GIAC
CERTIFICATIONS

Global Information Assurance Certification Paper

Copyright SANS Institute
Author Retains Full Rights

This paper is taken from the GIAC directory of certified professionals. Reposting is not permited without express written permission.

Interested in learning more?

Check out the list of upcoming events offering
"Network Monitoring and Threat Detection In-Depth (Security 503)"
at http://www.giac.org/registration/gcia

Under The Ocean of the Internet - The Deep Web

GIAC (GCIA) Gold Certification

Author: Brett Hawkins, hawkbluedevil@gmail.com
Advisor: Adam Kliarsky
Accepted: May 15, 2016

Abstract

The Internet was a revolutionary invention, and its use continues to evolve. People around the world use the Internet every day for things such as social media, shopping, email, reading news, and much more. However, this only makes up a very small piece of the Internet, and the rest is filled by an area called The Deep Web. This is not to be confused with "The Dark Web" or "The Darknet". Those terms refer to a part of The Deep Web. The majority of the population visits the "Surface Web", but what is under the surface? Located under the surface, The Deep Web isn't indexed by search engines such as Google, and you need specially configured software to access it, such as the Tor browser for example. Any individual who accesses The Deep Web is doing so for a reason. Somebody could be part of a dark market selling exploits or stolen credit cards, or they could be a law enforcement officer monitoring for illegal activity. Where exactly is The Deep Web and who accesses it? How can you access it? What is it used for and what content is available? In this paper, we will take a look under the ocean of the Internet, and into The Deep Web.

© 2016 The SANS Institute Author retains full rights.

1. Introduction

The Internet is like a big ocean. That ocean is filled with large continents and islands that people visit. A large continent would be Google, and an island would be the news site for your local newspaper. Every day the average person visits these continents and islands using their web browser, which acts as a boat navigating to destinations on the Internet. The reality though is that these continents and islands only make up 4% of the Internet. The rest of the Internet is made up of the Deep Web, which is located under the ocean (Epstein, 2014). The Deep Web is used for both good and bad, while some may assume its use is for illegal purposes. The use of the Internet continues to evolve, and the Deep Web is a big part of that.

1.1. Internet Usage

People all over the world use the Internet every day. There are currently over 3 billion people that use the Internet, more than 1 billion websites, and 3.5 billion Google searches a day. There are also 500 million tweets sent a day ("Internet Live Stats - Internet Usage & Social Media Statistics", 2016). These numbers have grown significantly in the past 10 years, and will continue to grow as the Internet evolves and its use expands.

The vast number of people using the Internet have similar vast needs for using it. According to *Top10Base*[1] , the top 10 tasks that the Internet is used for are:

1. Email
2. Music & Movies
3. Searching
4. Buying Tickets
5. Shopping

Social Media is right outside of the top 10 ("Top 10 Uses of Internet", 2016). Out of the listed tasks, how many of them do you regularly perform online? Your answer is most likely all of them. For these tasks you can either use them via a web browser, such as Google Chrome, or you have some piece of software to perform the task like

[1] Top10Base is a website known for collecting data relevant to popular categories, and reporting the 'Top 10' in these categories.

© 2016 The SANS Institute Author retains full rights.

iTunes to download music and movies. As the usage of the Internet continues to change and evolve, the design of the Internet continues to be tested as well.

1.2. Internet Design

The Internet was originally designed as an open-architecture network, which allowed communication and collaboration. As stated in "Brief History of the Internet":

"*In an open-architecture network, the individual networks may be separately designed and developed and each may have its own unique interface which it may offer to users and/or other providers. including other Internet providers. Each network can be designed in accordance with the specific environment and user requirements of that network. There are generally no constraints on the types of network that can be included or on their geographic scope, although certain pragmatic considerations will dictate what makes sense to offer*" (Leiner, 2016).

As you can see, the Internet was meant to be a free and open space upon its conception. It was designed in a way that multiple systems, no matter how different, could all communicate with each other. This architecture still stands today, and it is making sure that your packets get delivered to their destination.

The backbone of the Internet is comprised of many large interconnected networks. These networks are called Network Service Providers (NSP). The purpose of the NSPs are to deliver packets back and forth. Each NSP is capable of connecting to multiple Network Access Points (NAP), which allow the movement of network traffic between NSPs. NSPs can also connect with Metropolitan Area Exchanges (MAE), which serve the same purpose as a NAP. Another term used for NAPs and MAEs are Internet Exchange Points (IXP). An advantage of an IXP is that it makes it easier for a group of Internet Service Providers (ISP) to connect with each other. The ISPs purchase network bandwidth from NSPs. Then this bandwidth is delivered to their customers for Internet connectivity ("AS Computing - Unit 2 The Internet"). Figure 1 below shows the Internet design previously discussed at a high level.

© 2016 The SANS Institute Author retains full rights.

