CHAPTER 1

Understanding the Dark Web

Dimitrios Kavallieros, Dimitrios Myttas, Emmanouil Kermitsis, Euthimios Lissaris, Georgios Giataganas, and Eleni Darra

1.1 Introduction

Dimitris Avramopoulos, European Commissioner for Migration, Home Affairs and Citizenship, said:

> The Dark Web is growing into a haven of rampant criminality. This is a threat to our societies and our economies that we can only face together, on a global scale…

D. Kavallieros (✉)
Center for Security Studies-KEMEA, Athens, Greece

University of Peloponnese-Department of Informatics and Telecommunications, Tripoli, Greece
e-mail: d.kavallieros@kemea-research.gr

D. Myttas • E. Kermitsis • E. Lissaris • G. Giataganas • E. Darra
Center for Security Studies-KEMEA, Athens, Greece
e-mail: d.myttas@kemea-research.gr; e.kermitsis@kemea-research.gr; e.lissaris@kemea-research.gr; g.giataganas@kemea-research.gr; e.darra@kemea-research.gr

B. Akhgar et al. (eds.), *Dark Web Investigation*,
Security Informatics and Law Enforcement,
https://doi.org/10.1007/978-3-030-55343-2_1

Having in mind a number of similar statements around the world and before diving into the Dark Web, it is essential to make an introduction to the principal terms of the "digital world" as it was presented for the first time in the 1960s. Initially, it must be outlined that even though many people use the interchangeable terms Internet and World Wide Web (web), the two terms are not synonymous. The Internet and the web are two separate, but of course, related things.

The Internet is a massive network of networks, and it connects millions of computers together globally, forming a superset in which any computer can communicate with any other computer as long as they are both connected to the Internet. Information that travels over the Internet does so via a variety of "languages" known as protocols (Internet protocol, IP) and through satellite, telephone lines and optical cables forming the global electronic community. The Internet has no centralised governance in either technological implementation or policies for access and usage; each constituent network sets its own policies (Strickland 2014). On the contrary, the World Wide Web, or simply web, is a way of accessing information over the medium of the Internet. It is an information-sharing model that is built on top of the Internet. The web uses the HTTP protocol, only one of the languages spoken over the Internet, to transmit data. Web services, which use HTTP to allow applications to communicate in order to exchange business logic, use the web to share information, consisting of HTML text, images, audio, video and other forms of media. The English scientist Tim Berners-Lee invented the World Wide Web in 1989, as he wrote the first web browser computer program in 1990, and has been employed at CERN in Switzerland. The web browser was released outside CERN in 1991, first to other research institutions starting in January 1991 and to the general public on the Internet in August 1991.

Having in mind that the two terms are not synonymous and should not be confused, easily one can say that the web is just a portion of the Internet, although a large portion of it. Content on the World Wide Web can be broken down into two basic categories: structured and unstructured, while the web consists of several layers of accessibility. The first layer is called the clear web or surface web. Surface web is the portion of the web that is readily available to the general public and searchable with standard web search engines. This part is accessible through regular search engines and is where social media platforms reside also. The surface web has been part of the World Wide Web since the first browser was invented,

connecting users with websites that can be discovered through a regular Internet browser (e.g. Edge, Mozilla, Opera, etc.) using any of the main search engines (Google, Yahoo, etc.). This is what you use when you read the news; buy something, e.g. on Amazon; or visit any of your usual daily websites and is also the area of the web that is under constant surveillance by governments across the world. Surface web is made up of static and fixed pages. Static pages do not depend on a database for their content. They reside on a server waiting to be retrieved and are basically html files whose content never changes. Thus, any reference to surface web will be referring to common websites, that is, sites whose domains end in .com, .org, .net or similar variations and whose content does not require any special configuration to access.

On the other hand, the Deep Web was also part of the web at its conception, and in basic terms, it is the opposite of surface, because its search engines cannot find its content. This is the key difference between the two in real data terms. Sites on the surface Internet are indexed for search engines to find, but the Deep Web is not indexed. However, both are also accessible by the public; they just require different methods to access them – usually a specific password encrypted browser or a set of log-in details. A common image used to represent the meaning of surface versus Deep Web is that of an iceberg: the visible portion of the iceberg represents a very small part of the whole (the whole in this case being the whole of the Internet, surface and Deep Web) as Fig. 1.1 depicts.

The Deep Web contains all of our medical records, financial records, social media files and plenty other important information we want and need to keep secure. It is this need to keep secure files that gave rise to the need to keep a portion of the web secured away from being "googled" at the impulse of anybody at any time.

It is estimated that the Deep Web contains about 102,000 unstructured databases and 348,000 structured databases. In other words, there is a ratio of 3.4 structured data sources for every one (1) unstructured source. Figure 1.1 is the result of a sample of Deep Web databases conducted by Bin et al. (2007) (Fig. 1.2).

Finally, the Dark Web is part of the Deep Web, but it has one major difference. It is not possible to get to the Dark Web using a regular web browser. A special browser is needed, specifically designed for the task such as Tor or similar browser technology. These browsers work differently than conventional browsers, and they are by far the best and most

