The Mysteries of the Deep Web

David Ko, Matthew Schwartz, Farhan Islam, Andy Ehlert and Bastion Fennell

Abstract

This paper focuses on the history and the forthcoming of the dark entity known as the Deep Web. We discuss different sub-networks of the Deep Web and ways to access different parts of the Deep Web; private networks and software. The Deep Web also has many forms of anonymous communication that we will touch on, ranging from PGP Encryption to TorChat. The Deep Web is not without its fault, there are many dark and illegal activities that we cover such as drug trade and fire arm sales. Security is a big issue since everyone is anonymous there will be people who want to do malicious attacks on other unsuspecting people. Not all of the Deep Web is bad, there are Hactivist groups such as Anonymous that attempt to perform good deeds to the community through online means.

1. Introduction

MATT PUT THE INTRPDUCTION HERE

2. History of the Deep Web

The deep web was first realized when the Internet became so dependent upon search engines that people wanted to delve deeper. The name for the Deep Web originated from the term invisible Web in 1994 which was in reference to the large of amounts of websites that were not registered to any surface search engine. But because search engines only go as far as the people who made them to go, the deep web allows people to scavenge for contents that aren't queried in those search engines. The searchable web (search engines) is barely 1% of what is actually available throughout the whole World Wide Web [6]. People who excelled at extracting information and content from static web pages provided these search results at a surface level for people who are looking for it. There is very valuable information below the surface of what those people find on the surface with static web pages. This data and information is stashed in online databases and dynamic web pages and other forms that require a different way to search for content. There are other methods of casing data and content in the Deep Web such as using unlinked content (which prevents search engines from linking them to the surface), private websites, contextual websites, scripted content (accessible only through links produced by Javascript and content from Web Servers via Flash or Ajax solutions), and non-HTML/text content [7]. Information like this isn't necessarily dangerous or hidden. Such information could be very large data bases such as census data bases, a large collection of patents, weather and climate data and academic databases that have valuable information from research papers and etc (and many more). These databases are usually stored in a table format whereas journals, papers, books, and other research would be saved in a digital format that is either not accessible by a regular browser or behind a protect and/or paid gateway to prevent free access. These files that are behind protection must downloaded or acquired via a different method and then opened on a computer. The size of the Deep Web is practically impossible to measure due to a very good portion of it inaccessible, locked, or hidden making it hard to be accounted for. In 2001, there was an estimated calculation of the deep web's size by the University of California, Berkeley which saw the size to be around 7.5 petabytes. To visualize how big this is, 1 petabyte is 1,000 terabytes and 1 terabyte is 1,000 gigabytes, so 7,500,000 gigabytes of data with much more unaccounted for on hidden networks [6]. This was back in 2001, and now it is 2015. Imagine the size now with the rate of data expansion in the last decade.

3. Deep Web Networks

Besides these databases and dynamic web pages, there are also a domain of the internet (popularly considered the deepest and darkest corner of the Internet) that is the onion domain that is a subsection of the Deep Web. Onion websites can only be accessible through Tor, which is an a free software and an open network that helps defends against intruders and protects against network surveillance that would threaten personal (or national) freedom and privacy and other confidential information and activity. In addition, onion websites are called onion sites because of the fact that the extension is .onion rather than the traditional .com. In this domain of the internet, anonymity is of the upmost importance. These onion sites are sites that want to remain hidden (websites like these would be illegal porn, drug trade, political dissidents, terrorists, and etc). You can think of Tor as its own network separate from the rest of the web. Websites that circulate through Tor are inaccessible to search engines. This is because they do not have any real domain (like a .com, .org, or .net) nor do they have an IP address. To access these sites, you need the afore mentioned Tor software.