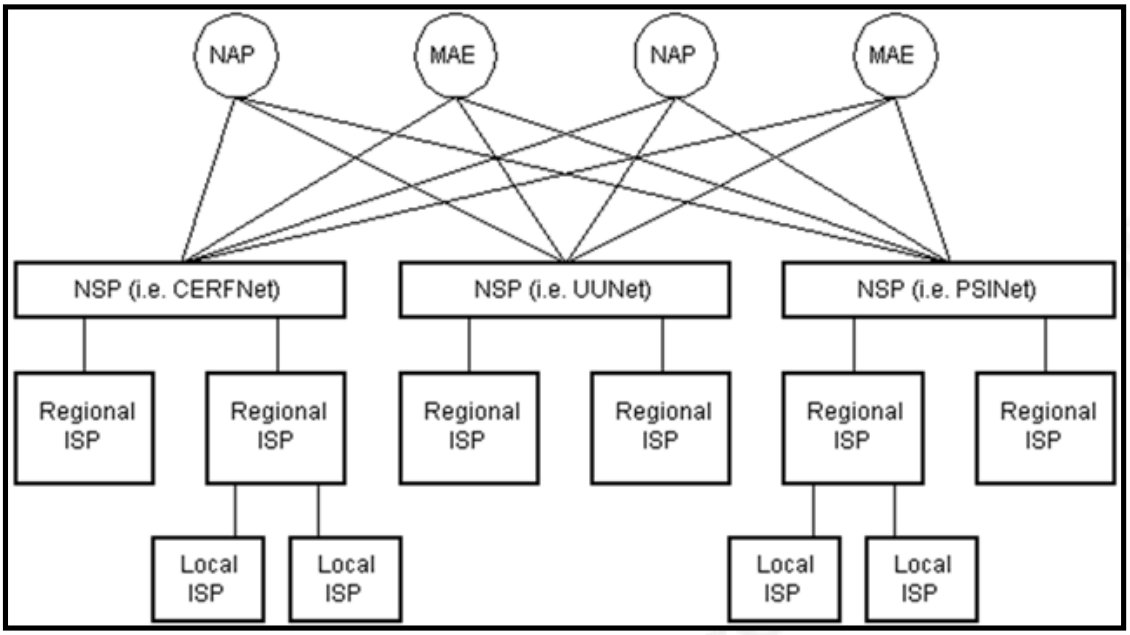


Figure 1 ("AS Computing - Unit 2 The Internet") - Diagram of Internet Design

As you can see, the design of the Internet has remained very similar throughout its history, starting in the late 1960's.

1.3. History of The Internet

The concept of the Internet was invented by Larry G. Roberts in the late 1960's, and was called the ARPANET (Leiner, 2016). This stood for the Advanced Research Projects Agency Network (Andrews, 2013). In the early 1980's, ARPANET started to use the Transmission Control Protocol and Internet Protocol (TCP/IP) model that was developed by Robert Kahn and Vinton Cerf. This is a model that defines the standards for how data is supposed to be transmitted between various networks (Andrews, 2013). This same model is still used today, and is shown in figure 2 below.

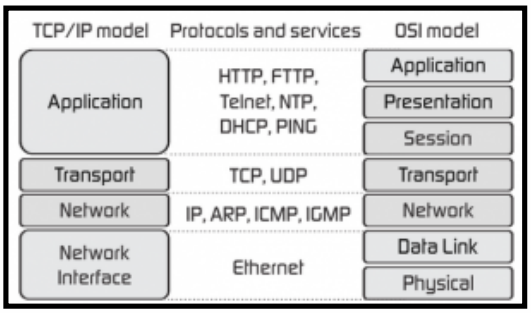


Figure 2 ("Advantages And Disadvantages Of Tcp/ip Model") - The TCP/IP Model

© 2016 The SANS Institute Author retains full rights.

Then in 1990, another key event took place in the history of the Internet. This was the invention of the World Wide Web (WWW) by Tim Berners-Lee. This is what the Internet is mostly recognized for today, and is what helped boost the popularity of the Internet to the general public (Andrews, 2013).

1.4. The Surface Web vs. The Deep Web

The Internet is comprised of two pieces. Those two pieces are the Surface Web and the Deep Web. The Surface Web is the area of the Internet that the average person visits, such as visiting Facebook, Google, Amazon, or YouTube. These areas can be accessed using a standard piece of software, such as a web browser. The other area of the Internet is called the Deep Web. The Deep Web is made up of the Dark Web, Deep Web Databases, and much more. You need specialized software or access in order to interact with the Deep Web. The distinction between these two areas of the Internet are very important.

1.4.1 The Surface Web

The Surface Web is an area of the Internet that is indexable by search engines, such as Google. Other names for this area of the Internet are Visible Web, Lightnet, Indexed Web, Clearnet, or Indexable Web ("Surface web", 2016). Just to put it in perspective, there are currently over 4 billion indexed web pages ("The size of the World Wide Web [The Internet]").

Let's take a look at how a search engine like Google indexes web pages. They use pieces of software called web crawlers, whose primary purpose is for the discovery of web pages on the Internet. You will know that a Google web crawler is crawling your site by seeing "Googlebot" in the user agent String. Of course, the user agent string could be spoofed by an attacker. Once the web crawler visits a page, it will look for any links on that page and visit those pages. Upon visiting these pages, data is gathered and sent to Google. Google has software that determines which sites are to be crawled, the frequency of the crawling, and the number of pages to be retrieved from sites. It will give additional attention to sites that are new, have changed, or are no longer available. These pages are officially indexed by Google, so

© 2016 The SANS Institute Author retains full rights.

that the pages can be retrieved in an efficient and appropriate manner when needed. A Google index is comprised of information about various keywords and where they are located. When you search for something, Google does a lookup of your search term in an index to find matching pages on the web ("Crawling & Indexing – Inside Search – Google"*)*. There is an area of the Internet though that Google cannot reach via a search, since it is not indexable.

1.4.2 The Deep Web

The Deep Web is an area of the Internet that is not indexable by search engines and not linked to pages on the Surface Web. Other names for this area of the Internet are Deep Net, Hidden Web, or Invisible Web ("Deep web", 2016). This part of the Internet makes up 96% of it, which is obviously significantly larger than the Surface Web. The Deep Web is 500 times larger than the Surface Web (Epstein, 2014).