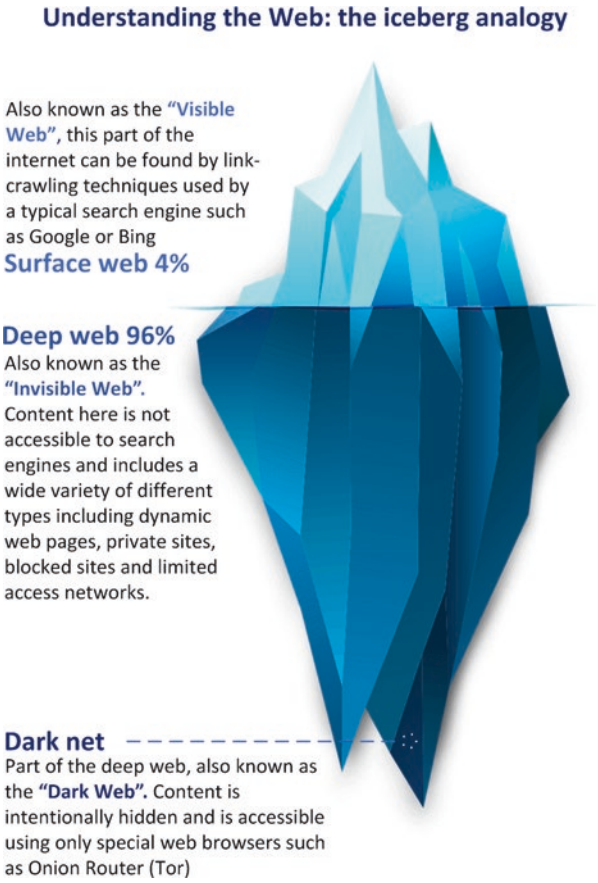

Fig. 1.1 A visual aid in understanding the web (EMCDDA–Europol 2016)

popular (with an estimated 2.5 million daily users). Named "The Onion Router", it was quickly coined with the shorter "Tor" with its name coming from application layer encryption within communication protocol stacks as many layers representing the layers of an onion. With Tor, you'll be able to reach not only the Dark Web but the even smaller subsection known as the Tor network.

	Sampling Results	Total Estimate	99% Confidence Interval
Deep Web sites	126	307.000	236.000 – 377.000
Web databases	190	450.000	366.000 – 535.000
- unstructured	43	102.00	62.000 – 142.000
- structured	147	348.000	275.000 – 423.000
Query interfaces	406	1.258.000	1.097.000 – 1.419.000

Fig. 1.2 Deep Web databases (Bin et al. 2007)

The technology to create the Dark Web was initially created by the US government in the mid-1990s to allow spies and intelligence agencies to anonymously send and receive messages. So easily one can realise that its anonymous nature makes it a good place for all kinds of things people wouldn't dare do on the surface web.

As of 2015, the term "the Dark Web" is often used interchangeably with the Deep Web due to the quantity of hidden services on the darknets. The term is often inaccurately used interchangeably with the Deep Web due to Tor's history as a platform that could not be search-indexed. Mixing uses of both these terms have been described as inaccurate (Bright Planet 2014). The darknet(s) as its name depicts recalls images of shadowy alleys, malicious, hard-faced individuals and socially damaging activities, covering a range from political protestors – rebels – to drug dealers, to terrorist and gun dealers, to paedophiles and everything in between.

Darknets are used for several legitimate purposes: to avoid identity theft, for marketing tracking, to circumvent censorship and to perform research on topics that might be sensitive in certain countries. Chapter 2 will describe both the legitimate and criminal stakeholders of the darknets, as well as their motives behind the use of this side of the web.

Finally, trying to present a definition, the one by Sherman and Price (2007) will be used (as cited in Lievrouw): "[The Dark Web is compromised by] websites that are outdated, broken, abandoned, or inaccessible using standard web browsing techniques". Specifically, the description of "inaccessible" is adequate for our understanding. As we will realise later, many of the sites on the Dark Web strive to be private or at least only accessible to those who know what they're looking for.

Having until now presented synoptically the three web layers, it is possible to present an easy view of the differences among them in Fig. 1.3:

Surface Web	Deep Web	Dark Web
Accessible	Accessible by password, encryption, or through gateway software	Restricted to special browsers
Indexed for Search Engines	Not indexed for Search Engines	Not indexed for Search Engines
Little illegal activity	Little illegal activity outside of Dark Web	Large scale illegal activity
Relatively small in size	Huge in size and growing exponentially	Unmeasurable due to nature

Fig. 1.3 Main web layer differences

1.2 Infrastructure of the Dark Web

As described in the previous section, the Dark Web is intentionally hidden from the general public/simple users. The Dark Web consists of overlay networks, known as darknets, which offer various hidden services. These networks can only be accessed using specific software, such as Tor and I2P (Brown 2016). In this section we will describe the technical infrastructure of the Dark Web through the analysis of the best known darknets and the tools used to access them (Hawkins 2016).

None of the aforementioned software (Tor and I2P) was created to provide safe passage and dissemination of illegal markets and products in neither surface nor Deep Web. Nevertheless, the anonymity and data encryption provided by using software like Tor made them as a tool used by people wanting to sell their illegal services/products (hacking for hire, 0-day vulnerabilities, guns, drugs, etc.), furthermore, for extremist and terrorists who want to disseminate their opinion (propaganda) and proselytise people as well as for people sharing illegal videos and images (e.g. child abuse material (CAM)).

1.2.1 The Tor Project (Tor): Overview

The Onion Routing Project or simply Tor is employing the third generation of the onion routing technique in order to provide anonymous surfing and communication. The onion routing technique (first generation) was developed in the mid-90s by the US Naval Research Laboratory and the Defense Advanced Research Project Agency (DARPA) to provide safe communication between operatives in the field and intelligence gathering (Tor website) by anonymising TCP-based applications. In 2002, the Naval

Laboratory released the source code of Tor, and the Electronic Frontier Foundation (EFF) undertook Tor's founding, becoming the cornerstone for the creation of the Tor Project organisation responsible for maintaining, upgrading and shaping Tor network as it is nowadays (Syverson 2015; Immonen 2016; Çalışkan et al. 2015).

Tor became the most famous tool, in terms of anonymity and privacy to access and publish material on the Dark Web among other tools (e.g. I2P), and thus Tor darknet is the most famous and holds the highest number of visitors and services (Jardine 2015). At the time of writing this chapter and based on The Tor Project (2019b), between 1,500,000 and 3,000,000 relay users (only the ones directly connected to the Tor network) existed, from 3000 up to 7000 relay nodes and more than 1000 bridge nodes, between 60,000/per day and 80,000/per day unique .onion addresses (only version 2) and around 200 Gbits/s bandwidth consumption, while the relays can support approximately 400 Gbits/s overall bandwidth consumption.

