13.19
9	Texas Instruments	12.99
10	FedEx	12.58

OWL Cybersecurity presented additional evaluations focusing on specific sectors, e.g., for IT companies. Moreover, based on the Fortune 500 evaluation, it analysed the US government to compare the results with the commercial sector. Key points of their conclusions are that the "U.S. Government scored worse than expected as compared to the largest U.S. companies. The U.S. Government averaged 1.6 points higher than the average Fortune 500 company, meaning that the government has a comparably larger amount of darknet exposure" (OWL 2017). The analysis identified that the US Navy has the most extensive footprint of all government agencies examined, and that

> military and defense groups overall are the largest target, closely followed by Cabinet agencies. A target's attractiveness stems from the desirability of its protected information. Whether personal or proprietary, it would appear that the groups more closely linked to defense have data that cyber criminals find attractive (OWL 2017).

To what extent these footprints represent a real threat to the company in question is not easy to estimate. That the footprints of state organizations are very large is fundamental here. Table 4 presents the force numbers by service branch for 2016, as published by the DoD in December 2017.

TABLE 4. FORCE NUMBERS BY SERVICE BRANCH AND RESERVE COMPONENT FOR 2016. SOURCE: DOD, DECEMBER 2017.

Branch	Employees
Army Active Duty	471,271
Army National Guard	344,862
Navy Active Duty	320,101
Air Force Active Duty	313,723
Army Reserve	306,272
Marine Corps Active Duty	183,501
Navy Reserve	108,864
Marine Corps Reserve	106,581
Air National Guard	105,887
Air Force Reserve	104,520
Coast Guard Active Duty	39,597
Coast Guard Reserve	8,123
Sum	2,413,302
Active	1,778,942

In addition to these numbers, associated authorities, civilian employees, etc. must be added to the reflected attack surface. In comparison, only Walmart has 2.2 million employees, far more than any other Fortune 500 company. Next are McDonald's (420,000), IBM (412,000) and Kroger (400,000), while the average number of employees at the Fortune 500 companies is about 50,000. Given this, leaks with elements affecting one or another employee of the governmental sector are likely and possibly adding to the footprint. Therefore, there may not be a *direct* risk for a company; but of course, there is always the risk of social engineering attacks.

With respect to the data available on the Dark Web, it can be assumed that an evaluation of the importance or possible impact is usually very difficult. While extensive reputation systems have been established in the area of illicit drug trafficking or trading in stolen credit card numbers, this is not so easy for the trade in leaked data. Typically, the data will often come from different sources and sellers will be unknown. Here, we can look at other areas of the Dark Web struggling with similar 'problems': the arms trade and hitman services. There are multiple Dark Web websites offering these services. Yet such is the nature of the Dark Web, many scams can be found: since a buyer of illegal weapons or the client to a murder can hardly go to the police after they have paid, but have not received what they were promised, scammers can earn easy money here.

A prominent example is the 'Besa Mafia' website. While the page was very well set up and many discussions focused on the question of whether it was real or not, eventually it was shown to be a scam. The scammers had been able to collect money from different potential customers, but never executed an assassination (Jeffries 2017). Also, according to federal investigators, Ross Ulbricht ordered six murders over the Dark Web; but five never happened, and the sixth turned into an indictment because the supposed hitman was actually a federal agent working undercover (Jeffries 2017).

The same applies for the illegal arms trade. While it is possible to buy a weapon on the Dark Web, it is actually quite difficult, as the case of the Munich shooting rampage has shown. A study analysed the role of the Dark Web in facilitating trade in firearms, ammunition and explosives (RAND 2017). After collecting one week of data during September 2016,[7] it was systematically analysed and discussed in workshops and interviews. RAND concluded that the Dark Web is an enabler of the circulation of illegal weapons but also highlighted the limitations of the study, especially "the impossibility to determine with certainty the nature of a vendor (scammer, law enforcement or real vendor)" (RAND 2017). Some verified examples like the Munich case are mentioned, but the number is very small. Moreover, in terms of the weapons trade, the activities of scammers and undercover cops supersede real offers by far. For example, Agora stopped selling guns altogether when it was the largest market on the Dark Web, because of "scamming by dishonest vendors" (Cox 2015). Of course, the trade in 3d-printing plans is much easier to do and can lead to increasing proliferation.