3.1 Tor

Though Tor covers a wide array of illegal activities, not all things on the Tor network is bad. A lot of the illegal activity on the deep web occurs with Bitcoin transactions. Since bitcoins are virtually untraceable and have no way of being linked to any owners, it is the perfect anonymous form of currency. One of the most popular transactions of bitcoins would be through the notorious Silk Road on the deep web. Silk Road was and probably is still the biggest online drug trade market in the world. Being able to provide virtually any product to anywhere in the world; just picture Silk Road as the illegal drug trade version of Amazon or Ebay. Besides drug trade, there is also hit men for hire, people who steal for money, human experiments, illegal fire arm trade, and other dark activities. Because Tor allows complete anonymity, it is a safe haven for people such as journalists, protesters and other organizations that need to be protected from the government that is trying to track their movements and activities. There are lists and databases of accessible onion websites that people can look through to find what they are looking for. There are also search engines that are restrictedly for the deep web and for Tor called ”Deep Search.” Other popular onion directories that are widely used are TorLinks, The Hidden Wiki, and Deep Web Links with each on containing certain types of onion links from drugs to political to media. A lot of these websites to not stay up permanently and have constant URL changes since they are constantly being taken down and seized and then popping back up again. A lot of these sites and related sites are spread through word of mouth and many surface web forums and chats such as /r/Onions [7].

There is also a related software project called Tor2Web that does practically the same thing. Tor2Web asks as a very strong proxy that connects you to Tor instead of directly. Besides Tor, there are other softwares and method to grant access to different areas of the Deep Web.

3.2 Freenet

Freenet is a free software darknet. Freenet allows the sharing of files, chats, free browsing and activity with any blockage or censorship. Freenet is limited to its own domain such as Tor is in a way; Freenet is associated with ”freesites” in a way that Tor is associated with .onion sites. Freenet is decentralized which makes it less vulnerable to attack and has a special mode called ”dark net mode” which allows the users to only connect to people they know therefore allowing them to be in a closed off network of people they trust reducing the chances of being victim of malicious activity very minimal [6].

3.3 Gopher Protocol

In the early stages of the internet, the Gopher Protocol was a popular method of distributing, searching, and retrieving documents over the Internet. The HTTP protocol was introduced and standardized, servers and internet users used Gopher protocol to distribute and server its clients information and data in simple text formats through terminals. Even though it is not as widely used as it was before, Gopher sites still exists and they still host content. Gopher was technically part of the Deep Web since the information and data hosted through Gopher are not indexed by search engines since it is outside the HTTP/S protocol [6].

4. Deep Web Communication

Through providing anonymity, the deep web holds significant value to individuals or entities that wish to communicate privately with other parties. Methods such as PGP encryption are implemented and used to achieve private and secure communication. Previous such implementations such as Tor Mail have been successful in providing users a secure and anonymous means of communication to other parties both within and outside of the deep web. While the various methods of secure and private communication offered by the deep web and encryption algorithms have obvious benefits, the anonymity of the deep web results in their nondiscriminatory distribution to anyone who wishes to utilize them. The frightening repercussion is that there are many groups and individuals on the deep web that are using anonymous and private communication to engage in a wide range illegal practices, some of which can be quite horrifying [3].

4.1 PGP Encryption

PGP (Pretty Good Privacy) encryption is a method of encryption that combines features from both conventional cryptography and public key cryptography to form an encryption process that aims to find the optimal balance between speed and security. It is used to provide cryptographic privacy and authentication for data communication by signing, encrypting, and decrypting communications such as texts, e-mails, files, etc.

4.1.1. How It Works Similar to public key encryption, in the PGP encryption process every party has both a public key and a private key. However, it begins to diverge from public key encryption through its generation and use of a session key. The session key acts as a one-time private key, generated by the sender, that is specific to the corresponding message. These keys are all vital pieces in that are leveraged to provide trust and ensure privacy in communications that utilize PGP encryption. PGP encryption begins by compressing the plaintext message in an effort to reduce the message transmission time and increase the cryptographic security. Next, the private session key is generated by the sender and used by a conventional encryption algorithm to encrypt the plaintext, producing ciphertext. Upon completion of the plain text encryption, the session key is then encrypted to the recipients public key and sent alongside the message. Upon receiving the message, the recipient will begin the decryption process by using their private key to determine the session key. The session key is then used to decrypt the message. Finally, the recipient will decompress the plaintext and will have a readable plaintext message [2].