As it was described before, services on the Dark Web are intentionally hidden, and thus the .onion sites do not have the formatting used in the clearnet, www.example.com. The .onion top-level domain (TLD) is specifically used to access hidden services hosted only on the Tor network, and it is not part of the Internet DNS root. Furthermore, the addresses under the .onion consist of 16 alphanumeric characters. A user needs to either know the exact address of a hidden service or to use a search engine specifically designed to work on the .onion. A few examples of such sites are below:

1.2.1.1 Search Engines and Introductory Points

Torch is a Tor search engine and can be accessed through http://xmh57jrzrnw6insl.onion/

Not Evil is another search engine designed for the Tor network and can be accessed through http://hss3uro2hsxfogfq.onion/

The Hidden Wiki is mainly used as a directory of .onion links. The links are categorised based on the service they offer like financial services, drugs, email/messaging and P2P file sharing among others. Nevertheless, one category that it is not included in the site is .onion links regarding terrorism and child sexual abuse material.

Chapter 4 will provide further in-depth details regarding services and markets of the Dark Web as well as the respective links and descriptions, while Chap. 3 will describe the activities of terrorist organisations in the Dark Web and the clearnet and how terrorists are moving from one part of the Internet to the other based on their goals, objectives and activities.

1.2.2 Tor Architecture and Routing

Data transmitted using the onion routing is encapsulated in multiple layers of encryption, resembling the layers of an onion (see Fig. 1.2). The number of layers is equal to the number of users acting as nodes, also known as relays, each time. This technique assists the user to remain anonymous and evade eavesdropping and traffic analysis techniques which could reveal the origin, destination and content of a message (The Tor Project 2019c). Tor users have to decide whether they will participate in the network as nodes or not; thus, it is in volunteer bases. As the first generation of onion routing technique, Tor is anonymising TCP packets, while offering significant improvements in comparison with the first generation such as *perfect forward secrecy*, *separation of "protocol cleaning" from anonymity*, *many TCP streams can share one circuit*, *leaky-pipe circuit topology*, *congestion control*, *directory servers*, *variable exit policies*, *end-to-end integrity checking*, *rendezvous points and hidden cervices* (Dingledine et al. 2004) (Fig. 1.4).

Tor consists mainly of (i) the Tor browser which offers the appropriate setting (e.g. proxy) in order to connect to the Tor network and (ii) the hidden services/sites hosted in the Tor network. Furthermore, Tor network consists of the Tor nodes and the directory servers, while the nodes share part of their bandwidth meaning that it will increase or decrease based on the number of the nodes; thus the higher the number of nodes in Tor, the faster it will be (The Tor Project 2019e).

1.2.2.1 Tor Nodes

Tor nodes, or relays, are created by users offering their computer(s), purely on a volunteer basis, in order to be used as a node. It is highly important

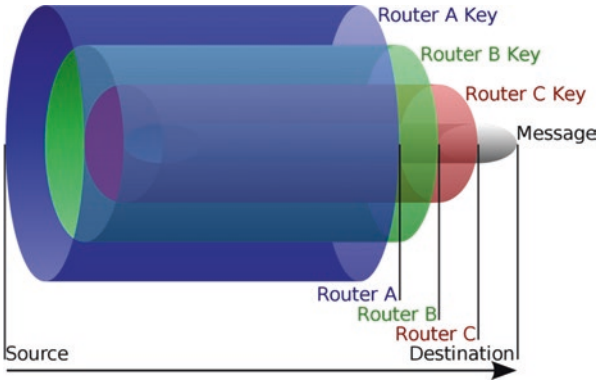

Fig. 1.4 Onion routing technique (Neal 2008)

to explain that the higher number of nodes means higher available bandwidth, increased robustness of the network against attacks and making it more difficult to analyse the traffic. Tor consists of three types of nodes based on (Erkkonen et al. 2007; Aschmann et al. 2017; The Tor Project a, b, c, d, e; Tor Challenge 2018):

- The guard/entry node: the guard node is the first node each user will hop to in order to connect to the Tor network and to the requested service/site. The selection of the guard nodes is done at the user level, and the selection is random in order to minimise the chances of "eavesdropping".
- Middle or internal node: middle nodes are the nodes that exist between the guard node and the exit node.
- Exit nodes: exit nodes are the last nodes before a user reaches the requested destination, and thus this type of node is responsible for sending the request either out of the Tor network or to a hidden service.
- Bridge node: the main difference with the aforementioned nodes is that bridges are not listed in the main Tor directory authority. Thus, it will be difficult for ISPs to block Tor traffic passing through these bridges. At the moment of writing this section, Tor has more than 1500 IPv4 bridges and around 250 IPv6 bridges (The Tor Project 2019d).

To run nodes might have legal impact, especially for the ones running exit node, as it can be used to identify the respective IP address if the exit node is used for illegal purposes. Furthermore, it is advisable not to run exit node using a home computer while it is best to notify the ISP (The Tor Project 2018a).

1.2.2.2 Tor Directory Authorities

The Tor directory authorities, which are ten (see Fig. 1.5) at the time of writing this chapter, are databases which contain information and the list of all the active nodes at the network. Thus, they have complete knowledge and view of the network's topology. Information relevant to the routers stored in the director authorities is encrypted with digital signatures. To further secure both the directory authorities as well as the entire Tor network, the administrator of each server will process and approve information regarding nodes in order to be published to users (Erkkonen et al. 2007).

