Issues in Information Systems
Volume 16, Issue III, pp. 230-237, 2015

INTO THE DEPTHS OF THE INTERNET: THE DEEP WEB

Marcus G. Tapia, Texas A&M University – Kingsville, marcus.tapia@gmail.com
Jack Shorter, Texas A&M University – Kingsville, jack.shorter@tamuk.edu

ABSTRACT

At the heart of the Deep Web lies a complex system of routing that maintains anonymity while accessing the recesses of the Internet. Most of the Deep Web is fairly innocuous, filled with data from prominent organizations such as the National Oceanic and Atmospheric Administration. These databases aren't normally accessible via the regular Internet. Black markets run rampant, with people from around the world distributing all manners of illegal goods. The advent and increasing popularity of Bitcoin helps energize these markets. Bitcoin is revolutionary in that it is completely digital, and not controlled by any regulating agency. These black markets are slowly being curtailed by joint task forces from cooperating federal agencies from around the world. Dozens of arrests have been made using evidence found by infiltrating the black market organizations. The anonymity offered by the Deep Web can work against the perpetrators of illicit activities. During civil unrest, the Deep Web has been useful in protecting backlash and exposing corruption and unfair political practices. Despite the appeal of Internet anarchy, some users of the Deep Web have exhibited humanitarian qualities by helping this little known aspect of the Internet become a safer place. Uneasy is the future of the Deep Web, and as more attention is drawn to it, it is very difficult to say whether or not it will be restricted or further embraced by those looking for a safer place to share information.

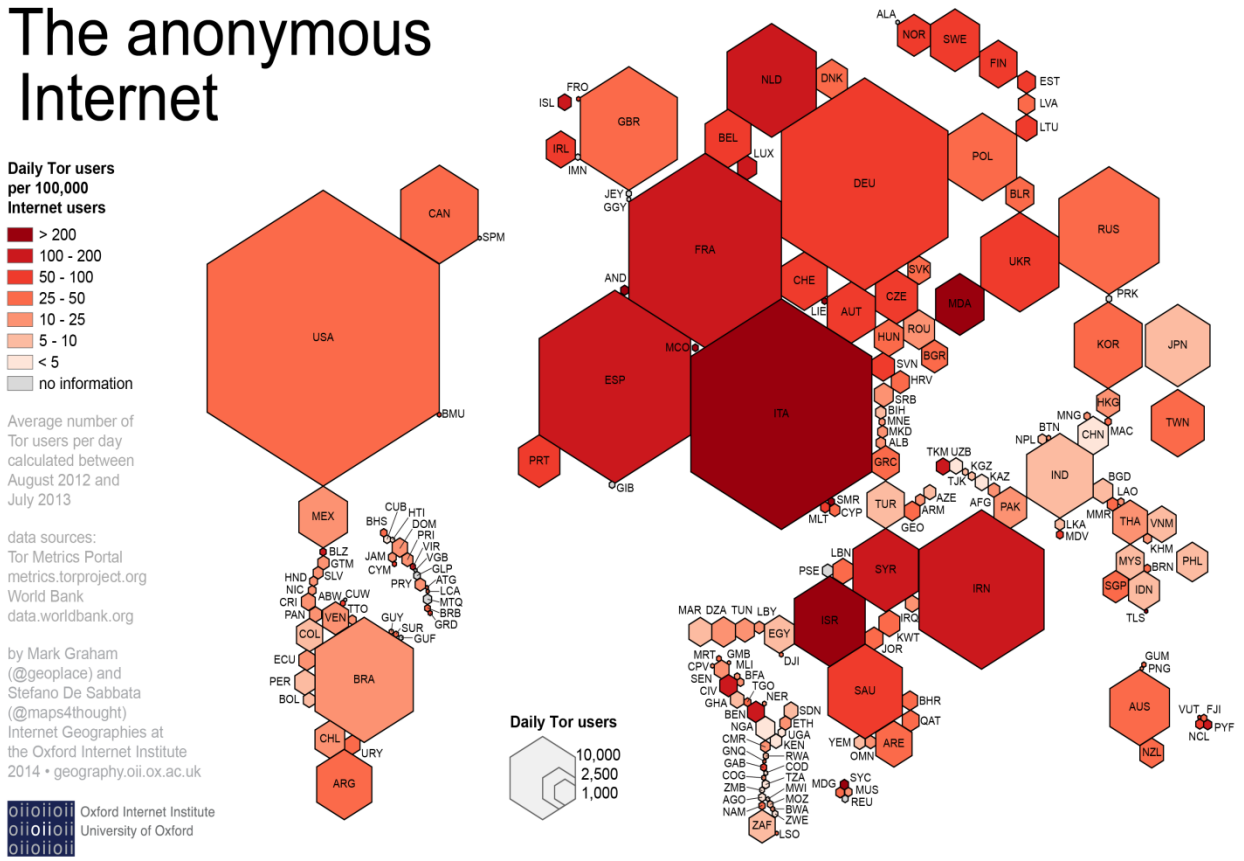
Keywords: Deep Web, Bitcoins, TOR Network, Routing/Proxy Nodes, Invisible Web