Taking a look again at data that may have an impact on military operations, direct and indirect effects have to be differentiated. For example, trading in mission plans or classified reports and evaluations, as well as access credentials to systems or services, can generate a direct impact; while personally identifiable information (PII) can generate an indirect impact.

However, based on the available reports and experiences, it can be assumed that trading data like mission plans and classified reports is not easy and not very likely on the Dark Web. Sales on the Dark Web are mainly financial data, login access, access to online services and identities including fake IDs like passports (Ablon 2014, McFarland 2015, Ray 2017). The same applies for the governmental sector: PII is the most compromised record type, counting 57.4 per cent of available data from breaches in the governmental sector (Huq 2015).

Although evidence can be provided about the authenticity of the data – for example, the provision of individual screenshots or excerpts from documents – due to its peculiarity (as opposed to the dumping of credit card numbers, etc.), the sale will be much more difficult, and will attract undercover agents. Rather, it can be assumed

[7] 19-25 September 2016.

that such data is traded outside of the Dark Web, in the traditional way. For instance, while access to some SCADA systems was offered for sale on the Dark Web in 2015 (Aharoni 2015), three years later, this is still a rare case and not yet a new trend. More likely is trade in credentials or PII as part of leaks, which may not even be directly affecting the military, but indirectly affects its personnel. Another indirect impact may also be generated by more inconspicuous services available on the Dark Web: namely, the proliferation of attack tools regarding knowledge, which then can be used to implement and execute attacks on military communication systems:

- Weaknesses, 0-days, 1-days
- Exploit code
- Malware frameworks
- Ransomware as a Service (RaaS), Crime as a Service (CaaS)
- Botnet access/rent for the execution of DDoS attacks
- Jamming devices

These categories may pose a special, indirect danger for military operations. While this is no direct trade in mission-critical information, specially crafted malware used in social engineering campaigns, or the offer to hack social media accounts can be a starting point to access a mission-critical environment. There are regular data leaks available; and hence, a lot of PII with which to identify potential targets: with numerous servicemen and women possibly affected, too. For example, the xDedic marketplace is offering easy access to legitimate organizational servers; different advertisements for hacking email or social media accounts can be found (Paganini 2017).

Based on this broad background – the technical functioning of Tor and the possibilities of user deanonymization, the activities which can be observed in Dark Web marketplaces and a realistic estimate of their importance compared to the surface and Deep Web – the actual risk to military operations from the Dark Web can now be discussed.

4. DISCUSSION

Several studies have been published highlighting the apparently predominantly illegal use and content of the Dark Web; but this only holds true at first glance. The actual numbers show that criminal activities committed on the Dark Web are only a very tiny portion, while a vast amount happens on the surface web and the Deep Web. In fact, the Dark Web page provided by Facebook at facebookcorewwwi.onion to allow users in countries with surveillance and repression to access the service is the most widely used site on the (Tor) Dark Web.

Dark Web marketplaces can have several hundred thousand dollars in sales per day, but the focus of trade is drugs and financial fraud, while a lot of PII is traded, too, which can be the enabler for social engineering and targeted attacks. Even more, CaaS with offers like hacking social media accounts are services which we must consider. Accordingly, a threat to military operations may result if social media or system accounts of soldiers are hacked in order to gain access to a target system. The trade in PII from data leaks can additionally support this. Nevertheless, the process is time-consuming and long, opening various options for detection and early warning.

The greatest threat seems to arise if PII is not made available for sale but publicly available. Automatically monitoring the relevant forums and pages is relatively easy for a tech-savvy user to do, so data deployed there can be used very quickly for (especially) social engineering attacks, often before those affected have heard of the original leaks. For example, the recent so-called Germany-Leaks, including details of German lawmakers up to Angela Merkel, were distributed by a hacker with the pseudonym 'Orbit' in December 2018, with subsequent comprehensive media coverage in the beginning of January 2019 (Times 2019). The original links are no longer available, but the material and alternative links still can be found quickly on corresponding websites on the Dark Web. In this context, it should also be mentioned that on the same forum where this data and other leaks were provided, no military-related record or post could be identified.