1.2.2.3 Tor Circuit

In order for a user to connect to the Tor network, the user's Tor client (Tor browser) will have to communicate with a Tor directory authority, holding a list of all the available Tor nodes. Once the user receives this list,

Nickname†	Advertised Bandwidth	Uptime	Country	IPv4	IPv6
dizum (2)	3.13 MiB/s	28d 1h		45.66.33.45	-
Serge (1)	1.45 MiB/s	6d 2h		66.111.2.131	2610:1c0:0:5::131
moria1 (1)	500 KiB/s	11d 20h		128.31.0.34	-
tor26 (1)	75 KiB/s	9d 6h		86.59.21.38	2001:858:2:2:aabb:0:563b:1526
bastet (1)	50 KiB/s	2d 3h		204.13.164.118	2620:13:4000:6000::1000:118
maatuska (8)	50 KiB/s	20d 15h		171.25.193.9	2001:67c:289c::9
dannenberg (1)	40 KiB/s	22d 6h		193.23.244.244	2001:678:558:1000::244
Faravahar (1)	40 KiB/s	8d 22h		154.35.175.225	2607:8500:154::3
gabelmoo (1)	40 KiB/s	26d 7h		131.188.40.189	2001:638:a000:4140::ffff:189
longclaw (1)	38 KiB/s	9d 21m		199.58.81.140	-
Total	**5.4 MiB/s**				

Fig. 1.5 Tor directory authorities (29/11/209) (Tor Metrics 2019)

the client will randomly decide the path of nodes that will be used in order to reach the destination server, which in turn publishes the hidden service the person is looking for (see Fig. 1.6) (Aschmann et al. 2017; The Tor Project 2018a, b). Each node has knowledge only for the previous and next node (one hop knowledge) of the network, exchanging a different encryption key each time. This method ensures that even if one node is compromised it will not be able to identify the entire path of the Tor network. In order for a route, or circuit, to be formed, the Tor browser will download the current list of register nodes, from the directory authorities, and it will randomly select a guard node. Then, it will select the rest of the nodes based on their bandwidth and stability (highest bandwidth and only stable nodes are selected). By default, when three nodes are selected, the circuit is formed, and the generation of encryption keys follows.

1.2.2.4 Setting Up Tor

Tor is available for Windows, Apple MacOS and GNU/Linux, and it can be downloaded from https://www.torproject.org/projects/torbrowser.html.en in sixteen (16) languages:

- English
- Arabic
- Deutsch

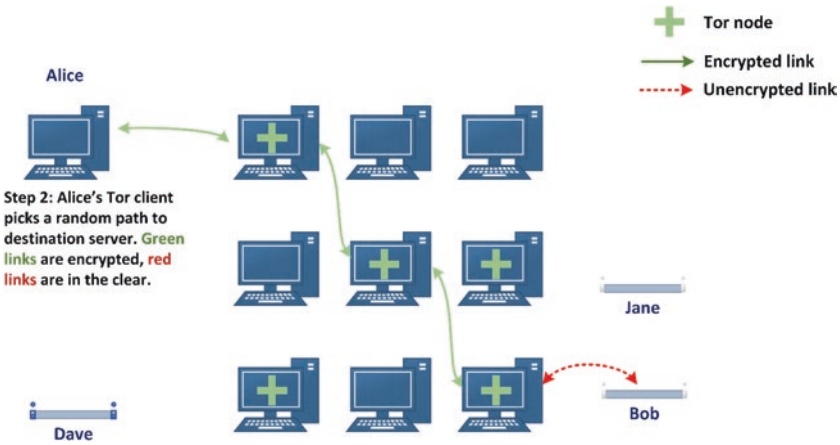

Fig. 1.6 Tor network-circuit setup (The Tor Project 2019a)

- Spanish
- Farsi
- French
- Italian
- Japanese
- Korean
- Dutch
- Polish
- Portuguese
- Russian
- Turkish
- Vietnamese
- Chinese

To install Tor in Windows machines is a straight forward procedure. The user has to download the appropriate file from the list, save the file and then open it. Choose "Run", choose your language of preference, and press install. When the installation is complete, press finish, run the Tor browser, and click "Connect". To install Tor in MacOS, the user needs to download the respective file, save it and drag the *.dmg* file into the application folder.

To install Tor in Linux/GNU, the user needs to first download the architecture file and then run the following commands from the terminal: tar-xvJf tor-browser-linux32-7.5.3_*LANG*.tar.xz (for 32-bit OS) or tar-xvJf tor-browser-linux64-7.5.3_*LANG*.tar.xz (for 32-bit OS). The next step is to navigate to the Tor browser using the following command: cd tor-browser_*LANG*, and run the Tor browser either from the graphical interface (click it) or by executing the following command: ./start-tor-browser.desktop from the terminal.

Finally, Tor browser can also be installed in android-based machines such as smartphones and tablets from Google Play.

1.2.3 The Invisible Internet Project (I2P): Overview

In this section we are describing the I2P network and its infrastructure. I2P is a decentralised, peer-to-peer overlay network that started in 2003, offering anonymity by employing the *garlic routing/encryption* techniques (Erkkonen et al. 2007). The design of the network is message based in order to run on top of IP, but communication can also be achieved on top of TCP and UDP based on the requirements of each application/service.

The I2P client software can act as a router once it is installed in a machine, providing connectivity to I2P websites (TLD, .i2p) in the darknet, or it can host a service (e.g. an .i2p website).

The garlic routing technique, which is a variant of onion routing, was coined back in 2000 and in the framework of I2P provides the following three attributes (I2P Garlic Routing 2018):

- Tunnel building and routing (in order to transmit data, each router creates one-way tunnels (inbound and outbound tunnels)).
- Data bundling to be able to evaluate the end-to-end message delivery status.
- ElGamal/EAS + SessionTags encryption algorithms are used to provide end-to-end encryption and minimise the possibility of traffic analysis attacks.

I2P can be downloaded from geti2p.net/en/download, and it is available for Windows, Mac OSX, GNU/Linux, BSD, Solaris, Debian, Ubuntu and Android.

1.2.4 I2P Network Database

In order for a client to connect to other clients and setup a circuit, it will have to ask the I2P netDB which contains all mandatory information regarding other user's inbound tunnels. The I2P netDB basically contains two important types of records, the RouterInfos, which is the contact information of the I2P routers (IP address, respective port and public key), and the LeaseSets, which contains the destination contact information (tunnel endpoints and the public key of the requested service). Furthermore, tunnels expire every 10 minutes; thus clients have to request the aforementioned information from the netDB, if they want to stay connected with the service (Egger et al. 2013; I2P, 2018b).