INTRODUCTION

An ocean is a huge expanse filled with various forms of life. While we tend to know a great deal about this life, there are vast amounts of life forms that have never been discovered. The unseen life forms are analogous to the unimaginable amounts of information stored on the Internet. There are enormous amounts of data that exist on the Internet that are not accessible to the casual browser. The ways websites become publicly available today are affected by search engines. Two prime examples are Google and Yahoo. Most search engines use a web crawler to identify and index linked pages and then make these pages available through their respective web sites. The search net cast by these services creates what many perceive to be the entirety of the Internet. The contents of the Deep Web are not included in the search results of conventional search engines. The crawlers of conventional search engines identify only static pages and cannot access the dynamic Web pages of Deep Web databases. Hence, the Deep Web is alternatively termed the "Hidden" or "Invisible Web." The term Invisible Web was coined by Dr. Jill Ellsworth to refer to information inaccessible to conventional search engines. [10] An exponentially larger amount of data is available that exists just out of reach of most users of the Internet. This assemblage of data repositories is referred to as the Deep Web. There are basically two types of data in the Deep Web, structured and unstructured content, with the former being found in databases which have been compiled by various organizations throughout the world. [5] The latter are individual sites that can only be accessed from a direct *.onion link. However, web crawlers are unable to access databases, which are only accessible by specialized queries and are said to make up about 54% of the Deep Web. [13] While most of the Deep Web is filled with legitimate information, there also exists a dark underbelly. The propagation of a new form of digital currency, Bitcoin, fuels illegal markets. Law enforcement agencies struggle to close the most prominent markets in order to keep some kind of order, but when one market is shut down, another one takes its place. Due to the high profile of government seizures, many think that users of the Deep Web are only there for criminal activities.

This portion of the Internet can also be used for positive purposes. The Deep Web has become an outlet for those who wish to expose the injustices in their country, and the world. Whistleblowers can release information freely without fear of retaliation, if they are careful not to implicate themselves. The danger of course comes when what you are exposing involves matters of national security. The measures that users of the Deep Web trust are very complex when you factor in the concept of anonymity. These sites are intermittent, and cannot be searched for or found by using a search engine. Because of the difficulty of accessing these resources, special measures must be

taken if one wants to browse the Deep Web. One website you can access to travel to the dark side is a deep web news portal called Hidden Wiki. You are instructed to install a Tor Browser from http://torproject.org/. This then gives you the ability to use search engines designed to search the Deep Web. [8] There are hundreds of URLs on this one page alone, which just boggles the mind when you think about the actual size of the Deep Web.