4.2 Bitmessage

Bitmessage is a communication protocol that is offered as a means to communicate privately by leveraging decentralized, encrypted, peer-to-peer, trustless properties that allow the messages to pass through the deep web anonymously and securely. Each users inbox is encrypted using public-key cryptography and replicates it within the person-to-person network, mixing it in with multiple inboxes of other users to make it difficult to determine which public key belongs to which user. This protects the identity of users from anyone attempting to eavesdrop and contributes to the decentralized functionality of the peer-to-peer network. Bitmessage has built a reputation for protecting the privacy of users from virtually any entity intent on monitoring or tracking activity, such as the NSA. This contributed to Bitmessages massive growth following Edward Snowdens release of information regarding the surveillance practices of the NSA in 2013 [2].

4.2.1. How It Works Bitmessage is able to provide such strong privacy through a combination of multiple properties and practices. The system replicates all messages that are sent through its person-to-person network in a very successful attempt to make it extremely difficult to track the source and destination of any given message being sent. Bitmessages generates all of its addresses cryptographically and uses a variation of public-key cryptography in which only the recipient can decrypt a sent message; this means not even the sender can decrypt the message once it has been encrypted with a public key. Another important property is that messages are sent without any destination address. The messages are replicated and sent to arbitrary users that will attempt to decrypt it. If a node, or user, fails to decrypt the message it will know that the message was not meant for it and will disregard it. Bitmessage uses an acknowledgement system that will let senders know when a message has been received by a recipient, but the sender does not have any way of identifying which specific network node is the recipient because there will be multiple users with the sent message in their inbox that were unable to decrypt it. Lastly, Bitmessage nodes only maintain received messages from the previous two days. After two days in an inbox messages are deleted so that the network does not maintain any record of previously sent messages. The two day period is to allow the message to find its way to the recipient within the person-to-person network. If the message is not successfully received within that timeframe, the sender will simply be notified by the acknowledgement system and will have the option to try sending it again [2].

4.3 TorChat

TorChat is an instant messenger that is decentralized and anonymous through the use of Tor hidden services as its underlying network. It is popular amongst deep web users and has been widely prior to the takedown of Tor Mail and the Edward Snowden leaks. It is known for being relatively easy to setup and use, as it can be run from an USB drive on Windows (updates have made it compatible with Linux and Mac machines). TorChat has a reputation for being secure and maintaining anonymity; even if your device browsers are already being monitored, those monitoring cannot see what you send, who you send it to, or the location or identity of any of your contacts.

4.3.1. How It Works Upon installation of TorChat, the user will be assigned an automatically generated 16-character alphanumeric address that corresponds to the .onion address of a hidden service. Users are authenticated by other users by proving that they are reachable through their randomly generated .onion address. It accomplishes this by ensuring that an outgoing message can only reach the .onion address that it has generated as its destination address. TorChat does not trust any incoming message; it responds by generating a corresponding outgoing message to the same address to verify that the incoming message came from the address it claims it came from [2].

4.4 Tor Mail

Tor Mail was a Tor hidden service that provided users with an e-mail application that could send messages to and receive messages from e-mail addresses that were both inside or outside the Tor network while maintaining the anonymity and privacy of the user. Tor Mail went offline in wake of an FBI takedown of Freedom Hosting, which hosted many .onion sites. The infiltration of the Tor Mail servers was accomplished through a JavaScript attack on the FireFox browser that Tor is built on top of. The FBI was able to inject some JavaScript code into the Tor Browser Bundle that was then exploited to enable them to send user IP addresses and Windows computer names to an FBI server in Virginia. This attack eventually led to the FBI admitting they had access to the Tor Mail servers [2].

