DURBAN UNIVERSITY OF TECHNOLOGY

NETWORK ANALYSIS OF DARK WEB TRAFFIC THROUGH THE GEO-LOCATION OF SOUTH AFRICAN INTERNET PROTOCOL ADDRESS SPACE

GOKHALE, CRAIG

2019

Network Analysis of Dark Web Traffic through the Geo-Location of South African Internet Protocol Address Space

By

Craig Gokhale

(21649731)

Submitted in fulfilment of the requirements of the

DOCTOR OF PHILOSOPHY IN INFORMATION TECHNOLOGY AND COMMUNICATION TECHNOLOGY

In the

DEPARTMENT OF INFORMATION TECHNOLOGY

In the

FACULTY OF ACCOUNTING AND INFORMATICS

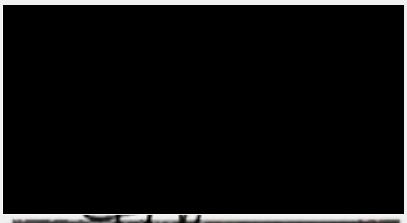
At

DURBAN UNIVERSITY OF TECHNOLOGY

2019

DECLARATION

I, Craig, Gokhale hereby declares that this dissertation is my own work and has not been previously submitted in any form to any other university or institution of higher learning by other persons or myself. I further declare that all the sources of information used in this dissertation have been acknowledged.

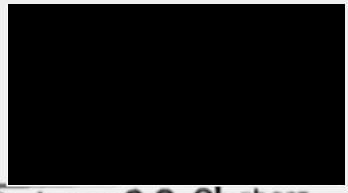


Craig, Gokhale

16/04/2019

Date

Approved for final submission

Supervisor: 

Professor O.O. Olugbara
PhD (Computer Science)

16/04/2019

Date

ACKNOWLEDGEMENT

Early in the process of completing this project, it became quite clear to me that a researcher cannot complete a thesis alone. Although the list of individuals I wish to thank extends beyond the limits of this format, I would like to thank the following persons for their dedication, prayers and support: Professor Oludayo Olugbara for helping, guiding and supporting me in this course from the beginning to the end. Dr. Kiru Pillay for all the help, guidance and motivation through this journey. A heartfelt thank you to Mr. Alan Walker for the countless hours he sacrificed in ensuring the experiment was a success. To my girlfriend, Kressantha for her undying love and support through this journey, thank you. To the BankSeta, Council for Scientific Industrial Research and the National Research Fund for funding this study. I am happy to have good friends and dear family with whom I have shared life, feelings and experiences. To my parents, family and friends for your support, help and understanding during this journey I undertook. Thank you all.

LIST OF ABBREVIATIONS AND ACRONYMS

CEM	Child Exploitation Material
FBI	Federal Bureau of Investigations
ISIS	Islamic States of Iraq and Syria
IP	Internet Protocol
I2P	Invisible Internet Project
NIC	Network Interface Card
SADC	South African Democratic Countries
SNA	Social Networking Analysis
TOR	The Onion Router

LIST OF FIGURES

LIST OF TABLES

TABLE OF CONTENTS

ABSTRACT

This research was supported financially by the BankSeta, the Council on Scientific and Industrial Research and the National Research Foundation with the aim to log The Onion Router (TOR) traffic usage in South Africa. The recent public disclosure of mass surveillance of electronic communications, involving senior government authorities, has drawn the public attention to issues regarding Internet security privacy. For almost a decade, there has been several research efforts towards designing and deploying open source, trustworthy and reliable electronic systems that ensure anonymity and privacy of users. These systems operate by concealing the true network identity of the communicating parties against eavesdropping adversaries of which TOR is an example of such a system. Clients that use the TOR network construct circuits (paths) which are utilised to route multiple network streams. A circuit is considered secure if there is one non-malicious router in the circuit. Such systems have served as anti-censorship and anti-surveillance tools.

The implementation of TOR allows an individual to access the Dark Web, an area of the Internet that is said to be of a much larger magnitude than the Surface Web. The Dark Web which has earned a reputation as a sort of immense black market, associated with terrorist groups, child pornography, human trafficking, sale of drugs, conspiracies and hacking research, has received significant national and international press coverage. However, to date little or no research has been conducted on the illicit usage of the Dark Web and no research has been conducted in the use or misuse of the Dark Web in South Africa. There has not been any study which characterises the usage of a real deployed anonymity service. Observations obtained are presented by participating in the TOR network and the primary goal of this study is to elicit Dark Web traffic by South Africans. Past researchers undertook Dark Web crawling focusing only on specific web content such as explicitly focusing on child exploitation and terrorist activity. The experiment design of this study further builds on experiments conducted in previous studies. The deanonymisation methodology utilised in this study will allow for the detection of exit routing traffic and the logging of all Dark Web traffics areas omitted from the previous studies.

This study does not confine the declassification of onion addresses to specific content types and aims to log all exit routing traffics, undertake a comprehensive declassification of websites visited by clients and obtain the Internet Protocol Addresses (IP) of these clients. The analysis of the sample results reveals that in the South African context, Dark Web traffic is mainly directed to social media websites. There are however causes for concerns as there are illicit activities occurring that include the sale of drugs, visiting of child pornographic websites, and the sale of weapons. Finally, the study presents evidence that exit routing traffic by the TOR node is limited to a large number of different countries some of which have serious Internet censorship laws.

CHAPTER ONE

1.1 Background Information

With the proliferation of cyber-attacks and the increasing creativity of cyber criminals, organisations and nation states are increasingly finding themselves vulnerable and at risk (Ablon, 2014). The Internet technology is the medium for the new forms of criminal activities these days. Network communication systems that hide an individual's Internet Protocol Address will prevent the determination of the source or destination of messages (Ablon, 2014). The technology stack makes it difficult for site owners and technology experts to identify and trace the client systems. Grabosky (2014) stated that there is little or no differentiation between virtual crime and crime committed in the real world. "'Virtual criminality' is basically the same as the terrestrial crime with which we are familiar, whilst some of the manifestations are new a great deal of crime committed with or against computers differs only in terms of the medium. While the technology of implementation and particularly its efficiency may be without precedent the crime is fundamentally familiar. It is less a question of something completely different than a recognisable crime committed in a completely different way."

Individuals are familiar with the online world where Google, Facebook and other common websites exists, however beneath this layer lies a network of sites where individuals openly engage in illicit activities (Weimann, 2016). Weimann (2016) notes the terms Deep Web or Dark Web as the layer of the Internet that conventional search engines like Google cannot index and the Internet could be dissected into different layers namely the Surface Web which is accessible by conventional search engines and the Dark Web which is not accessible by conventional Internet browsers like Google (Weimann, 2016).Wright (2017) noted the Dark Web as being a portion of the Deep Web where illicit materials and Dark markets exist and can only be accessed by a specialised browser, hence the Dark Web contains materials such as weapons and drugs for sale, child pornography, leaked information and money laundering. The size of the Dark Web is almost impossible to measure as majority of the content is hidden, however researchers such as Barker (2016) estimate the Dark Web to be in excess of 400 times larger than the Surface Web with only 16% of the Surface Web being indexed by Google.

Standard search engines do not index Dark Web content and therefore a specialised browser such as The Onion Router (TOR) or The Invisible Internet Project (I2P) has to be utilised. TOR was initially developed to allow for anonymous communication by the creators, the United States (U.S) Naval Research Laboratory (Fineka, 2016). Communication on the Dark Web can be undertaken by individuals using platforms such as chat rooms and emails hosted on TOR whilst maintaining their anonymity, thus making it extremely difficult to track and trace the client systems (Finklea, 2016). Like any technology that was previously discovered, anonymity can either be used for good or bad or both, therefore users turn to the Dark Web for freedom of speech as they fear political retribution for their actions on the Surface Web (Chertoff, 2015). Chertoff (2015) further notes the urgent need for the global community to evaluate the impact of the Dark Web due to the impact it has on nation states and the urgent need for its inclusion in Internet governance.

Previous studies Burch (2014), McCoy (2014), Westlake (2017) and Adler (2012), firstly reduced the set of relays to monitor, then identified the source of the anonymous traffic. A study by Burch (2014) and Bauer (2014), reveals that efforts have been made to decrease the number of relays to be searched and then explored various methods to identify an anonymous user in the network connections utilising the routers and relays. Few researchers have thus far tried to comprehensively implement a complete attack which involves identifying the source of the anonymous communication and the illicit activities being engaged in on the Dark Web. Furthermore, to date, there is little or no evidence of any studies which characterise the use of an anonymity service. Thus, this study aims to identify the illicit trade conducted by South Africans on the Dark Web. The study also provides results of how the Dark Web is being used and misused by clients. Finally, the study presents the exit routing traffic by country with the geo-location of the IP addresses and discovered the utilisation of TOR by countries that have serious Internet censorship laws.

1.2 Motivation of the Study

Studies undertaken by Moore (2016), Westlake (2017) and Chen (2014) added value to forensic investigations, however a potential gap exists between aggregate studies and studies engaging in illicit activity. There is no experimental architecture that

efficiently records illegal activities across various crime types. This thesis is focused upon enhancing techniques for the monitoring of illicit activities on the Dark Web and exit routing traffic by country.

There have been a few prior studies that aimed to ascertain the content type that is located in an onion ecosystem. These include studies that measured and analysed the Dark Web drug marketplaces (Christin 2013; Soska and Christin 2015b), a study that have attempted to exploit flaws in TOR hidden services in order to try improving its performance (Biryukov, Pustogarov, and Weinmann 2013; Biryukov et al. 2014; Owen and Savage 2016). Systems such as DeepDive (Niu et al. 2012), have been used to analyse Dark Web content, however the content needs to be available before conducting a crawl and an analysis. Christin (2015) undertook an analysis of the sellers of Silk Road (Christin 2015). Soska (2015) thereafter conducted a longer-term study that measured vendor activity across various marketplaces. Biryuokov (2014), attempted to exploit flaws in TOR's hidden service protocol and tried to measure how popular the various onion services were and thereafter deanonymise them.

This study unlike the prior efforts utilises strategies such as the setting up of a TOR exit node on the TOR network to enable Dark Web crawling and to undertake a DNS trace in order to capture onion addresses that conform to TOR ethical research guidelines (Tor Project 2015). The experiment would result in the declassification of all onion addresses obtained by the exit node and not only those confined to certain content types such as drug trade and child pornography as conducted by Christin (2015, b), Westlake (2017) and Chen (2014). The study methodology would allow for determining exit routing traffic by country and obtaining the geo-location of the IP addresses, an area omitted from all related research previously done on this topic. This study will therefore for the first time undertake Dark Web crawling, the declassification of all extracted content and the monitoring of exit routing traffic by country utilising South Africa as the country of origin.

1.3 Aim and Objective of the Study

The overarching aim of this study was to log traffic usage and determine the use or misuse of the Dark Web in South Africa. There is a need for further research in the deanonymisation of anonymous communication networks and Dark Web crawling,

where the content is not limited to only a few themes as research conducted by Chen (2014), Moore (2016) and Westlake (2017). This study will present a high-level experiment methodology building upon experiments as conducted by McCoy (2014), Dingeldine (2014) and Murdoch (2012) that will allow for the logging of TOR traffic usage utilising South Africa as the country of origin. Based on the research, for the first time the usage of illicit trade within this network will be identified utilising South Africa as the country of origin. The output of results will not be limited to certain content types as presented in studies by Chen (2012) and Westlake (2017). All websites that are visited by clients will be logged and analysed and a further analysis of illicit trade engagements will be undertaken. All exit routing traffic to the various countries will also be logged and these countries will be further analysed into countries associated with extremist groups, and Internet censorship countries. The acquisition of individuals Internet Protocol addresses (IP) will also be obtained but not presented in the study. In collaboration with changes in the South African legal landscape (implementation of laws and structures to govern South African Internet use) it is likely that the study will provide an independent analysis and justification for some of the controls proposed.

1.4 Statement of the Problem

In recent years there has been an increase in the number of cyber-attacks and the creativity of cyber criminals has drastically developed over the years. This has led to organisations and nation states finding themselves increasingly vulnerable and at risk. The increase in uptake on the usage of anonymous network communication systems makes it extremely difficult for site owners and technology experts to identify and trace the client systems. The technology stack makes it extremely difficult and costly to deanonymise anonymous communication networks as the TOR network utilises three layers of encryption which makes it extremely difficult to deanonymise. There is an increasing amount of pressure on organisations and nation states to improve their Internet governance policies that consider the usage of anonymous communication systems such as TOR (Weimann, 2016). The use of the Dark Web as a tool by extremist groups to communicate and recruit has gained much publicity off late. Extremist groups such as Islamic State of Iraq and Syria (ISIS) have made it publicly aware that they utilise the Dark Web as a means of communication and recruiting (Weimann, 2016). Many of these extremist groups' websites have been taken down by

the hacking group, Anonymous. The capture of Ross Ulbricht in 2013, the brains behind the infamous Silk Road, where an estimated 144,000 Bitcoins, nearly 100 million dollars at today's exchange rates was seized by the Federal Bureau of Investigation (FBI) (King, 2016). The capture of Peter Scully who was most notorious for the snuff movie "Daisy's Destruction", where it was estimated that he sold a copy of the movie to clients for 10,000 dollars each. The list of events such as the above has led to increased pressure from various countries to have the Internet regulated and new policies developed in ensuring technology stacks such as TOR are taken into consideration.

1.5 Research Questions

There has been minimal research that has been undertaken to truly analyse the usage of the Dark Web by clients in various countries. The research in this domain is limited to a few studies that truly attempt to deanonymise anonymous communication systems such as Murdoch (2012), Dingeldine (2014) and McCoy (2014). This research against this background research attempts to answer the following questions:

1. How is TOR primarily being used in South African?
2. How is TOR being misused in South Africa?
3. Who is using the TOR network in South Africa for what purposes and to what extent?

1.6 Research Methods

To simplify the complexity of defining the main experiment for the study tool, which would enable the researcher to answer the research questions, there were three stages of the sub experiment that were performed as shown in figure 1.1. The network will consist of a TOR exit node which will run the latest version of TOR. Some of the hardware components to be utilised in building the TOR exit node will comprise of a Front-End Dark web pair routing device and a Client Management System. A TOR licence has to be purchased in Bitcoins in order for the node to be placed on the TOR network. A TOR license will enable the logging of exit routing traffic, logging of onion addresses and the obtaining of IP addresses from clients accessing the Dark Web. The implementation of the geo-based firewall will ensure that only South African Dark Web

traffic usage will be logged. The data collected from study will be deployed in a big data analysis engine for pattern detection and analysis utilising open source and proprietary software. Gephi will be used to determine Networking patterns and basic statistics about the network.

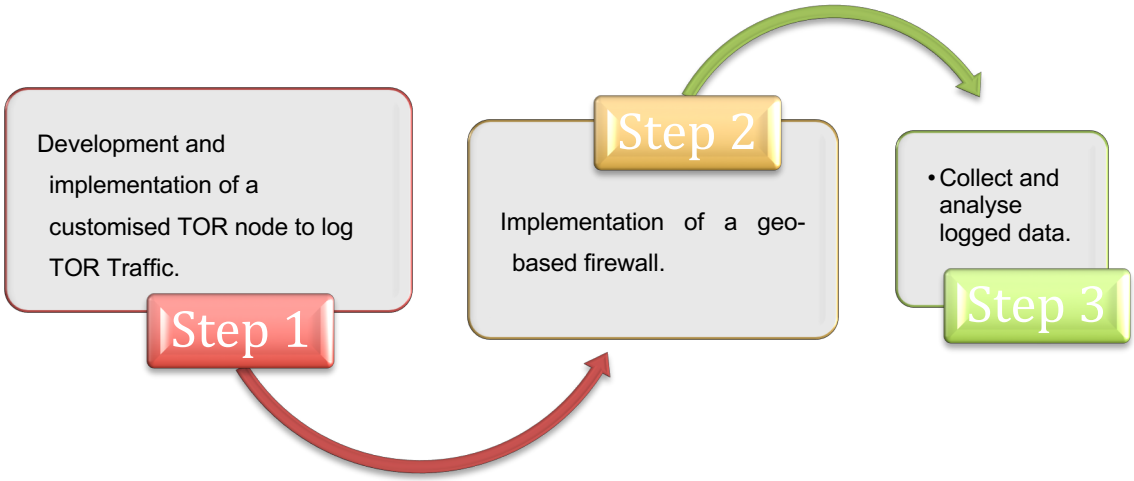


Figure 1. 1: Experimental Design

Figure 1.2 is a graphical representation of the experimental architecture for the logging of TOR traffic and determining Dark Web usage. TOR has three layers of encryption and to access Dark Web content a client has to pass through three points on the TOR network. The three stages of the sub experiment begin with:

1. The Entry Guard. The point of entry into the TOR network. The client system which represents an end user will try to gain access to the Dark Web utilising a TOR browser and the client system connects to the entry guard to access the TOR network. At this point the experiment will log the users IP address.
2. Middle Relay. The middle relay is the relay which receives traffic and passes it on to another relay on the TOR network. The client will connect to the TOR network through the entry guard and thereafter passing through a middle relay. The purpose of this middle relay is to try and hide the clients IP address as TOR has three layers of encryption, this is the second layer of encryption.
3. Exit Relay. This is the final relay that TOR traffic passes through reaching its destination. The vulnerability in TOR is the exit relay which reconnects the user back to Internet services. It is at this point the exit node will be placed to allow for the logging of exit routing traffic, onion addresses visited and the geo-location of IP addresses. This will also allow for the logging of the onion addresses that

TOR users visit on the Dark Web. The utilisation of declassification tools will assist in categorising these websites into specific content type.

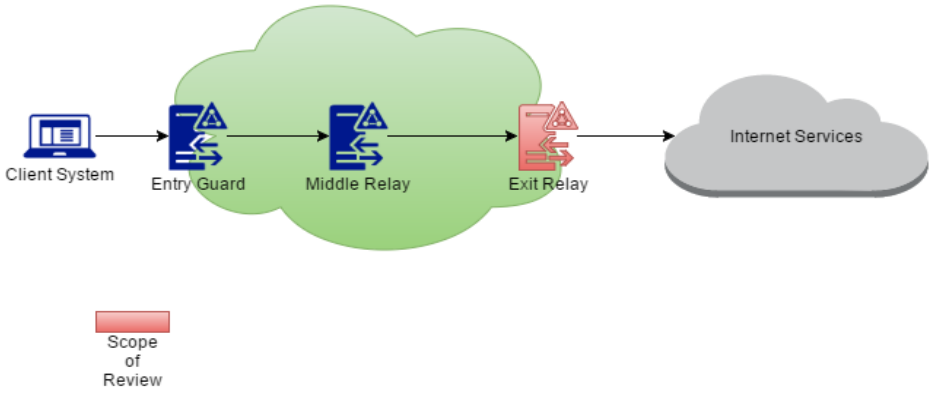


Figure 1. 2: Experimental Architecture

The declassification of websites will be undertaken by utilising both open source and proprietary website declassification tools. These tools will allow for the declassification and categorisation of websites against certain content types. The declassification tools utilised in the study confirm a 99% accuracy rate when declassifying websites. For the deanonymisation of communication networks and the logging of exit routing traffic, Gephi will be utilised. Gephi is a software that is open source and is widely used for network visualisation and analysis. For the exit routing traffic by country, Gephi will be utilised to analyse the data set and to present basic statistical analysis of the data.

The methodology of this study comprises of three main components which are the focus group1, experimentation and focus group 2. Figure 1.3 shows the overall process followed to execute the method of the study.

1. Focus Group 1: In the focus study, three specialists in cyber security were engaged to solicit the configuration settings for the TOR node. These specialists had prior experience on TOR logging and Dark Web traffic analysis as this was a task, they had previously undertaken for the South African Police Service and South African Government. The author is of the opinion that the research reported in this thesis would benefit from the experiment
2. Experiment: In the experiment the TOR exit node was conceptualised, designed, implemented and ran for a period of 20 days as conducted by Bauer and McCoy (2008) and Dingeldine (2015). The first week of the experiment was utilised as a

test period and the data collected from the test period was not utilised in the study. Bauer and McCoy (2008) and Dingeldine (2015) ran the testing period for a period of three days. The reason for a longer test period in this study was attributed to the TOR node being utilised not only for logging of TOR traffic but also attaining the geo-location of IP addresses and a comprehensive Dark Web crawl which the studies mentioned above did not take into consideration. There were concerns from the security specialists in focus group 1 with regards to undertaking the experiment for that period of time as the exit node will come under attack by ransomware and other viruses, therefore this test period will allow for the constant monitoring of TOR traffic and any issues that arose, contingency measures could be put into place.

3. Focus Group 2: The results of the experiment were presented to the South African Development Community (SADC) cyber security group, which comprised delegates from 14 countries who debated topics on the cyber security landscape in various countries. The South African contingent comprised members from the South African State Security Agency, Chief Director of Information Security Services, Council for Scientific and Industrial Research and the General of the South African Army.

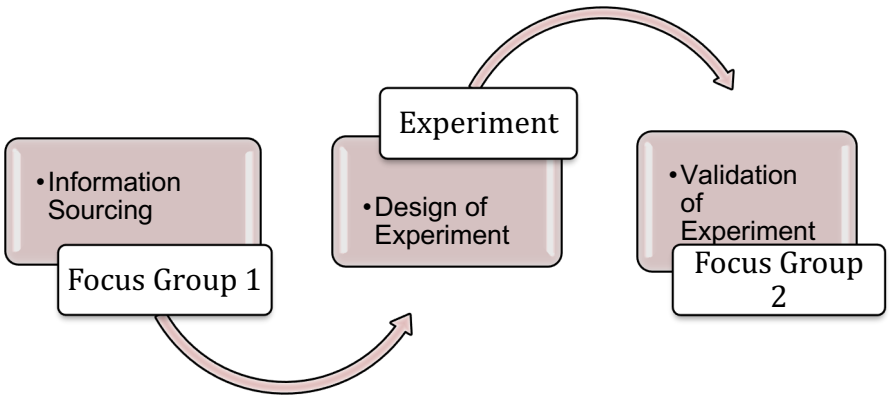


Figure 1. 3:Overview of Research Methods

1.7 Contribution

This study will contribute to literature on previous studies by providing a comprehensive overview of Dark Web usage in South Africa. Many nation states have asked for some controlling policies that govern the usage of TOR. The results obtained from the experiment could contribute to some of the controls governing TOR usage.

The unique contributions when compared to the previous related studies conducted by McCoy (2008), Dingeldine (2015), Chen (2014) and Westlake (2017) of this study are enunciated as follows:

1. An experiment architecture was developed to enable the logging of exit routing traffic by country, log the IP addresses of the clients and log Dark Web Traffic. The architecture of the experiment conceptualised and designed in this study will further improve on a similar experiment conducted by McCoy (2014).
2. Determining the exit routing traffic by country. A comprehensive analysis of all exit routing traffic through the exit node was undertaken and this provided an overview on the countries that South Africans are connecting to on the TOR network.
3. Determination of the primary usage of the Dark Web by South Africans. For the first time to reveal if the usage of the Dark Web is purely for illicit activities as per previous research conducted by Dingeldine (2014), Murdoch (2014) and Westlake (2017) or the Dark Web in South Africa is purely used for non-illicit activities such as the accessing of educational resources and engaging in educational forums.
4. For the first time, the implementation of a Dark Web crawl to ascertain if South Africans are engaging in any illicit activities on the Dark Web.

1.8 Structure of Thesis

Chapter one provides a brief overview of the TOR browser and the increase in the number of illicit activities individuals engage in on the network. The chapter also provides an explanation on the need and aim for the study. This is built on citing previous studies that have undertaken and the identification of the shortcomings of those studies, hence leading to this current study. The chapter further presents the research questions that the experiment will seek to address and the contribution of the work that will assist in building the current literature on the Dark Web. Chapter Two is a comprehensive literature review, which will present a research taxonomy and a theoretical framework for the study and the reasons for the chosen taxonomy and theoretical framework. The chapter will also provide an in-depth description on the uses of the Dark Web, how to access it and the need for policy development and controls

for anonymous communication networks. The chapter also outlines studies related to this thesis and the need for a new model.

The new model proposed is built on shortcomings of the existing frameworks in related research that has been conducted. Chapter Three will present a comprehensive research methodology for the study. The chapter will present a comprehensive overview of the experiment design and the reasons for the chosen framework. The data collection methodology will also be presented and the tools that were utilised in the declassification of the websites will be discussed. Screen captures of the actual server logs will be presented so as to provide an insight into the large volumes of traffic that passed through the exit node. In Chapter Four, the results of the experiment are analysed and presented in table of figure form. The data presented will depict exit routing traffic by country, the websites that were visited by clients during the experiment and a Gephi diagram that represents the social networking analysis for the experiment. Chapter Five presents an analysis of the data. The analysis chapter aims to link the results of the experiment to previous literature as presented in the literature review. The chapter will also present an analysis of findings and possible shortcoming of previous studies undertaken. The chapter will further present the focus group responses that were received and analysed during the South African Development Community (SADC) cyber security conference Chapter Six provides recommendations based upon the findings for future research. The chapter will emphasise the need for future studies based on the findings. Illicit usage will reinforce the need for state security agencies to undertake further analysis of TOR usage.

CHAPTER TWO

LITERATURE REVIEW

2.1 Introduction

The discussion of related studies presented in this chapter shows originality, relevance and relatedness of this study to the previous research and to justify the suitability of the study methodology. The Dark Web is a layer of the Internet that is not visible to predominant search engines such as Google, Internet Explorer and others. There are various activities that goes on in the Dark Web and these vary from the freely available educational resources to illicit activities such as the selling of drugs, firearms and the huge databases of child pornography that is easily and readily available, hence the Dark Web is closely associated with criminal underground activities. A conventional browser such as Google cannot access Dark Web content and in order to access the Dark Web, specialised browsers such as TOR and the Invisible Internet Project (I2P) has to be utilised. TOR is a browser that allows individuals to browse the Internet whilst remaining totally anonymous.