(https://en.wikipedia.org/wiki/Onion_routing)

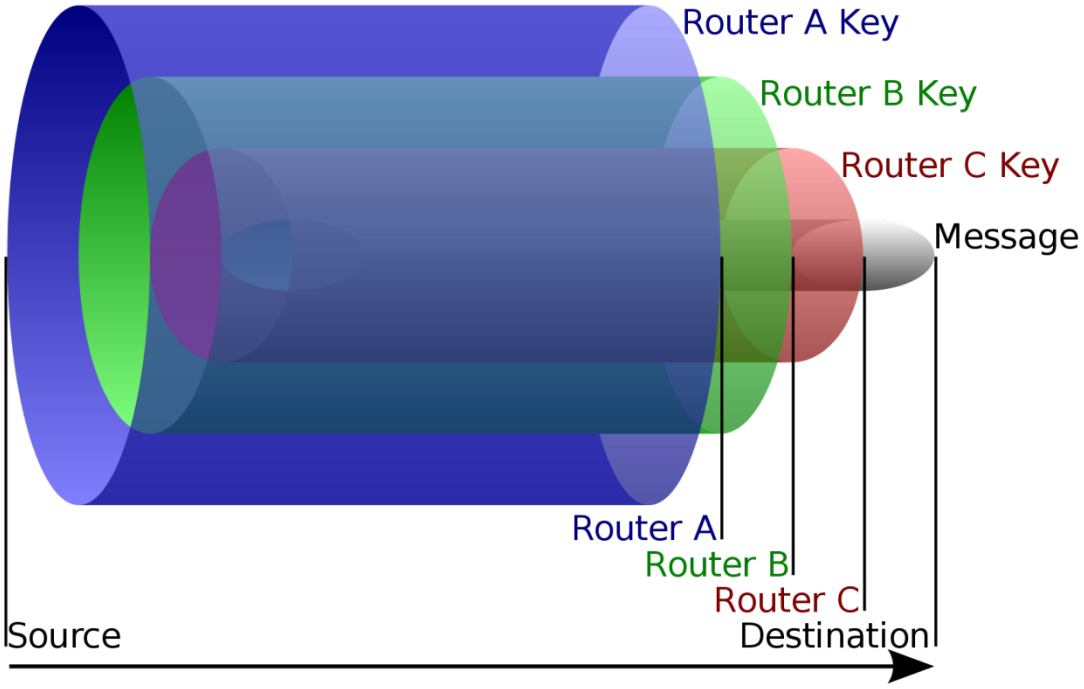
THE TOR NETWORK

The key element when attempting to access the Deep Web is to utilize a special category of web browser, called the TOR browser. The Onion Routing program, or TOR, was originally developed for use by government programmers. Tor enables users to surf the Internet, chat and send instant messages anonymously, and is used by a wide variety of people for both licit and illicit purposes. It was developed by the US Naval Research Laboratory to create a discrete method of sharing classified information among authorized groups. It has since been embraced by those requiring a more secure and anonymous connection, whether it be for positive or negative purposes. The download of the TOR browser, had almost doubled in August of 2014 [2].

Normal web browsers are deemed too insecure because they create a direct connection between the user and the information. The Onion Routing program creates a truly anonymous connection between the user and the data that will be accessed. TOR uses a series of nodes, known as the TOR network, which tends to obfuscate the connection. The two types of nodes used by TOR are the routing nodes and the routing/proxy nodes. Routing/proxy nodes exist on the fringes of the network and are the entry points for users into the Deep Web. The use of proxies makes this connection method compatible with almost all protocols such as HTTP. Because the connection is more complex, it is resistant to surveillance and cyber-attacks. The nodes have the capability to delay, reorder, and cushion the data with encryption to further complicate the web traffic. The reason for the peculiar name "onion" is the encapsulation process that the packets endure. Much like onions the packets have multiple layers. The layers consist of encryption