There are many reasons that a web page is not crawlable. The web page could be password protected, which would prevent a web crawler from accessing it. Another scenario could be that the web page is only allowed to be accessed a certain number of times, then it becomes unavailable. If that threshold is met before a crawler reaches the web page, then it wouldn't be crawled. Another way that a web page cannot be crawled is if the site's robots.txt file explicitly says not to crawl it. A robots.txt file is located in the root of a web site, and will let web crawler's know which directories are not allowed to be crawled on its site and which user agent's the rule applies to. The last scenario that would cause a web page to be uncrawlable, is if the page is simply hidden or not linked on any other page of the website. For a "hidden" page, somebody would need previous knowledge of the path to the page in order to visit it ("The Ultimate Guide to the Invisible Web", 2013). The average Internet user is not going to use the Deep Web, so its use should be considered suspicious.

© 2016 The SANS Institute Author retains full rights.

2. The Deep Web

The Deep Web is a complex and mysterious area of the Internet. There are many reasons that its content can be accessed or used for legitimate or illegitimate purposes. There is plenty of content available in the Deep Web, such as Dark Web Hidden Services and Deep Web Databases to name a couple. Special software, such as Tor, is required to access the Deep Web. The details on all of these facets of the Deep Web will be covered in the sections to follow.

2.1. Why Go Under Water?

The use of the Deep Web can be split into two categories. These categories are legal activity and illegal activity. An example of illegal activity would be the selling of stolen credit cards. An example of legal activity would be using the Wayback Machine to see a previous version of a web page. Regardless whether the usage is legal or illegal, the act of accessing the Deep Web is an intentional action.

2.1.1 Legal Activity

Believe it or not, there is plenty of legal activity that goes on in the Deep Web. The Deep Web can be a very useful resource for a plethora of information. For instance, there are plenty of search engines that allow you to search databases not indexed by the Google's and Bing's of the world. These databases can contain virtual academic libraries or old versions of web pages ("The Ultimate Guide to the Invisible Web", 2013).

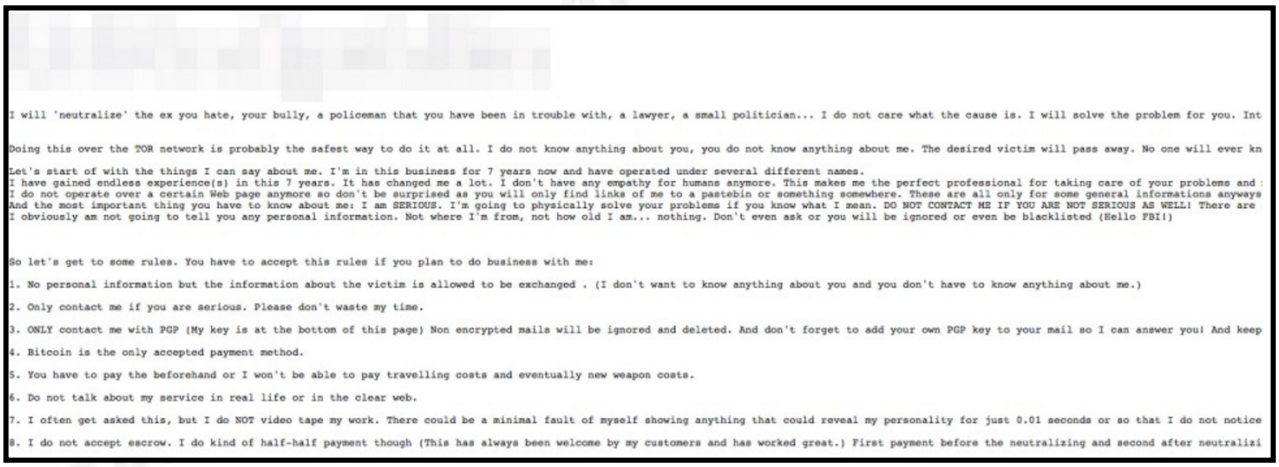
There are several tasks that are perfectly legal to perform on the Deep Web, and you might not realize that the data being accessed actually resides there. When somebody performs a background check on an individual, it searches several databases on the Internet for information. This information is actually being searched for on the Deep Web. Another use for the Deep Web is if an adopted person wanted to try and search for their natural parents. The databases that house this adoption information are on the Deep Web. You can also use the Deep Web to perform veteran research, or lookup your genealogy history. Legal research is also conducted on the Deep Web for cases. If you are a student, you can use the

© 2016 The SANS Institute Author retains full rights.

Deep Web too. There are several academic databases that you can search through for topics, such as scientific journals for example (Dube, 2014). Although there are legal activities to perform on the Deep Web, there are illegal activities as well.

2.1.2 Illegal Activity

This is no surprise, but there is definitely illegal activity that happens on the Deep Web. The Deep Web is a place to go where you can be anonymous. As such, this is a popular destination for criminals to buy and sell information. Some information that can be sold on the Deep Web are social security numbers, medical records, credit card numbers, and other Personally Identifiable Information (PII). You can buy enough information on somebody to easily steal their identity (Ingevaldson, 2015). The Deep Web is also used to sell drugs, display child pornography, trade weapons, and hire hitmen (Reporter, 2013). In figure 3 below, it shows an example of a hitman posting on the Deep Web.