2.2 Related Studies

Network traffic analysis is an attack technique that aims to identify the various patterns of communication amongst system entities (Charavarty, 2014). Traffic analysis may also be defined as a process of examining and intercepting messages that aims to obtain information from the patterns in communication and that anonymous traffic exiting the exit node could be a traffic analysis attacker (Charavarty, 2014). The traffic analysis attacker may be able to correlate traffic patterns and also observe traffic flowing to the server and the exit node (Charavarty, 2014), hence a way to identify traffic on anonymous communication systems such as TOR will be to identify the clients of the exit node that eavesdrop on traffic exiting the exit node. By undertaking a networking analysis on TOR, the geo-location of IP addresses may be determined. This traffic will be logged at the exit node of the experiment.

TOR is a browser that allows individuals to browse the Internet whilst hiding their IP addresses therefore remaining totally anonymous. TOR was chosen for this study as it was the first Dark Web browser created and the Dark Web originated from its creation. The browser is also the most popular for those wanting to access the Dark Web. The network analysis for this study will concentrate on the TOR network (Dingeldine, 2014). This analysis will ascertain the exit routing traffic by country and who is accessing the Dark Web. The exit routing by traffic will also determine the Geo location of the Internet Protocol address space.

Previous research conducted in this field was first undertaken by Murdoch (2012) where the deanonymisation of anonymous communication was undertaken in two stages; the initial stage involved the identification of the anonymity set. Murdoch (2012) further noted that the second stage, the observer monitors the main set and the other set having the anonymised identities, however there is no way to determine the relationship between the two sets. A flaw in the experiment methodology utilised by Murdoch (2012) found the monitoring of all networks for traffic identification was not feasible for adversaries that were powerful.

McCoy (2014) attempted to log traffic passing through the TOR network and also attempted to log Dark Web usage for a period of 20 days. The researcher setup a 1GB/s network link connected to a router and logged only 20bytes in order not to breach any laws. The setup of the experiment was conducted separately, and the researcher utilised both an exit router and non-exit router. The logging of traffic was undertaken for 15 days and logged both entry and middle node traffic. A tcpdump was utilised for logging exit routing traffic over the router and the first 150bytes of the packets packet was logged. For protocol analysis, Ethereal was used and for exit routing traffic the router relayed 709GB of traffic however only the first 150bytes of the packets was logged, this will limit the number of websites that could have been classified. The experiment conducted did not log any exit routing traffic to any country. The experiment was undertaken in two stages, where the entry and middle routing traffic was logged for 15 days the exit routing traffic for only 5 days. A 1GB/s network, therefore limiting the amount of logged web traffic.

A study by Murdoch (2015) also attempted to log TOR traffic only and not attempt to classify any Dark Web activity. The experiment conducted in the study revealed that through simulations, a set of IXes, can observe traffic exiting and entering

several TOR relays within the UK. Murdoch (2015) further revealed through simulations, that NetFlow, a traffic monitoring system installed in commodity routers, could be used to launch analysis attacks against TOR. The results were however, mostly based on simulations involving data obtained from by observing a single TOR relay. The results were based on simulations and not actual data received and this was a flaw in the study. The experiment undertaken only focused on entry and exit nodes within the U.K and there were no attempts to log exit routing traffic to any other country. The experiment did not log any Dark Web usage in the U.K and no classification of Dark Web activities were analysed.

Houmansadr (2014) conducted an experiment of the TOR network with the aim to provide performance improvements on TOR. The application layer software, namely, Skype was executed in a virtual machine and a Linux machine. Virtual machines were then connected to virtual distributed ethernet. The researchers constructed a plugin which could modify packet contents and drop packets at variable rates. A central switch that had switches connected to it and provided DHCP connectivity to the Internet. The study purely focused on providing performance improvements of TOR by developing an algorithm. There was no logging of exit routing traffic nor the attempt to log any Dark Web activities.

A study conducted by Chaabane (2012) undertook a traffic analysis to monitor traffic of TOR in five different countries. The experiment utilised HTTP and BitTorrent protocols. The researcher designed and monitored multiple TOR nodes implemented in France, USA, Taiwan, Germany and the US with the allocation of 100kB bandwidth for a period of 23 days. A total estimate of 20GB of data was received by the servers on each day at the exit and the entry relays. The bandwidth allocated to the server was minimal and this had an effect on the amount of traffic that was logged. The greater the bandwidth allocated to the server, the greater the amount of traffic will pass through it. The experiment did not classify any Dark Web usage by undertaking a Dark Web crawl.

In a study undertaken by Bai (2014), the capturing of TOR traffic was attempted. The study aimed develop a methodology for the capturing of TOR traffic. The traffic that passed through the server was dummy traffic. The experiment utilised 8 computers with one running Java Anonymous proxy and the other TOR. Only the capturing of dummy traffic was collected and analysed from the 6 computers utilising Ethereal. The use of Winsock Packet Editor recorded packets that were generated by specific

applications, with the test duration being 120 minutes that included five repetitions. The total test period was 120 minutes, therefore the time allocated to the experiment was insufficient. The traffic was not actual TOR traffic but dummy traffic that was allowed to run through the testbed.

A study conducted by Barker (2014) developed a TOR setup in order to try to analyse TOR network analysis and Dark Web usage. The use of a Selenium browser executed 170 simulations and as a result managed to access 30 websites. A total of 15 relays were configured for the experiments, with an additional three directory servers. The private TOR network collected and analysed both HTTPS and HTTP traffic. The amount of website traffic that was logged was minimal as only 30 websites were logged during the experiment. There was also no attempt to log any exit routing traffic to any countries nor the attainment of the IP addresses of the individuals accessing the network.

In an experiment undertaken by Tang (2014), a node of his own, which acted as the middle node was setup. The aim of the experiment was to develop an algorithm that will help improve the performance of TOR. From the directory of TOR nodes, the researcher selected both entry and exit nodes and the use of a PlanetLab testbed was not required as the nodes were provided with only 100KB/s. The download of the target file was conducted by using Webfetch from the web server. The experiment undertaken in this study did not attempt to log any Dark Web usage nor exit routing traffic. The primary aim of the study was to develop an algorithm to improve TOR's performance. There was no entry and exit node utilised in the experiment, entry and exit nodes enable the geo-location of IP addresses and the ability to undertake a Dark Web crawl.

Alsabah (2014) aimed to enhance TOR's performance by utilising traffic classification in real time. The experiment collected offline data off 200 circuits that were from three different application traces across the world. These three applications which comprised of a stream client, web browsing and BitTorrent client was configured on a machine to only use a specific TOR entry node. The collection period for data was 24 hours over 6 weeks with various intervals. The experiment only utilised an entry node to monitor traffic and no exit node was utilised. The utilisation of an exit node will enable the logging of Dark Web usage. The experiment conducted in the study only aimed to log traffic entering the TOR network so as to provide recommendations on TOR's performance improvements.

A study conducted by Winter (2014) aimed to ascertain how China is blocking citizens from using TOR. The researcher deployed a relay in Russia then two bridges which were located in Sweden and Singapore. The hosting of a relay hosted on Amazon cloud was undertaken in Singapore. In Sweden and Russia, a bridge and relay were hosted by an organisation and a data centre and in order to obtain vantage points in China and a Linux sever with 32 SOCKS proxies was used. (Winter, 2014) undertook an investigation into China blocking TOR relays and bridges and the researcher showed TOR bridges were blocked by IP addresses and not by Tuple. Winter (2104) further noted that the researchers' experiment showed that the implementation of fragmentation would bypass China's firewall and there were no adversaries that conducted traffic fingerprinting for domestic traffic. This experiment undertaken was an attempt to ascertain how China is blocking TOR and there was only an attempt to log traffic to or out of China. The study did not attempt to ascertain China's TOR usage.

The study conducted by McCoy (2008) will provide the basis of this study. Further improvements will be made on the study conducted by McCoy (2008) by critically evaluating the experiment design and implementation. The proposed improvements to the experiment conducted in the study will entail but not limited to the following:

1. An entry node exit node and a man in the middle attack. This will enable all traffic that passes through the node to be logged
2. Greater Bandwidth allocation to the server which will enable the logging of more traffic.
3. Readjustment on the time allocated to capturing traffic passing through the entry node and exit node.

2.3 Dark Web Traffic

There are multiple means to obtain freely accessible information from the web, with the Internet just comprising a small portion of this information. Beneath the Internet lies multiple layers which are said to be much larger than the Surface Web.

Chertoff (2015) noted the use of specialised browsers in order to access this content and individuals are familiar with conventional browser such as Google. Chertoff (2015) noted that the "Dark Web refers to a class of content on the Internet that, for

various technical reasons, is not indexed by search engines and thus would not be accessible through a traditional search engine". Chertoff (2015) furthermore noted that the hiding of the Dark Web was intentional and the ability to share information whilst remaining anonymous is one of the key reasons why users access the Dark Web (Chertoff, 2015).

The web is growing exponentially with the US accounting for 100,000 new domains daily. In the same vain around 70 000 domains are taken offline daily, therefore an approximate 30,000 domain names are added daily. Due to content in the Dark Web not being indexed, accessing this content cannot be undertaken using conventional browsers. The Dark Web cannot be accessed by traditional search engines as the content in this layer of the web is not indexed and there is no static information found on the Dark Web as there is on the Surface Web and the emergence of new tools is making it easier to access the Dark Web (Dingeldine, 2015). There is no clear indication on the portion of Dark Web content is used for legal or illegal activities. The Internet as we know is broken into three categories namely the Surface Web, the Deep Web and the Dark Web as shown in the figure 2.1.

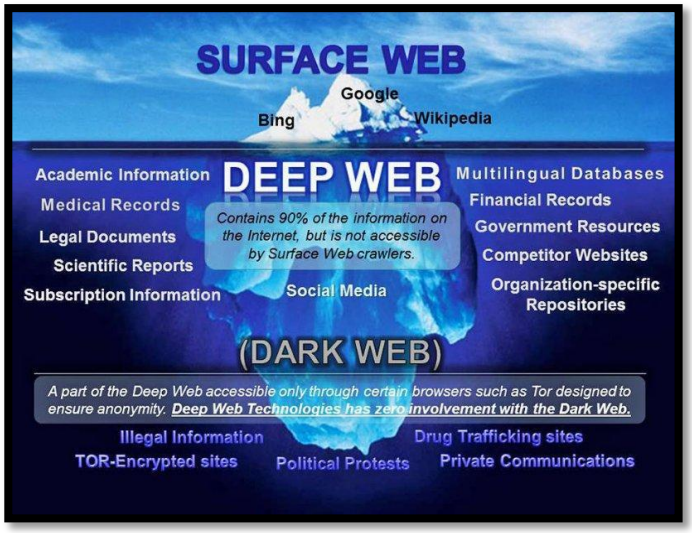


Figure 2. 1: Layers of the Internet (Source: Dingeldine 2015)

Individuals have confidence in engaging in illegal activities on the Dark Web due to the level of anonymity that TOR provides. The various illicit trade that occurs on the Dark Web is presented below.

2.3.1 Malware Trade

The Malware trade is a flourishing market on the Dark Web, and this is justified by the increase in Ransomware attacks. The total damage caused by the recent WannaCry virus was estimated at $4 Billion Dollars with the Malware trade ranging from data-stealing Trojans, Ransomware, RATs to ATM malware (Yaneza, 2014). The Dark Web and Malware are similar in relation to command-and-control infrastructure (Yaneza, 2014). Yaneza (2014) also noted that the use of cryptography and other tools by TOR and I2P is implemented so as to hide their server location and therefore it is extremely difficult for forensic investigators and researchers to investigate these applications using traditional methods (Yaneza, 2014). Utilising the various services and sites is an easy task and therefore a large number of cybercriminals access C&C using TOR (Hacquebord, 2015).

2.3.2 VAWTRAK

VAWTRAK like a CryptoLocker is also a Trojan horse, however a VAWTRAK is mostly distributed to users via Microsoft Word documents. VAWTRAK are mostly associated with phishing attacks, but the Trojan is capable of logging keystrokes, taking screenshots and also hijacking webcams. VAWTRAK malware are generally spread via phishing emails (Sancho, 2015). Sancho (2015) noted that a client will communicate with a C&C server where the IP addresses were retrieved by downloading a file *(favicon.ico)* on a TOR site, thus the anonymous advantage for the criminal server, however not for the already infected user. Sancho (2015) noted a *favicon.ico* file with

a C&C server and individual may search the sites from a list from the system, thereafter downloading the C&C server address everyday as shown in figure 2.2

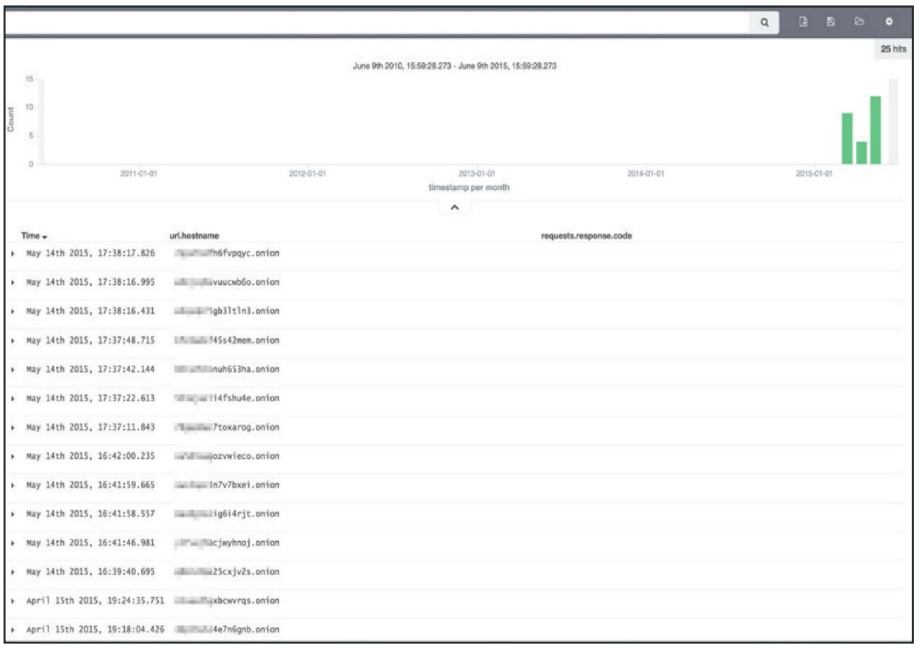


Figure 2. 2: VAWTRAK (Source: Sancho 2015)

2.3.3 CryptoLocker

CryptoLockers have gained much popularity over the years. A CryptoLocker is similar to a Trojan horse in that searches for files on your computer and then encrypts them and in order for one to decrypt these files, a ransom needs to be paid in order to obtain the private key (Sancho, 2015). A major malware that could be found in the Dark Web is a CryptoLocker, which is similar to a ransomware that also encrypts victims' personal documents, the decryption key will only be provided to the victim once payment has been made. Sancho (2015) further stated that a CryptoLocker can make adjustments to the payment page to the victims' home language automatically. Sancho (2015) notes a TorrentLocker utilises TOR in order to host various payment sites where Bitcoin is the only means of payment. This therefore reveals the reason cybercriminals appeal to the Dark Web as it provides a platform to make their infrastructures more robust and nearly impossible to takedown (Sancho, 2015). Figure 2.3 are screenshots that a Dark Web Analyser captured, with both soliciting different languages aimed at their potential targets. This provides an indication on the magnitude of the Ransomware attacks on the Dark Web where they are not confined to only a few countries.

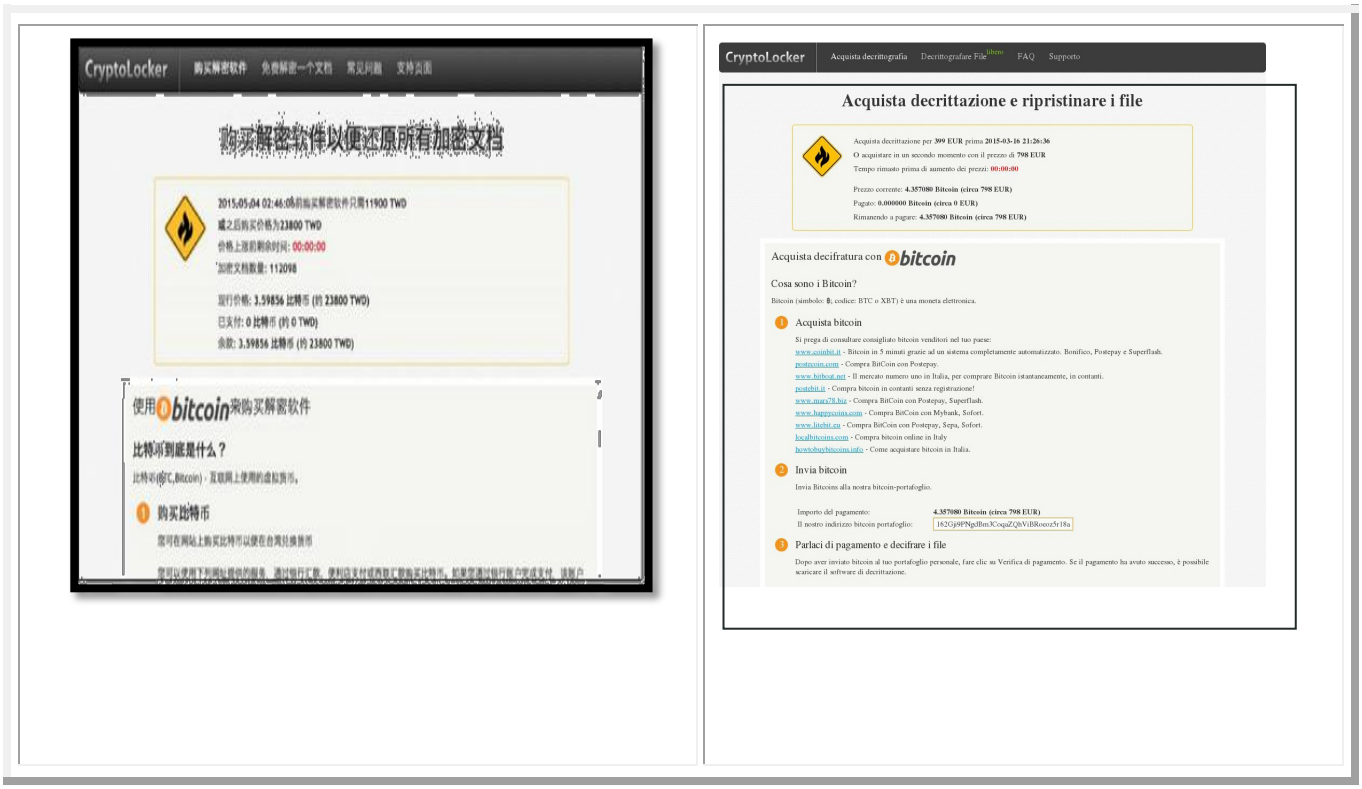


Figure 2. 3: CryptoLocker (Source: http://ndvgtf27xkhdvezr.onion)

2.3.4 Illicit Drugs

In every report on the Dark Web it is very common to talk about the free availability of weapons and illegal drugs. After the capture of Ross Ulbricht (Vinto, 2015), noted that procuring drugs on the Dark Web seems relatively trivial. Vinto (2015) further noted that the variety of drugs available on the Dark Web, with these ranging from the sale of conventional tobacco to the sale of more extreme items like cocaine. Figure 2.4 provides a screenshot of the two infamous online drug stores on the Dark Web. Silk Road realised 22 million dollars in annual revenue just by drug sales, whilst Silk Road operators earned 143000 dollars of commission per month, (Christin, 2015). With the closure of Silk Road emerged Silk Road 2 and Dolliver (2015) undertook research into Silk Road 2 and found approximately 1834 products for sale off which 348 were aligned to drugs and 145 vendors that shipped drugs to various countries with the primary destination and origin being the U.S. Aldridge and De´cary-He´ tu (2015) describe the results as being flawed and further research conducted by Van Buskirk et al. (2015)

found that Silk Road 2 discovered 9103 various drug items for sale from 519 different vendors.

With the closure of Silk Road emerged a new Dark Web site that deals in illegal drugs called Grams (Vinto, 2015). With a logo similar to Google, Grams is now the main site for those who wish to purchase such goods after the seizure and closure of Silk Road (Vinto, 2015). There are also TOR sites that possess information on cannabis growing houses which show live feeds on the temperature and moisture levels over time (Vinto, 2015).

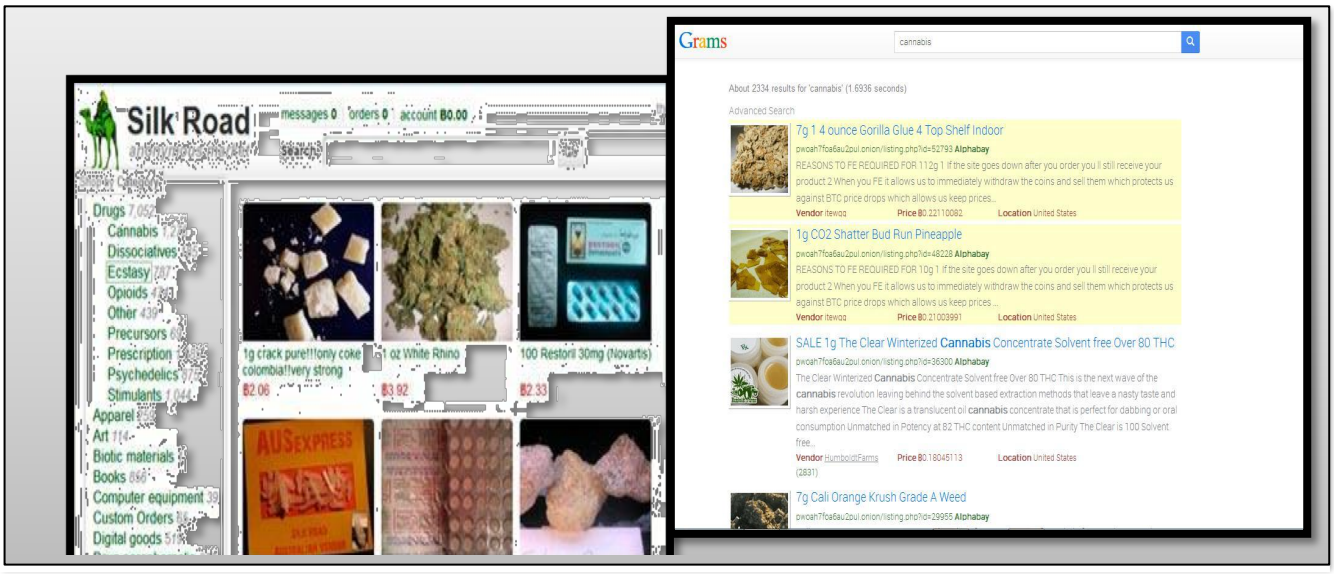


Figure 2. 4: Drug Trade (Source: Sancho 2015)

2.3.5 Money-laundering

In today's black markets individuals have numerous institutions in order to facilitate trade and to evade law enforcement. The emergence of cryptocurrencies such as Bitcoin has made transactions untraceable. Hardy (2016) noted the need for a new creation of new black markets due to the emergence of the Dark Web and with the anonymity that it provides, being detected by the law is extremely difficult, hence engagements between the respective individuals can be undertaken with a level of confidence.

Bitcoin (BTC) is the currency of choice on the Dark Web and virtual wallets allow for the storage of Bitcoin with exchanges being made through anonymous transactions

(Briere et al., 2015). The mining of BTCs, utilise computational power in order to solve mathematical problems and this currency is segregated from central banking policies and by design cannot be inflated due to the nature of its design. BTC is an anonymous currency and hence can be used to engage in illegal trade (Lavrinc, 2015). Lavrinc (2015) noted additional services that have emerged that will further add to the anonymity of the system, therefore making it almost impossible to trace. One can achieve this by mixing ones Bitcoins and transferring it pass a spidery network of smaller transactions then returning to the source (Lavrinc, 2015). One ends up with the exact amount of money, however the transactions are extremely difficult to trace. Figure 2:5 shows a bitcoin laundering service found on the Dark Web and the presence of Bitcoin-laundering services assists with the anonymity of moving money (Sancho,2015). Sancho (2015) noted that Bitcoin users will either extract the money in traditional mediums or utilise it for payment means.

Figure 2. 5: Bitcoin-laundering (Source: http://easycoinsayj7p5l.onion)

2.3.6 Stolen Accounts

There has been an increase in the number of Dark Web markets that sell stolen accounts. These may vary from PayPal accounts to counterfeit money to other accounts such as Netflix. It is unclear though how many of these accounts are truly valid. Figure 2:6 shows a website on the Dark Web showing sites that offer counterfeit

US Dollars. The stolen accounts markets are not merely restricted to the Dark Web but also on the Surface Web in criminal underground forums (Goncharov ,2015).

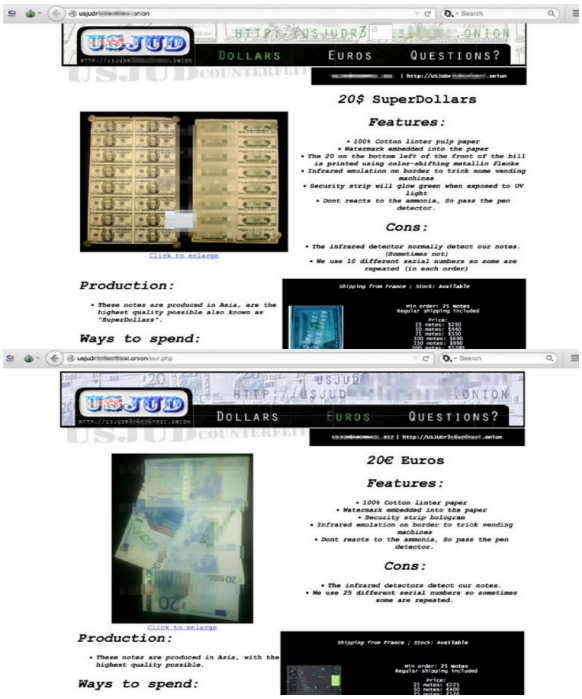


Figure 2. 6: Counterfeit US$ (Source: http://usjudr3c6ez6tesi.onion)

Goncharov (2015) notes a price variation across sites on the Surface Web, however stolen accounts such as PayPal have a higher monetary value, and these will be sold either as verified accounts with present balances or in bulk as shown in figure 2.7. Goncharov (2015) also noted the sale of these accounts are categorised and that the expensive accounts are found in category one as the likelihood of a return of investment is higher whilst the second category is said to be far cheaper.