1.2.5 I2P Routers and Tunnels

I2P routers use two pairs of one-way tunnels in total from which the one handles the inbound traffic and the other the outbound traffic as Fig. 1.7 depicts. Thus, for one message and the respective reply, the router will build four tunnels each time. To clarify how the I2P tunnels work, we first need to understand the philosophy of how the inbound and outbound tunnels are built. The creator of a tunnel decides the number of the peers

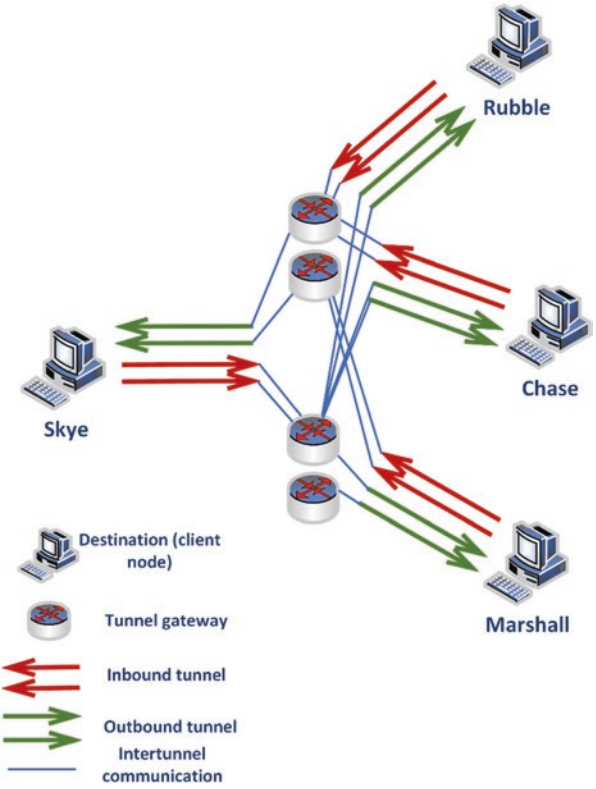

Fig. 1.7 I2P tunnels

(number of hops) and which peers will participate in the tunnel in order to strengthen the security of the tunnel and minimise any chances of either third parties or other tunnel participants to identify the total number of hops the tunnel has and if they belong in the same tunnel (Erkkonen et al. 2007; Egger et al. 2013; I2P 2018a).

1.2.6 Freenet: Overview

Freenet developed in 1999 and released in 2000 is most likely the third ranking darknet after Tor and I2P in terms of the number of users (Brown 2016). It is a peer-to-peer network designed for sharing, storing and retrieving files as well as publishing Freenet websites, called "freesites"

providing high anonymity to the users. Furthermore, it is not a centralised platform; thus it has stronger resistance in attacks. From the release of 0.7.5 version, the architecture of the Freenet has changed, strengthening user's privacy and enhancing the security of the Freenet nodes against malicious attacks – see Freenet's architecture and design – (Clarke et al. 2001, 2010). As it is a decentralised network, Freenet relies on its users to store, insert, edit and request files anonymously. To achieve that, it is "mandatory" for all Freenet users to contribute in terms of hard drive space (portion of their own hard drive) and their bandwidth. These files can be anonymous or pseudonymous static websites, forums, microblogs and regular files. The five main goals of the design followed by Freenet designers based on Clarke et al. (2001) are:

1. Anonymity for both producers and consumers of information
2. Deniability for storers of information
3. Resistance to attempts by third parties to deny access to information
4. Efficient dynamic storage and routing of information
5. Decentralisation of all network functions

Freenet is free for everyone and can be downloaded from freenetproject.org/pages/download.html and is available for Windows, GNU/Linux, Mac OSX and Posix.

1.2.6.1 Freenet's Architecture and Design

As discussed in the previous section, Freenet is a peer-to-peer overlay network consisted from as many nodes as its users. Each node provides a part of their hard drive and bandwidth for storing, retrieving and editing Freenet files. These files are stored after they are divided into encrypted blocks, distributed among multiple nodes, while the holders of the files are not familiar with the content of the files (Aschmann et al. 2017). The stored files are associated with a key (or address) which is based on a string, given by the user. This key has two purposes, to locate where in the Freenet network the file is stored and to authenticate this file.

Up until the release of 0.7.5 version, Freenet was designed to choose the edges and the nodes of the network to be used based on the best optimisation route. Version 0.7.5 introduced the darknet mode which allows the nodes and edges to connect only to nodes the user trusts (friends list), with whom they have previously exchanged public keys as Fig. 1.8 depicts. Thus, the new architecture offers two choices to the users, either to use

Opennet mode	Darknet Mode
Easy to block	Very hard to block
Limited anonymity	Good anonymity
Somewhat centralised	Fully decentralised

Fig. 1.8 Darknet vs opennet mode

the darknet mode and stay hidden by connecting only to trusted nodes (creation of private network) or to also connect to nodes operated by strangers (opennet mode) (Aschmann et al. 2017; Clarke et al. 2010). Figure 1.8 summarises the key differences between Freenet's darknet and opennet mode (Freenet Project, 2018b):

Freenet nodes are designed to cache as many of the files they transfer as possible; thus the node's storage is getting fuller easier and faster. To this extent, Freenet is using a best-effort algorithm in order to randomly remove files from a node when its storage is full (Fig. 1.9).

1.2.6.2 Sharing, Requesting and Accessing Data

As it was described in the previous section, each file (or part of it) is associated with a key. All nodes have a routing table depicting the addresses and their respective keys. Thus, if a user is "looking" for a file, he has to send a request based on the key of the file. To locate the file, the request hops from node to node (creating a route) based on which neighbour node is the most desirable, from those available. When this is not feasible, the request goes back to the previous node in order to restart the route as depicted in Fig. 1.10. When the file is located within a node, then the file is backtracked to the previous node which forwards it back to the node that initially requested the file. This process is terminated either because the file has been found or because the nodes gave up looking for the file (Clarke et al. 2010).