I will 'neutralize' the ex you hate, your bully, a policeman that you have been in trouble with, a lawyer, a small politician... I do not care what the cause is. I will solve the problem for you. Int

Doing this over the TOR network is probably the safest way to do it at all. I do not know anything about you, you do not know anything about me. The desired victim will pass away. No one will ever kn

Let's start of with the things I can say about me. I'm in this business for 7 years now and have operated under several different names.
I have gained endless experience(s) in this 7 years. It has changed me a lot. I don't have any empathy for humans anymore. This makes me the perfect professional for taking care of your problems and
I do not operate over a certain Web page anymore so don't be surprised as you will only find links of me to a pastebin or something somewhere. These are all only for some general informations anyways
And the most important thing you have to know about me: I am SERIOUS. I'm going to physically solve your problems if you know what I mean. DO NOT CONTACT ME IF YOU ARE NOT SERIOUS AS WELL! There are
I obviously am not going to tell you any personal information. Not where I'm from, not how old I am... nothing. Don't even ask or you will be ignored or even be blacklisted (Hello FBI!)

So let's get to some rules. You have to accept this rules if you plan to do business with me:

1. No personal information but the information about the victim is allowed to be exchanged . (I don't want to know anything about you and you don't have to know anything about me.)
2. Only contact me if you are serious. Please don't waste my time.
3. ONLY contact me with PGP (My key is at the bottom of this page) Non encrypted mails will be ignored and deleted. And don't forget to add your own PGP key to your mail so I can answer you! And keep
4. Bitcoin is the only accepted payment method.
5. You have to pay the beforehand or I won't be able to pay travelling costs and eventually new weapon costs.
6. Do not talk about my service in real life or in the clear web.
7. I often get asked this, but I do NOT video tape my work. There could be a minimal fault of myself showing anything that could reveal my personality for just 0.01 seconds or so that I do not notice
8. I do not accept escrow. I do kind of half-half payment though (This has always been welcome by my customers and has worked great.) First payment before the neutralizing and second after neutralizi

Figure 3 - Example posting of hitman

As described here, it's clear there is an abundance of illegal activity, but how does one detect and identify that activity on their network?

The most common way to communicate with the Deep Web is through Tor. Therefore, you will want to be able to detect Tor traffic on your network. When analyzing Tor traffic in a packet capture, it will look like normal HTTPS traffic. However, if you take a look at the certificates used in Tor traffic, you will see the issuer and subject ID are using randomly named domains, indicating they are suspicious certificates. This is shown in figure 4 below.

© 2016 The SANS Institute Author retains full rights.

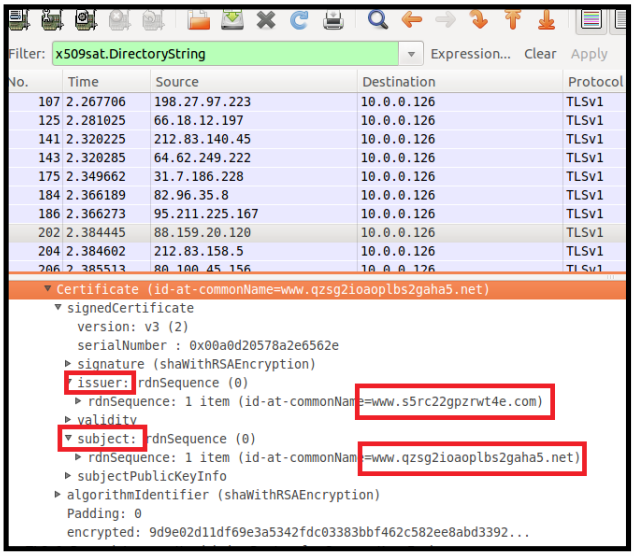


Figure 4 (Reese, 2016) - Sample packet capture of Tor traffic

As you can see, in order to detect Tor traffic, it is vital to be able to identify random certificates. There is a Bro script that allows you to do this, which can be found at https://raw.githubusercontent.com/sethhall/bro-junk-drawer/master/detect-tor.bro. Another detection method for Tor traffic is by alerting on communication with known Tor servers. However, you will not want to just rely on detecting suspicious traffic via Tor servers only, because those servers can be used for multiple purposes, some of them being legitimate purposes. You can find a list of known Tor servers at https://www.dan.me.uk/torlist/. Using the combination of the Bro script to detect random certificates and alerting on the communication with known Tor servers, will allow the accurate and efficient detection of Deep Web activity on your network (Reese, 2016).

Detecting the traffic is one thing, but what about identifying the traffic when analyzing a packet capture? In order to be able to successfully identity Tor traffic when analyzing packets, you need to be able to perform a statistical analysis of the protocol in order to see differences in various SSL implementations. A tool that can be used for this is CapLoader, which utilizes its "Port Independent Protocol Detection" feature. When loading a pcap file into CapLoader, it is able to identify the Tor protocol, based on its statistical method for protocol detection, which ignores port numbers (Hjelmvik, 2013). Figure 5 below shows this.

© 2016 The SANS Institute Author retains full rights.

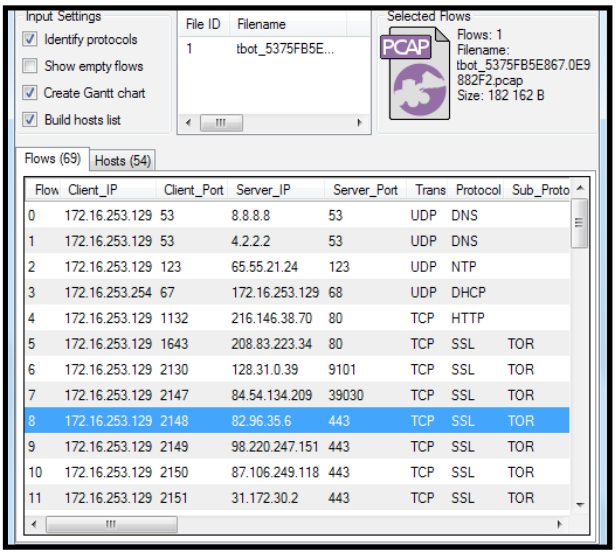


Figure 5 (Hjelmvik, 2013) - Usage of CapLoader to identify Tor traffic

The various methods shown can help in the detection and identification of Deep Web traffic. Now let's take a look at what content is available on the Deep Web.