Figure 2. 7: Stolen PayPal Accounts (Source: http://uigtyr4c5eg3tesi.onion)

Credit cards can be readily found on the Dark Web; however, these will not appear on the Surface Web as shown in figure 2.8. Ciancaglin (2015) noted that there is a probability they exist on the Surface Web; however Dark Web sites have a more professional look and feel to them.

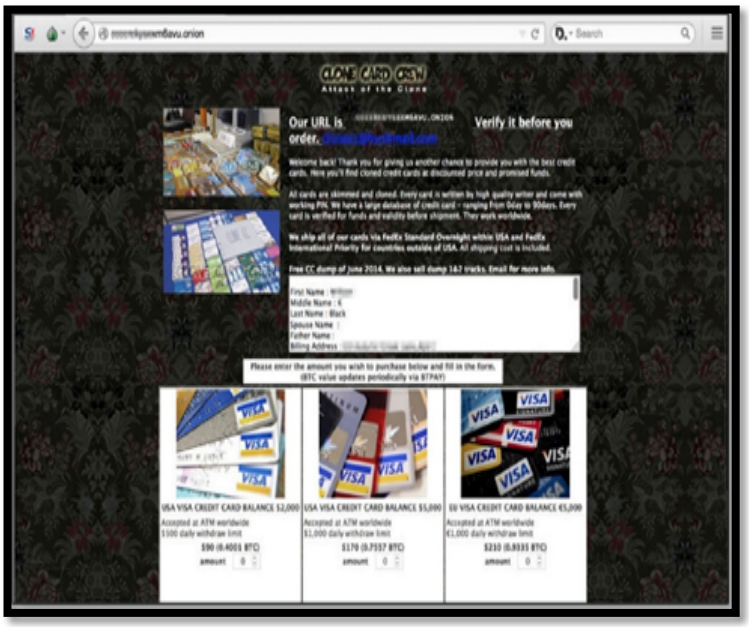


Figure 2. 8: Replica credit cards (Source: http://ccccrckysxxm6avu.onion/)

2.3.7 Passports for Sale

Passports and Identity Documents are documents that are both powerful and unique, hence the need for fake ones are extremely valuable as a form of identification which can assist in the attainment of crossing borders and opening of bank accounts. Ciancaglin (2015) notes that thus the acquisition of these documents is extremely valuable. Figure 2.9 are screen captures from Dark Web markets that appear to sell various forms of official Identity Documents at varying prices, depending on the country and seller. Many of these services are scams with the intention of preying on desperate individuals who are seeking citizenship in other countries (Ciancaglin, 2015). This is a very lucrative business as many individuals find it difficult to obtain Visas or green cards that allow them access to work in other countries and as a last hope will turn to such markets to purchase these Identity Documents.

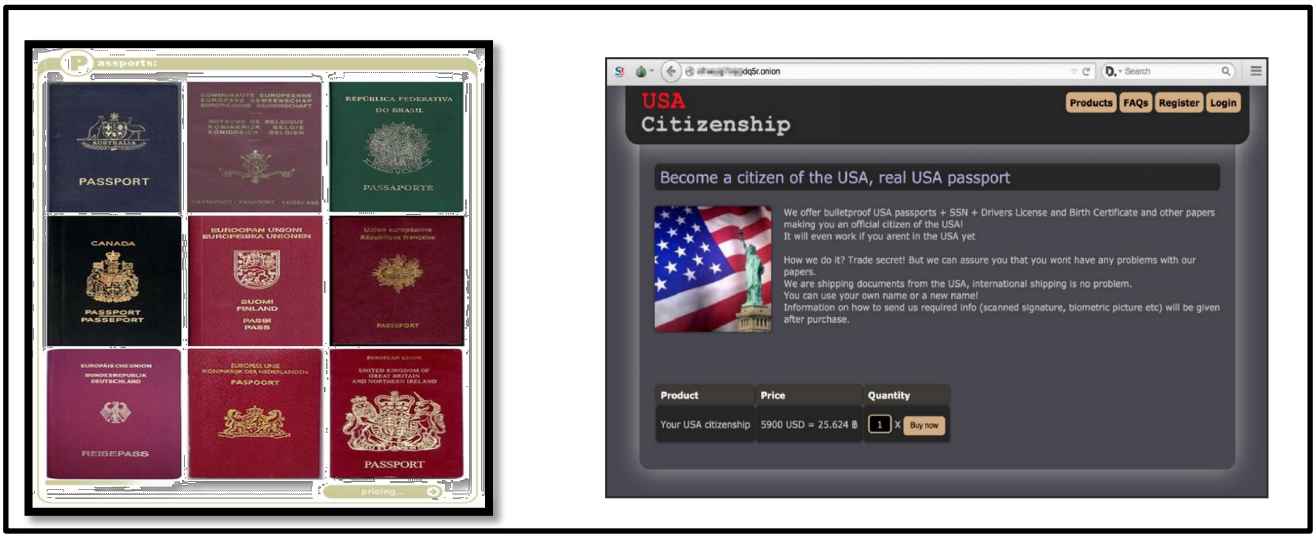


Figure 2. 9: Fake Identity Documents (Source: http://xfnwyig7olypdq5r.onion/)

2.3.8 Leaked Details

There are numerous forums on the Dark Web where one can gain access to information that cyber attackers have leaked. These may vary from credit card details, passports, passwords and personal details. It is common for both hackers and gamers to form a close alliance as they are both like-minded individuals (Ciancaglin, 2015). Falling outs occur regularly due to the nature of the activities they undertake (Ciancaglin, 2015). The utilisation of Doxing as a practice thereafter becomes common practice when this occurs. Ciancaglin (2015) noted that Doxing is utilised by hackers to publicly disclose individual's private information as shown in Figure 2.10. The utilisation of social engineering, hacking and the combination of generally accessible information is utilised to undertake this task. Doxing is not limited to the hacking community, hackers may exploit and expose normal citizens as well and these may include celebrities and companies (Ciancaglin, 2015). Ciancaglin (2015) notes that when a company gets breached, this is not merely undertaken by a hacker but culprits within the organisation may also be involved as with the WikiLeaks's case.

Figure 2. 10: Cloudnine Doxing Site (Source: http://cloudninetve7kme.onion)

Ciancaglin (2015) further notes the difficulty in identifying the details of the leaked information to be factual. The leaked details vary from dates of birth to physical addresses to phone numbers. Figure 2.11 shows a website called Cloudnine that presents doxing information that includes the likes of FBI agents, political figures like to celebrities and actors. The leaked information varies from president Obama's email address to Kim Kardashians leaked information.

Figure 2. 11: Leaked Details (Source: http://cloudninetve7kme.onion)

The capture of Ross Ulbricht and the takedown of Silk Road by the Federal Bureau of Investigation led to sympathy by certain hackers and some have intentionally targeted individuals pertaining to the case (Sancho, 2015). An alleged doxing as shown in figure 2.12 where Judge Katherine Bolan Forrest personal information was published on Doxing sites.

Figure 2. 12: Assassination Services (Source: Ciancaglin 2015)

An area for concern on the Dark Web is the potential to hire hit men (Ciancaglin,2015). These services exist on the Dark Web with acknowledgement of the highly secretive nature of their business (Ciancaglin, 2015). Some of these sites allow for bets to be placed on who the next target is, and the winner receives all the pool money. Greenberg (2015) noted a different take on such services is “crowdsourced assassination.” Greenberg (2015) noted one site, Dead Pool, as shown in figure 2:13 enables users to recommend potential targets and the contribution needs to be made in BTC. Greenberg (2015) further notes that individuals may then anonymously predict the death of their targets and the correct prediction on how the target dies, will result the total pool money being allocated to that individual; however, it is unclear if these services are truly administered or these websites are just merely fake (Ciancaglin,2015).

Ciancaglin (2015) also noted a site stating that no proof of their prior work could be made publicly available and that there was importance to keep them all secretive. The utilisation of these services all involve payment in Bitcoins in order for the job to be undertaken(Ciancaglin,2015). Ross Ulbricht was sentenced for running Silk Road forum, he also attempted to assassinate five of his partners that he had fallen out with (Greenberg, 2015).

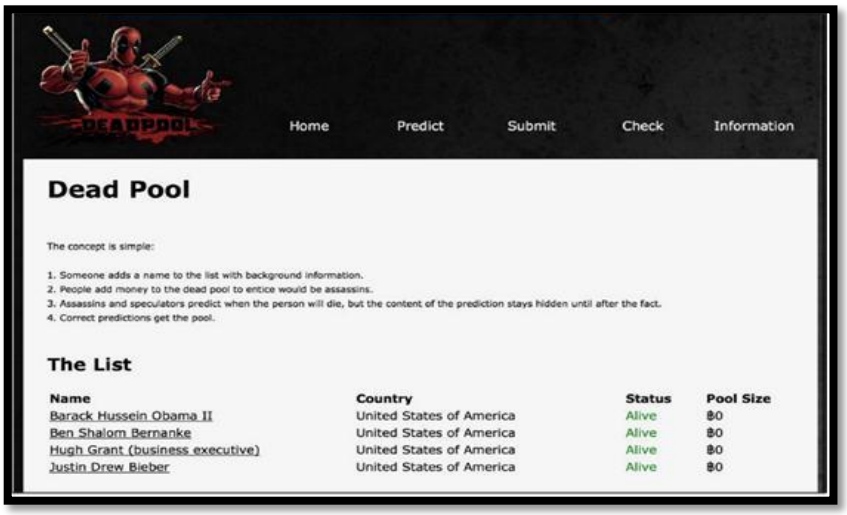


Figure 2. 13: Assassin for Hire (Source: http://cthulhuuap7ch47k.onion)

2.3.9 Social Networking

Extremist groups utilise the Dark Web to engage in social media activity and these also include closed forums where various topics are discussed. The anonymous chat sites make it ideal for engagement if illicit activity. Terrorist organisations on the Dark Web are a major concern for nation states due to the threat that they pose (Weimann, 2016). Weimann (2016) further noted that extremist groups have utilised the anonymity of the Dark Web, with organisations such as Islamic State of Iraq and Syria (ISIS) utilising it as a means for recruitment and spreading the word of their beliefs globally. Weimann (2016) noted the use of traditional marketing styles by terrorist organisations that are more aligned to the younger generation so as to attract attention. Paganini (2017) further notes that terrorist organisations have successfully used social media platforms to recruit individuals to their organisations with marked success. Weimann (2016) stated that by using the Dark Web there are no restrictions in maintaining their presence. Terrorist organisations, such as ISIS', utilise the Dark Web as a means for

communicating and providing information to their followers to note their activities. These materials may provide the necessary knowledge in making their own weapons, provided they have access to the Dark Web.

2.3.10 Pornography

More than 80% of pornographic material found on the Dark Web is related to child pornography. The infamous Destruction of Daisy movie was made by the notorious Scully which he sold to certain clients for up to $10 000. The movie featured the brutal rape of a number of girls aged between 12 years to 18 months. One can probably find questionable pornography on the surface we or perhaps the Dark Web such as incest, bestiality or (fake) rape videos (Norton, 2016). Norton (2016) further noted that some countries prohibit violent pornography. A study conducted by the University of Portsmouth by Norton (2016) revealed that over 80% of TOR network visits are related to pedo sites. The portion of TOR users that purchase drugs or leak information to journalists is far less than those who search for child abuse.

2.3.11 Cryptomarkets

The black market on the Internet contains a group of websites that can be accessed on TOR; with anonymity in conjunction BTC users can make secure and anonymous purchases over the Internet. Cryptomarkets generate large sums of money each year with expanding operations globally (Kruithof, 2016). Kruithof (2016) noted that monthly revenue in the region of $25million dollars was generated in January 2016. There are currently in the region of 80 platforms that either exist or existed on the Internet black market. These platforms have come under serious attacks with some succumbing to scam activities resulting in 20 different platforms exiting today (Branwen, 2015). There are no marked differences between Dark Market websites and conventional websites that you see on the web today, with typical home pages and a listing of all goods and services that may be acquired on the site (Norgaard et al., 2017).

2.3.12 Criminal underground

Figure 2.22 shows the criminal underground associated with the Dark Web. Some of the countries have serious Internet censorship laws, (Paganini ,2017). Paganini (2017) also noted the criminal underground activities for each country as follows:

1. North America: The North American underground is not a locked vault and is openly visible to everyone.
2. Brazil: The country is seen with the quickest path to cybercrime as anyone can gain popularity here with the utilisation of a few tools.
3. Germany: This is a niche market, boasting of wares that are uniquely German.
4. China: The underground may be classified as hub with the sale of latest software, hardware and services.
5. Russia: Primary focus on automation in the criminal underground, where there is competition on which player strives to go to market first.
6. Japan: The criminal underground in this country is not traditional and caters for those who lookout for the taboo.

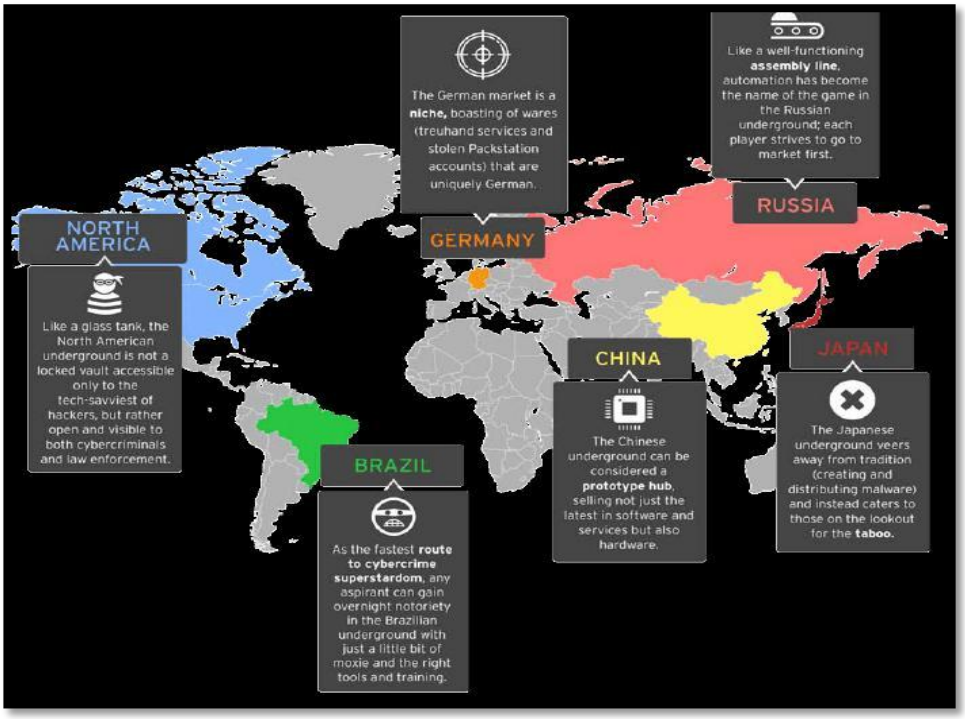


Figure 2. 14 Criminal Underground (Source: Paganini 2017)

2.3.13 Communication

West African cybercriminals typically operate in tight knit groups and each group member constantly communicates with peers sharing targets, compromised email accounts, tools, and best practices. Although cybercriminals use attack tools usually restricted to technical communities or only commonly seen in underground markets, the communication tools they use are less clandestine in nature (Flores, 2016). According to an INTERPOL survey, most West African cybercriminals use email and social media to communicate with one another although some still use IM applications such as Yahoo! Messenger (Flores, 2016). West African cybercriminals generally utilise various forms of social media when conducting cybercriminal activity (Flores, 2016). Figure 2.15 shows which platforms they utilise the most to try and engage in such activities. The utilisation of emails known as phishing emails is the predominant method (Flores, 2016). The utilisation of social media to scam individuals has grown immensely over recent times.

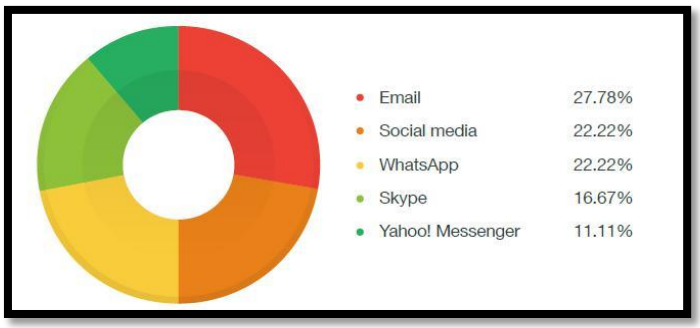


Figure 2. 15: Communication Tools (Source: Flores 2016)

A platform commonly utilised is for scamming activities is Facebook. A study by PlainSite (2019) revealed more than 50% of Facebook accounts are fake. A typical Facebook messenger screenshot from a fraud campaign is shown in figure 2.16.

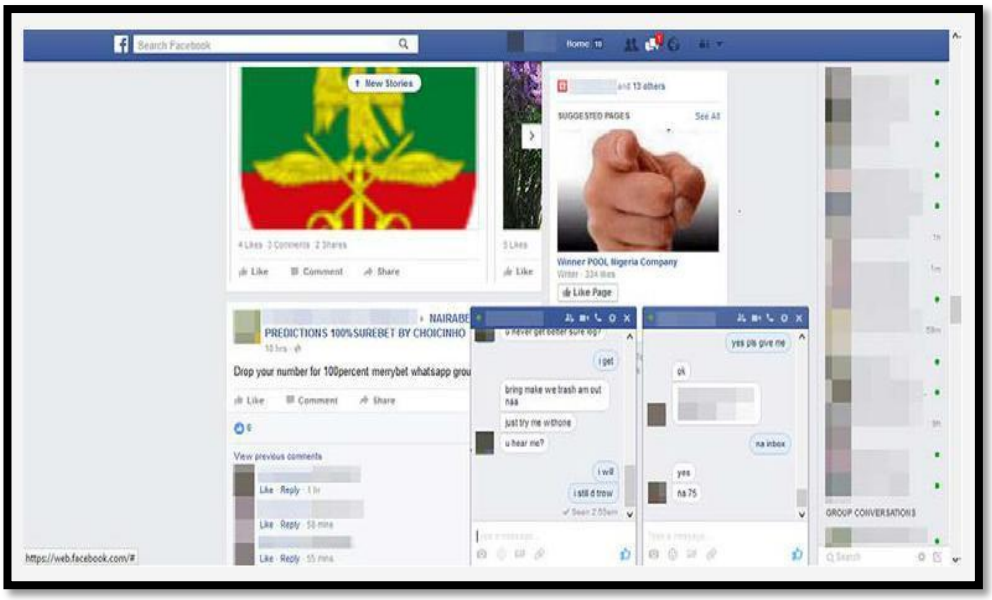


Figure 2. 16: Facebook Messenger (Source: Flores 2016)

Figure 2.17 provides an illustration on the portals utilised by cybercriminals to target businesses. When targeting businesses, West African cybercriminals prefer to use emails although online-dating websites, social media, and IM applications are also used when engaging a target (Flores, 2016). The emails directed towards the business community are phishing emails, where there is an attempt to log username and passwords.

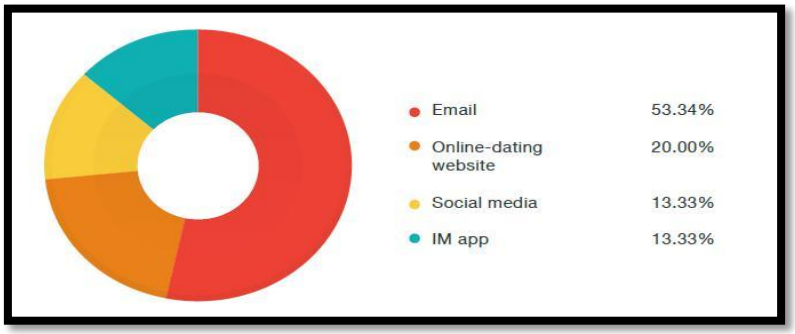


Figure 2. 17: Portals used to Target Business (Source: Flores 2016)

Flores (2016) further stated that when targeting individuals, however, they prefer to use social networking and online-dating websites, which were primarily created for personal use as shown in figure 2.18. The primary scam activities are aimed at

individuals using social media sites such as Facebook, Twitter, Instagram and Snapchat (Flores, 2016)

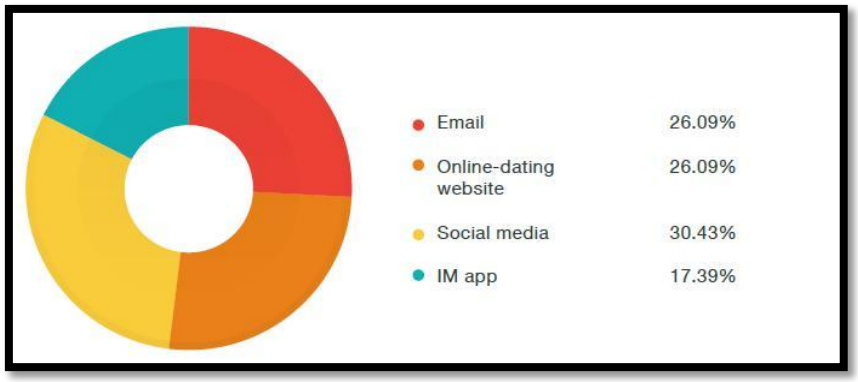


Figure 2. 18: Portals used to Target Individuals (Source: Flores 2016)

The mediums used to target individuals differ and the targeting of businesses is primarily email-based, whilst when targeting individuals is through social media and online dating sites (Flores, 2016). For cybercriminals to maintain their anonymity, much of their social media activity occurs on the Dark Web. The level of anonymity that the Dark Web provides, instils confidence to engage in these scam activities (Flores, 2016)

2.3.14 Terrorism

Terrorists and extremist groups are growing in numbers on the Dark Web and the level of anonymity that is afforded by the Dark Web makes it an ideal area where terrorist groups can undertake their terrorist activities. Terrorists online activities have been present since the late 1990s however the risk of the Surface Web is the lack of anonymity that it provides therefore making individuals easily to track and trace (Weimann, 2016). Weimann (2016) notes that counter-terrorism agencies monitor terrorist websites on the Surface Web and many of these have been shut down and hacked. The Dark Web, with the anonymity it provides enables these platforms to be hidden and difficult to detect. The attacks on Paris by ISIS during November 2015 has led to ISIS migrating to the Dark Web where the extremist group has been spreading news and propaganda in an attempt to safeguard its followers and itself from hackers (Ghaffar, 2016). Ghaffar (2016) further notes the move comes after hundreds of websites associated with ISIS were taken down as part of the Operation Paris

(OpParis) campaign launched by the amorphous hacker collective Anonymous. ISIS's media outlet, Al-Hayat Media Centre, posted a link and explanations on how to get to their new Dark Web site on a forum associated with ISIS. Moore (2016) further notes that the Dark Web provides terrorists organisations with new opportunities and terrorists have used the Internet in various ways, ranging from utilising it as a communication tool, to spreading propaganda to coordinating attacks. These activities are no longer found on the Surface Web as these illicit activities are now purely found on the Dark Web.

2.4 Navigating the Dark Web

The Dark Web may be accessed through decentralised, anonymised nodes on various networks including TOR (TOR, 2016) or I2P (Invisible Internet Project). TOR, which was released as the 'Onion Routing'' project in 2002, (Dingeldine, 2015). Dingeldine (2015) noted the creator of TOR was the U.S. Naval Research Laboratory, which aimed to utilise it as a means to communicate anonymously. TOR's connects to websites via a number of virtual tunnels instead of making a direct connection, therefore allowing for the sharing of information whilst maintaining anonymity (TOR, 2016). The web traffic is routed through other users' computers ensuring the traffic remains untraceable to the original user. TOR essentially establishes layers (like layers of an onion) and routes traffic through those layers to conceal users' identities and according to Clark (2014), to pass through layers, TOR has established "relays" on various computers across the globe which information passes. Clark (2014) noted that information is encrypted between relays, and TOR traffic passes through three different layers before reaching its destination. The final relay is the "exit relay," and the source of the traffic is defined by this IP address. The utilisation of TOR hides an individual's IP address and hence disguises the traffic as conventional HTTPS traffic (Clark, 2014). There is no clear indication of the size of the Dark Web in relation to the Surface Web, however statistics are available on the number of daily TOR users. Statistics obtained from the TOR metrics reveals that the largest user of TOR is the U.S.A closely followed by Germany and then Russia.(TOR, 2016).

2.4.1 The Onion Router Browser

The Onion Router also known as TOR is the most popular anonymous communication browser. The browser has three layers of encryption thus making it ideal to utilise when engaging in illegal activities. The purpose of Onion Routing is to maintain anonymous communication between all the entities on the network, with the exchange of information resisting against and attacks or traffic analysis (Goldschlag, 2015). In order to accomplish this, there are various layers of encryption that constantly change the paths between a set of routers (Goldschlag,2015).

TOR is a distributed, anonymous network run by diverse organisations and individuals making their bandwidth freely available (TOR 2016). TOR (2016) noted that the application as being open source, hence individuals may provide recommendations on improving TOR with the project being maintained by Free Haven Project. McCoy (2014) noted that TOR routing is undertaken in the protocol stack and supports only TCP applications which gain access through the Socks (which is a general-purpose proxy server) interface, thereafter entitling all applications that support Socks to utilise TOR for anonymous communication.

Chaum (2015) noted the Invisible Internet Project being developed parallel to TOR, as also being a peer-to-peer network that also utilises garlic routing technology. There are many of differences in and similarities on how the network is organised compared to TOR. Chaum (2015) further noted an argument that the use of anonymous browsers will assist criminals in engaging in future crimes without revealing themselves. TOR argue that whilst the statement is true there is no reason why an individual should not be afforded the opportunity to communicate anonymously. Criminals have the means to remain anonymous and online crimes can be planned to use encrypted communication at public places such as libraries with the hijacking of stole phones and computers by Trojans and other malware are other possibilities (Chaum, 2015).

Goldschlag (2015) stated that identity theft has increased in recent years, in order for criminals to hide their true identity. Individuals have at least a singular reason for remaining anonymous whilst not engaging in any illicit activities. The utilisation of TOR by governments are primarily used to gather intelligence on other countries whilst

in China and other nations with no Internet freedom, TOR is utilised as a communication tool with other freedom seekers (Goldschlag, 2015).