that is added by the routing/proxy nodes. The structure of the onion relies on the path that the onion has to travel in order to get to its destination. Each node is only aware of neighboring nodes. Only the previous node and the subsequent node are known by routing nodes. Past one hop, the nodes are literally in the dark about other nodes on the network, whether it be the number of nodes or location of said nodes. The onion loses layers of encryption as it travels, but with each step a random string the size of the removed layer gets added to the end of the packet, so that the path cannot be determined by the size of the onion. [7] Onion traffic is made uniform by the initial proxy by adding the appropriate amount of padding based on the portion of encryption that was removed. Depending on whether it's going forward through the node or backwards, an appropriate sized bit string is added. Virtual circuits are created and data is transferred freely between sender and receiver. Although this complex routing process makes the connection very secure, the connections within the Deep Web are not untouchable even from denial of service attacks. The nodes themselves could be altered to put a halt to all connections that pass through the node. The highest liability occurs if an affected node is a proxy/routing node. Proxy/routing nodes are critical for the security of the connection, and can reveal the entire route if they are compromised. If the pathway is exposed, then the connection becomes more akin to a regular Internet connection. It is theoretically possible to find patterns within the connections, but the sheer amount of data that needs to be monitored makes this method impracticable. The Deep Web is a complex system of connections that creates an extremely secure circuit. It maintains both the anonymity of the sender and receiver as long as no nodes have been compromised along the way. Although the negative use of the Deep Web is what receives the spotlight, most of it is fairly innocuous.

That being said, using TOR does not guarantee security. Jon Fingas writes in an article in April of 2015 that your Tor-based email isn't as secure as you think. An email service SIGAINT has had to warn their users that someone launched a sustained attempt to break inot its email servers and was looking at messages. The attack was not successful, but the perpetrator tried to set up a malicious exit node in order to send data to the normal internet in an attempt to try to spy on messages the moment they left Tor. [6]



In this example onion, the source of the data sends the onion to Router A, which removes a layer of encryption to learn only where to send it next and where it came from (though it does not know if the sender is the origin or just another node). Router A sends it to Router B, which decrypts another layer to learn its next destination. Router B sends it to Router C, which removes the final layer of encryption and transmits the original message to its destination. (https://en.wikipedia.org/wiki/Onion_routing)

THE INNOCENT SIDE OF THE DEEP WEB

The majority of the information on the Deep Web is simply a repository for legitimate purposes. Not only does the NOAA store their information in the repositories that comprise this Web, but universities may store their research and businesses may store their employee information, such as payroll and internal information, in databases stored in the Deep Web. This makes it a target due to the fact that others may want access to this employee information. Also stored in the Deep Web are Web pages that contain linked pages that are protected behind a username and password. Additional data that is passed over by web crawlers in the Deep Web would be any information that requires a subscription to access and anything that relies on a CAPTCHA to authenticate. CAPTCHA is the acronym for Completely Automated Public Turing test to tell Computer and Humans Apart.

These structured databases of information are immune to web crawlers used by search engines, and are only accessed by the institutions that own the servers on which they are stored. The reason that the web crawlers are ineffective is because they act like a spider in the sense that they crawl along the web by following linked pages. These pages are linked by the hyperlinks that connect them. Another reason pages end up being lost in the Deep Web is at times the web crawlers ignore URLs that contain a large amount of characters that are not found in traditional addresses. This is most, if not all accessible pages within the Deep Web. Because of the modular nature of databases, they are not linked by normal means. They require queries that are designated by the database administrators. These queries are able to retrieve data based on a specific search that the user requests. These pages are unable to be detected.

Another thing to consider is the fact that TOR itself is not just useful for illegal purposes, its inherent security capabilities allow users to bypass traffic monitoring used by third party marketing companies. These companies utilize this data to forecast prices and monitor traffic patterns to adjust their own marketing plans. Many view this as a violation of privacy, especially when most of the time, the fact that companies are tracking activity is freely offered information.