2.2. What is in The Ocean?

There is plenty of content available on the Deep Web. Some of it is good and some is bad. The bad content involves things such as Dark Web Hidden Services, the Hidden Wiki, and Silk Road. Although Silk Road was taken down, it was one of the most popular sites to visit when it was still up and running. These pieces of content are available through the Dark Web.

2.2.1 The Dark Web

The Dark Web is an area that resides on the Deep Web. Several people confuse the Deep Web and the Dark Web thinking they are the same thing. This is definitely not the case. The Dark Web is mainly accessed via a software client called Tor, which will be discussed in more detail later in this paper. Tor is a special browser that allows you to navigate the Dark Web. One popular use of the Dark Web is in relation to malware. Large amounts of malware are using the Dark Web to communicate with their Command & Control (C&C) servers. An example of a piece of malware that does this is SkyNet (Cox, 2015).

© 2016 The SANS Institute Author retains full rights.

SkyNet is a trojan that has the capabilities of performing a DDoS attack or mine Bitcoins. It uses Hidden Services provided by Tor to communicate anonymously with its C&C servers. An advantage of using these Hidden Services for C&C communication is that the traffic is encrypted, so it masks the origin, destination, and payload. Another advantage is that the owner of the C&C servers can move them around, since they can just re-use the private key for the Hidden Service ("Skynet, a Tor-powered botnet straight from Reddit", 2012). The use of Hidden Services on the Dark Web are very powerful.

2.2.2 Dark Web Hidden Services

Hidden Services on the Dark Web are used to provide a variety of services to users of the Dark Web, while the users identities remain anonymous. Some of the categories of services that are offered are Financial, Communications, Commerce, News, Pornography, Search Engines, File Storage, and Hidden Service Directories and Portals. There are specific services associated with these categories. For example, if you wanted to use some Financial Hidden Services, you could use Bitcoin Fog or BitBlender. The use of Communications Hidden Services could be taken advantage of by using TorChat or RiseUp. There are several Hidden Services for Commerce as well. This is usually associated with the Darknet Market. An example of a couple of Commerce Hidden Services are Assassination Market and AlphaBay Market. If you are using News Hidden Services, you could use DeepDotWeb or Wikileaks. A couple of Search Engines Hidden Services available are The Pirate Bay and Sci-Hub. Free Haven is one of the most popular Hidden Services for File Storage, and for the most popular Hidden Service Directory, you can use The Hidden Wiki ("List of Tor hidden services", 2016). There are tons of Hidden Services available, and the details of how they work are complex.

In order to publish a Hidden Service, you need to make it available on the Tor network, so that users can connect to it. For the first step, the owner of the service will need to pick an introduction point and build Tor circuits to them. An introduction point is a Tor relay, which is essentially a router. Your Hidden Service can choose up to 10 introduction points. The more popular your Hidden Service is,

© 2016 The SANS Institute Author retains full rights.

the more introduction points it will need ("Hidden Services need some love", 2013). After picking introduction points, you will need to advertise your Hidden Service as “something.onion”. The Hidden Service will create a descriptor, which will include its public key and a summary of the introduction points used by it. The Hidden Service will sign this descriptor with its private key. That descriptor gets sent to a distributed hash table, also known as the database. Once this happens, the Hidden Service is officially setup and users can access it by requesting it at “something.onion” ("Tor: Hidden Service Protocol”).

Now that the Hidden Service is setup, let’s take a look at how you can access it. First, you need to know that the specific “.onion” address is in existence, similar to when you need to know “google.com” exists before visiting it. Once you have a “.onion” address that you want to access for its Hidden Service, you will attempt to connect to it via a software client, such as Tor. The user will attempt to download the descriptor for the Hidden Service from the distributed hash table. This descriptor will tell the user the introduction points and public key that needs to be used. During this process, the user is also creating a Tor circuit to a random Tor relay that will be used as a “rendezvous point”. Once you have a descriptor and the “rendezvous point” has been established, the user will send a message that is encrypted by the public key of the Hidden Service, via a Tor circuit to one of the introduction points, that includes the “rendezvous point” location and a one-time secret. Once the Hidden Service has received the message, it will decrypt it. Then it will create a Tor circuit to the “rendezvous point” and send the one-time secret. Lastly, the “rendezvous point” notifies the user of a successful connection. Once this happens, the user can communicate with the Hidden Service via their Tor circuits to the “rendezvous point” ("Tor: Hidden Service Protocol”). One of the most popular services used is the Hidden Wiki.

2.2.3 Hidden Wiki

The Hidden Wiki is a site that contains links to various Hidden Services available on the Dark Web. As you can see in the screenshot taken below in figure 6, this is a snippet of the main page for the Hidden Wiki.

© 2016 The SANS Institute Author retains full rights.