2.4.2 Hidden Services Browser

Hidden services allow a user to set up a web page or message board thereafter allowing for usage to anybody without the individual knowing its location or who created it and vice versa. Charavarty (2014) noted that with TOR's domain name, the host name can only be identified using the TOR network. Charavarty (2014) noted that this as a potential flaw in that the applications do not pass the DNS lookups through the Socks server. The hidden services on TOR may be accessed by all participating on the TOR network, and its design is aimed to resist censorship, Dodos and physical attacks.

The question then arises is: How is it possible for a server that everyone accesses, hide its location on the network and physical location? Charavarty (2014) further notes rendezvous points are used by TOR in order to reach hidden services. According to Charavarty (2014) in order to set up a hidden service, the individual develops a public/private key and chooses a few onion routers, known as introduction points, where tunnels are located and the service in conjunction with the public key is announced. Charavarty (2014) thereafter notes the service is announced to clients or end users through various means and when a client wishes to access the service, they need to locate the introduction points through the Service Lookup Server. According to Charavarty (2014) a client picks a router through where a tunnel is developed and for a connection, the entry point is communicated on the rendezvous point that flows onto the client owner. Charavarty (2014) then noted that the service owner constructs a tunnel to the rendezvous point, then the connection among the user and the service is complete. Cryptographic methods protect these changes with the implementation of a public key.

2.4.3 Garlic Routing Browser

Garlic routing was developed on Onion Routing technology with a few changes: Onion Routers can join multiple messages with the routing information independently run on every level onto the next onion for the next node (Pseudonym, 2014). Pseudonym

(2014) states that onion messages may have various options such as the request not to pass the message onto the next node, whilst it is disassembled and reassembled into new onions. Pseudonym (2014) also noted that the onions may include padding to hide the true number of cloves there actually is. Pseudonym (2014) further noted that I2P was developed in 2003, with the primary objective to allow for anonymous communication. Pseudonym (2014) noted the end to end encrypted communication and like TOR has layered network layer like the garlic routing network layer found in TOR. Pseudonym (2014) further noted that the I2P developers remain anonymous and are only known by their pseudonyms; an example will be the lead developer calls himself "jrandom". The project is currently still under development stages and is not ready for broad use as yet.

2.4.5 The Invisible Internet Project Browser

The Invisible Internet Project or I2P is an anonymous browser just like TOR that is utilised to gain access to the Dark Web. Pseudonym (2014) notes that there are some major differences between the two specialist browsers, however like TOR the browser allows for anonymous browsing. First of all, in I2P data is sent in packets and is a transport protocol similar to IP, whilst TOR randomly chooses a tunnel path for a connection, whereas in I2P the tunnels are one way (Pseudonym, 2014). Pseudonym (2014) further notes that every node has numerous tunnels to various peers and the sending of a packet, is addressed to the receivers ingoing tunnel and thereafter to the outgoing tunnels. Pseudonym (2014) noted that the sender does not know off the pathway after the outgoing tunnels endpoint. Pseudonym (2014) further noted that this is executed at the packet level and not the connection level and there are advantages, as the disappearing node will not interrupt the connection and bandwidth connection will be distributed for greater throughput. The service is supports streaming services in a similar way that TCP uses IP. The difference in the layout of the network in that TOR has a central directory of servers, I2P is more distributed that requires a bootstrapping operation in order to locate a peer to join the network (Pseudonym, 2014). The presence of a file with published nodes on the I2P homepage and once a router is located, thereafter a query can be made for additional routers. Every node has private statistics pertaining to latency and the behaviour of the known routers (Pseudonym,2014). This is thereafter utilised to categorise the routers into four

different categories: fast and high capacity, high capacity, not failing, and failing (Pseudonym, 2014). Pseudonym (2014) notes that routers located in the lower categories are utilised to find alternative paths on the network.

HTTP proxies are developed that allow traffic to leave I2P and reach conventional web pages, however the node driving the proxy needs to be known (Goldsmith, 2016). Goldsmith (2016) stated that TCP/IP applications have to be modified and cannot be used directly or utilise software such as I2PTunnel to connect to other I2P hosts. There are various I2P applications, with I2Phex and I2PSnark being the most common. I2P concept of hidden services had the concept similar to that of TOR before its design phase, hence a similar setup to TOR (Goldsmith, 2016).

2.5 Dark Web Crawling

Dark Web crawling is the process of utilising TOR or any other Dark Web browser to access the Dark Web, thereby allowing for the collection and classification of .onion addresses. TOR or Dark Web crawling can be undertaken by reconfiguring a proxy to route traffic through TOR enabling the logging of websites. These logging of websites can be further declassified to particular content types. Websites on the Dark Web are hidden on the TOR network and therefore need some specific protocol in order for them to be accessed. The websites found on the Dark Web all have domain names that end in. onion. Moore (2016) undertook a web crawl of TOR and thereafter created a taxonomy of twelve-categories whilst other crawls undertaken by authors aimed to provide a brief overview of TOR and TOR services. There was no specific interest or use cases and the authors restricted their crawler to only identify certain textual materials.

Research into the Dark Web in regard to law enforcement generally focuses upon certain content types. Chen (2014) study established a schema focusing only on terrorist groups. Westlake et al. (2017) implemented a web crawler focusing only on child exploitation materials (CEM). This study utilised a variety of keywords in identifying CEM in addition to certain domains previously identified as not being of interest, and the crawler successfully limited the content to three ‘categories’ (Westlake,2017). Both studies relied on manual labelling the focus was only aligned to specific content types (Westlake, 2017).

2.6 Dark Web Usage in South Africa

South African TOR usage is relatively small in comparison to global usage as shown in figure 2.19, with on average there being around 8 000 South Africans that directly navigate through the gateway per day, however the nature of their activities remains largely unknown (TOR, 2018). The uptake of TOR from a South African perspective is not large due to a few possible reasons, one being that individuals are unaware of its existence and the other is their fear of utilising it due to the negative publicity.

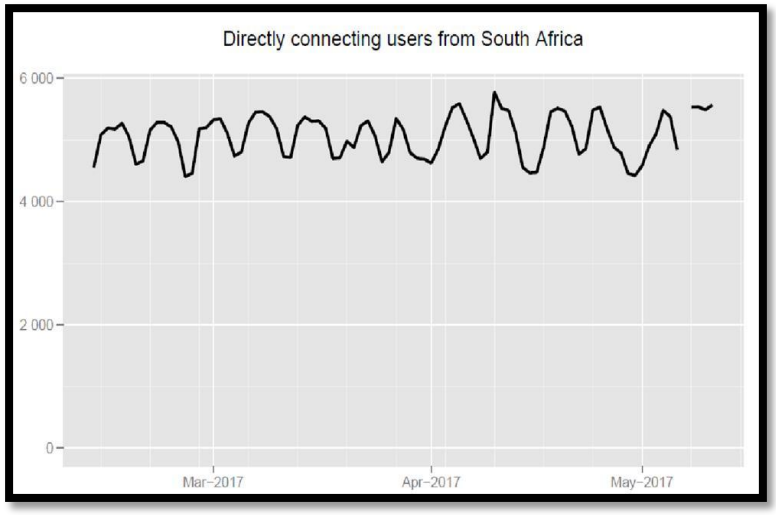


Figure 2. 19: Daily TOR users in South Africa (Source: TOR 2018)

Governments often control international Internet connections either directly or via regulation and that some of the issues censored are aligned to military and militant websites, drugs, music, sex and human rights (as in the case with China) (Murdoch, 2015). Murdoch (2015) also noted that there is software to resist censorship such as TOR and citizens in these countries utilise TOR to engage in illicit activities. Figure 2.20 provides a graphical representation of the countries with the highest Internet censorship laws.

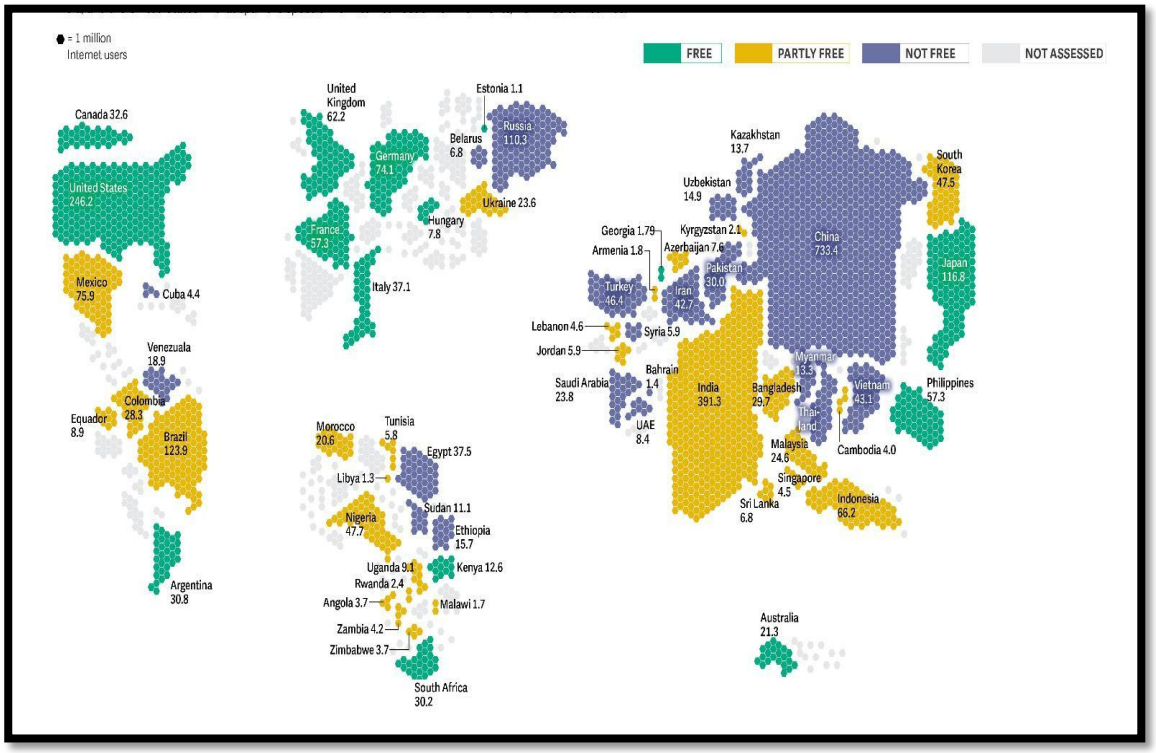


Figure 2. 20: Internet Censorship Countries (Source: Murdoch 2015)

There are many countries that impose serious Internet censorship laws, some off which even bock the use of Google. In order to access the Internet and remain anonymous, many individuals in these Internet censorship countries utilise TOR to access various content on the Internet. The Russian government blocks websites critical of the government and websites associated with militant or extremist organisations. It also pervasively blocks gambling, pornography, and alcohol and drug websites, (Warf, 2016). Clark (2017) further notes that China has the second highest Internet censorship laws and that Internet accessibility is the government's willingness to allow their citizens to access cyberspace and that oppressive governments fear the power the Internet has as it allows individuals to freely express themselves. Clark (2017) noted the Theorisations of Internet censorship can draw fruitfully on contemporary geographic discussions of the state, power, and discourse and that Foucauldian perspectives loom large in this regard. (Crampton, 2016) and (Dobson, 2016). Clark (2017) noted that such work has demonstrated that clearly the Internet can be made to work against people, as well as in their favour. Clark (2017) further stated that the Internet can be utilised to strengthen hegemonic powers, instil a sense of fear and ultimately prevent or minimise dissent and that there are ample motivations for Internet censorship in various forms and types ranging

from human rights activists to include software piracy to ultimately cultural restrictions that exist as part of the oppression of ethnic minorities. The imposing of censorship by governments utilise excuses such as protecting the public or the excuse or from ostensible sins such as pornography, although off late the major factor has been the combating of terrorism (Clark, 2017). Vague notions of national security and to ensure social stability are just some of the notions that governments hold onto when it pertains to censorship and nation states therefore have a choice in the extent to which they wish to implement Internet censorship (Clark, 2017).

Figure 2.21 shows the patterns of censorship across the 26 countries in which there is evidence of filtering (Clark, 2017). Clark (2017) notes the diversity of filtering practices that is evident from the heavy interventions imposed by countries such as China, Saudi Arabia and Iran, to the more limited interventions by Singapore, Hungary, and Lebanon. From all the anonymous communication networks, TOR is the most popular with a worldwide usage of over 50,000 Internet users every day (Carvalho ,2017). Carvalho (2017) further notes that “over half of TOR users are located in Europe, which is also the region with the highest penetration, as the service is used by an average of 80 per 100,000 European Internet users”. Carvalho (2015) noted “Italy in particular, accounts for over 76,000 users a day, which is about one fifth of the entire European TOR daily user base and that Italy is second only to the United States in terms of average number of users, as over 126,000 people access the Internet through TOR every day from the United States”.

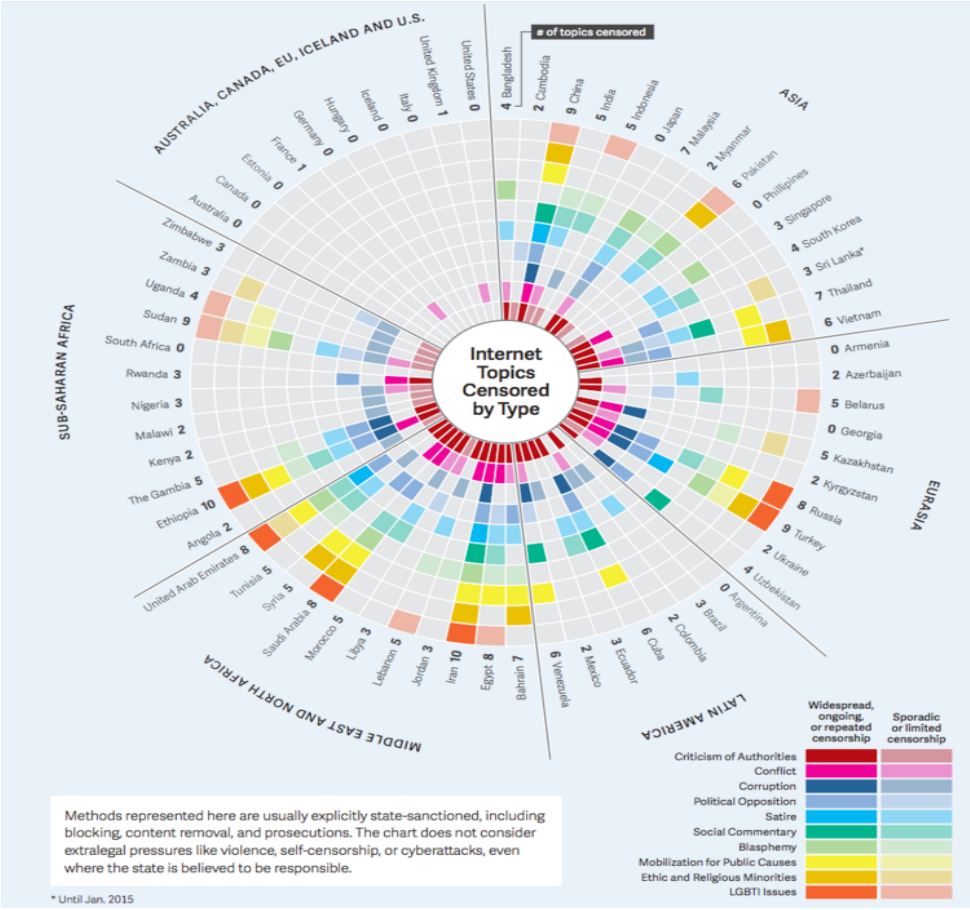


Figure 2. 21: Censored Countries (Source: Clark 2017)

Carvalho (2017) notes the popularity of TOR throughout Europe, with a high popularity index in Moldova as well as in other smaller countries with huge traffic through TOR from small countries such as Andorra and San Marino. Carvalho (2017) noted the number of TOR users as a percentage of their population, countries like North Africa and the Middle East have the second highest usage rate with their usage estimated to 60 per 100,000 Internet users utilising the service, which accounts for more TOR users than India, whilst having less than 4% of its Internet users.

2.7 The Onion Router Architecture

Anonymous communication systems will provide users with the tools to hide their IP address from communication peers, hence creating difficulty for network eavesdroppers to determine the source or destination of the messages. Majority of the systems depend on sending traffic through one or multiple proxies, hence encrypting

traffic further so as to hide the source or destination of messages (Dingeldine, 2014). These may be classified as, low-latency or high-latency communication systems (Dingeldine, 2014). Low-latency systems are effective for semi-interactive applications, whilst high-latency systems are aligned to delay tolerant applications. Low-latency network anonymisation systems can be classified on the routing paradigms they use– that are created from Onion Routing (Dingeldine, 2014), and those that are based upon Crowds (Reiter, 2014). TOR and other anonymous browsers employ deterministic routing, where proxies through which the traffic flows are known by session initiator (Dingeldine, 2014).

Figure 2.22 shows the steps for TOR communication where a list of relays is obtained by the client from a directory service 1 and thereafter establishes a circuit with multiple TOR nodes 2, thereafter sending the traffic to the newly created circuit (Bauer, 2014), and One Swarm (Isdals, 2015) use probabilistic traffic routing related to Crowds. The traffic forwarding relay in the system choose randomly where to send the traffic be it either to the destination or another relay in the system. The Onion Router or most commonly known as TOR is an open source software that allows for anonymous browsing. “The technology stack has three layers of encryption, making it extremely difficult for site owners to identify the end user as TOR is closely modelled on the Onion Routing (Isdals, 2015), (Reed, 2014) paradigm, and is one of the most widely used low latency anonymity networks, with an estimated user base of more than 500,000 users as of May 2016 (TOR, 2016)”. TOR (2016) noted that TOR’s main aim is to protect the anonymity of the users by relaying TCP streams over a network of overlayed nodes that are run by volunteers and the TOR overlay network which consists of over 2500 proxies.

Dingeldine (2014) noted user traffic being relayed through circuits, that are formulated by TLS connections between various nodes. TOR circuits consist of three nodes: the initial one is the entry node, the second is the middleman and the third is the exit node. Dingeldine (2014) further noted that “when establishing a circuit, a TOR client negotiates shared secret keys with the relays that it chooses for the circuit and thereafter, the client uses these keys to encrypt transmitted messages in multiple layers of encryption, starting with the key it shares with the exit node, thereafter each of the nodes then first “peels off” one layer of encryption and then forwards the message to the next node on the circuit”. The exit node decrypts the final layer of encryption, which

reveals the original plain-text message of the user and forwards it to its actual destination through a regular TCP connection. Thus, if in-transit traffic is intercepted by eavesdroppers, they cannot determine the actual source and destination of the traffic.

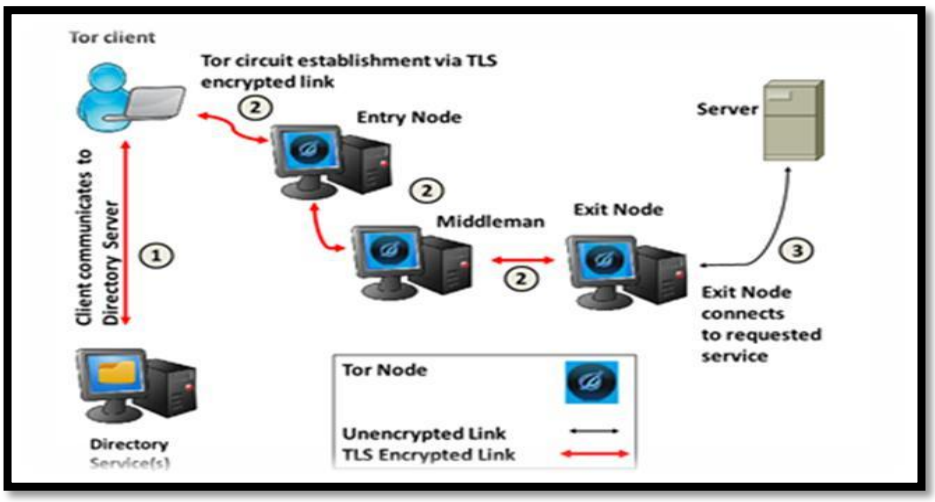


Figure 2. 22: Communicating through TOR (Source: Bauer 2014)

Murdoch (2015) states that “each relay decrypts one layer of encryption and forwards the resultant cells to the next relay along the path and the exit node sees the original packet (M) and finally transmits it to the server”. As depicted in Figure 2.23, the basic steps for the creation of a new TOR circuit consists of three onion routers:

1. “The TOR client queries the directory service to obtain a list of the available TOR relays.
2. It then establishes TLS connections to a node that it selects as the entry node. The entry node may already have established TLS connections to other nodes, which act as middlemen node and these nodes again may already have connections to nodes that may act as exit nodes. If there are no TLS connections between the entry node and the middlemen and middlemen and exit node, then those connections are established.
3. The TOR client thereafter establishes TOR circuits using the circuits, through the TLS connections established in the previous step. The process involves, amongst other activities, the establishment of shared secrets with the relays.
4. The client then selects one of the circuits and establishes the TCP connection to its communication peer (the server), through this circuit”.

Murdoch (2015) further notes that “TOR provides access control features to exit node operators and usually, TOR exit nodes are configured to allow traffic forwarding for only a small set of TCP services. The supported services are defined by the operator of the exit node through the specification of an exit policy”. Murdoch (2015) stated that “TOR supports responder anonymity through Hidden Services. Responder anonymity allows a server to provide a TCP service without revealing its IP address and Hidden Services prevent against attacks that require IP address of the server”. In this section of the study, an overview of how Hidden Services work is explained (Murdoch, 2015).

Dingeldine (2014) stated that “generally, service URI to IP address translation is done using the Domain Name System (DNS) and for a Hidden Service, the regular DNS name used within the TCP/IP model, is replaced by a pseudo-random string derived from the long-term public key of the server ending with “.onion” domain name”.

Murdoch (2015) further noted that “a query to resolve a service URI ending in “. onion”, can be resolved only within the TOR network and this new URI and the long-term public-key, representing the service, is published by the server, the first time it joins the TOR network, only a TOR user can thereby access the service through an anonymous TOR circuit connecting himself to Hidden Service”. This is reflected in figure 2.23 where there is an establishment of a connection between the client and the server known as a Rendezvous Point. A Diffie-Hellman key exchange occurs between the Rendezvous Point and the Introduction Points and as a result of this exchange, there is a connection between the client and server circuits, therefore there is an ensuing of communication between the client and server (Dingeldine, 2014).

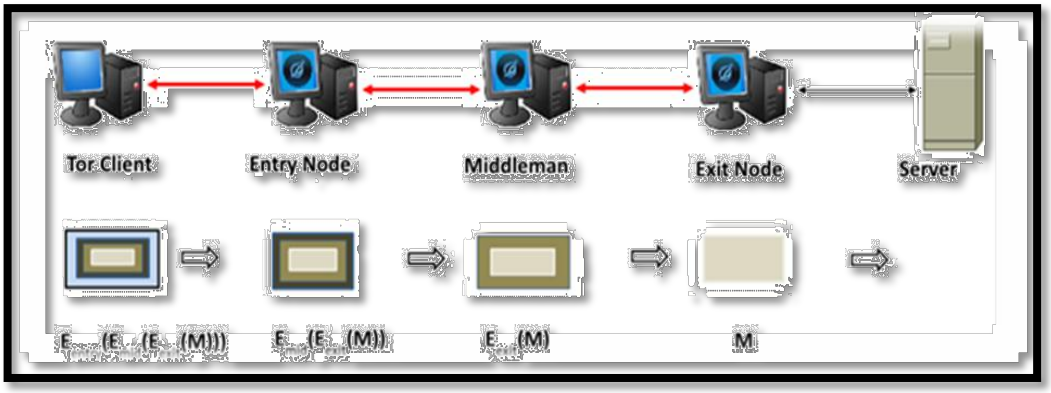


Figure 2. 23: Layered Encryption used by TOR. (Source: Bauer 2014)

Figure 2.24 reflects the various rendezvous points that the client and server create. A client queries to resolve an onion address and is returned the introduction points information. Lastly Rendezvous Points connect to introductory Points commence communication with the hidden server 4 (Dingeldine, 2014).

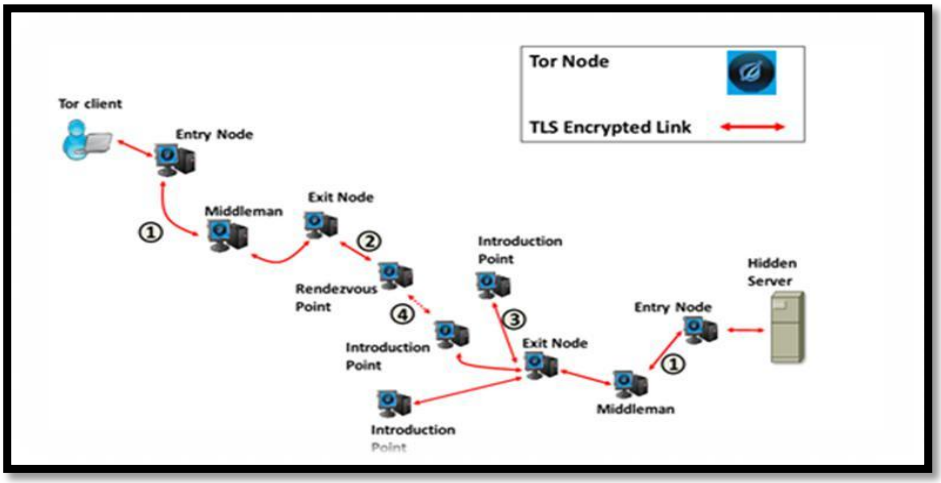


Figure 2. 24: TOR Communication (Source: Bauer 2014)

2.8 Dark Web Policy Issues

An open Internet with anonymity poses huge risks in conjunction with some associated benefits, hence the creation of policies for the Dark Web requires an understanding of the two. Sweeping legislation will encroach on civil liberties however not taking action on Dark Web Illicit activities will result in individuals continuously engaging in them. The creation of a governance policy to address this requires knowledge on the advantages and disadvantages of an anonymous Internet (Chertoff, 2015). Legislation

that has not been properly thought out has the ability to encroach on civil liberals and will be a nightmare to enforce. Chertoff (2015) further notes that addressing the illicit activities on the Dark Web is of vital importance to nation states.