On the Deep Web, individuals are free to post without any fear of censorship. The right to privacy is constantly being challenged. It is unfortunate that many, who specifically visit the Dark Web, use it for its abundance of illegal marketplaces that exist there.

RISKS OF ILLEGAL MARKETS

Why use such a complicated method of creating a secure connection? There are all manners of nefarious things that can be found on the Deep Web. The appeal of a truly anonymous Internet connection makes it an apt place to conduct criminal activities. From counterfeit money to drugs, the Deep Web offers many online storefronts in which purchasing a pound of marijuana is as simple as buying a book from Amazon. Dozens of drug markets exist on the Deep Web. One of the biggest markets was recently shut down for the second time by the FBI, known as 'Silk Road'. Foxnews.com reports that the first iteration generated an estimated $1.2 billion dollars from 2011 to 2013, $80 million of which went back to the site owners as commission. [14] Officials were unsure of what to do with such a huge amount of bitcoins. They had confiscated over 30,000 bitcoins from the site closure. The bitcoins were valued at $17 million the day it was sold. It was meant to be sold off in nine different blocks. However, one bidder, venture capitalist Tim Draper, had the winning bid. The amount that was paid for the bitcoins was not released. [9] Counterfeiting has also found a home in the Deep Web. Counterfeit passports, birth certificates, and even currency can be purchased for numerous countries. In a time when there is a heightened fear of terrorist attacks, the ease that counterfeit documents can be purchased is exceptionally troubling. There are many examples of brazen assassins offering their services on the Deep Web. One Dailymail.co.uk article gives examples of mercenaries for hire. $10,000 can buy an assassination in the United States, and in Europe, it can be contracted for $12,000. The option of making the assassination look like an accident or a suicide is available as well. Another

assassination site offers a betting pool that users can access to bet on contract completion, with the closest guess winning the pot. [4] The pressure put on by the closure of Silk Road sadly has done nothing to curb the instances of these services being offered elsewhere online. Another prevalent market place on the Deep Web deals with the firearms industry. Because of the ease of acquiring firearms in the US, this doesn't really create a concern in the United States. However, in other countries, such as Great Britain, this could cause an influx of illegally obtained weapons. In most countries, anyone with an Internet connection can receive weapons in the mail.

Hackers for hire are available as well. With bitcoins or Western Union wire transfers, anyone can hire a very powerful hacker to infiltrate and compromise secure information, such as one's email account or online banking account. The worst of the worst, however, is child pornography. Thankfully, the FBI has committed itself to pursuing these criminals to the full extent of the law, and despite all the security that the Deep Web offers, many prominent proliferators of this horrible crime have been jailed. The Deep Web seems like a version of the Wild West, where with enough money, one could buy almost anything. Because anonymity is crucial for transactions that involve illicit activities, a new unregulated, unrestricted method of payment is desired.