5. Deep Web Security

MATTS SECTION - FILL IN - USE SUBECTIONS IF U NEED

6. Darknet

The deep web is a fairly nebulous term, wrapping several criminal networks into one. One of these networks is the ominously named DarkNet. The DarkNet is the online black market, with a heavy emphasis on drugs. If you want to purchase illegal items from the comfort of your own home, it is only a click away. When talking about the DarkNet, it is important to know what it is, the risks and tradeoffs involved, and how all of the tools we talked about make it possible for such a place to exist. [15]

6.1 What is the DarkNet?

While the FBI and NSA give the DarkNet a bad name, it is not all bad. At its core, it isjust people who want to buy and sell goods without regulation or government interference. There are some people who cannot afford the medication they need, for example, so they buy it through the DarkNet. DarkNet users in countries with better or free healthcare can get the prescriptions much cheaper and sell them at a more affordable rate online. Of course, in the eyes of the law, all of the sales on the DarkNet are illegal, but a lot of the time it is more benign then news stories let on. For the most part, it is not terrorists funding their terror armies by selling their terror drugs. It is usually just people who want affordable and safe drugs to use at home [8].

The implications of a vast international underground black market are immense. When people are able to do whatever they want with no government overhead, shady organizations will quickly move in and use the space. With the anonymity of the DarkNet, it is almost impossible for law enforcement to catch anyone, or prosecute them when they are caught. The NSA themselves have already said that they, will never be able to deanonymize all Tor users all the time [14]. If the government cannot stop a system that is using their own postal services, then imagine what kinds of things people can sell and buy.

As we said, the DarkNet has a heavy emphasis on drugs. Most people are looking to buy or sell drugs in a safer and subtler way than a traditional drug dealer. In a recent forum post, vendors listed what they were selling, which included narcotics, prescription drugs, steroids, weapons, MDMA, amphetamines, false identities, and fake credit cards. $(http:$ $//www.reddit.com/r/DarkNetMarkets/comments/308joe/what_to_buy_we$ This is just what is openly advertised. The DarkNet does still have an even darker subnet. If you know where to look, you can buy weapons and explosives, powerful exploits, child pornography, or even hitmen. Just last year, someone running a child pornography site was caught. He said that when his site was shut down, it had over 40,000 user accounts on the site. We used to get sometimes 500 page views per second. If you want something, the DarkNet is selling it. It is no surprise that the FBI and NSA are trying so hard to bring it down [9].

6.2 The Risk of the DarkNet

Law enforcement make the DarkNet sound like a criminal haven, where people can do whatever they want without fear of retribution. A lot of its denizens also treat it as such, some vendors will even openly advertise or even flaunt their wares. While it may seem perfect for criminals, it is full of risks from both criminals and law enforcement. If you do not do everything right, you are bound to get doxxed by someone and made to pay for it. It might be as small as malware making your computer part of a botnet or a bitcoin mining network, or as serious as years of jailtime and hitmen being sent after you. Recently, Ross Ulbricht was convicted on several charges and will be facing a minimum of 30 years in prison [11]. Ulbricht was one of the masterminds behind the Silk Road, one of the first, and biggest, DarkNet marketplaces. Use of the DarkNet puts a target on your back for law enforcement. They are actively trying to bring down these sites, so when they are brought down the next step is to try and find the people who have been using it [3].

While it may seem like the biggest problem is law enforcement, and everything would be better if they just kept out of it, that is not necessarily the case. Recently, Evolution, the largest DarkNet marketplace currently in operation, pulled a huge exit scam. A couple of weeks before the scam, users stopped being able to withdraw their bitcoins from the site. The site said that it was having technical difficulties but was trying to fix it. A week after that, the site went down and all of the funds, the equivalent of millions of dollars, disappeared with it. Someone claiming to have admin access to the site came out in a forum post and said that the admins were, preparing an exit scam with all the funds. The admins managed to get away with millions, and most DarkNet users are furious with them. One user has been confirmed to have killed himself after losing all of his money, several people have pulled together to put a bounty out on them for someone to doxx them. This is a serious problem with serious consequences, and it is not the first time it has happened[2].