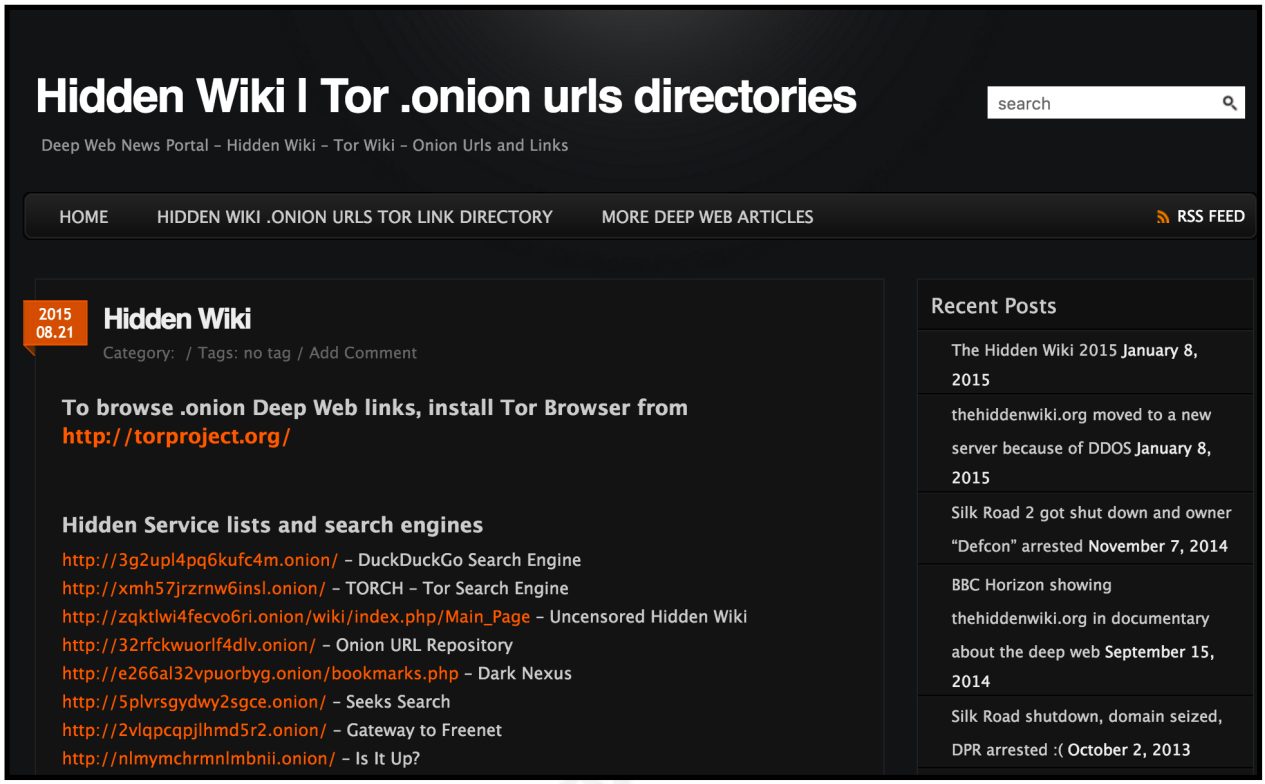


Figure 6 - Main page of Hidden Wiki

The original Hidden Wiki was created a little before October of 2011 and was only accessible via Tor. It was ran through a .onion pseudo-top-level domain. A pseudo-top-level domain is a domain that doesn't participate in the official DNS. Around August of 2013, the site became hosted on Freedom Hosting, which was one of the largest web hosting services used by these Hidden Services at the time. In March of 2014, the Hidden Wiki got hacked and was redirected to a site called Doxbin. Doxbin is a site used to disclose PII. Once this happened, content from the Hidden Wiki started to get mirrored to several other locations. Because of this, there is no single Hidden Wiki any more ("The Hidden Wiki", 2016). A popular site that the Hidden Wiki links to is Silk Road.

2.2.4 Silk Road

Silk Road was an online marketplace on the Dark Web that was used to sell drugs, guns, personal data, malware, and more. It was pretty much the Amazon of the black market and was created by Ross William Ulbricht. Silk Road was active for around 2 years before it was taken down in 2013. The site had a similar feel to

© 2016 The SANS Institute Author retains full rights.

Amazon. You could shop by categories, search for products, and communicate with sellers. Once you found something you wanted to buy, you would add it to your cart and checkout. The currency used to pay for the items was Bitcoin, which is a currency still in use today. Bitcoin is a virtual currency used that is created and stored electronically with no paper trail, and is essentially untraceable. Each Silk Road user was required to have a Bitcoin address. These addresses were stored on Silk Road's servers in a "wallet". As far as getting your suspicious purchases delivered, it was very inconsistent. Your packages may or may not have gotten intercepted by law enforcement, but that was the risk that was taken (Albanesius, 2013).

2.2.5 Deep Web Databases

There are a number of databases on the Deep Web that can be utilized for useful information. Let's start with people research databases. These databases can be used for background searches, and are also behind some popular sites such as pipl.com and spokeo.com. A screenshot taken of pipl.com is shown in **Figure 7** below.

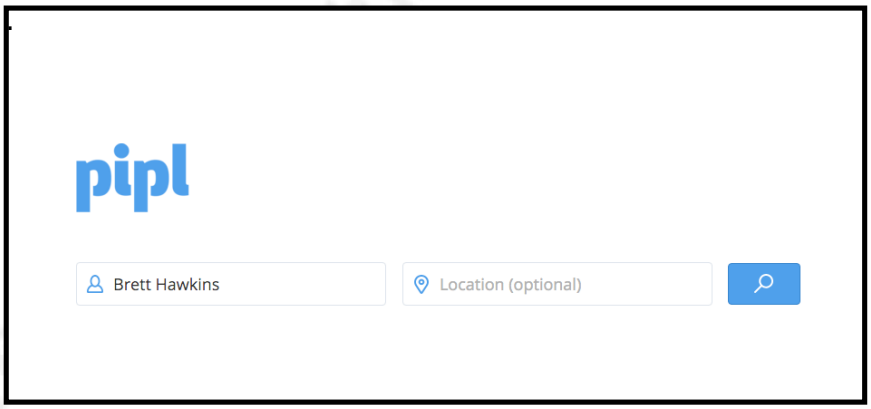


Figure 7 - Main page on pipl.com

These types of sites can be used to search for information on a person, such as their phone number, address, relatives, and more. Another database that is used on the Deep Web is an Adoption Research database. This database can be searched by adopted children that want to find their natural parents. Some other databases that can be searched on the Deep Web are genealogy databases, cemetery records,

© 2016 The SANS Institute Author retains full rights.

historical society databases and file archives (Dube, 2014). Now that we have seen what is in the ocean, let's take a look at how to get there.