BITCOIN: THE MOST VOLATILE OF CURRENCIES

Bitcoin is the currency that runs the Deep Web. This revolutionary form of currency is one hundred percent digital. It was created in 2009 by Satoshi Nakamoto, but Bitcoin was first envisioned by Wei Dai in 1998. At this point there are several developers working on it. The design of the Bitcoin specification is entirely open source, which means that anyone can view the code and adapt it to suit their needs. The changes that can be made must conform to a set of rules agreed upon by all the developers that are working on it. The strength of Bitcoin depends on the cooperation of all the developers and any dissonance makes it unusable. Bitcoin.org makes the comparison between the protocols behind email and Bitcoin, in that no one person owns the rights to the protocols themselves. The user is free to allocate their bitcoins according to their needs, while each transaction is stored in something called a "block chain". The block chain is publicly available while the signatures of sender and recipient remain anonymous. Authenticity is preserved by digital signatures that accompany the transactions. Processing transactions includes the reward of bitcoins. In order to "mine" for bitcoins the user has to allow their computer to help with the load of transactions by running specialized hardware on their personal machine. This is not the only way to receive bitcoins; there are Bitcoin exchanges and anyone can accept Bitcoin as a form of payment for a service online. The growth of Bitcoin has expanded exponentially since its inception. [9] The businesses that have adapted to accept it is not limited to the Deep Web, various real world businesses have begun to accept Bitcoins as payment. Bitcoin.org estimates that the value of all bitcoins in circulation exceeded $1.5 billion dollars US in 2014. There are many advantages to using Bitcoin as a type of currency. The security and ease of use in regards to no limitations on when or where money can be sent make it an ideal form of payment to utilize in the Deep Web. One enormous downside of using bitcoins is its volatility. Because of the limited amount of bitcoins in circulation, small or large transactions, either buying or selling bitcoins can severely impact the value. For example, the auctioning off of the bitcoins recovered from the shutdown of Silk Road caused the value to increase substantially. While this form of payment is ideal for transactions on the Deep Web, it of course adds to the danger of what one can purchase with money. [9] Like TOR, Bitcoin is not completely untraceable. Recent history shows that the transparency of Bitcoin once one end of the transaction is compromised allows criminals to be exposed.

WHAT ARE SOME WAYS TO FIND INFORMATION IN THE DEEP WEB?

Experts who study and research the Deep Web say that the typical search, lets choose Yahoo! And Google are only able to gather approximately 1 to 2 percent of the available information on the Internet. The rest is hidden in the deep web. Some people also refer to this area as the invisible web. So how is it possible for you to find the other 98 or 99 percent of the information on the Internet? The answer is Meta-Search Engines such as SurfWax, Academic Index, Dogpile, Ujiko, and Freebase. Meta-search engines use the resources of many different search engines to gather the most result possible. Another way is to use Semantic search tools and Databases such as Hakia, Zotero, DBpedia, Evri, and Boxxet. Semantic Search Tools depend on replicating the way the human brain thinks and categorizes information to ensure more relevant searches. You can also use Academic search engines and databases such as Google Scholar, WorldCat, Microsoft Libra, BASE –Bielefeld Academic Search Engine, and Intute. The world of Academia has many databases not accessible to search engines such as Google and Yahoo!. So using the just mentioned search engines should help if you are looking for Scholarly information. [11]

A SPOTLIGHT ON THE DEEP WEB

What has brought a lot of attention to the Deep Web was the high profile seizure of a marketplace, called Silk Road. Silk Road version 1 alleged to be the brain child of Ross Ulbricht. He was arrested in October 2013 and was charged with not only creating and running the site, but also attempted murder. He had hired a contract killer, who was actually an undercover FBI agent. The government still had around $144,000 in bitcoins after the auction of the confiscated bitcoins. Several arrests followed, as the FBI gained access to the site's sellers. The perpetrators were tied to the distribution of methamphetamines, cocaine and heroin. [1] Along with the list of sellers, the FBI also gained access to extensive customer feedback that created some legitimacy to the sellers that made Silk Road thrive. Because of Bitcoin's transparency, officials were also able to track buyers as well. Another case that exposed the Deep Web was the seizure of Silk Road 2.0. This operation was a joint operation between Europol and the FBI. The mastermind behind this iteration of Silk Road was said to be Blake Benthall. Like Ulbricht, Benthall faces a litany of charges. These include but are not limited to conspiracy to traffic narcotics, conspiracy to commit computer hacking, and money laundering. He faces life in prison. Silk Road 2.0 was reported to be generating $8 million dollars a month. Benthall sealed his own fate by unwittingly giving moderator privileges during the initial phases of the site. Officials traced the server down to another country and created an image. He was further implicated when Google surrendered evidence that directly tied his email account to the compromised server. Called "Operation Onymous", it was hinted that further arrests would be made that would "disrupt global activity" in the depths of the Deep Web. [3]