6.3 Staying Anonymous in the DarkNet

The entirety of the DarkNet is possible because of the technologies we explored earlier. Users are provided confidentiality through onion routing with Tor and secure operating systems like TAILS, authentication of buyers and sellers with PGP, nonrepudiation through tracking numbers and shipping information and bitcoin logs, availability through peer to peer protocols and back ups, and integrity through the technical prowess of many of the DarkNet's denizens. With all of these in place, it is really difficult for law enforcement or other users to affect you through the DarkNet. It essentially fulfills all of the DoD's aspects of information needing protection. As long as you use it correctly, it is very difficult to trace anything back to you, and even more difficult to prove anything even if they do [1].

It may sound simple to just use all of the software correctly and do all of the mailing and packaging correctly, but it is all very complex. There are a lot of precautions you have to take and it is easy to forget any of the steps. If you neglect to check the PGP signature, you could easily have a law enforcement agent spoof an email and get all of your information along with incriminating evidence. If you do not use Tor or TAILS, then you can easily get doxxed by law enforcement or criminals. The easiest way to protect yourself is to know what you are getting into and what to look out for before you even start. There are a lot of sources on the web and the DarkNet to help new people get into the marketplace. They are surprisingly helpful to new people, especially because they are trying to push an image of being safer than buying drugs on the streets. Just like any other application of security, you have to know what you are getting into and be aware of all the risks, so you know what to avoid.

A lot of people on the web try to push anonymity as a great thing that everyone should want. Internet users expect a certain level of privacy for their browsing, but it can be hard to determine how much privacy a user should have. Should someone be able to buy cocaine on the internet without fear of being caught? Should it be possible for tens of thousands of people to have access to and trade child pornography around the world? It brings up a lot of the same questions that antiterrorism laws do. It is difficult to decide where the line between public safety and personal security should be drawn. Anonymity can be powerful, but with great power comes great responsibility [1].

7. Hacktivists

Hacktivism is a term to describe the phenomenon where the networks and computers are used in order to send a message through various forms of cyber-attacks. These messages could be expressions of protest or promotion of a political ideology. Hacktivism is also a very new term spread throughout the internet so the interpretation could vary among actions committed by hackers who carry out cyber-attacks in order accomplish a goal. The many goals that hacktivists have could range from simply defacing a website with images or text to a denial-of-service attack [3]. A primary component of performing operations to complete a hacktivist attack is the deep web. Hacktivism relies heavily on anonymity because it is allows for the most pure of free speech to be expressed. This removes the fear of persecution by any institution that has the identity of the expresser. With all these factors coming into play, a new context has given rise in which activists are more powerful in the pursuit of spreading information and ideas [9].

The Deep Web is an attractive medium through which hacktivists can get organized and execute their operations. The common notion of why this medium is used by hacktivists is primarily because of the level on anonymity it can guarantee, though it is also because the users of the deep web who inherently have the technical skills and common knowledge of cyber space. This creates the opportune meeting space for technically skilled individuals who would like to see change which ultimately ends up brewing into acts of hacktivism. Operations that take place on the Deep Web consists of users who are in active mode [10] when it comes to their approach to the Deep Web rather than passive mode users who are only there to gather information. Since communication is vital in order to carry out operations, hacktivists on the Deep Web use a micro blogger service that is similar to Twitter called TorStatusNet. This is only one known platform that hacktivists use to communicate to each other. There are potentially many unknown services that are shared between users that are exclusive to a hacktivist group [8].

7.1 Anonymous

It is very difficult to discuss the topic of Hacktivism without mentioning one of the most popular hacktivist groups, Anonymous. Anonymous is an international collective of hacktivists that have no leader. The symbol that Anonymous is identified with is a man standing in front of a globe with a question mark on his head which represents an organization without one sole leader. They have been known to commit cyber-attacks on government institutions, religious institutions, and corporations. The members of Anonymous are called Anons and are usually wearing a Guy Fawkes mask when they are in the public. Anonymous has been a very influential force in the cyber-world as well as the real-world. Time had even called Anonymous on of the 100 most influential people in the world [13].