In addition to the requirement to first find respective leaked data, the question is also whether a targeted attack against a *particular* mission will be feasible – or whether 'only' an endangerment of a 'random' mission may arise. Moreover, the past few years have repeatedly shown police operations in which Dark Web marketplaces were shut down and those responsible were held to account. Studies on the Dark Web also continue to regularly show that a high proportion of the nodes involved are run by governmental agencies, research laboratories and universities; and numerous monitoring measures are implemented. For example, there are also fingerprints for the website and distribution 'TAILS' in the xkeyscore monitoring program of the NSA: if an attacker succeeds in manipulating the Tor Browser or a relevant distribution during the download – for example, inserting a backdoor – anonymity can be broken from the beginning. The numerous incidents and attacks which are known about, and extensive research on the topic of deanonymization, all make it questionable if someone is willing to sell sensitive data which is important for the success (or failure) of a military operation on the Dark Web – and equally, whether another party is willing to buy it there.

On the other hand, it should also be noted that the security of onion services will increase significantly in the near future – and thus the effort to deanonymize the services or their users will become more challenging. This is due to the recent introduction of onion services version 3. Currently, companies providing tracking and intelligence services for the Dark Web benefit highly from design weaknesses of long-used hidden services of version 2. For example, placing malicious HSDirs is a very popular and heavily used technique to identify new services in the (hidden service v2) Dark Web. Recently, researchers found more than 100 of these malicious HSDirs, reflecting the intense activities of companies providing Dark Web intelligence, as well as researchers and public authorities. With the availability of the new onion service v3, the exploited design shortcomings of the predecessor are fixed. Several design decisions and measures guarantee much better protection of users than before, and thus a much higher degree of anonymity. This is realized, among other things, by the following properties (Tor 2013, Tor 2017):

- Use of stronger cryptographic building blocks: SHA3/Ed25519/Curve25519 instead of SHA1/DH/RSA1024 in version 2
- Improved directory protocol with less metadata leaked to directory servers
- New pseudo random variables to prevent predictable Tor uses
- Better onion address security against impersonation: new addresses with 56 characters instead of 16 characters in version 2
- A cleaner, more modular code base

Therefore, tracking opportunities for the companies mentioned above decrease significantly, while attacks on services are more challenging. With the new name space of the services and the protocol adaptions, finding new, as yet unknown pages on the Dark Web will become much more difficult. This could again lead to much greater use of the Dark Web for criminal activities, but the question is: what kind of activities?

When talking about data which can pose a risk to military operations, there are two scenarios: a ‘random’ hack or a ‘targeted’ hack. If a hacker obtains the data more or less by chance, they will also offer it more visibly in order to make money; available contacts to interested parties are not to be expected here. This increases the likelihood of detecting traces of sensitive data, even on the Dark Web, in a timely manner. However, in the case of a targeted attack, possibly even controlled by a state, the interested party is clear; and a particularly visible offer is unnecessary and unlikely.

In the case that mission-critical information is available on the Dark Web, another thought must also be taken into consideration: finding and recognizing it may not be enough, or may be too late with respect to a current mission. While early detection of a new set of credit card numbers available for sale on the Dark Web can be used to disable and exchange the affected cards, protecting customers from financial damage even before the data can be exploited, this can be much more difficult with respect to an ongoing operation. Therefore, another approach can be beneficial too: the deliberate introduction and monitoring of honeydata: consciously placed, realistic looking records.

Based on these considerations, a comprehensive data management strategy must include the following elements:

1. Continuously tracking the surface and Deep Web as well as the Dark Web for the appearance of new leaked and stolen data. This requires the creation of fingerprints (hashes) for sensitive files, which then can be used to search for leaked data on the surface web as well as the Deep and Dark Web. Here, services like PwnedList can be integrated too.
2. The implementation of honeydata to increase detection probabilities.
3. The preparation (and testing!) of action plans and guidelines for fast, accurate handling of detected data leaks, including procedures to initiate the deletion of data from typically used platforms like pastebin.

Another aspect involves using the Tor network in essentially the way it was invented for – to hide the communication and identity of agents. Offensive actions may be executed by using anonymization networks like Tor; but as the analysis has shown, it is quite easy to monitor the Exit Nodes and very easy to blacklist them. Therefore, monitoring the IPs of the Exit Nodes can be used for an early warning if someone is willing to execute an attack over the Tor network.