2.3. Getting Under the Ocean

There is plenty of content on the Deep Web, but you need to have a way to get to that content. There are two ways you can get to the Deep Web. You can either use the Surface Web to access content through Deep Web search engines, or you can use a piece of software called Tor.

2.3.1 Tor

The Onion Router (Tor) is a free web browser, which is a variant of Firefox. You can run it on all the common platforms such as Windows, Mac OS X, and Linux. Tor is used to connect to the Deep Web, while maintaining anonymity. Simply visit https://www.torproject.org/download/ to download Tor. The purpose of Tor is to provide a networking protocol that can keep the data being transmitted across it anonymous. When using Tor, your packets go through several servers encrypted before reaching their destination. These servers are called Tor relays, which function as routers. There are thousands of these servers across the world. When your packets get sent across the Tor network, it removes pieces of the header that contain information that could identify where the packet is coming from or where it is going. As your packets go from relay to relay, it decrypts just enough data to know which Tor relay the packet came from and where the next hop is. It does not decrypt any additional information (Scharr, 2013). Another way to access content on the Deep Web is through Deep Web Search Engines.

2.3.2 Deep Web Search Engines

Deep Web Search Engines are used to access content on the Deep Web from the Surface Web. Keep in mind, the content you can access with these search engines is limited, compared to the content you can access with Tor. Some search engines that you can use for this are The WWW Virtual Library, SurfWax, and IceRocket. A screenshot of The WWW Virtual Library is shown in figure 8 below.

© 2016 The SANS Institute Author retains full rights.

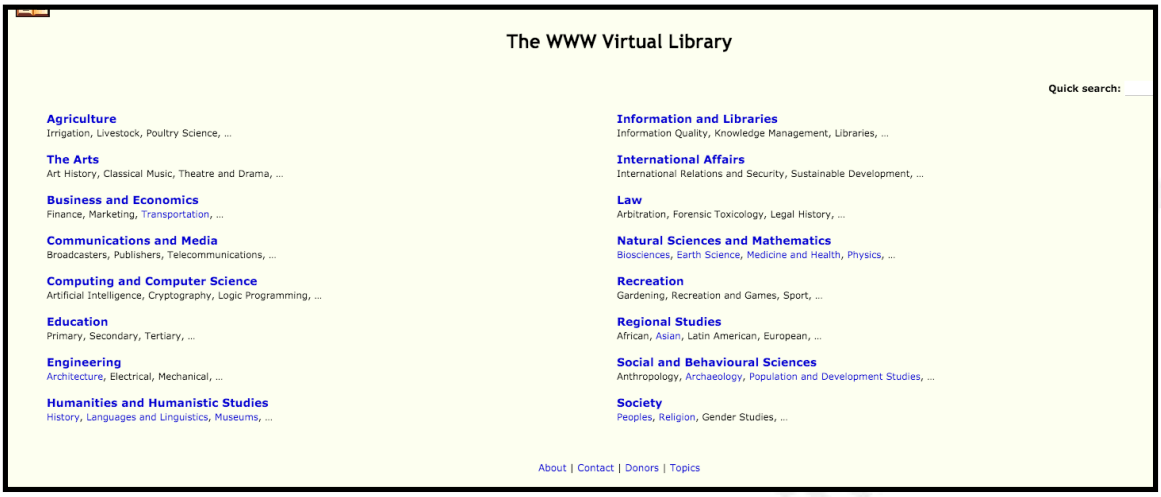


Figure 8 - Main page of The WWW Virtual Library

These search engines have the capability to talk to the Deep Web Hidden Service via Tor and its relays, resolve the .onion address, then return the content to a normal web browser being used on the Surface Web (Krishnan, 2016). There are both good guys and bad guys that use one of these to methods to access the Deep Web.

2.4. The Deep Web Population

There are several different types of people that access the Deep Web. They all have different motives and reasons. However, they can be split into 2 categories. They are either a “good guy” or a “bad guy”.

2.4.1 The Good Guys

There are non-criminals that do access the Deep Web. One example is somebody who purely just wants to access the web anonymously. They have no criminal motives or intentions, but they just don't want to be tracked by the public or private sector. Some others that use the Deep Web are military, police, and journalists ("Everything You Need to Know on Tor & the Deep Web", 2013). One of the most notable Information Security journalists that will use the Deep Web in part of his research for articles is Brian Krebs. Brian’s website is krebsonsecurity.com, and is a very notable online blog for Information Security news. Military and police will use the Deep Web to monitor for suspicious activity,

© 2016 The SANS Institute Author retains full rights.

such as the selling of drugs, PII, and guns. Law enforcement will also interact with the sellers of these items as undercover agents in order to catch them in the act. The sellers of these items are classified as the "bad guys".

2.4.2 The Bad Guys

Those that use the Deep Web with malicious or criminal intent are the "bad guys". Some of the most common activities they will participate in are the selling of illegal items, such as drugs, malware, and more. Cyber-criminals also use the Deep Web to communicate anonymously with each other and to sell stolen information, such as credit card numbers and health records.