To share and store a file, a user will have to "upload" the file, and then it will be cut down to smaller parts in order to be distributed among the nodes, following the same logic with the one followed when a user is requesting a file. First, the file will be associated with a key, and then it will create a route, and when it is terminated, the file will be forwarded to the node that terminated the route. The user that uploaded the file does not have to stay online as the file has been stored in other nodes on the network (Balduzzi and Ciancaglini 2015).

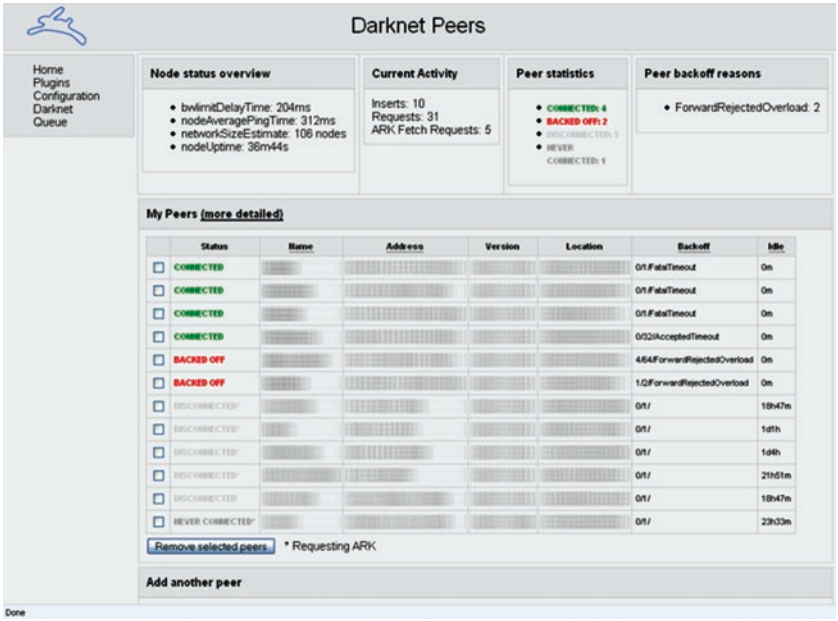

Fig. 1.9 Freenet darknet mode peers (Freenet Project 2018a)

A user can search and access documents or freesites by either using a Freenet search engine such as Freegle or by typing the key associated with the data/freesites in the following format, http://localhost:888/[Freenet Key].

For example, if you have installed Freenet to your pc, you can insert the following key, http://127.0.0.1:8888/freenet:USK@tiYrPDh~fDeH5V7NZjpp~QuubaHwgks88iwlRXXLLWA,yboLMwX1dChz8fWKjmbdtl38HR5uiCOdIUT86ohUyRg,AQACAAE/nerdageddon/-49/ in order to download the Nerdageddon freesite which includes most of the freesites on Freenet, excluding most of the pornographic freesites. Furthermore, freesites do not contain dynamic content (e.g. scripts and databases), and they are constructed in HTML.

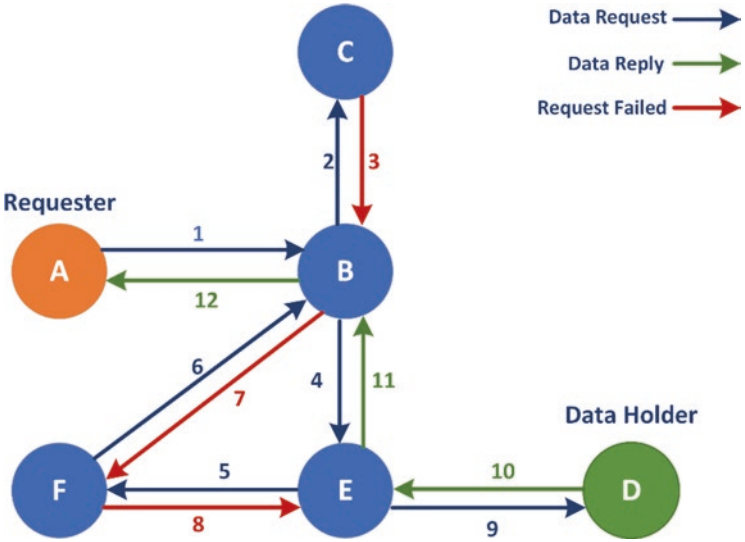

Fig. 1.10 Data request (Balduzzi and Ciancaglini 2015)

1.3 Law Enforcement Agency Investigating the Dark Web

The above-mentioned definitions guide us to the area where the exact problems are encountered, especially in the Deep and Dark Web during a law enforcement agency investigation. The use LEA's make of the Dark Web itself for investigations is usually not specified in the openly available literature. It is nevertheless generally known that the activities on the Dark Web are an object of investigation, within two operational contexts, as listed below:

- Proactive investigation for intelligence, employing the use of Tor and often manual searches of the encountered content; in a few cases, LEAs are experimenting the use of novel automatic tools to crawl portions of the Dark Web and index their content.
- Reactive investigations, for example, to collect information on the Dark Web on a suspect or organisation. These operations are also challenging, mainly because of the difficulty to de-anonymise digital traces on the Dark Web.

The basic conclusion that derives from all the above is that an anonymity-granting system such as Tor, along with other similar technologies, is just a tool. The endurance and impact of operations concerning the Dark Web are often mentioned in the media and in LEA's official reports, so in the following paragraphs, we will indicatively focus on "how policing the clear web influences the Dark Web" and on the "policing dilemma the Dark Web poses".

1.3.1 How Policing the Clear Web Influences the Dark Web

Online policing is also as beneficial as offline policing. The anonymity of Tor does not necessarily slow down law enforcement efforts. There are limits to the effectiveness of online policing. One limitation is that online criminals can be global, even while most law enforcement agencies (Interpol not included) are local. Another limitation is that cybercrime is rapidly increasing, which threatens to overwhelm any and all available policing capacity of nations.