King (2016) notes that regulating TOR is the primary Dark Web policy issue as the Dark Web will not exist without the specialised browser. The anonymity that the Dark provides is the biggest differentiating factor from the Surface Web, hence policy pertaining to anonymity and the usage of TOR is the most relevant (Stevens, 2016). Stevens (2016) noted the presence of two challenges pertaining to Dark Web policy creation: protection of remaining anonymous and global agreement, with Dark Web policies needing global agreement without encroaching onto people's ideals.

Stevens (2016) notes the biggest challenge in that using TOR is not illegal and one cannot clearly differentiate the criminals from the innocent users and the other challenge is that the Internet is global and therefore coordinating regulations would be very challenging.

There is no controlling body that governs the Internet and various countries would like different controls and procedures when it comes to Internet traffic as the development a 'great firewall' will have little or no success. Stevens (2016) further notes the Internet regulation cannot be implemented country by country as it will ultimately destroy the value the Internet provides to all countries.

Both the Russian and Chinese government have passed regulation in individuals accessing TOR in their countries. The Austrian government recently also passed regulations in eliminating TOR traffic. Both China and Russia have recently passed some of the most oppressive laws with regards to the use of the Internet and the Dark Web (King, 2015). Russia's Federal Security Service embarked on a project to confiscate the encryption keys from various Internet service providers, with a non-refusal to hand them will over resulted in a hefty fine (Vitaris, 2016).

The Russian government attempted to break the anonymity of the Dark Web with limited success, (Clemmit, 2016). Clemmit (2016) noted the entering of a contract with a service provider called Rostec, however the organisation was unable to penetrate the TOR network by the given due date.

One of the most dramatic policy actions was taken by Austria in 2014, (Chertoff, 2015). Chertoff (2015) noted that Authorities had arrested a man who had made his

computer a TOR relay and held him accountable for 'contributing to the completion' of a cybercrime committed by another TOR user who was not involved with the arrested man, beyond the fact that the cybercriminal's traffic was routed through the Austrian man's computer the verdict has set a precedent that it is potentially illegal to operate a TOR exit relay in Austria.

Chertoff (2015) noted Chinas action of blocking all of TOR's relays IP addresses, however the creation of special relays now known as bridges may now access the network whilst hiding their IP addresses. Chertoff (2015) noted the lengthy application process, however TOR users in China are willing to undertake this lengthy process just so as to obtain access to the network and this presents a stumbling block to the Chinese government in their quest to block all known IP addresses.

Chertoff (2015) notes that the Dark Web is the largest part of the Internet, yet the majority of the population doesn't even know about it, or even accesses it. It can be used for good and for bad, legal and illegal activity. Chertoff (2015) notes that it is important to understand that it is not all bad. There is plenty good about the Dark Web, which includes the right of privacy when surfing the Internet. The understanding of the Dark Web and its capabilities is vital to the future of the Internet.

2.9 Summary of Related Studies

Table 2.1 outlines a comprehensive overview of the related studies. The overview has been carried out according to two components, which are TOR's traffic identification and Dark Web Crawling. This study will further build on the shortcomings of the studies below in order to provide a more comprehensive overview of Dark Web usage.

Table 2. 1: Related Studies

Author	Problem Solved	Method used	Strength	Weakness
TOR's Traffic Identification				
Bai (2014)	The capturing of TOR traffic utilising a Testbed.	Utilised an experiment to capture TOR traffic Setup eight Pc's one using TOR to monitor dummy traffic.	Successful utilisation of Java anonymous proxy to monitor traffic	The test period was only 120 minutes, therefore the time allocated to the experiment was insufficient. The traffic was not actual TOR

		Theoretical framework of the study was based on Network Theory of Crime Utilised application software such as Java Anonymous proxy in order to monitor traffic		traffic but dummy traffic that was allowed to run through the testbed
Barker (2014)	Collected TOR network traces by developing a complete TOR setup and also attempted to log Dark Web usage	Utilised an experiment to capture TOR traffic and crawl Dark Web traffic. Theoretical framework of the study was based on the Social Learning Theory and Cybercrime and the Network Theory of Crime. Utilised the Selenium browser testing framework	Using unsupervised machine learning techniques over packet.	Captured only thirty websites, duration of test was only one hundred and twenty minutes.
Alsabah (2014)	Enhanced TOR's performance by the developing of an algorithm	Utilised an experiment to undertake an attack on TOR. Undertook this by analysing TOR traffic Theoretical framework of study was based on Network Theory of crime. Tested Bayesian Networks and decision tree classifiers utilising a proposed classification method	Developed an ML classifier to differentiate web traffic from bulk download traffic	The utilisation of an exit node will enable the logging of Dark Web usage. The experiment conducted in the study only aimed to log traffic entering the TOR network so as to provide recommendations on TOR's performance improvements.
Houmansadr (2014)	Differentiated the traffic of anonymous networks from other network traffic and claimed that mimicking other traffic is an obsolete way for anonymity	Undertook an experiment in order to attempt to differentiate traffic. This was accomplished by undertaking an attack on TOR. Theoretical framework of study was based Network Theory of Crime Devised a number of passive and active attack strategies to breach anonymous networks and suggested the use of partial	Passive and active attacks to bypass traffic imitation technique	The study purely focused on providing performance improvements of TOR. There was no logging of exit routing traffic nor the attempt to log Dark Web activities.

		imitation and use of new strategies by incorporating popular protocols like HTTPS email etc.		
Chakravarty (2012)	Provided improvements on TOR's performance	Undertook an experiment based on an attack on TOR. Theoretical Framework was based on the Differential Theory. Utilised LinkWidth successfully. LinkWidth. LinkWidth sends a train of pulses comprising of alternate TCP-SYN and TCP-RST packets and capacity is computed at the receiver end by estimating packet dispersion	Observed bandwidth fluctuations through compromised node by utilising Virtual Machines and CensorSpoofer	The study concentrated on identifying TOR relays that participated in a circuit. There was no attempt to log traffic of individuals in the U.S participating on the network. The experiment only identified a portion of the network location of participating TOR relays.
Winter (2014)	To ascertain how China is blocking TOR	Undertook an experiment to analyse how China is blocking TOR. The Theoretical Framework for the study was based on the Differential Theory of crime. Deployed two bridges in Sweden and Singapore attempts were made to connect to TOR clients in China.	Utilised fingerprinting and the sue of tuples to connect to TOR relays and bridges	This experiment undertaken was an attempt to ascertain how China is blocking TOR. There was no attempt to log traffic to or out of China nor the illicit activities that individuals in China are engaging in.
Chaabane (2012)	Monitored TOR traffic in 6 countries	Undertook an experiment to monitor TOR traffic. Theoretical Framework of the study was based on the Social Learning Theory and Cybercrime. Traffic analysis of TOR using HTTP and Bit Torrent protocols	Utilised multiple countries, six in total	Allocation of limited amounts of bandwidth to server, only 20gigs. No website classification.
Tang (2014)	Development of an algorithm to help	Utilisation of an experiment to undertake an attack on	Utilisation of Webfetch to	The experiment undertaken in this

	TOR's performance improvements.	TOR in order to improve TOR's performance. Theoretical Framework for the study was based on the Social Learning Theory and Cybercrime. Setup a node, which acted as the middle node. Utilised WebFetch to download the target file	download the target file from the web server.	study did not attempt to log any Dark Web usage nor exit routing traffic. The primary aim of the study was to develop an algorithm to improve TOR's performance. Only a middle node was setup, for the identifying the source and destination of TOR traffic an entry and exit node is imperative.
McCoy (2014)	Attempted to log traffic passing through the TOR network and also attempted to log Dark Web usage	Undertook an experiment to log traffic and Dark Web usage. Theoretical Framework of the study was based on the Space Transition Theory. Setup of a router that was connected to a 1 GB/s of network link with a rank of top 5% TOR routers and flagged as Running.	Logged traffic passing through TOR and Dark Web usage.	For exit routing traffic the router relayed 709GB of traffic however only the first 150bytes of the packets was logged, this will limit the number of websites that could have been classified. The experiment conducted did not allow for the logging of exit routing traffic to any country. A 1GB/s network line will hamper the amount of traffic that passes through the node.
Murdoch (2015)	Observed traffic entering and leaving several TOR entry and exit nodes within the UK and also attempted to log TOR traffic.	Utilised simulations and with the utilisation of NetFlow to monitor traffic. Theoretical Framework of the study was based on the Network Theory. Utilised entry and exit nodes to undertake the experiment.	Successful implementation of NetFlow to monitor TOR usage.	The experiment undertaken only focused on entry and exit nodes within the U.K. There were no attempts to log exit routing traffic to any other country. The experiment did not log any Dark Web usage in the U.K

		Only used the UK to log traffic through several exit and entry nodes.		
Dark Web Crawling				
Moore (2016)	Creation of 12 category taxonomy	Undertook an experiment based on Dark Web Crawling. Theoretical Framework of study was based on Social Learning Theory and Cybercrime. Reconfiguration of a proxy to route traffic through TOR enabling the logging of websites. Restricted the crawling of Dark Web usage. Accomplished this by crawling only certain textual materials. Creation of 12 category taxonomy that enabled the classification of Dark Web traffic.	Using peer-to-peer applications with compromised exit relays to deanonymise users.	Restricted the crawler to textual materials due to the high risk of inadvertently accessing illegal materials such as child pornography and terrorist publications.
Westlake (2017)	Created a web crawler specifically focused upon locating child exploitation materials	Undertook an experiment and created a web crawler to focus on specific content. Theoretical Framework of study was based on Space Transition Theory. Reconfiguring a proxy to route traffic through TOR enabling the logging of websites Crawler created only searched for child exploitation materials.	Passing duplicate cells through compromised entry relay	Only concentrated on child pornography on the Dark Web.
Chen (2014)	Established a schema for categories of use for the web (also applied to 'dark webs'), focused	Undertook an experiment to categorise specific content only. Theoretical Framework of study was based on Space Transition Theory.	Development of a TOR exit node to log web traffic	Only concentrated on Terrorism on the Dark Web.

	upon terrorist organisations	Reconfiguring a proxy to route traffic through TOR enabling the logging of websites		

The studies presented in the table either attempted a network traffic analysis of TOR or a Dark Web Crawl. The only study to undertake an analysis of Dark Web traffic and a Dark Web crawl was a study presented by McCoy (2008). This study will therefore be based on the study by McCoy (2008) and will further improve on the experiment utilised by McCoy (2008). This study will therefore attempt to:

1. Log exit routing traffic by country and determine the countries South Africans are connecting to on the TOR network.
2. Obtain the geo-location of the IP addresses and obtain the IP addresses of the clients connecting to the TOR network.
3. Undertake a comprehensive Dark Web Crawl and reveal the Dark Web traffic that South Africans are engaging in.

The Theoretical Framework to be implemented in this study will be based on the Space Transition Theory as utilised by McCoy (2008), Westlake (2017) and Chen (2014). This theory deals with cybercrime and the anonymity of a user when engaging in cyber-criminal activities, which forms the basis for this study. The improvements to previous studies as outlined above will ensure the reliability and validity of the results when the experiment is conducted. The successful conceptualisation, design and implementation of the experiment will ensure the determination of the use and misuse of the Dark Web by South Africans. The suggested improvements for this study are as follows in table 2.2:

Table 2. 2: Improvements to Existing Theoretical Framework

No.	**Study Improvements**	
	Previous Studies	***Proposed Improvements***
1.	No validation of experiment architecture	Architecture of experiment validated by cybersecurity specialists, this will ensure reliability of results
2.	No validation of experimental results	Validation of results conducted by presenting to focus group and utilising

		application software that ensures a 99% accuracy rate in declassifying of websites.
3.	Bandwidth allocated to the node was only 1mib/s. The classification of the web addresses was undertaken manually.	The utilisation of a 5mib/s network line and not a 1mib/s. The greater the amount of bandwidth allocated to the node, the greater the amount of traffic will pass through the node. This study will utilise Zvelo for the declassification of web addresses. Zvelo guarantees 99% accuracy rate for the classification of web addresses, therefor it is an ideal application to be implemented into the study.
4.	McCoy (2008) undertook the experiment in stages whereby entry and exit traffic was logged for 15 days and exit routing traffic for 5 days	This study will log entry and exit routing traffic simultaneously for a period of 30 days utilising a 1-week period as a testing period.
5.	No meet in the middle attack	This study will undertake a man in the middle attack. A meet in the middle attack is like eavesdropping where the attacker secretly relays and alters the communication between two parties who believe they are directly communicating with each other
6.	Categorisation of Dark Web content done primarily manually	Utilisation of open source and proprietary software for the categorisation of Dark Web Content.
7.	No Social Networking Analysis conducted	Social Networking Analysis conducted by utilising Gephi

The next chapter will present a comprehensive overview on the experimental architecture to be implemented in the study. The chapter will further present the social networking analysis tools that will be utilised for creating network visualisations and basic statistical computations. The chapter will also determine the theoretical framework to be utilised in the study.

CHAPTER THREE

GEO-LOCATION METHOD

3.1 Introduction

The methodology presented in the chapter will further reinforce the validity, significance and relevance of the study. TOR is privacy system which is designed to protect the privacy of users from traffic analysis attacks (Dingeldine, 2015). Since its initial development, researchers have analysed the system's performance (Wendolsky 2015, Tang 2014, Houmansadr 2014) and security properties (Goldberg, 2014, Alsabah, 2014, Dingeldine, 2014). There is yet to be a study aimed at logging exit traffic with the intention to geolocate all IP addresses and declassify all Dark Web content. The idea of geo-location method is to provide a mechanism to investigate how the Dark Web is currently being used, misused and who is using it and or what purposes and to what extent. To accomplish this task, it is important to:

1. **Determine exit routing traffic**

To determine the exit routing traffic by country based on the study introduced in chapter 1.5 where the experiment analyses the data thorough the application layer header through the exit node in order to ascertain the IP distribution in the network. The analyses of the application layer will allow for the logging of exit routing traffic by geo location of IP addresses.

2. **Explore Dark Web usage**

To explore how the Dark Web is currently being used, the study will examine both malicious router and client behaviours on the TOR network. The client or end user exits the TOR node and connects to an onion address. This will reveal a comprehensive log on all Dark Web activities that South Africans are engaging in

3. **Determination of illicit Dark Web usage**

To explore how the Dark Web is currently being misused all Dark Web traffic will be logged and stored in a big data analysis engine. The use of open source and proprietary software in declassifying the onion addresses will reveal the illicit activities that South Africans are engaging in on the Dark Web.

3.2 Theoretical Framework

A theoretical framework makes specific relevance to the relevant scholarly literature pertaining to the study. The theoretical framework will allow the researcher to demonstrate an understanding of the various theories and concepts relevant to the study. A theoretical framework is important and imperative as it will allow to conceptualise the study in a broader context. For the determination of a theoretical framework, one has to have knowledge on the various categories of cybercrime and possess the ability to categorise offenses accordingly. There are two distinct categories of cybercrime namely; Violent and Non-violent cyber-crimes (Blond, 2014). The majority of cybercrime committed is off a non-violent nature due to there being no physical contact between the target and the offender. Examples of nonviolent cybercrimes are cyber-fraud and cyberstalking (Blond, 2014). Blond (2014) notes Violent Cyber Crimes as “those crimes that pose a physical danger to some person or persons and non-violent cyber-crimes do not cause any physical damage to persons; instead they cause financial loss, psychological disorders and social harm”. A theoretical framework for the study is presented in Figure 2.1. The figure represents the theoretical framework of cybercrime which is derived from the sociological theorists and the Sociological Theories.

3.2.1 Classical Sociological Theorists

Sociological Theorists are commonly known as the founders or fathers of the sociological theories. Figure 3.1 shows the three different categories of Sociological Theorists. These theorists can be classified as Classical Sociological Theorists, Modern Sociological, Theorists or Postmodern Sociological Theorists.

Comte’s (1865) positive philosophy is his contribution to social and political philosophy. Comte (1865) noted positivism as the last stage of intellectual development and reason and objectivity is the basis of positivism of his philosophy. Comte (1865) believed even cyber-crime has a logical basis as it involves analysis, objectivity, technical knowledge, experimentation and observation. Comte (1865) also believed that social development was the outcome of the development of the human mind. As the human mind progresses, society also develops criminal tendencies and new types of crime emerge. Cyber-crime is no exception to this development (Comte, 1865).

(i) **Durkheim1893(1933)** emphasises the fact that as society advances and industrialisation progresses, division of labour not only becomes important but also inevitable. Cyber-crime is a feature of organic solidarity where heterogeneity and complexity occurs and according to Durkhiem (1893,1933) there is no society without crime but instead, every society has crime. In Durkhiem (1893, 1993) noted that crime occurs when an individual diverges from the collective norms and exhibits a criminal character. Cyber-crime is no doubt borderless and it occurs in abstraction, without any face-to-face interaction. It has brought together governments of various countries on a common front to make laws to fight against cyber-crimes.

Weber(1991) introduced 'rationalisation' to explain societies of the West who have shifted from traditional orientation to rational and scientific orientation. Weber (1991) noted that cyber-crime can be committed for known as well as unknown persons. Cyber criminals know that it is easier to commit 'e-fraud' in comparison to committing fraud in the physical space. They have calculated ends and means and they commit crime in such a manner that leaves negligible chances to be caught (Weber,1991) .

Marx's 1932(1844) concept of Alienation can be aptly used as a tool for understanding contemporary society. Technology has become a part and parcel of present society and Marx (1932, 1844) has referred to production as a technical process as it involves technology. Marx (1932, 1844) notes that even though many types of cyber-crimes such as cyber bullying, cyber defamation and cyber blackmailing occur in virtual environment they also do have an effect in real life and it is very difficult to control these online crimes in the physical world because of the lack of adequate knowledge and expertise which is required to deal with online crime. This causes alienation in present generation who although are techno savvy and active on various social networking sites feel powerless to deal with cyber-crimes **(**Marx's ,1932,1844).

Pareto(1961) states clearly that every individual performs both logical and non-logical actions and everyone tries to justify even non-logical actions as logical. Pareto (1961) also believes that every social phenomenon has two aspects: one is reality and the other is its form and whereas the former involves the actual insistence of the thing, and form is the way in which phenomenon presents itself to the human mind (Pareto,1961). The former Pareto (1961) calls for objective and later subjective aspect and also believes that all actions of the individual have two aspects; one being the end

and the other being a means to an end (Pareto,1961). Similar is the case in technology, where many actions seem logical by the actor (subjective) cannot be logical in reality (objective). For example hackers view hacking as useful and profitable and the viewers of porn sites never call it as harmful. Nevertheless, cyber-crime is a severe form of deviant behaviour but deviants justify it as logical and acceptable.

In views of **Veblen(2003)** process of social change is more or less constant, and one change results in another change. For Veblen (2003) social change indirectly reflects our technological advances and vice versa. The use of the Internet has established online communities which have brought new kinds of social relationships (Veblen,2003). It is through the use of technology that people learn more about worldly affairs and the use of mobiles and the Internet for instant communication have become a commonplace. Technology has also given rise to a new type of crime i.e. cyber-crime (Veblen,2003). Online crime can be conducted from anywhere and at any time with a computer and a network connection and it leads to easy victimisation. Pornography has degraded social and moral values of youngsters and even children (Veblen,2003).

Tonnies(1991) vividly mentions the concepts of Gemeinchaft and Gesselschaft. In Gemeinchaft (community) each person has some sort of relationship with others (without a choice), while in Gesselschaft (association) members enter into interaction according to their individual desires for achieving some specific purpose. The emerging world society falls in the category of Gesellschaft where communication is virtual and individuals enter into relationships on their will for some specific motives. Such interactions occur on social networking sites where individuals freely enter or leave a relation as per convenience. However, it has led to emergence of negative consequences such as cyber bullying, cyber defamation and cyber stalking (Tonnies,1991).

In his essay, 'On liberty' **Mill (1859)** elaborated the notion of state as Paternalistic. If Mill (1859) is correct in suggesting a need for a balance standard between the society and state, there emerges a series of complex issues and questions. If law is an instrument to preserve the "civilised community through imposing morals and virtues, then where is the means by which such assessment is made" (Mill, 1859). Such is the current situation with respect to the presence of sexually explicit material online in society. It is very easy to find thousands of porn sites on the Internet which can be viewed, downloaded, transferred at a fast pace without any legal or social

restrictions (Mill, 1859). Government has not been able to implement a ban on porn sites although law prohibits its viewing and downloading. The law is present but it is difficult to be implemented in the cases of online crimes.

3.2.2 Modern Sociological Theorists

Modern Sociological Theories give a lucid overview of the core concepts that sociological theories must address and attempt to reconcile. These theories explain the major contributions to the analysis of each concept by classical and contemporary theorists, and links these ideas to current sociological issues such as change and globalisation, feminism and sociological theory and the return to cultural analysis.

Merton(1938) opines the gap between approved goals and the means creates strain. In contemporary society, success is primarily measured in terms of material achievements and social standing. Merton (1938) used the anomie theory to apply specifically to deviant behaviour in various societies. Cyber criminals come from a very diverse background (Merton,1938**)**, with those who are in high schools or colleges are most likely to fit into these theories. They may see how they put in hard work into their studies and development of skills and yet realise that it is unlikely that they could achieve the financial success and as a result they may see crime as a means to achieve enormous financial success. Any individual would see computer crime as a way and means to make large sums of illegitimate money and modernity recognises the advantages of technology and sees risk as its inevitable feature (Merton,1938**)**. The point is how this risk can be prevented, minimized or channeled. In classical modernity the ideal was equality while in contemporary modernity, the ideal is safety.

Giddens (1991) describes the modern world as a 'Juggernaut 'which is a runaway engine of enormous power which, collectively as human beings ,we can drive to some extent but which also threatens to rush out of control and which could render itself asunder". The Internet is a product of modern technology moving along through time and over physical space (Giddens, 1991). Digital netizens are the agents who steer it in their directions. There are two types of disembedding mechanisms that play a key role in modern societies. First are symbolic tokens and second are expert systems (Giddens, 1991). 'Symbolic tokens' are money, which allows for time space distanciation. We are able to engage in transactions with others who are widely

separated. This has also given spurt to online frauds and hacking through which cyber criminals can make transactions from any part of the world and at any time (Giddens, 1991). The use of credit cards online exposes the user to the risk of identity misuse and eventual fraudulence. Second mechanism i.e. 'Expert Systems' involve professionals like lawyer, physicians and engineers; common things like cars, gadgets are also created and affected by Expert Systems (Giddens, 1991). They provide guarantees (but not without risks) of performance across time and space. Personal Computers are always prone to viruses while using the Internet and trust is very important in modern securities dominated by abstract systems. Online trust on someone with whom one is transacting or forming any relationship plays an important role but because of physical proximity, it becomes easier for an individual to break the relationship at his own will and at any time because he or she is not answerable to anyone (Giddens, 1991). Cyber experts are limited in number because of the technical skills required in this field which are complicated and difficult to learn. Besides this, cyber-crime is conducted from any part of the globe and at any time and by anyone. Giddens (1991) also talks about new and dangerous risks associated with modernity that always threatens our trust. Risk is global in intensity (cyber war). The type of risks on the Internet has acquired varied forms such as e-fraud, e-theft, cyber bullying, stalking, cyber terrorism, pornography etc. These risks have far reaching social effects and have given us the feeling of a runway juggernaut and fill us with ontological insecurity.

Beck(1992) calls the modern world as a 'Risk Society'. The emerging new modernity and new technologies are associated with the risk society. The contemporary world has elements of both. Beck (1992) recognised a strange paradox in late modern society; risk is increasing due to technology and science rather than being abated by technological progress. In the case of cyber- crime one is unaware of the risk that can occur with a single click of the mouse. Hacking, fraud and online schemes are some of the risks to which users are exposed. The 'Information Society' is thus creating risks for people thereby exhibiting a 'Risk Society'. In the case of cyber risks, the weapons are software and knowledge, the environment in which the attack occurs is virtual; the possible attacker is unknown and is able to hide himself effectively (Beck, 1992).

Ritzer's (2008) impetus is on the McDonalization of society which represent a contemporary paradigm of formal rationality. Ritzer (2008) provided four dimensions to formal rationality:

1. Efficiency i.e. the search for the best means to an end
2. Predictability i.e. a world of no surprises
3. Rational systems tend to emphasize quantity, rather than quality.
4. Reliance on non-human technologies

Ritzer (2008) noted the ample amount of information on the Internet that is easily accessible and with a simple search one may find practically anything they wish for. It is highly predictable as one knows what to get and from which site (Ritzer, 2008). Ritzer (2008) argued that even though there is ample information on the Internet, the authenticity of this information is questioned. The theories of modernity thus see the Cyber society as a threat to the world system and the advancement of technology has given rise to associated Mcdonalization also applies to contemporary Cyber risks. The Internet has definitely made life and work easier for people but has also exposed them to certain threats like online crimes (Ritzer, 2008).

3.2.3 Postmodern Sociological Theorists

Postmodern society has seen the dawn of sentiments and emotions. In the postmodern era, the disadvantages of technology are recognised. While in modern society individualism was important, in postmodern era, emphasis is on collectivity or groups.

Baudrillard (1984) believes that there was a time when signs stood for something real and now they refer to little more than themselves. Distinction between what is real and what is fabricated is the cornerstone of the postmodern world and according to Baudrillard (1984) 'We live in the age of simulation' which leads to "reproduction of objects or events". Software Piracy or the counterfeiting and distribution of products is intended to pass for the original is done by illegal downloading. In photo morphing, a face can be morphed with someone else's body; it is difficult to distinguish real from the duplicate. Baudrillard describes the postmodern world as hyper reality (Baudrillard,1984). For Example, stealing the IP address or identity of another computer or to obtain access to the other computers on the network

(Spoofing). Only a decade ago people had to trek down from one place to another for shopping. Now with a single click of mouse, these are available at one website from which we can order online. Shopping, banking, games, movies, entertainment are available online and in rationalising these forms of re-enchantment, they are by definition disenchanting them (Baudrillard,1984). E-commerce has given rise to a number of online cyber-crimes. Identity theft and online transactions are examples, in which fraud can be conducted without physical circumstances.