A LIGHT SHINING IN THE DARKNESS

The temptation of pursuing illegal activities on the Deep Web is difficult to overcome. That doesn't mean everyone uses the Deep Web for illegal purposes. Installing the TOR browser does not make you a criminal. For those who believe that modern day patriots come in the form of whistleblowers, the deep web can be considered a safe haven to expose corruption in high levels of government and business. There are countless forums dedicated to preserving free speech. Citizens of the United States may take free speech for granted, but in other countries where free speech is not met with a positive reaction, the Deep Web can be an effective platform for ideas and criticism of the status quo. During the revolution in Syria in 2011, the Deep Web was utilized to host videos of the atrocities of the Syrian people. One can argue that the Deep Web was "vital in the Arab Spring uprising, by allowing dissidents to communicate and unite without being detected." [15] Most recently in Turkey, the Prime Minister completely prohibited the use of Twitter during the civil unrest that occurred there. In these instances, where anonymity is literally life or death, the Deep Web can be used as a place where justice can be safely advanced.

The right to anonymity and privacy is a hot button issue for many, especially with the news that NSA may be monitoring much more than is being reported. The exposure that the NSA surveillance has received is due mainly in part to individuals that felt that this organization was invading the privacy of millions around the world. Though not as prominent as patriots fighting for their rights, some people use the deep web to help others. One hacker took it upon himself to take over the Deep Web's unofficial start page The Hidden Wiki. At first glance, this may seem like a typical illegal activity that is found on the Deep Web. The Hidden Wiki contains links to various sites where child pornography is available, so the hacker held the Wiki hostage, and removed all the links that lead to these sites. It is very honorable for someone who has the capability to expose those committing heinous crimes such as child pornography to do so. Even on the Deep Web. Another Deep Web do-gooder provides his service as a "trained harm reduction specialist" in order to educate the masses who purchase drugs illegally on marketplaces such as Silk Road. He takes his time to teach the dangers of mixing certain drugs, and the effects of prolonged use of drugs when suffering from mental illness. Even in the illegitimacy of the Deep Web, some choose to help.

THE FUTURE

Despite the negative use of the Deep Web, users are drawn to the security that the TOR browser offers. Large corporations track user habits and sell that knowledge to the highest bidder. Users of the Internet are going to the deep web to fulfill their desire for a private browsing experience. Cyber-attacks of corporations are more pronounced. It is believed by one survey that 94% of organizations have had their security compromised in the past year, up 9% from previous surveys. [12] For corporations that have a strong web presence, this can be a massive blow. If consumers lose faith of shopping online, a large stream of revenue could disappear. When each successful

cyber-attack can cost a company millions of dollars, including loss of consumer trust, there is a need for additional security measures. Unfortunately mix network configurations, such as onion routing can be very slow and unreliable due to the haphazardness of available nodes. It is simply not feasible for some businesses to have the anonymity that is offered by TOR. Despite its inherent security flaws, TOR remains for the most part very robust, and as long as there are people who want to hide their web activity, there will be a place for the Deep Web. With the crackdown of the FBI, many view the Deep Web as no longer being as secure as it once was, so many are speculating that the Age of the Deep Web is coming to an end.

CONCLUSIONS

The Deep Web is a topic that has so many areas for future research; the problem is narrowing down what aspect of it to study next. Possible directions for future research are: the advantages and disadvantages of the AIW (Academic Invisible Web) to Universities and Government organizations, database security and the Deep Web, privacy issues and the Deep Web, the astonishing amount of information that normal search engines miss that are stored in the Deep Web, criminal enterprise's lurking in the Deep Web, cyber wars and terrorism and the role of the Deep Net in fighting those battles, and a myriad of other topics too numerous to mention here.

Although the Deep Web has a very dark side, Internet users are still drawn to it because of the security that truly anonymous web browsing brings. The majority of the Deep Web is actually filled with databases stuffed full of information stockpiled by academic organizations.