Summing up these arguments, we can conclude that the new, more secure anonymous onion services will certainly lead to an increase in the popularity of illegal exchanges, but sensitive data important for military operations will still not be the focal point. More dangerous is the overall trade of data from breaches and leaks, which *may* contain details connected to the military; and in the broader sense, to military operations. For example, data records from dating agencies or sports applications may be assigned to soldiers, which can make them targets for social engineering, blackmailing or just make them (and therefore, their unit) trackable. While such information can be an element in a much broader mission to eventually influence a military operation, the risk factor is significantly lower than in the case of directly trading data on such operations. For the military, this means that a threat intelligence capability, monitoring potential risks associated with data breaches, is increasingly important. The main focus remains on the surface and especially the Deep Web; but monitoring the Dark Web is also beneficial.

5. CONCLUSION

Anonymization networks like Tor can be used to hide someone's identity or trade illegal goods on the Dark Web. Numerous data-related incidents and the trade of the corresponding records represent an increasing challenge. The availability of specially crafted malicious software or CaaS over the Dark Web can also generate new risk potential.

On the one hand, a closer look at the Dark Web, its technical base and the available data identifies no direct endangerment of armed forces capabilities. Scammers, law enforcement and surveillance opportunities do not make the Dark Web a reliable vector for sophisticated attackers. Therefore, monitoring the Dark Web does not play a superior role; the main activities, which can pose a risk for military operations, take place on the surface and the Deep Web. On the other hand, due to the multitude of available PII, which can also affect servicemen and women when being used for, say, social engineering campaigns, timely detection of sensitive information is of particular importance. While such data cannot be routinely targeted against an operation or military capability, it can open access to somewhere in the system and thus be the beginning of a longer attack path. Accordingly, it is important to monitor all parts of the web continuously through a holistic strategy, and develop and regularly practise emergency plans for rapid response to recognized data loss.

REFERENCES

Ablon, L., Libicki, M., and Golay, A. 2014. *Markets for Cybercrime Tools and Stolen Information: Hackers' Bazaar*. Tech Rep. RAND Corporation.

Aharoni, I. 2015. SCADA Systems Offered for Sale in the Underground Economy. Available online: http://www.infosecisland.com/blogview/24608-SCADA-Systems-Offered-for-Sale-in-the-Underground-Economy.html accessed on 30. June 2018.

Biddle, P., England, P., Peinado, M. and Willman, B. 2002. The Darknet and the Future of Content Protection. In *ACM Workshop on Digital Rights Management*, Springer-Verlag Berlin Heidelberg; pp. 155-176, ISBN 3-540-40410-4.

Biryukov, A., Pustogarov, I., Thill, F., and Weinmann, R. 2014. Content and Popularity Analysis of Tor Hidden Services. In *Distributed Computing Systems Workshops*, IEEE 34th International Conference; pp. 188-193.

Chaum, D. 1981. Untraceable Electronic Mail, Return Addresses, and Digital Pseudonyms. In *Communications of the ACM*, vol. 24 no. 2.

Cox, J. 2015. Scams and Undercover Cops Are Denting the Dark Web Gun Trade. Available online: https://motherboard.vice.com/en_us/article/wnx88q/scams-and-undercover-cops-are-denting-the-dark-web-gun-trade accessed on 30. June 2018.

Cox, J. 2016. Confirmed: Carnegie Mellon University Attacked Tor, Was Subpoenaed by Feds. *Motherboard*.

Cox, J. 2016. The FBI Used a 'Non-Public' Vulnerability to Hack Suspects on Tor. *Motherboard*.

DARPA. 2014. Memex Tools and Components. Available online: https://github.com/darpa-i2o/memex-program-index accessed on 9 March 2019.

Dingledine, R. 2002. *Pre-Alpha: Run an Onion Proxy Now!* SEUL Project Archives.

Dingledine, R. and Mathewson, N. and Syverson, P. 2004. *Tor: The Second-Generation Onion Router*. Naval Research Lab Washington DC.

Dingledine, R. 2014. Tor Security Advisory: "Relay Early" Traffic Confirmation Attack. Tor Blog.

Eddy, M. 2019. Hackers Leak Details of German Lawmakers, Except Those on Far Right. *The New York Times*.

Ferner, J. 2017. Betreiber eines TOR-Exit-Nodes kann für Urheberrechtsverletzungen haften. Available online: https://www.ferner-alsdorf.de/urheberrecht__betreiber-eines-tor-exit-nodes-kann-fuer-urheberrechtsverletzungen-haften__rechtsanwalt-alsdorf__56543 accessed on 09. March 2019.