(ii) **Jameson (1991)** recognises that postmodernism is usually associated with a radical break. Jameson (1991) describes this new form as a 'cultural dominant' and according to Jameson (1991) four elements are basic to the postmodern society:

Postmodern society is characterised by superficiality and depthlessness. It truly depicts the use of Internet and its activities and the emergence of social networking sites coincide with superficiality as the people who are connected through these networks lack the basic connection and sentiments which are required for long lasting relations. Jameson (1991) used the term "simulacrum" in which one cannot distinguish between the original and the copied.

1. In Postmodern society, emotions or intensities have faded. It causes alienation and anomie. Jameson (1991) prefers to call it as 'Intensities'. The new electronic media gives rise to postmodern intensity. In cyberspace, humans lose their real self and dream about illusions in a virtual environment, which gives rise to unusual fantasies. Pornography is an example of such a phenomenon.
2. In Postmodern society, there is a loss of historicity (pastiche) i.e. the past cannot be traced and all one has is the access to our texts or pictures. Even on the net one has to trust the information posted on it.
3. Postmodern society has impressions based on reproductive technologies, especially electronic media like TV, computer, and Internet. The postmodern era gives birth to new and varied cultural products than the explosive, expanding technologies of the modern era did. Cyber culture has produced a new type of culture in society which has both positive as well as negative aspects in it. For instance, Infringement of copyrights (Plagiarism) has increased on the Internet.

Thus one realises that post modernity is full of technical advancements but it carries within itself the forces of destruction. Cyber war is no exception to that and

developed as well as many developing countries rely on networks and servers for important services and activities (Jameson, 1991). Jameson (1991) noted that economic strength depends on these which are a pillar for a good society. Besides the theories of modernity and post modernity, many other theories need to be examined in the light of dynamic issues relating to cyber-crime and social networking.

Sociological Theories aim to understand what we know as the modern world. This is approached through the understanding of the transition from pre-modern or traditional societies to modern societies. The various sociological theories shown in figure 3.1 can be classified into six different categories and this study will be based on one such theory.

3.2.4 Social learning Theory

Tarde (1903) explains the fact that individuals learn deviant behaviour and it is not biologically inherent. Tarde (1903) observes that there are four main requirements in which social learning occur; First individuals must have a close contact with those they are imitating which can be family members, close friends or teachers; second, individuals must engage in imitation of their superiors; third, is that they must understand their behaviour i.e. they need to know what the behaviour is like. The individual must be a role model to the person who is imitating the behaviour. This theory takes into account the fact that the behaviour learnt could be negative as well as positive.

3.2.4.1 Space Transition Theory

"Space Transition Theory" as proposed by Jaishanksr (2007), explains the behaviour of the persons who bring out their conforming and non-conforming behaviour in the physical space and virtual space. Jaishanksr (2007) further states that the virtual space provides an individual with such space where he can express his feelings and even vent out his outrage against anyone. Cyber stalking and Cyber defamation are instances where offenders use online space because of its anonymity and widespread approach (Jaishanksr,2007). Cyber stalking and cyber defamation also argues that people behave differently when they move from one space to another. One of the

important postulates of the Space Transition theory is that ‘People with repressed criminal behaviour in the physical space have a propensity to commit crime in cyberspace, which, otherwise they would not commit in physical space, due to them status and position” (Jaishanksr, pp 5, 2007). Criminology views the emergence of cyberspace as a new locus of criminal activity and therefore, some new theories are needed to explain the occurrence of cybercrime.

3.2.4.2 Network Theory

Dingeldine (2014) states that Network theory focuses on a wide range of micro to macro structures, since links occur at the large-scale, social-structural level as well as at micro level. While sociologists talk about ‘strong ties’, network analysts talk about ‘weak ties. Dingeldine (2014) further notes that Social networking sites preserve the culture of maintaining weak ties, since weak ties prevent isolation and allow individuals to better integrate through social networking sites. However, such integration also gives rise to deviant behaviour on social networking sites, because networks are transitive. If there is a tie between A and B and B and C, there is likely to be a tie between A and C; this link between A and C is weak and therefore could lead to some form of crime in cyber space, like Identity theft and hacking of the users’ account. It is also observed by (Ritzer, 2015) that networks have a dynamic quality with the structure of the system changing with shifting patterns of coalition and conflict.

Social Network Analysis (SNA) is a research technique that focuses on identifying and comparing the relationships within and between individuals, groups and systems in order to model the real-world interactions at the heart of organisational knowledge and learning processes, (Davies, 2014). The SNA process in this study will be undertaken as a mass surveillance tool involving the following:

1. Collecting information about relationships within a defined group or network of people and identifying the target network.
2. Collecting data and mapping out the network visually, mapping using a software tool designed for the purpose and generating a baseline through the analysis of data.

3.2.4.3 Rational Choice Theory

Gottfredson, (1990) argued in his Rational choice theory that an individual decision to commit a crime is based on cost – benefit proportion. 'Rationality' means that an individual balances cost and benefit to arrive at action that maximizes personal benefit. Cyber stalkers commit a crime after weighing the prospective rewards against the potential risk and stalking via the Internet allows the offender to do it from a relatively remote distance. The offence inflicts the same type of fear and harassment as in the case of victims who are in direct face-to-face situation with cyber stalkers. Rational choice theory is accepted by many people because it assumes that people act in a manner that is rational and it assumes that many of the cyber criminals are very talented and well educated, not necessarily in formal manner but they have an ability to think rationally. They attack the victims whom they believe would give them the greatest amount of financial gain with least chance of getting caught with the high tech cybercriminals hardly being caught because of the skill to cover their tracks and move through proxy servers so that they are undetected. They commit large fraudulent schemes and remain undetected online.

3.2.4.4 Differential Association

This theory is most widely accepted theory in criminology. It was first proposed by Sutherland (1924) and it was originally created to explain the rise in white collar crimes. The basic idea behind this theory is that criminal tendencies are learned in interaction with other deviant persons. It is through interaction with others that one engages in illegal acts. This theory could explain why normal law-abiding individuals can turn into criminals or deviants depending on the circumstances that they may be put into. This theory considers social environment as a means to explain why some individuals engage in criminal behaviour. This is seen in poor socio- economic conditions which encourage disobedience of law and authority. The main premise of this theory is that criminal behaviour is learnt through social interactions.

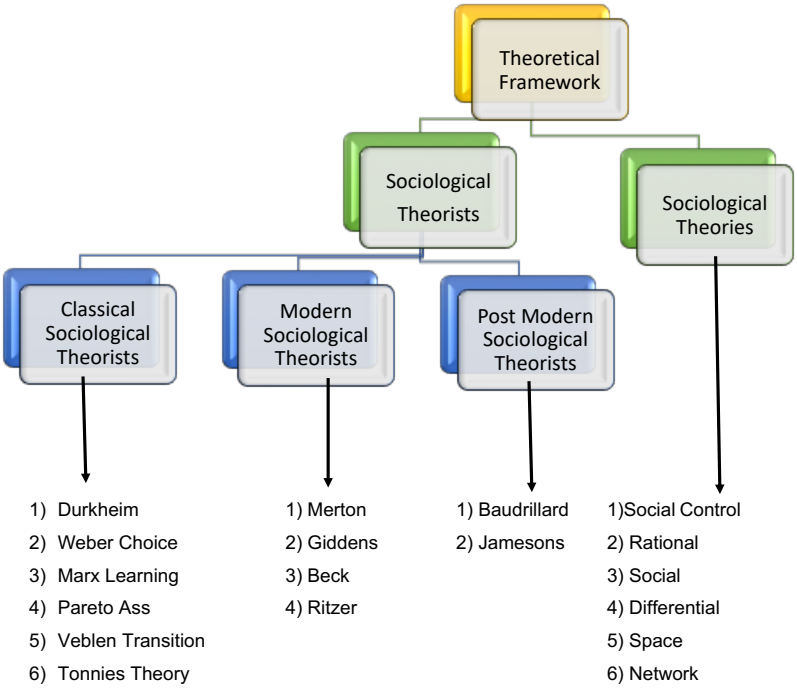


Figure 3. 1: Theoretical Framework (Source: Dingeldine 2014)

3.3 Research Taxonomy

Table 3.1 presents the various research taxonomies used from previous research conducted on this topic. Based on the studies conducted and the taxonomies used, a relevant research taxonomy was identified for this study.

The Theoretical Framework for this study will be based on the Space Transition Theory as proposed by Jaishanksr (2007), where he explains the behaviours of the persons who bring out their conforming and non-conforming behaviours in the physical space and virtual space. Jaishanksr (2007) also noted that in the Space Transition Theory people behave differently when they move from one space to another and that individuals that repress crime in a physical space, would commit such crimes in the virtual space. Jaishanksr (2007) also noted concepts such as Cyber stalking and Cyber as defamation and where offenders make use of online space because of its anonymity and widespread approach. Related studies undertaken by McCoy (2008), Chen (2014) and Westlake (2017) utilised the Space Transition Theory as a Theoretical Framework where attempts were made to log TOR traffic and determine Dark Web traffic. This study will utilise the Space Transition Theory as utilised by McCoy(2008), Chen (2014)

and Westlake (2017). The rationale for utilising the theory is that this study is similar to the studies mentioned above as the implementation of TOR in accessing the Dark Web ensures a degree of anonymity that will provide individuals in South Africa with the confidence to engage in illicit activities.

Table 3. 1: Research Taxonomy

Author	Study Conducted	Research Taxonomy Used
Bai (2014)	The capturing of TOR traffic utilising a Testbed.	Theoretical framework of the study was based on Network Theory of Crime
Barker (2014)	Collected TOR network traces by developing a complete TOR setup and attempted to log Dark Web usage	Theoretical framework of the study was based on Social Learning Theory and Cybercrime and the Network Theory of Crime.
Alsabah (2014)	Enhanced TOR's performance by developing an algorithm	Theoretical framework of study was based on Network Theory of crime.
Houmansadr (2014)	Differentiated the traffic of anonymous networks from other network traffic and claimed that mimicking other traffic is an obsolete way for anonymity	Theoretical framework of study was based Network Theory of Crime
Chakravarty (2012)	Provided improvements on TOR's performance	Theoretical Framework was based on the Differential Theory.
Winter (2014)	To ascertain how China is blocking TOR	The Theoretical Framework for the study was based on the Differential Theory of crime.
Chaabane (2012)	Monitored TOR traffic in 6 countries	Theoretical Framework of the study was based on the Social Learning Theory and Cybercrime.
Tang (2014)	Development of an algorithm to help TOR's performance improvements.	Theoretical Framework for the study was based on the General Theory of Crime.

McCoy (2014)	Attempted to log traffic passing through the TOR network and also attempted to log Dark Web usage	Theoretical Framework of the study was based on the Space Transition Theory.
Murdoch (2015)	Observed traffic entering and leaving several TOR entry and exit nodes within the UK and attempted to log TOR traffic.	Theoretical Framework of the study was based on the Network Theory.
Moore (2016)	Creation of 12 category taxonomy	Theoretical Framework of study was based on Social Learning Theory and Cybercrime.
Westlake (2017)	Created a web crawler specifically focused upon locating child exploitation materials	Theoretical Framework of study was based on Space Transition Theory.
Chen (2014)	Established a schema for categories of use for the web (also applied to 'dark webs'), focused upon terrorist organisations	Theoretical Framework of study was based on Space Transition Theory.

3.4 Exit Routing Traffic

The TOR network is a second-generation onion routing system that aims to provide anonymity, privacy, and Internet censorship protection to its users. As an open-source project, TOR is run by a non-profit organisation, routes encrypted TCP traffic through a worldwide network of over four thousand relays run by volunteers across the world. Despite attacks against it, TOR remains one of the most popular and secure tools to use against network surveillance, traffic analysis, and information censorship. The implementation of the TOR network consists of an end user (client), the various relays (internal TOR communication) and the exit nodes.

The architecture of TOR is based on three layers of encryption with the three different layers comprising of an entry node, middle node and exit node. The exit node is the point of exit to the desired onion address that a client wishes to access. To better understand South African TOR usage, the study sets up a TOR exit node on a 5 GB/s network link. This node joined the currently deployed network during the period of 27th

September – 5th November 2017. The first week was utilised as a testing period for the exit node and the data gathered during this period was not utilised in the analysis of this study. The experiment was conducted at a local cyber-security company, Walcom Security. The reasons for the utilisation of their premises was due to the following reasons:

1. The company has in their possession a proprietary anti money laundering system (as trusted and utilised by most of the global finance providers). This allowed for a currently unpublished vector of attack against the anonymising network which is currently protected by international patents and European data privacy laws as defined within the Bundesdatenschutzgesetz. The Bundesdatenschutzgesetz is a data protection act that with the data protection acts of Germany governs the disclosure and exposure of personal data that are stored in Information Technology systems.
2. The allocation of an external Internet Protocol Address is against university policy and therefore the exit node could not be placed on the Universities network.

The software and hardware components for the construction and implementation of the TOR exit node comprised the following:

1. Front end Dark Web pair routing devices that had the following specifications: 16Gigabytes of Ram, 256 Gigabytes of Solid Slate Drive, 4 Terabytes attached RAID, Secondary Network Interface Cards (Software MAC allocation). The network interface cards (NIC) utilised in the experiment did not possess any MAC addresses. These are exclusive NIC that cannot be purchased over the counter.
2. A Client Management System that had the following specifications:8 Gigabytes of RAM, 512 Gigabytes Hard Drive
3. A TOR licence that enabled the exit node to be placed on the TOR network. This had to be paid in Bitcoin.
4. Open source and proprietary application software such as Shallas Secured Services and Zvelo. The two software applications allowed for the declassification of the websites. Zvelo claims a 99% accuracy rate due to the large database of Uniform Resource Locator they possess.

The TOR Network is an overlay network whose infrastructure is run entirely by volunteers resulting in the largest public anonymity network in the world. Clients that use the TOR network construct circuits (paths) which are utilised to route multiple network streams. A circuit is considered secure if there is one non-malicious router in the circuit. The relay monitor (arm) provides real time statistics for bandwidth, CPU, and memory usage, relay's current configuration, logged events and connection details. TOR with its three layers of encryption is seen as a mix network. The reason behind mixed networks is that they aim to route IP traffic through an overlay of chain mixes as shown in figure 3.2 thus hiding their relationship between origin and destination. Our investigation introduced a compromised entry and exit point to establish that two disparate systems are using the TOR network and establishing a timing analysis is able to link the two routing nodes as shown in figure 3.2.

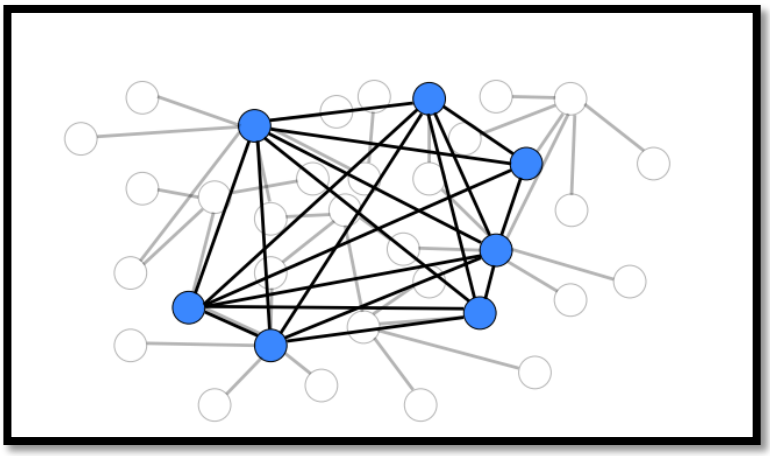


Figure 3. 2: Time Analysis

In addition to this a compromised exit node allowed us to perform meet in the middle interception attacks to identify the nature of all outbound (TOR to standard INTERNET) traffic as shown in Figure 3.3. This in collaboration with publicly available address classification sites ensured a relative degree of external address classification. As the bandwidth in South Africa is incredibly limited and latency to TOR nodes in Europe, Middle-East and Africa was considered to have higher latency, it was possible to identify TOR nodes utilising South Africa as the country of origin.

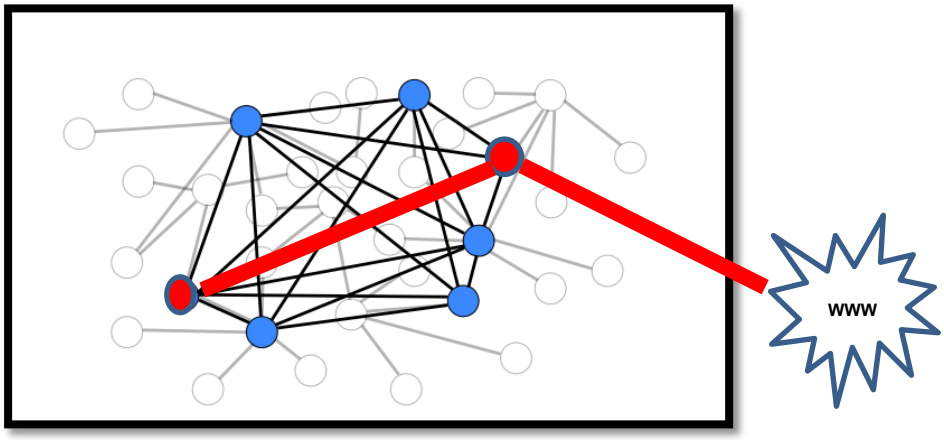


Figure 3. 3: Meet in the Middle Attack

3.4.1 Geolocation Method

The geo-location method will help to obtain the exit routing traffic by country and the geo-location of the IP addresses. This is accomplished at the third phase of the experiment where the end user will exit the TOR exit node and connect to an onion address in the Dark Web. It is at this point that TOR has its greatest vulnerability and allows for the logging of TOR traffic.

At a high level, the exit router traffic detection technique is based on the running of a packet sniffer on a local network. Tcpdump is configured to perform reverse DNS queries on IP addresses they monitor, therefore if one has access and control to the authoritative DNS server for a set of IP addresses, then one may trace reverse DNS queries back to the exit node. The detection methodology for the experiment is presented in figure 3.4:

1. An authoritative domain name server (DNS) that maps domain names to a vacant block of IP addresses. This server the researcher controlled.
2. Utilising of a TOR client, a circuit will be established using an exit router.
3. Once a circuit has been established, a SYN ping is sent to one of the IP addresses for which the researcher provides a domain name resolution. This procedure as illustrated in figure 3.4 is repeated.

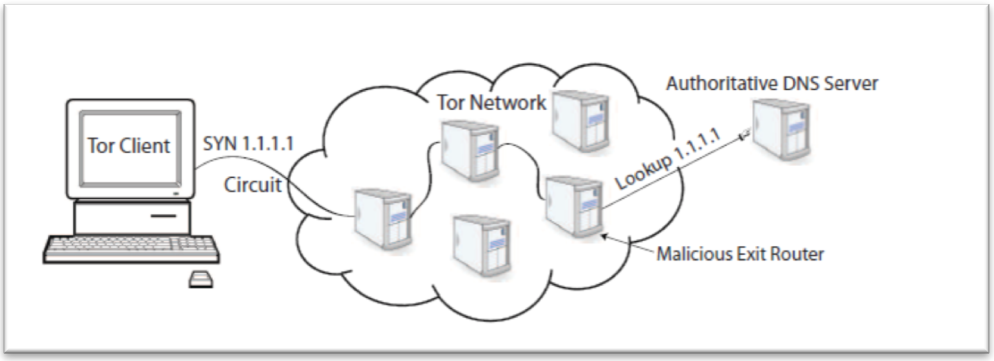


Figure 3. 4: Experiment Procedure

The review included the publication of a TOR relay as well as a compromised exit node. The TOR relay is able to be configured by modifying the TOR runtime settings to allow all traffic to be logged. As some of the initial TOR traffic is via the Internet it is possible to start recording the origination TOR traffic as well as the possible next-hop as well. With the implementation of a compromised exit node via the installation of a standard HTTP and HTTPS transparent proxy it is possible to record all Internet facing websites visited. Although the usage of traffic in South Africa did identify some nefarious usage, the majority was attributed to proxy avoidance of work and school systems. The review of the relay traffic indicated a combination of both the source and destination traffic as well as the encrypted TOR node endpoints. Based on this initial relay and next hop, the following diagram indicated the next hop traffic flows as shown in figure 3.5:

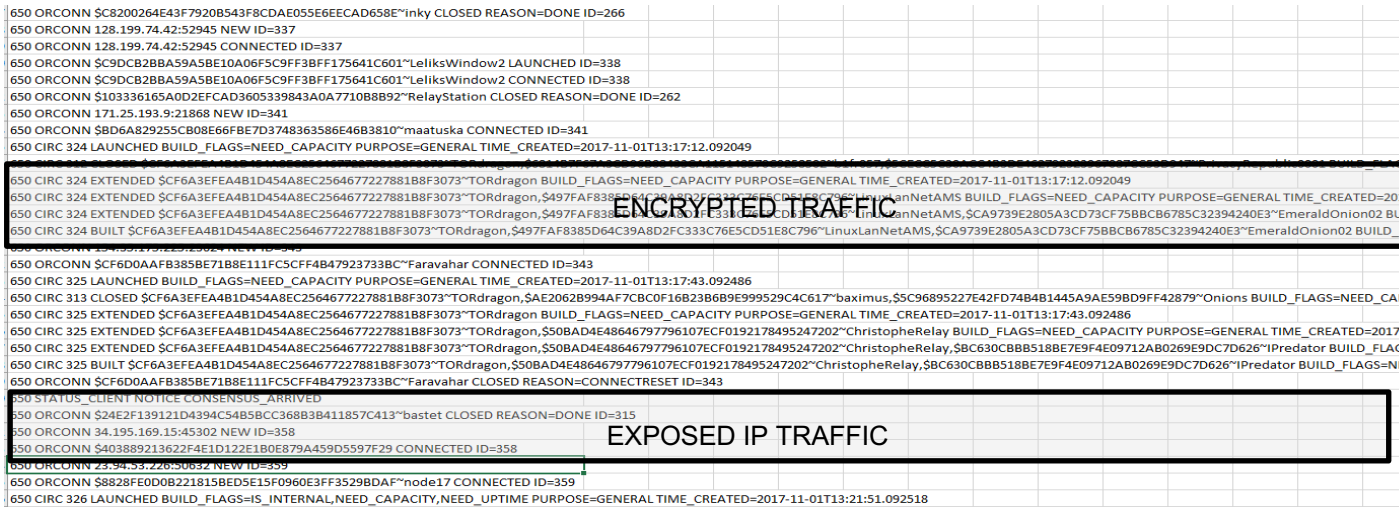


Figure 3. 5: Encrypted Traffic

Figure 3.6 shows the log file from the server and provided an indication of the encrypted traffic passing through the node, however the IP address of the clients were clearly exposed. This allowed for the attainment of the geo-location of the IP addresses. With the implementation of a monitored exit node, it was also possible with the transparent proxies to intercept, record and view all traffic. This in combination with the usage of traffic classification sites (e.g. Shalla's blacklist, http://www.shallalist.de/) it became possible to classify the sites from the Squid proxy logs.

Size	URL
2.3 M	outlook.office365.com:443
2.2 M	ow2.res.office365.com:443
2.2 M	ow2.res.office365.com:443
4.2 M	mail.google.com:443
4.0 M	outlook.office365.com:443
2.1 M	outlook.office365.com:443
2.6 M	ow2.res.office365.com:443
2.5 M	http://adl.windows.com/appraiseradl/2019_01_31_03_04_x64.cab
9.6 M	http://www.mithastravel.co.za/images/packages/5313681615_baa27725ea_o.jpg
5.7 M	http://www.mithastravel.co.za/images/packages/52784.jpg
3.4 M	outlook.office365.com:443
2.9 M	www.google.com:443
9.6 M	encrypted-tbn0.gstatic.com:443
4.2 M	www.capehalal.com:443
2.9 M	r4.res.office365.com:443
3.1 M	r4.res.office365.com:443
2.5 M	http://adl.windows.com/appraiseradl/2019_01_31_03_04_x64.cab
3.4 M	r4.res.office365.com:443
3.1 M	outlook.office365.com:443
3.4 M	mail.google.com:443
2.5 M	http://adl.windows.com/appraiseradl/2019_01_31_03_04_x64.cab
2.3 M	r4.res.office365.com:443
3.8 M	mail.google.com:443
2.5 M	http://adl.windows.com/appraiseradl/2019_01_31_03_04_x64.cab
4.6 M	ow2.res.office365.com:443
2.5 M	http://adl.windows.com/appraiseradl/2019_01_31_03_04_x64.cab
3.1 M	outlook.office.com:443
4.7 M	d1y6jrbzotnyjg.cloudfront.net:443
4.1 M	r4.res.office365.com:443
5.1 M	attachments.office.net:443
2.8 M	outlook.office365.com:443
2.8 M	outlook.office.com:443
5.6 M	outlook.office.com:443
3.7 M	settings-win.data.microsoft.com:443
3.7 M	settings-win.data.microsoft.com:443
3.7 M	settings-win.data.microsoft.com:443
3.0 M	i.pinimg.com:443
2.1 M	outlook.office.com:443
52.2 M	pctrainingbu-my.sharepoint.com:443
3.1 M	ow2.res.office365.com:443
2.6 M	r4.res.office365.com:443
3.7 M	outlook.office.com:443

Figure 3. 6: Logged Traffic

3.4.1.1 Network Analytical Tool

Table 3.2 is an analysis of previous studies conducted by McCoy (2008) and Barker (2014). The related studies did not use any network analytical software such as Gephi to portray exit routing traffic by country. This study therefore aims to further improve on the studies mentioned above by utilising Gephi to provide a graphical representation of exit routing traffic. Gephi architecture allows it to work with complex large data sets.

The data logged during the experiment exceeded 20 Terabytes therefore the use of Gephi would allow for the large data set to be efficiently and effectively analysed.