3. Conclusion

The Deep Web is the largest part of the Internet, yet the majority of the population doesn't even know about it, or even access it. It can be used for good and for bad, legal and illegal activity. It is important to understand that it is not all bad. There is plenty good about the Deep Web, which includes the right of privacy when surfing the Internet. The understanding of the Deep Web and its capabilities is vital to the future of the Internet, and hopefully this paper helps accomplish that goal.

© 2016 The SANS Institute Author retains full rights.

References

Advantages And Disadvantages Of Tcp/ip Model. (n.d.). Retrieved March 22, 2016, from http://www.whatisnetworking.net/tag/advantages-and-disadvantages-of-tcpip-model/

Albanesius, C. (2013, October 3). *What Was Silk Road and How Did It Work?* Retrieved March 30, 2016, from http://www.pcmag.com/article2/0,2817,2425184,00.asp

Andrews, E. (2013, December 18). *Who invented the internet?* Retrieved March 22, 2016, from http://www.history.com/news/askhistory/who-invented-the-internet

AS Computing - Unit 2 The Internet. (n.d.). Retrieved March 22, 2016, from http://www.multiwingspan.co.uk/as2.php?page=internet

Crawling & Indexing – Inside Search – Google. (n.d.). Retrieved March 23, 2016, from https://www.google.com/insidesearch/howsearchworks/crawling-indexing.html

Cox, J. (2015, October 20). *The Dark Web Is Becoming a Safe Haven for Malware.* Retrieved March 29, 2016, from http://motherboard.vice.com/read/malware-is-using-the-dark-web-to-stay-hidden

Deep web. (2016, March 26). Retrieved March 27, 2016, from https://en.wikipedia.org/wiki/Deep_web

Dube, R. (2014, October 31). *Journey Into The Hidden Web: A Guide For New Researchers*. Retrieved March 29, 2016, from http://www.makeuseof.com/tag/journey-into-the-hidden-web-a-guide-for-new-researchers/

Epstein, Z. (2014, January 20). *How to find the Invisible Internet*. Retrieved March 21, 2016, from http://bgr.com/2014/01/20/how-to-

© 2016 The SANS Institute Author retains full rights.

access-tor-silk-road-deep-web/

Everything You Need to Know on Tor & the Deep Web [Web log post]. (2013, December 17). Retrieved March 31, 2016, from http://www.whoishostingthis.com/blog/2013/12/17/tor-deep-web/

Hjelmvik, E. (2013, April 6). *Detecting TOR Communication in Network Traffic* [Web log post]. Retrieved April 24, 2016, from http://www.netresec.com/?page=Blog&month=2013-04&post=Detecting-TOR-Communication-in-Network-Traffic

Ingevaldson, D. (2015, March). *What's Lurking in the Deep End of the Internet?* Retrieved March 29, 2016, from http://www.wired.com/insights/2015/03/whats-lurking-deep-end-internet/

Internet Live Stats - Internet Usage & Social Media Statistics. (n.d.). Retrieved March 21, 2016, from http://www.internetlivestats.com/

Krishnan, R. (2016, February 10). *Deep Web Search Engines to Explore the Hidden Internet*. Retrieved March 31, 2016, from http://thehackernews.com/2016/02/deep-web-search-engine.html

Leiner, B. M., Cerf, V. G., Clark, D. D., Kahn, R. E., Kleinrock, L., Lynch, D. C., . . . Wolff, S. (2016). *Brief History of the Internet*. Retrieved March 22, 2016, from http://www.internetsociety.org/internet/what-internet/history-internet/brief-history-internet

List of Tor hidden services. (2016, March 30). Retrieved March 30, 2016, from https://en.wikipedia.org/wiki/List_of_Tor_hidden_services

Nex. (2012, December 3). *Skynet, a Tor-powered botnet straight from Reddit* [Web log post]. Retrieved March 29, 2016, from https://community.rapid7.com/community/infosec/blog/2012/12/06/skynet-a-tor-powered-botnet-straight-from-reddit

© 2016 The SANS Institute Author retains full rights.

Reese, S. (2016, January 16). *Detecting Tor traffic with Bro network traffic analyzer* [Web log post]. Retrieved April 24, 2016, from https://www.rsreese.com/detecting-tor-traffic-with-bro-network-traffic-analyzer/

Reporter, D. M. (2013, October 11). *The disturbing world of the Deep Web, where contract killers and drug dealers ply their trade on the internet.* Retrieved March 29, 2016, from http://www.dailymail.co.uk/news/article-2454735/The-disturbing-world-Deep-Web-contract-killers-drug-dealers-ply-trade-internet.html

Scharr, J. (2013, October 23). *What Is Tor? Answers to Frequently Asked Questions*. Retrieved March 31, 2016, from http://www.tomsguide.com/us/what-is-tor-faq,news-17754.html

Surface web. (2016, March 12). Retrieved March 23, 2016, from https://en.wikipedia.org/wiki/Surface_web

The Hidden Wiki. (2016, January 11). Retrieved March 30, 2016, from https://en.wikipedia.org/wiki/The_Hidden_Wiki

The size of the World Wide Web (The Internet). (n.d.). Retrieved March 23, 2016, from http://www.worldwidewebsize.com/

The Ultimate Guide to the Invisible Web. (2013, November 11). Retrieved March 27, 2016, from http://oedb.org/ilibrarian/invisible-web/

Top 10 Uses of Internet. (n.d.). Retrieved March 22, 2016, from http://www.top10base.com/top-10-uses-internet/

Tor: Hidden Service Protocol. (n.d.). Retrieved March 30, 2016, from https://www.torproject.org/docs/hidden-services.html.en

© 2016 The SANS Institute Author retains full rights.