Policy actions might assist LEAs to overcome the aforementioned limitations, but before considering them, it is important to highlight that policing actions on the surface web influence the Dark Web. An example is given by the Internet content regulation from a drug-policy perspective: measures such as the Australian compulsory Internet filtering regime to block drug contents on clear web websites would likely drive drug discussions to the Deep or the Dark Web, where digital spaces are not affected by Internet filtering and where governments are unable to regulate Tor website content (Barratt et al. 2013). On the other side, this measure might also push violent online extremism into the Dark Web, where monitoring of content is much more difficult and less debate takes place (Hussain and Saltman 2014).

1.3.2 The Dark Web Also Poses a "Policing Dilemma"

Based on Jardine (2015), anonymity can be the shield for people doing "good" and for people that without it surfing the web could be impossible. As it was previously discussed, Tor, I2P, Freenet, etc. are just technologies, tools designed neither of ill nor good use. Thus, it is the person and the reason behind them who are responsible for the deeds.

Based on the available literature, there are few examples regarding the difference between Tor technology and the Tor hidden services, with the

latter based on Guitton (2013) promoting unethical and illegal content and strongly believing that a stop should be put on the development of Tor hidden services. Other sources (Jardine 2015) argue that shutting down anonymity networks will damage greatly those people that need this technology and use it within legal boundaries. Furthermore, based on Stevens and Jardine (Stevens 2009; Jardine 2015), actions and methods should be employed to raise awareness regarding technologies granting anonymity and how people around the world benefit from them. After all what matters is not what the technology is, but how it is used and what the net effect turns out to be. Based on Brink et al. (2016):

> When it comes to the Dark Web similar attitudes appear across law enforcements worldwide: authorities tend to focus on attacking the offender. The effectiveness of this approach is questioned in the academic literature. Two points are also highlighted: the difference between various cybercrimes perpetuated through the Dark Web might call for different type of measures; moreover, measures developed to exercise control on phenomena on the Clear Web might bring weak if not counterproductive effects if applied to the Dark Web, as already mentioned.

LEAs face a great challenge in the fight against criminal activities on Tor hidden services (THS). As previously explained, Tor is designed so that no single entity in the circuit, including the ISPs, has knowledge of the complete circuit, meaning that it is highly difficult to monitor the website a user is visiting and the user's behaviour in general (Griffiths 2013). However, Dark Web anonymity can be used by LEAs in order to maintain anonymous communication with their sources and undercover agents and to conduct online sting operations and surveillances. Furthermore, LEAs were able to achieve significant wins against cybercriminals in the Dark Web (Nath and Kriechbaumer 2015; Brink et al. 2016):

- Malicious software attacks in August 2013, computers of people accessing THS hosted by servers of the company Freedom Host were infected by malicious software.
- Silk Road shutdown in October 2013, the FBI arrested the alleged operator of the illegal marketplace Silk Road, which was operated as a THS.

- Operation Onymous about a month after Silk Road was taken offline, the THS Silk Road 2.0 was launched to continue the criminal market activities.

There hasn't been any publicly available (formal) reports discussing and explaining the strategy, methods and tools used to achieve the aforementioned shake downs and apprehension of the cybercriminals. Two methods that may have been employed are:

- Exploitation of user mistakes
- Exploitation of Tor technical limitations and security issues

It is also worth mentioning the operations of other actors, who autonomously infiltrate, disrupt and eventually take down Dark Web websites and services (Weimann 2015; Brink et al. 2016):

- The self-named "Operation Darknet" by Anonymous. In October 2011 they released in the public around 1500 user credentials from a CAM website in the Dark Web.
- The self-named "Operation Paris", by Anonymous, took down hundreds of ISIS websites on the surface web. ISIS had to move to the Dark Web after this operation, and al-Hayat Media Center posted a link and guidelines on how to reach their new site in the Dark Web.

1.4 Concluding Remarks

The exponential growth of sites offering illegal goods (e.g. guns, drugs, etc.), services (e.g. hacking as a service), closed forums and sites relevant to radicalisation, extremism and terrorism combined with the anonymity the darknets provide poses continuing challenges to the LEAs, since Dark Web sites proliferate at a rate far greater than LEAs have been able to intervene. Also, current tools seem to not be fully adequate to counter activities taking place in the Dark Web, while cybercriminal's quiver regarding tools, techniques and methodologies is evolving continuously. The right of anonymity every person has alongside with the jurisdiction challenges posed by the nature of the Internet multiplies the level of difficulty for those responsible to safeguard society. Finally, it might be difficult to justify the cost and the person months of operations aimed at

regulating and monitoring the Dark Web, while at the same time, LEAs have to also fight against other forms of cybercrime.

In response, LEAs are constantly looking for technological tools to assist the identification and analysis of illegal material on the Dark Web, as well as the proactive identification of radicals, extremists and terrorists, while at the same time respect the anonymity of legitimate users. The TENSOR platform designed to meet the aforementioned objectives consists of a number of tools that enable the proactive identification, collection and analysis of terrorist-related content.

References

M. Aschmann, L. Leenen, J. Jansen van Vuuren, *The Utilisation of the Deep Web for Military Counter Terrorist Operations*. (Academic Conferences and publishing limited, s.l., 2017)

M. Balduzzi, V. Ciancaglini, *Cybercrime in the Deep Web* (Black Hat Europe, Amsterdam, 2015)

M.J. Barratt, S. Lenton, M. Allen, Internet content regulation, public drug websites and the growth in hidden internet services. Drugs **20**(3), 195–202 (2013)

H. Bin, M. Patel, Z. Zhang, Accessing the deep web: A suervey. Commun. ACM **50**, 94–101 (2007)

Bright Planet, *Clearing Up Confusion – Deep Web vs. Dark Web*. [Online] (2014), Available at: https://brightplanet.com/2014/03/clearing-confusion-deep-web-vs-dark-web/