Table 3. 2: Related Studies

Study	Network Analytical Software	Reason
Barker (2014)	None	None
McCoy (2008)	None	None

3.4.1.2 Network Analytical Tool Graphs

Description Nodes represent the country of the exiting routing traffic, namely, South Africa. A directed edge is created when there is exit routing traffic to another country. The layout of the network is created using Gephi's ForceAtlas2 (FA2) algorithm, which created an easy-to-interpret graph. The FA2 algorithm is continuous and optimized for speed (suitable for dynamic graphs, as it will efficiently update in real time) and offers various options to help fine-tune the results. A directed graph or digraph in Gephi which is an ordered pair where D = (V, A) with:

• V represents a set whose elements are called vertices or nodes, and

• A represents a set of ordered pairs of vertices, called arcs, directed edges, or arrows.

• An arc a = (x, y) is considered to be directed from x to y; y is called the head and x is called the tail of the arc; y is said to be a direct successor of x, and x is said to be a direct predecessor of y. If a path leads from x to y, then y is said to be a successor of x and reachable from x, and x is said to be a predecessor of y. The arc (y, x) is called the arc (x, y) inverted. A directed graph D is called symmetric if, for every arc in D, the corresponding inverted arc also belongs to D. A symmetric loop less directed graph D = (V, A) is equivalent to a simple undirected graph G = (V, E), where the pairs of inverse arcs in A correspond 1-to-1 with the edges in E; thus the edges in G number |E| = |A|/2,

or half the number of arcs in D. A variation on this definition is the oriented graph, in which not more than one of (x, y) and (y, x) may be arcs.

3.4.1.3 Degree Distribution

The node degree is the number of relations (edges) of the nodes. However, in the case of the directed networks, the study distinguished between in-degree (number of incoming neighbours) and out-degree (number of outgoing neighbours) of a vertex.

Degree sum formula (also sometimes called the **handshaking lemma**),

$$\sum_{v \in V} \deg(v) = 2|E| \qquad 3.1$$

for a graph with vertex set *V* and edge set *E*. Both results were proven by Leonhard Euler (1736) in his famous paper on the Seven Bridges of Königsberg that began the study of graph theory. For a vertex, the number of heads ends adjacent to a vertex is called the *in-degree* of the vertex and the number of tail ends adjacent to a vertex is its *out-degree* (called "branching factor" in trees).

Let G = (V, E) and v∈V. The in-degree of v is denoted deg− (v) and its out-degree is denoted deg+ (v). A vertex with deg− (v) = 0 is called a source, as it is the origin of each of its incident arrows. Similarly, a vertex with deg+ (v) = 0 is called a sink. If a vertex is neither a *source* not a *sink*, it is called an *internal*. The *degree sum formula* states that, for a directed graph,

$$\sum_{v \in V} \deg^+(v) = \sum_{v \in V} \deg^-(v) = |A|. \qquad 3.2$$

If for every vertex v∈V, $\deg^+$ (v) = $\deg^-$ (v), the graph is called a balanced directed graph. Degree has generally been extended to the sum of weights when analysing weighted networks and labelled node strength, so the weighted degree and the weighted in- and out-degree was calculated (Barrat, 2004, Newman 2001, Opsahl, 2010).

3.5 Determination of Dark Web Usage

The results of the experiment are crucial to law enforcement agencies as it will provide an insight into any illicit Dark Web usage that individuals are engaging in. With the development and deployment of the experiment, there were concerns on the TOR exit node as such nodes generally come under serious attacks from various forms of Ransomware and viruses. There was constant monitoring undertaken in ensuring the exit node was protected at all times, hence the test period of one week and the minimal bandwidth allocation to the server resulting in minimal traffic as shown in figure 3.7. There were attempts however to breach the server.

Figure 3.7 is a screen shot of the amount of routing traffic passing through the exit node. During the test period only 2.5mib/s of bandwidth was allocated to the server, hence the low traffic on Friday 22 from 6pm to around 10am. Once this test period was over the bandwidth was increased to 5mib/s and this resulted in a huge spike in traffic around 12pm.

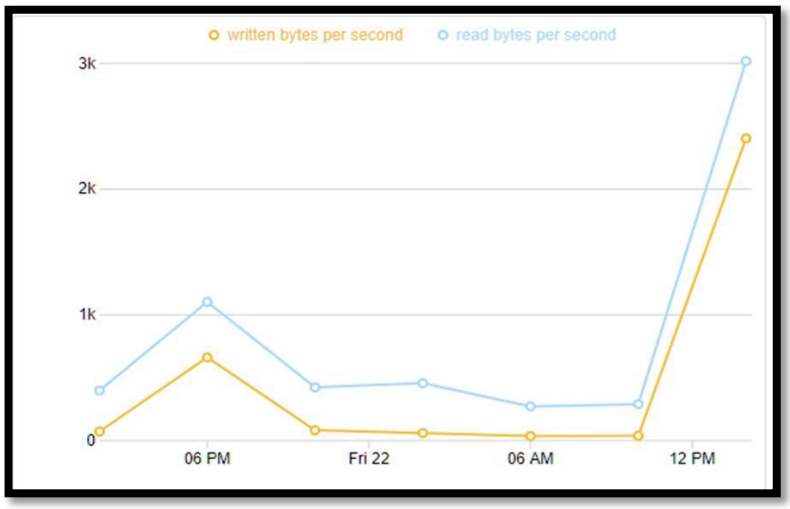


Figure 3. 7: Bandwidth Allocation

The nature of TOR makes it extremely difficult to find the source of the technology stack. TOR traffic on a network appears as normal web traffic as presented in figure 3.8. TOR is a browser that not merely provides encryption it is also designed to look like normal website traffic. Figure 3.8 is a screenshot of the connections on the server and the TOR onion addresses look no different than that of a real HTTPS

session. The utilisation of software such as CapLoader provides the tools to differentiate the different types of SSL traffic.

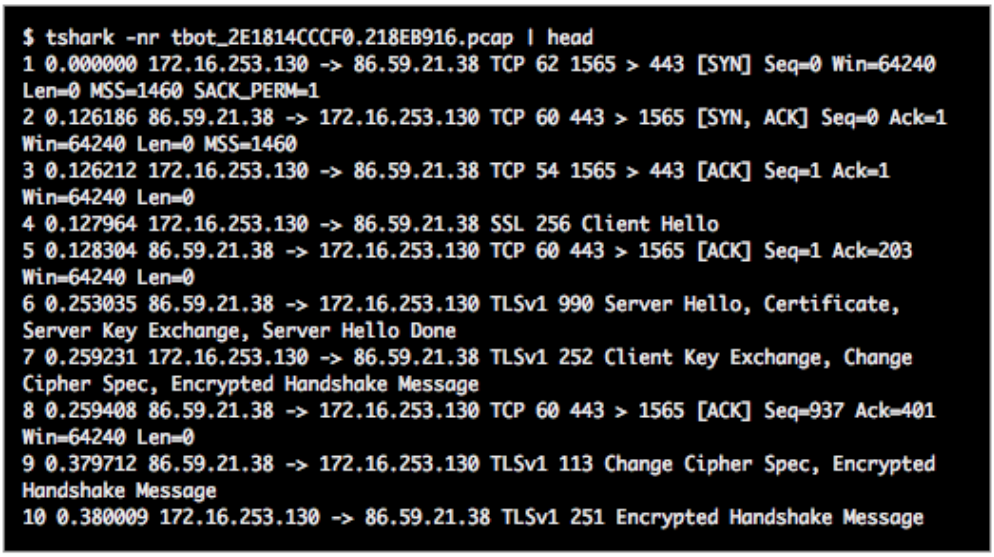


Figure 3. 8: HTTPS Connections

The one objective of the study was to ascertain what websites are being visited by the end user. The experiment was first conducted over a test period of four weeks. During this time there was minimal bandwidth allocation to the TOR exit node. The data collected during this period was not utilised for the study. A screen capture from the actual TOR logs on the server is presented in figure 3.9. The logs from the server provide an indication of the websites that were being visited at that point in time. All logs from the TOR network passing through the exit node was captured and analysed. From the figure it is evident that there were numerous updates being downloaded as a result of individuals possibly at work utilising TOR to bypass the proxy server.

```
http://85.25.218.149/Getintopc.com/Adobe_Dreamweaver_CC_2017_v17.0.1.9346.zip?
oneclient.sfx.ms:443
http://188.138.68.26/Getintopc.com/Adobe_Dreamweaver_CC_2017_v17.5.0.9878.zip?
http://download1568.mediafireuserdownload.com/51iwtx5408tg/344j4w3ddn4gw79/shqdespicbleme3webdl72.Ganool.ee.mkv
http://download1568.mediafireuserdownload.com/51iwtx5408tg/344j4w3ddn4gw79/shqdespicbleme3webdl72.Ganool.ee.mkv
http://download1568.mediafireuserdownload.com/51iwtx5408tg/344j4w3ddn4gw79/shqdespicbleme3webdl72.Ganool.ee.mkv
http://download1568.mediafireuserdownload.com/51iwtx5408tg/344j4w3ddn4gw79/shqdespicbleme3webdl72.Ganool.ee.mkv
http://download1568.mediafireuserdownload.com/51iwtx5408tg/344j4w3ddn4gw79/shqdespicbleme3webdl72.Ganool.ee.mkv
http://download1568.mediafireuserdownload.com/51iwtx5408tg/344j4w3ddn4gw79/shqdespicbleme3webdl72.Ganool.ee.mkv
http://download1568.mediafireuserdownload.com/51iwtx5408tg/344j4w3ddn4gw79/shqdespicbleme3webdl72.Ganool.ee.mkv
http://download1568.mediafireuserdownload.com/51iwtx5408tg/344j4w3ddn4gw79/shqdespicbleme3webdl72.Ganool.ee.mkv
http://download1568.mediafireuserdownload.com/51iwtx5408tg/344j4w3ddn4gw79/shqdespicbleme3webdl72.Ganool.ee.mkv
http://download1568.mediafireuserdownload.com/51iwtx5408tg/344j4w3ddn4gw79/shqdespicbleme3webdl72.Ganool.ee.mkv
http://download1568.mediafireuserdownload.com/51iwtx5408tg/344j4w3ddn4gw79/shqdespicbleme3webdl72.Ganool.ee.mkv
http://download1568.mediafireuserdownload.com/51iwtx5408tg/344j4w3ddn4gw79/shqdespicbleme3webdl72.Ganool.ee.mkv
http://download1568.mediafireuserdownload.com/51iwtx5408tg/344j4w3ddn4gw79/shqdespicbleme3webdl72.Ganool.ee.mkv
```

Figure 3. 9: TOR Logs from Server

The network history of the TOR server when the bandwidth allocation was increased is presented in figure 3.10. The network history provides an overview of the amount of data being sent and received through the server. It also provides an indication of the amount of bandwidth consumed at that point in time and a graphical representation in the form of a bar graph is presented to clearly provide an indication of the increase/ decrease of traffic passing through the server. The extra bandwidth allocated to the node automatically resulted in large amounts of traffic, resulting in large amounts of data been consumed. There was a total of 385Gigs downloaded within the space of 20 minutes. An increase in bandwidth will result in an increase in TOR traffic resulting in more engagement of activities on the Dark Web.

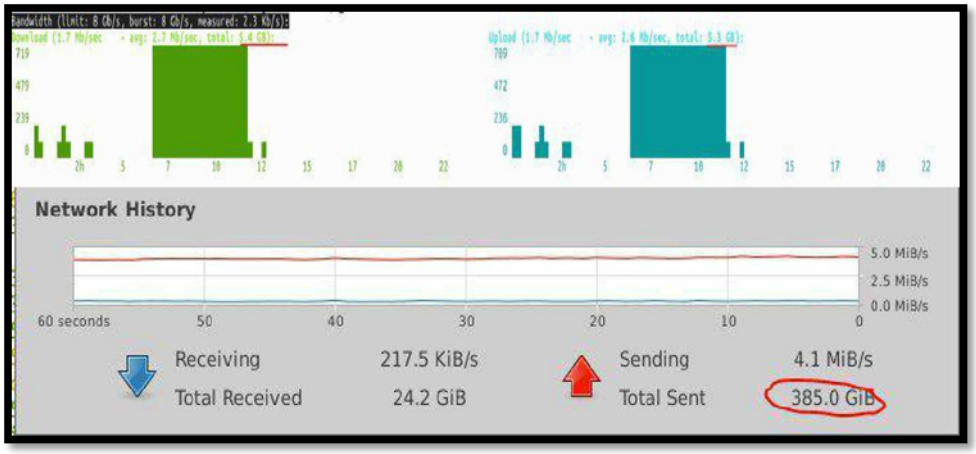


Figure 3. 10: Network History

3.5.1 Classification of Illicit Dark Web Usage

The classification of Dark Web analysis is a two-class problem that classifies Dark Web usage into illicit and non-illicit usage. During the first week of the pilot study there were no configuration issues experienced on the server. The first three days of the testing period, a minimal amount of bandwidth was allocated to the TOR exit node and the bandwidth was increased over the forthcoming days of the testing period. An increase in the bandwidth allocation to the exit node saw excessive amounts of traffic passing through the node, however the node remained stable. The only issue encountered over the testing period were the cyber-attacks the node came under.

3.5.2 Previous Classification Methods

Table 3.3 shows the tools used in conducting a Dark Web Crawl. The core component after conducting the crawl is the declassification of the websites. The framework for the classification of Dark Web traffic for this study was developed on previous methods used in past studies. A new classification architecture was developed based on previous methods. The previous methods could not log large amounts of Dark Web traffic and hence the tools used in previous could not be utilised in this study. The experiment implemented in this study aimed to log all entry and exit routing traffic. The exit routing traffic directs a client to the onion address they wish to access, and it is at this point that exit routing traffic will be captured. All data captured was stored in a big data analysis engine that had a number of opensource and proprietary application

software for the classification of onion addresses. Some of these tools utilised were Shallas Secured Services and Zvelo

The reason for the utilisation of the chosen tools is that they have a large database of URL's and ensure a 99% accuracy rate.

Table 3. 3: Classification Methods

Study	Method used for analysis	Reason
Barker (2014)	Selenium browser testing framework	Only aimed to capture 30 websites and for a time period of 120 minutes. The application will therefore be sufficient to log traffic.
McCoy (2008)	Utilised a reverse DNS approach with a SYN ping	Primary focus was to log exit routing traffic to countries.

3.6 Geo-Location Analysis Algorithm

The TOR Network is a network that is solely run by volunteers across the world who make freely available their bandwidth. Clients that use the TOR network construct circuits (paths) which are utilised to route multiple network streams. A circuit is considered secure if there is one non-malicious router in the circuit. The anonymising relay monitor (arm) will provide for real time statistics for:

1. Bandwidth, CPU and memory usage. The amount of bandwidth allocated to the TOR exit node and the total amount of bandwidth the TOR traffic is consuming.
2. Relay's current configuration. The current configuration setup for the exit node.
3. Logged events. The logging of exit routing traffic to other countries with determining the geo-location of the IP addresses and Dark Web traffic.
4. Connection details. Provide status on the performance status of the exit node.

The flowchart of figure 3.11 represents a comprehensive overview on the stages of logging Dark Web traffic and presents a comprehensive overview on the integration of the various TOR configuration settings that allowed for the logging of exit routing,

obtaining the geo-location of the IP addresses and determining Dark Web traffic in South Africa. The defaults and interface properties are configurable via a configuration file with the following configuration settings:

1. **CIRC** – This is information on newly created, already existing and closed TOR nodes. This module is responsible for ensuring the connection to the TOR network is maintained.
2. **STREAM -** Information on status of application streams including which circuit is used for the connection (for example, HTTP based connection data).
3. **ORCONN -** Determines newly established and closed connections to TOR nodes.
4. **BW –** Stands for **Bandwidth** utilised by the TOR node.
5. **STREAM_BW** - Bandwidth used by the various streams within the TOR node configuration.
6. **DEBUG, INFO, NOTICE, WARN, ERR** – These are various Information messages related to the running of the TOR node. INFO - meaning information, WARN - meaning Warning and ERR - meaning Error.
7. **ADDRMAP** – This is an address map. It relates to a Domain-to-IP mapping that is cached by TOR client to determine the actual Internet address of the connected client. This will allow for determining the geo-location of the IP address.
8. **NEWDESC, AUTHDIR_NEWDESCS, DESCCHANGED** – These relate to the various TOR directory services.

   The comprehensive configuration file of the TOR node is attached as appendix A.

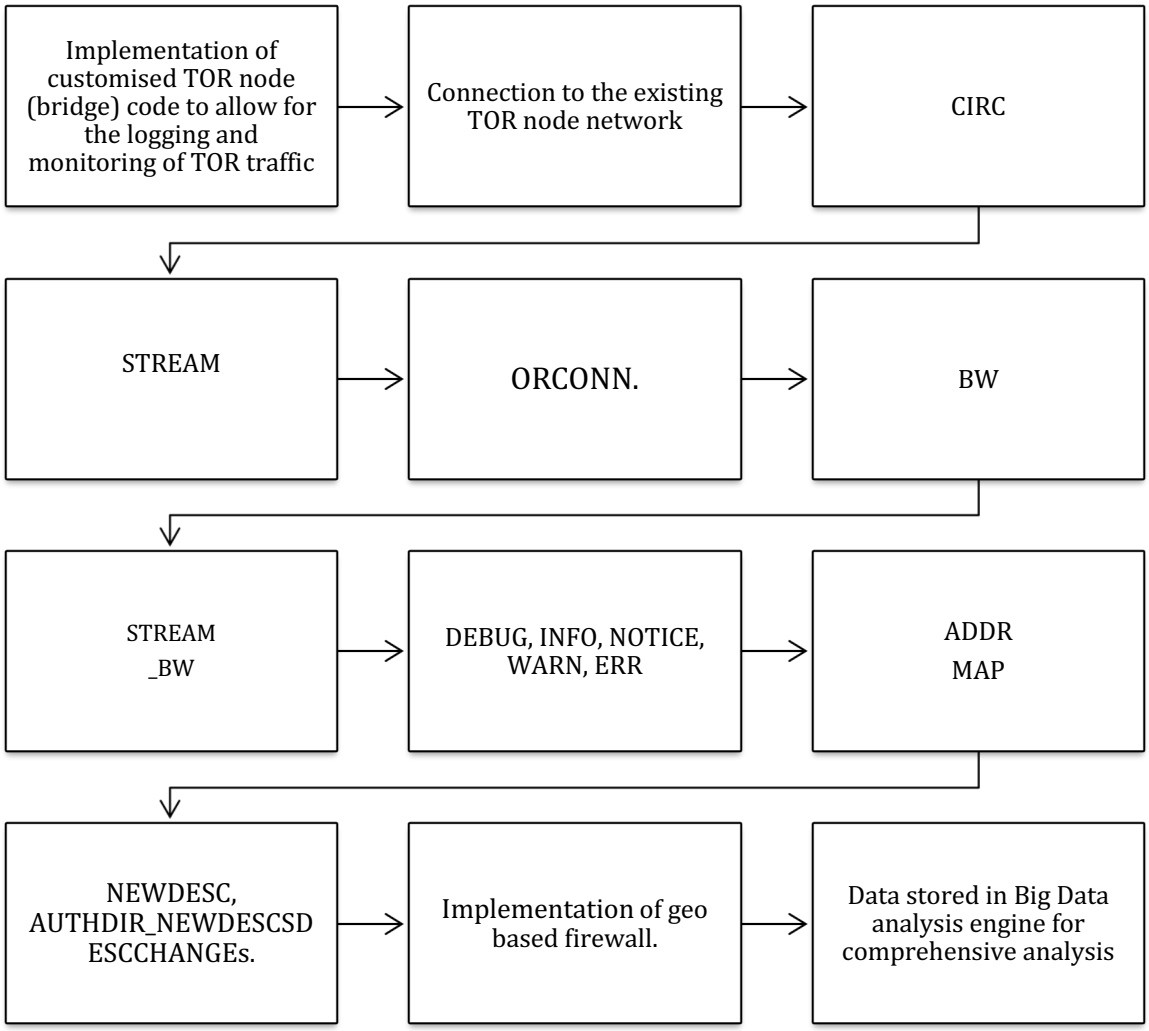


Figure 3. 11:Flowchart: TOR Configuration

An algorithm for the detection of exit routing traffic and Dark Web traffic is presented in the flowchart 3.12. This led to the successful logging and analysing of Dark Web usage in South Africa. The logging of TOR traffic will follow the process as presented in the flowchart 3.12:

1. For the TOR node to be setup on the TOR network the configuration file will have to be manipulated and thereafter configured onto the server as presented in figure 3.11.
2. The TOR exit node will log all the IP addresses of the clients and will be able to identify the exit routing traffic and the geo location of the IP address.
3. If the traffic identified is TOR traffic, the IP address and Dark Web traffic will be logged. The logged data will be stored in a big data analysis engine.

4. If conventional web traffic passes through the node, then neither the IP address nor the web traffic will be logged.

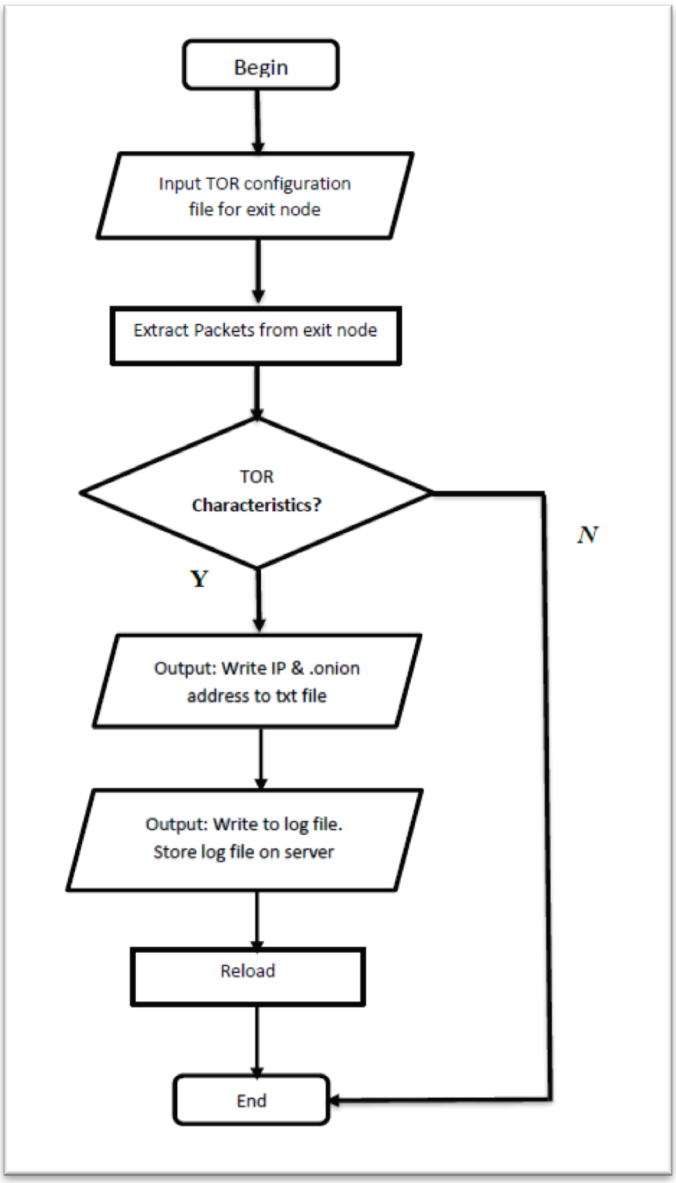


Figure 3. 12: Traffic Detection Algorithm

3.7 Summary

The aim of this study was to better understand how the Dark Web is currently being used or misused in South Africa. The methodology presented in this chapter provided further insight into the complexity of the experiment architecture to be implemented in the study. The chapter further provided an insight into the hardware and software requirements of the experiment that ensured for its successful implementation. A

sample TOR configuration file was illustrated with the entire configuration settings presented as an appendix. The chapter further dwelled into the Social Networking Analysis tool that will be utilised for network visualisations and basic statistical computations.

The next chapter will further provide evidence if the algorithm and experiment architecture presented in this chapter was successfully tested and implemented. The implementation of the experiment will lead to the collection of big data that will be analysed and the results will be presented in the next chapter. The data gathered and presented in the next chapter will provide an insight into Dark Web usage in South Africa.

CHAPTER FOUR

EXPERIMENTAL RESULTS

4.1 Introduction

This study aimed to ascertain Dark Web usage and the illicit trade that individuals are engaging in on the Dark Web. In order to achieve the aim of the study, a high-level experiment had to be undertaken which was based on studies by McCoy and Bauer (2008), Westlake (2017) and Chen (2014) who conducted an experiment to ascertain Dark Web usage. The studies by Westlake (2017) and Chen (2014) was conducted over periods of no longer than three days, whilst McCoy and Bauer (2008) conducted the experiment over a five - week period. For this study the author aligned the duration of the experiment to that implemented by McCoy and Bauer (2008) as the author believed that conducting the experiment over a longer duration will result in the collection of sufficient data that will provide a comprehensive analysis of South Africa's Dark Web usage. The methodology presented in the previous chapter gave a comprehensive overview on the experiment methodology that needed to be conceptualised, designed and implemented in order to obtain the data that will be presented in this chapter. The experimental results obtained was as a result of the successful implementation of the algorithm presented in chapter 3. The big data collected from the successful implementation of the experiment was critically analysed and the presentation of the data will outline the exit routing traffic by country with the geo-location of the IP addresses, provide a comprehensive overview on the websites visited by clients and a further declassification of the illicit websites that were visited during the conducting of the experiment. The social networking analysis tool, Gephi will present data on the network visualisation of the data set. This will result in a comprehensive overview of Dark Web usage and the illicit trade individuals are engaging in.

4.2 Dark Web usage in South Africa

An aim of the study was to declassify websites visited by individuals on the Dark Web. This study built on studies by Chen (2012) and Westlake (2017), where the researchers

restricted web content only to a few categories. This study will present all website categories, thus providing a comprehensive overview on the illicit trade that individuals are engaging in on the Dark Web. TOR is a browser that is utilised in order to access the Dark Web. The browser is not very familiar amongst many across the world and its popularity has only grown of late with the recent public disclosure of the numerous illegal activities occurring on the Dark Web. The South African daily TOR usage was relatively low as compared to America and Europe. The daily TOR usage averaged between 4000-6000 during the month of October and November 2017 when the experiment was undertaken.