Goulet, D. 2017. *Ongoing DDoS on The Network - Status*. The Tor Project.

Hug, N. 2015. *Follow the Data: Analysing Breaches by Industry. Trend Micro Analysis of Privacy Rights*. Clearinghouse.

Hargreaves, S. 2015. IWF Annual Report 2015. Internet Watch Foundation.

Hargreaves, S. 2016. IWF Annual Report 2016. Internet Watch Foundation.

Hargreaves, S. 2017. IWF Annual Report 2017. Internet Watch Foundation.

Janssen, D. and Janssen, C. 2018. *Deep Web*. Techopedia 2018.

Jeffries, A. 2017. People Keep Falling for this Murder-for-Hire Dark Web scam. Available online: https://theoutline.com/post/932/people-keep-falling-for-this-murder-for-hire-dark-web-scam accessed on 30. June 2018.

Kadianakis, G. and Loesing, K. 2015. *Extrapolating Network Totals from Hidden-Service Statistics*. The Tor Project.

Koch, R., Golling, M. and Dreo, G. 2016. How Anonymous is the Tor Network? A Long-Term Black-Box Investigation. *IEEE Computer* no. 3; pp. 42-49.

Lacey, D. and Salmon, Paul M. 2015. It's Dark in There: Using Systems Analysis to Investigate Trust and Engagement in Dark Web Forums. In *Engineering Psychology and Cognitive Ergonomics*, Springer International Publishing, Cham, Switzerland; pp. 117-128, ISBN 978-3-319-20372-0.

Mathewson, N. 2014. Some Thoughts on Hidden Services. Tor Blog.

Mavroudis, V., Hao, S., Fratantonio, Y., Maggi, F., Kruegel, C. and Vigna, G. 2017. On the Privacy and Security of the Ultrasound Ecosystem. In *Proceedings on Privacy Enhancing Technologies*; De Gruyter Open; pp. 95-112.

McCandless, D. 2018. World's Biggest Data Breaches. Available online: http://www.informationisbeautiful.net/visualizations/worlds-biggest-data-breaches-hacks/ accessed on 30. June 2018.

McFarland, C., Paget, F. and Samani, R. 2015. *The Hidden Data Economy - The Marketplace for Stolen Digital Information*. McAfee LLC.

Moore, D. and Rid, T. 2016. Cryptopolitik and the Darknet. In *Survival* Vol. 58 Nr. 1 2016, Taylor & Francis; pp. 7-38.

Noubir, G. and Sanatinia, A. 2016. *Honey Onions: Exposing Snooping Tor Hsdir Relays*. DEFCON 24.

OWL Cybersecurity. 2017. The OWL Cybersecurity Darknet Index: Reranking the Fortune 500 using Darknet Intelligence (DARKINT). OWL Cybersecurity, Denver, Colorado.

OWL Cybersecurity. 2017. DARKOWL Press Room. Available online: https://www.darkowl.com/news/ accessed on 30. June 2018.

Paganini, P. 2017. Digging into the Darkweb. CTI - EU Cyber Threat Intelligence ENISA.

Persi Paoli, G., Aldridge, J., Ryan, N., and Warnes, R. 2017. *Behind the Curtain*. RAND Corporation.

Ray, V. 2017. Exploring the Cybercrime Underground: Part 4 - Darknet Markets. Available online: https://researchcenter.paloaltonetworks.com/tag/cybercrime-underground/ accessed on 30. June 2018.

Reed, M., Syverson, P. and Goldschlag, D. 1998. Anonymous Connections and Onion Routing. *IEEE Journal on Selected Areas in Communications* vol. 16 no. 4 1998; pp. 482-494.

Soska, K. and Christin, N. 2015. Measuring the Longitudinal Evolution of the Online Anonymous Marketplace Ecosystem. *USENIX Security Symposium*; pp. 33-48.

The Tor Project. 2013. 224-rend-spec-ng.

The Tor Project. 2015. Tor Project: FAQ. Available online: https://www.torproject.org/docs/faq.html.en\#ChangePaths accessed on 9 March 2019.

The Tor Project. 2017. Tor 0.3.2.2-alpha is Released.

The Tor Project. 2018. *Welcome to Tor Metrics!*

Zhang, X. 2003. *System/Application Designs, Optimization and Implementations on Overlay Networks*. High Performance Computing and Software Lab; Ohio State University.