P. Brink et al., *MEDIA4SEC Report on State of the Art Review*. (MEDIA4SEC, s.l., 2016)

B. Brown, *Threat Advisory: 2016 State of the Dark Web*. (AKAMAI Threat Advisory, s.l., 2016)

E. Çalışkan, T. Minárik, A.-M. Osula, *Technical and Legal Overview of the Tor Anonymity Network*. (NATO Cooperative Cyber Defence Centre of Excellence, s.l., 2015)

I. Clarke, O. Sandberg, B. Wiley, T.W. Hong, Freenet: A distributed anonymous information storage and retrieval system, in *Designing Privacy Enhancing Technologies*, ed. by H. Federrath, (Springer, Berlin, Heidelberg, 2001), pp. 46–66

I. Clarke, O. Sandberg, M. Toseland, V. Verendel, *Private Communication Through a Network of Trusted Connections: The Dark Freenet*. (Network, s.l., 2010)

R. Dingledine, N. Mathewson, S. Paul, *Tor: The Second-Generation Onion Router* (Naval Research Lab, Washington, DC, 2004)

C. Egger, J. Schlumberger, C. Kruegel, G. Vigna, in Practical attacks against the I2P network, ed. by A. S. C. V. W, S. J. Stolfo. *Research in Attacks, Intrusions, and Defences* (Springer-Verlag Berlin Heidelberg, Berlin, 2013), pp. 432–451

EMCDDA–Europol, *EU Drug Markets Report: In-Depth Analysis (European Monitoring Centre for Drugs and Drug Addiction and Europol)* (Publications Office of the European Union, Luxembourg, 2016)

H. Erkkonen, J. Larsson, C. Datateknik, *Anonymous Networks.* (Computer communication and distributed systems, s.l., 2007)

Freenet project, *Documentation.* [Online] (2018a), Available at: https://freenetproject.org/pages/documentation.html. Accessed on 26 Mar 2018

Freenet project, *Freenet Help.* [Online] (2018b), Available at: https://freenetproject.org/pages/help.html. Accessed on 26 Feb 2018

M. Griffiths, *Monitoring Internet Communications* (POST – Parliamentary Office of Science and Technology, London, 2013)

C. Guitton, A review of the available content on Tor hidden services: The case against further development. Comput. Hum. Behav. **29**(6), 2805–2815 (2013)

B. Hawkins, *Under The Ocean of the Internet – The Deep Web* (SANS Institute-InfoSec Reading Room, 2016)

G. Hussain, E.M. Saltman, *Jihad Trending: A Comprehensive Analysis of Online Extremism and How to Counter It* (Quilliam, s.l., 2014)

I2P, *i2p Tunnel Implementation.* [Online] (2018a), Available at: https://geti2p.net/en/docs/tunnels/implementation. Accessed on 23 Mar 2018

I2P, *The Network Database.* [Online] (2018b), Available at: https://geti2p.net/en/docs/how/network-database. Accessed on 23 Mar 2018

I2P Garlic Routing, *Garlic Routing and "Garlic" Terminology.* [Online] (2018), Available at: https://geti2p.net/en/docs/how/garlic-routing. Accessed on 22 Feb 2018

V.V. Immonen, *Alice in Onion Land: On Information Security of Tor* (ITA-SUOMEN YLIOPISTO, s.l., 2016)

E. Jardine, The Dark Web Dilemma: Tor, Anonymity and Online Policing. *Global Commission on Internet Governance Paper Series,* Band 21 (2015)

C. Nath, T. Kriechbaumer, *The Darknet and Online Anonymity* (POST – Parliamentary Office of Science and Technology, London, 2015)

H. Neal, *Wikimedia commons.* [Online] (2008), Available at: https://commons.wikimedia.org/wiki/File:Onion_diagram.svg. Accessed on Apr 2018

C. Sherman, G. Price, *The Invisible Web: Uncovering Information Sources Search Engines Can't See* (Information Today, Medford, 2007)

T. Stevens, Regulating the 'Dark web': How a two-fold approach can tackle peer-to-peer radicalisation. RUSI J. **154**(2), 28–33 (2009)

J. Strickland, *Who owns the Internet?* [Online] (2014), Available at: https://computer.howstuffworks.com/internet/basics/who-owns-internet.htm

Syverson P, *Basic Course on Onion Routing* (U.S. Naval Research Laboratory, s.l., 2015)

The Tor Project, *The Legal FAQ for Tor Relay Operators*. [Online] (2018a), Available at: https://www.torproject.org/eff/tor-legal-faq.html.en. Accessed on 20 Nov 2019

The Tor Project, *Tor Project*: *FAQ*. [Online] (2018b), Available at: https://www.torproject.org/docs/faq.html.en#EntryGuards. Accessed on 20 Nov 2019

The Tor Project, *The Solution: A Distributed, Anonymous Network*. [Online] (2019a), Available at: https://www.torproject.org/about/overview.html.en#thesolution. Accessed on 21 Mar 2019

The Tor Project, *Tor Metrics*. [Online] (2019b), Available at: https://metrics.torproject.org/. Accessed on 20 Nov 2019

The Tor Project, *Tor Project*. [Online] (2019c), Available at: https://www.torproject.org/. Accessed on 20 Nov 2019

The Tor Project, *Tor: Bridges*. [Online] (2019d), Available at: https://www.torproject.org/docs/bridges.html.en. Accessed on 21 Mar 2019

The Tor Project, *Tor Relay Guide – Tor Bug Tracker & Wiki*. [Online] (2019e), Available at: https://trac.torproject.org/projects/tor/wiki/TorRelayGuide. Accessed on 22 Feb 2019

Tor Challenge, *What Is a Tor Relay?* [Online] (2018), Available at: https://www.eff.org/torchallenge/what-is-tor.html. Accessed on 21 Mar 2018

Tor Metrics, *Relay Search-Flag: Authority*. [Online] (2019), Available at: https://metrics.torproject.org/rs.html#search/flag:authority. Accessed on 29 Nov 2019

G. Weimann, Going dark: Terrorism on the dark web. Stud. Confl. Terror. **39**(3), 195–206 (2015)