Anonymous gained a major following after a video of Tom Cruise being interviewed about Scientology was posted on the internet. The video was ridiculed by many online hackers because it shed The Church of Scientology in a bad light which spurred the church to issue legal threats against any site that hosted the video. This began a debacle between Anonymous and the Church of Scientology in which Anonymous led a campaign of cyber-attacks against the Church. These actions rallied many online hackers to join Anonymous which led to many other future attacks against numerous institutions world-wide till this day. A very notable action that Anonymous took was causing a denial-of-service attack on PayPal and Visa in reaction towards PayPal and Visas statement proclaiming that the will no longer host the service for donations of Julian Assanges WikiLeaks which exposed many United States Classified documents [16].

7.2 Operation DarkNet

Anonymous has very close ties with the Deep Web since it is one of their major forms of communication. Specifically Anonymous members prefer communicating in chatrooms such as IRC (internet relay chat), on the deep web. It is here where Anonymous members began Operation Darknet which aimed to take down websites that were hosting child pornography. In their pursuit, Anons had discovered a hidden wiki named Hard Candy that consisted of links to nearly all child pornography sites. Anonymous had figured out the hosting service behind the links and had demanded that the hosting service cease any hosting to these sites. When the hosting service did not comply, Anonymous swiftly took down the Hosting server and has rendered the service offline after repeated attacks at shutting down the server. Anonymous efforts towards ending child pornography have been generally recognized as a positive effect of hacktivism.

7.3 Assured Anonymity

For Hacktivism to be possible, anonymity must be assured. How does one become anonymous in the cyber-world so that they may partake in Hacktivism? Many may think that the effort of remaining anonymous in the cyber-world is purely technical, but it seems that the psychological aspect plays a big part as well. The Tor network remains one of the only options to remain anonymous in cyber-space, though one must surf cautiously while using the Tor network is they wish to remain anonymous. To be truly anonymous, one must assume that there is an intelligence with the resources to have surveillance over the internet, so certain precautionary steps must be taken [5].

7.4 Software and OS Usage

A first step towards using the Tor Network is not to use a Windows operating system because security vulnerabilities have been exploited by intelligence agencies. Encrypting all storage devices is another step that must be taken if you want to keep your files safe. Always keeping security of software related to your TOR network updates is a vital step so that you may be protected from the newest security vulnerabilities. Scripting for websites should be disabled as well as having cookies and local data being constantly deleted. The machine you browse the Tor network should be a laptop that way it may be quickly destroyed in any instance that your identity is compromised. You must make sure you never use the network at your own home, or at any one location more than once. You should also not have any device with you while browsing that has the ability to be tracked by GPS such as your cell phone. These are all just precautionary steps one takes to make sure their anonymity is secure, though one must then delve into the psychological aspect of remaining anonymous [12].

7.5 Online Facades

Many users who truly wish to remain anonymous in Cyber-space usually create an online personality that is very different from theirs so that there may be no indication that they are one in the same. There should never be any mixing when it comes to online accounts with these two identities. It must be acknowledged by the user that these two identities are entirely mutually exclusive when going forward through using the Deep Web. This is a mental discipline that users who are still truly anonymous have mastered. Much paranoia is involved when stepping in the dark waters of anonymity because it is usually a clear indication that you are doing something that may jeopardize your freedom. The paranoia that users have are usually justified because there are examples of how identities of members in hacktivist groups were compromised due to not being entirely cautious [14].

8. References

List and number all bibliographical references in 9-point Times, single-spaced, at the end of your paper. When referenced in the text, enclose the citation number in square brackets, for example [1]. Where appropriate, include the name(s) of editors of referenced books.