Many are drawn to the Deep Web because of TOR. TOR's advantages are a complicated routing system that makes traffic monitoring almost impossible. Disadvantages include the exposure of routing once the path has been exposed. Once the route is realized, the traffic can be monitored, as no additional encryption exists along the path.

It is truly appalling that the reason most utilize the Deep Web involves the suffering of people, whether it is child pornography, illegal arms deals, or contract killers. No wonder many honest caring people despise the Deep Web. When cyber-terrorism is on sale to the highest bidder, it raises concerns about all security measures currently in place. Are they every going to be enough? The Deep Web is a place where one can go to fight persecution in all forms. While the future of the Deep Web seems uncertain, with large court cases drawing more and more attention to it, the increasing popularity of the TOR browser, with its promise of privacy and security, ensures the Deep Web will not be going away in the foreseeable future, if ever.

REFERENCES

1. Associated Press. (2013) Alleged mastermind of online drug bazar due in court following worldwide arrests. *Fox News* [online]. Available: http://www.foxnews.com/us/2013/10/09/alleged-silk-road-mastermind-due-in-court-amid-arrests-worldwide-alleged-users/
2. Chirpse, S. (2014) Ultimate Guide to the Deep Web. Available: http://www.sickchirpse.com/deep-web-guide/
3. Cook, J. (2014) FBI Arrests Former SpaceX Employee, Alleging He Ran The 'Deep Web' Drug Marketplace Silk Road 2.0. Available: http://www.businessinsider.com/fbi-silk-road-seized-arrests-2014-11
4. Daily Mail Reporter. (2013) The disturbing world of the Deep Web where contract killers and drug dealers ply their trade on the internet. Available: http://www.dailymail.co.uk/news/article-2454735/The-disturbing-world-Deep-Web-contract-killers-drug-dealers-ply-trade-internet.html
5. Deep Web: A Primer. (2014) Available: http://www.brightplanet.com/deep-web-university-2/deep-web-a-primer/
6. Fingas, Jon. (2015) Your Tor-based email isn't as secure as you think. Available: http://www.engadget.com/2015/04/26/tor-email-service-faces-attack/
7. Goldschlag, D., Reed, M., & Syverson, P. (1996) Hiding Routing Information. Available: http://www.onion-router.net/Publications/IH-1996.pdf
8. Hidden Wiki .onion Urls Tor Link Directory. (2015) Available: http://thehiddenwiki.org/
9. Hill, K. (2014). Silk road bitcoin auction winner Tim Draper won't say how many millions he paid. Available: http://www.forbes.com/sites/kashmirhill/2014/07/02/tim-draper-silk-road-bitcoin-auction/

10. Iffat, R. and Sami, L. (2010) Understanding the Deep Web. *Library Philosophy and Practice 2010.* ISSN 1522-0222. Available: http://www.webpages.uidaho.edu/~mbolin/iffat-sami.htm
11. Miller, Alisa. (2015) 100 Useful Tips and Tools to Research the Deep Web. Available: http://www.online-college-blog.com/features/100-useful-tips-and-tools-to-research-the-deep-web/
12. Mobile Enterprise. (2014) Nearly 100% of Organizations Cyber Attacked. (2014, October 28). Available: http://mobileenterprise.edgl.com/news/Nearly-100--of-Organizations-Cyber-Attacked96205
13. Pagliery, J. (2014). The Deep Web You Don't Know About. Available: http://money.cnn.com/2014/03/10/technology/deep-web
14. Rogers, J. (2014) Dark nets: Murky recesses of the hidden web. Available: http://www.foxnews.com/tech/2014/10/25/darknets-murky-recesses-hidden-web/
15. Yeung, P. (2014) . A Tour of the Best, Entirely Legal Hangouts on the Deep Web. Available: http://motherboard.vice.com/read/the-legal-side-of-the-deep-web-is-wonderfully-bizarre