The experiment was conducted during the last week of September 2017 and ended during the first week of November 2017. Figure 4.1 shows the TOR usage of South Africans during the implementation of the experiment. During this period, the total number of TOR users globally ranged between 275000 – 300 000 users and South Africa contributed only 2% of the total number of TOR users globally, which is disproportionately small to countries such as United States of America and Germany (Dingeldine, 2015).

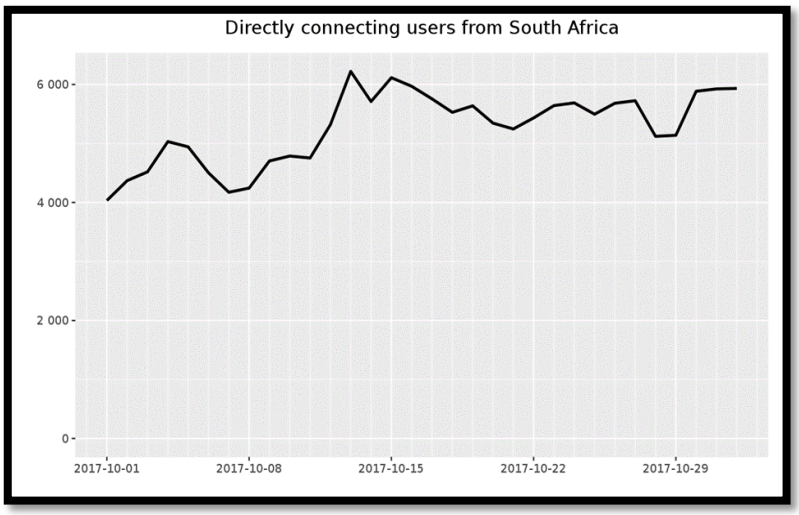


Figure 4. 1: TOR Users in South Africa

4.2.1 Bandwidth Usage

The construction of a TOR exit node on the TOR network is a time consuming and costly process. The purchase of a TOR licence may only be purchased using Bitcoin and the exit node requires large amounts of bandwidth to work effectively. The more

bandwidth allocated to the node, the more traffic will flow through it, thereby allowing for sufficient data to be gathered for analysis. A graphical representation of the bandwidth usage of the TOR exit node is presented in figure 4.2. There was a rapid increase in the number of users accessing TOR from Friday around 8pm, resulting in an increase in the total bandwidth being consumed. These results directly coincide with the research conducted by McCoy and Bauer (2014) where it was reported that web traffic makes up the majority of the connections and bandwidth, but non-interactive protocols consume a disproportionately larger amounts of bandwidth, when compared to interactive protocols. While HTTP traffic comprises an overwhelming majority of the connections observed, this traffic comprises of either interactive web browsing or non-interactive downloading.

There was a rapid increase in exit routing traffic when the bandwidth allocated to the exit node was increased to 5.0MiB/s as shown in figure 4.2. With the increase in bandwidth usage, large volumes of traffic passed through the exit node. A total of 385 Gigabytes was consumed by the exit node in the space of under three hours. The rapid increase in bandwidth usage was the result of the exit node being utilised in the building of a movie Torrent site on the Dark Web.

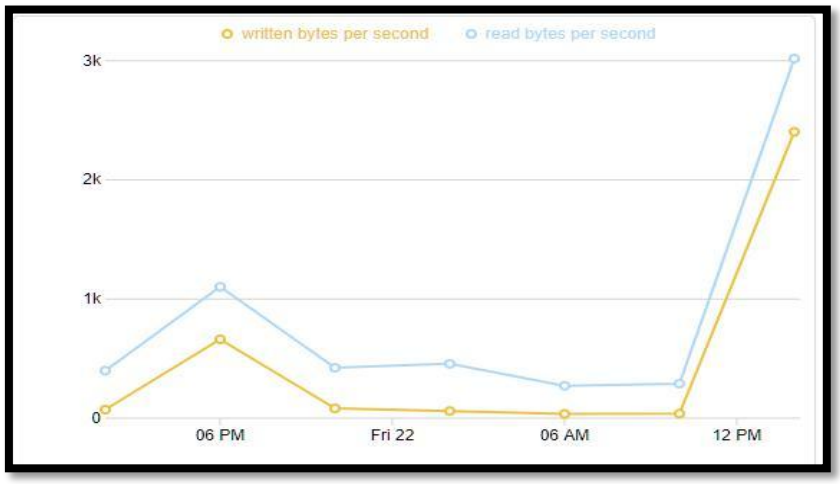


Figure 4. 2: Bandwidth Usage

4.3 Dark Web Misuse in South Africa

The Dark Web is synonymous with the engagement of illegal activities and is closely associated with a variety of black markets where one can purchase practically anything.

Warf (2016) noted that people wanting to engage in illegal activities will utilise the Dark Web as it provides the end user with some sort of anonymity.

Figure 4.3 is a presentation of the illicit activities that South Africans are engaging in on the Dark Web. Figure 4.3 gives us an indication that pornographic material comprised the majority of the illicit Dark Web usage in South Africa. Pornographic material is freely available both on the Internet and the Dark Web however, Norton (2016), Dingeldine (2015) and Chen (2014) noted in their studies that 80% of pornographic content on the Dark Web is assigned to child pornography. This is further entrenched by Vitare (2014) and Trendlabs (2017) where it was stated that 60% of child pornographic content is found on European servers whilst another 37% is found on servers found in the USA where the majority of the exit routing traffic was logged. From the statistics provided from the studies above, it can be assumed that a large portion of the pornographic material accessed by South Africans will be of a child pornography nature.

The sale of drugs on the Dark Web was synonymous with the infamous Silk Road and the majority of the trade that Silk Road engaged in was in the U.S and European countries as noted by Vinto (2015), Dolliver (2015), Van Buskirk (2015) and Christin (2015). There was no study that logged any drug transactions to the African continent by Silk Road or any other online drug store. The second largest amount of illicit traffic that South Africans engaged in pertained to drugs and there have not been any studies that associate South Africa with the purchasing of drugs online. This is an area of concern for the State Security Agency as there is clearly a presence of online drug trade occurring in South Africa.

There has been massive public disclosure on the crime rate in South Africa and just about every news headline is assigned to murder, rape or corruption. There have been many questions posed by authorities on how criminals have access to weapons that are only available to militant groups (Carvalho, 2017). The graph outlines the gun trade occurring in South Africa and the visitation to violent websites. These violent websites would be aligned to extremist websites, online stores selling violent products and forums that are related to violence. Weimann (2016), Ghaffar (2016), and Moore (2016) closely associated militant groups with violent activities and the access to weapons they have freely available on the Dark Web. This finding is therefore in accordance with the studies by Weimann (2016), Ghaffar (2016) and Moore (2016) and

pose a serious concern to law enforcement agencies in South Africa as the weapon trade in South Africa is rife.

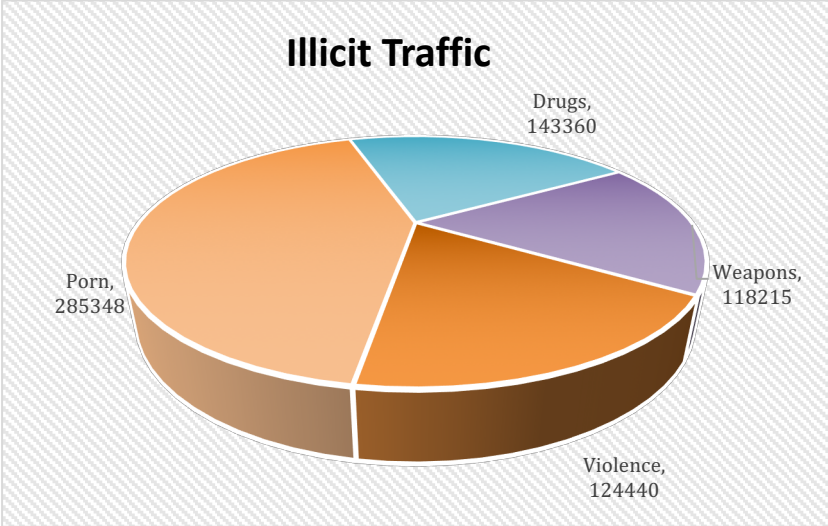


Figure 4. 3: Illicit Traffic

The recent WannaCry Ransomware sent shockwaves throughout the world with the monetary damage caused by its implementation running into the hundreds of millions of dollars (Hacquebord, 2015). Figure 4.4 presents a comprehensive analysis of all the malicious activities that South Africans engage in on the Dark Web. There is a large presence of hackers on the Dark Web and some utilise their skills to teach others in return for payment as noted by Moore (2016), King (2016) and Sancho (2015). The visitation of hacking websites will be individuals wishing to upskill themselves in that particular field.

A cause for concern is the visitation to websites that associate themselves with the selling or creating of Spyware. In recent times many South African governmental and organisational websites have come under serious attack. There is a possibility that the individuals visiting these government websites have malicious intent. This is in line with a study conducted by Goldsmith (2016), Carvalho (2017) and Norton (2016), where the findings on the usage of the Dark Web was to engage in illicit activities such as drug trade, child pornography and have access to large databases of ransomware and viruses.

Warez is software that has been stripped of its copyright protection and is made available on the Internet for downloading. Illegal movie downloads, leaked information

and pirated software is commonly found on the Dark Web and these can either be purchased or made freely available on torrent sites as noted by Yaneza (2014) and King (2016). There were approximately one hundred and fifty thousand visitations to websites to the Warez category. The logging of this malicious activity is therefore in aligned to studies of Yaneza (2014) and King (2016), and therefore it can be concluded that software piracy is rife in South Africa.

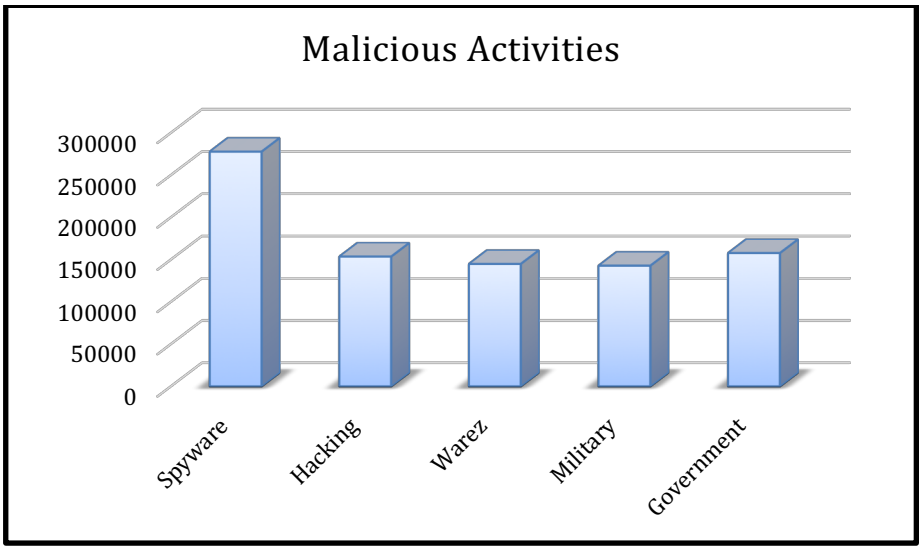


Figure 4. 4: Malicious Activities

Figure 4.5 shows some of the Dark Web activities that South Africans are engaging in, with these ranging from online gambling to accessing dynamic content and the visiting of celebrity/model websites.

Legislation in many countries does not permit gambling in an online environment and individuals find alternative means to participate in the online gambling world (Warf, 2016). Online gambling in South Africa is illegal and Section 11 of the National Gambling Act states: “A person must not engage in or make available an interactive game except as authorised in terms of this Act or any other national law”. The Dark Web has a variety of online gambling websites, thus making it a haven for online gamblers (Warf, 2016). The degree of anonymity that the Dark Web provides will encourage individuals to engage in online gambling as shown in figure 4.4. South Africans are therefore utilising the Dark Web to engage in online gambling due to its anonymous nature. This is in accordance with Warf (2016), King(2016) and Katyal

(2016) where they noted the use of the Dark Web to engage in illicit activities such as pornography and gambling due to its anonymous nature.

Dynamic content are dynamic pages that when accessed each time, the content on the pages continuously change, such an example of dynamic pages is Facebook, Reddit and online news providers (Sancho, 2015). Reddit is an infamous website that constantly gives updates on new feeds, fun stories, videos and memes (Sancho, 2015). Other dynamic web pages on the Dark Web include Facebooks .onion address and other social media platforms. The accessing of these web pages further reinforces the primary usage of TOR is to access social media platforms on the Dark Web, a contrast to the findings of Dingeldine (2015), Bauer and McCoy (2014), Sancho (2015), Chen (2014) and Westlake (2017) where they logged illicit traffic as the primary usage of the Dark Web.

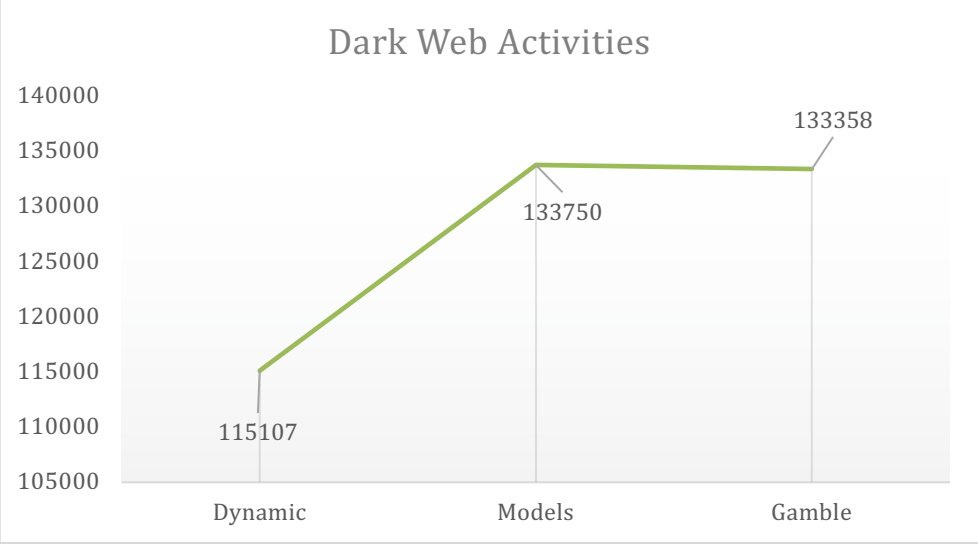


Figure 4. 5: Dark Web Activities

Figure 4.6 provides an overview on the various Dark Web services that South Africans are engaging in on the Dark Web. These services vary from the conventional Web TV services to the more radical services such as web radio, podcasts and news mediums where freedom of speech is encouraged. Oppressive governments see information as a powerful weapon and the movements of citizens are closely monitored to avoid the spread of a revolution as noted by Clark (2017), Goldsmith (2016), Crampton (2016) and Dobson (2016). The Dark Web promotes freedom of speech due

to its anonymity and individuals are free to express themselves on platforms without fearing any persecution. Platforms such podcasts and Web Radio provide individuals such as journalists to freely express themselves. The utilisation of these platforms is strongly used on the Dark Web by South Africans either for expressing their freedom of speech or engaging in discussion of sensitive topics that are prohibited on the Internet. This is in accordance with the findings from Flores (2016), King (2016) and Weimann (2016) where one of the primary uses of the Dark Web was to serve as a platform for journalists and other individuals to speak freely whilst remaining totally anonymous and to engage in forums and discussions that one might be prohibited to on the Surface Web.

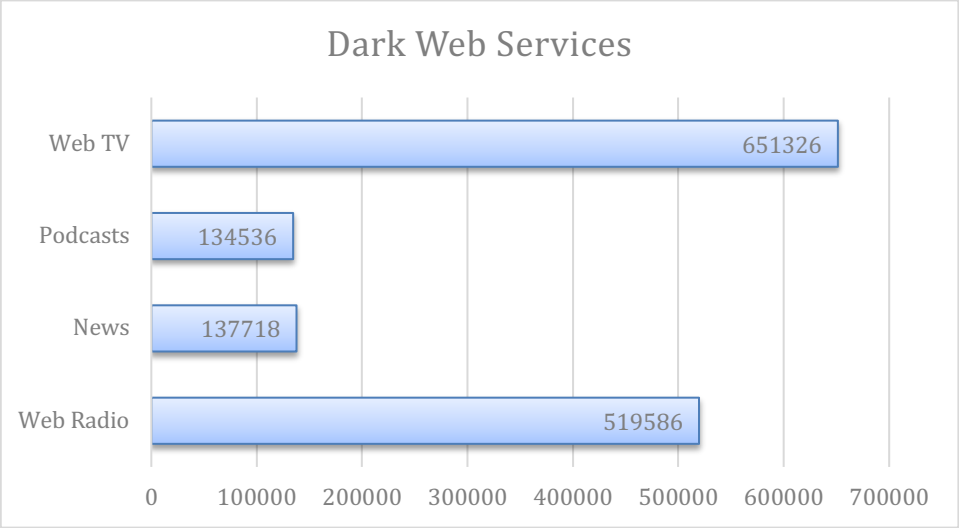


Figure 4. 6: Dark Web Services

Figure 4.7 shows the movie websites visited in relation to the number of downloads completed. A study by Sancho (2015) presented findings that there is a large presence of torrent sites on the Dark Web which leak new movie releases (Sancho, 2015). In this instance the TOR exit node of the experiment was utilised to build a movie torrent site on the Dark Web. The construction of the torrent site resulted in large bandwidth consumption and resulted in the issuing of a lawyers' letter from Paramount Pictures. This finding is therefore in accordance with Sancho (2016), where the study found one of the illicit activities that individuals engage in on the Dark Web is to view and supply leaked movie releases.

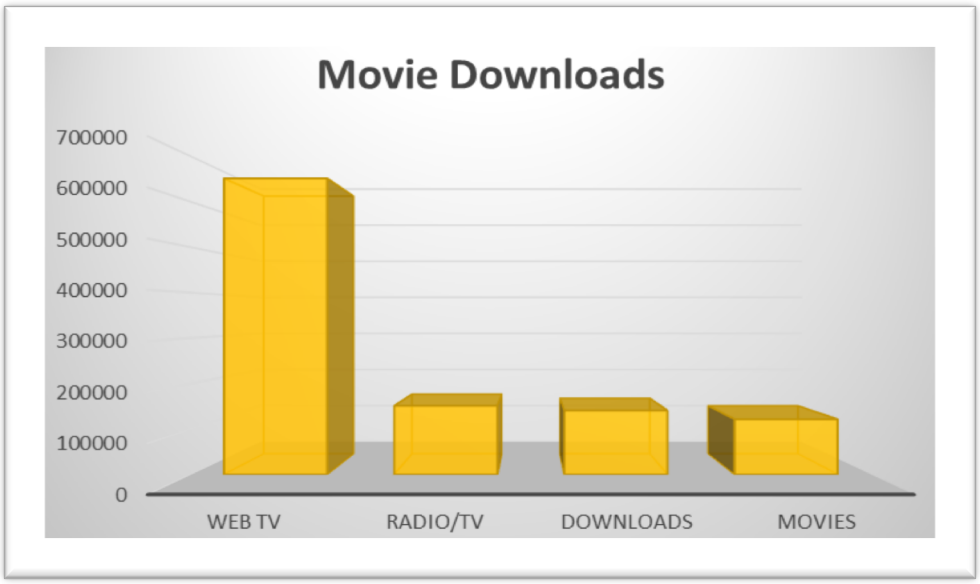


Figure 4. 7: Movie Downloads

4.4 Uses of the Dark Web in South Africa

The experiment was conducted over a five-week period and yielded a total of eleven million seven hundred and sixty-three thousand, four hundred and eighty hits on various websites as shown in table 4.1. These were thereafter categorised into fifty-two different categories which varied from social media to the visiting of extreme websites such as pornography, weapons and hacking. The declassification of the websites presented above were undertaken utilising application software such as Shallas Secured Services and Zvelo.

Table 4. 1:Website Classification

URL Classification	Hits
Social Networks	2208750
Job Search	718900
Web TV	651326
Advertising	567840
Web Radio	519586
Update Sites	389474
Porn	285348
Spyware	277695
Sex	180235
Forum	169024
Automobile	163080
URL Shortner	158928
Tracker	157724
Government	157718
Alcohol	156536
Hacking	153504
Radio/TV	151008
ISP	150282
Homestyle	148720
Music	147828
Webmail	147318
Web Phones	144823
Warez	144807
Drugs	143360
Military	142740
Chat	140990
Downloads	140685
Science	139113
Finance	139000
Ringtones	138320
News	137718
Religion	136956
Dating	136398
AnonVPN	135300
Podcasts	134536
Models	133750
Gamble	133358
Recreation	133342
Aggressive	132990
Image Hosting	132225
Search Engines	129870
Politics	129840
Remote Control	127534
Violence	124440
Hospitals	124230
Shopping	123903
Education	123615
Hobby	122089
Library	121847
Movies	121555
Weapons	118215
Dynamic	115107
Total URL Hits	**11763480**

An area the Dark Web is not notorious for is the large amounts of educational resources that are freely available. The Dark Web contains a constantly updated torrent of raw, unchecked information, surging with complex technical terms, (Stobing, 2018). There are many documents that one can find ranging from NASA mission data to conventional academic paper databases. The Surface Web only amounts to 5% of the information available for open search, hence the acquisition of academic resources will ideally be found on the Dark Web (Stobling, 2015). Figure 4.8 shows that individuals in the academic environment utilise TOR to access academic content and utilise the forums as platforms for engaging in academic discussions. The science forums provide individuals with a “space” where they can discuss unconventional theories, post

revelations and engage in constructive arguments on conspiracy theories (Stobling, 2016). Figure 4.8 also provides us with an indication that the Dark Web in South Africa is being utilised for educational purposes as per Stobling (2016) findings.

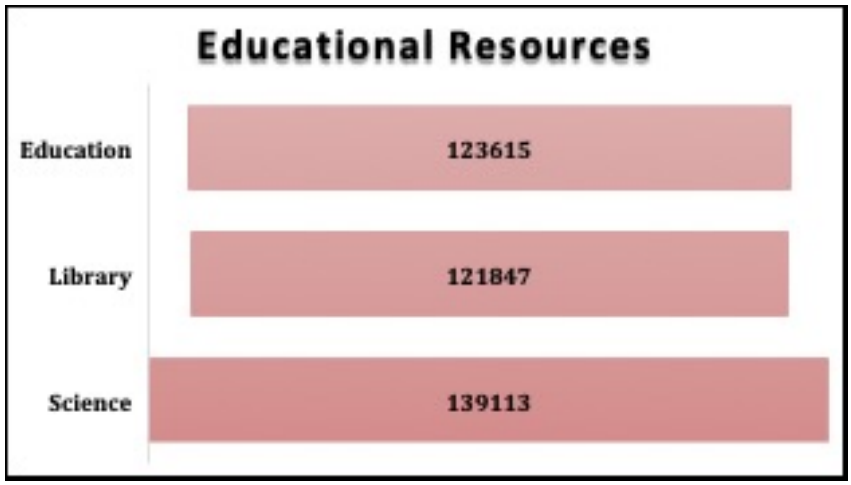


Figure 4. 8:Educational Resources

4.5 Purpose of Using Dark Web in South Africa

The primary purpose of utilising the Dark Web in South Africa is to access social media sites. The majority of the Dark Web traffic that was logged was during working hours, hence it can be concluded that the purpose of using the Dark Web in South Africa is to bypass a proxy server at work so an individual can gain access to social media websites. PlainSite (2019) however noted that over 50% of the accounts on Facebook are fake. Studies by Flores (2016) noted the use of social media websites such as Facebook by Western Africans was to engage in scam activities.

The majority of websites visited during the experiment was directed towards social media platforms as shown in figure 4.9. This is further reinforced by the server logs, which revealed a large number of application software downloads and the watching of Web TV. The bypassing of a proxy server utilising TOR will enable the automatic updates on that network. There is also a presence of large amounts of traffic directed towards Web TV and Job search which further reinforces that employed individuals are utilising TOR to watch movies during working hours and to apply for jobs whilst maintaining their anonymity. This is in contrast to previous research by Dingeldine (2015), McCoy and Bauer (2014), Chen (2014) Westlake (2017), Weimann (2016), Flores (2016) and Paganini (2017) where their studies found the majority of

web activity on the Dark Web to be associated with illicit activities. There were no previous studies that logged social media as the main Dark Web activity.

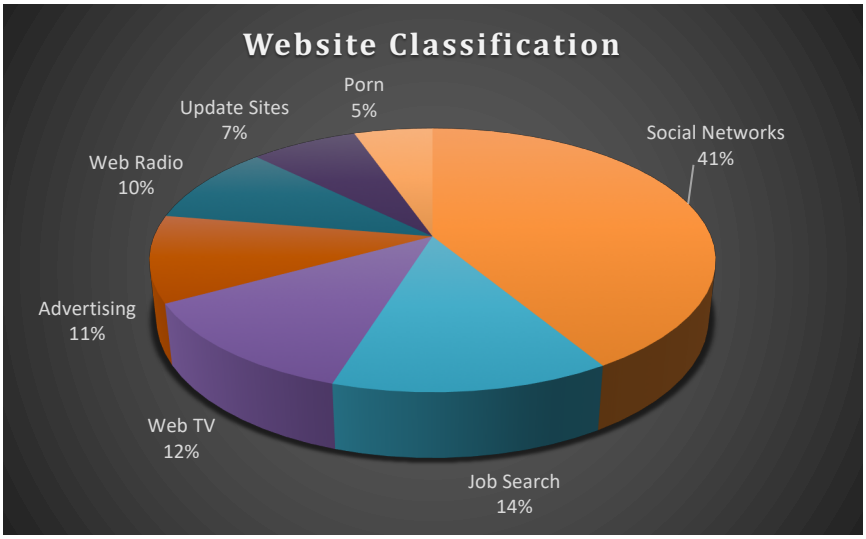


Figure 4. 9: Website Classification

4.6 Extent of Using Dark Web in South Africa

Previous studies did not consider exit routing traffic by country. This study aimed to fill that gap and provide data on the exit routing traffic by country utilising South Africa as the country of origin. Figure 4.10 represents the log file of the TOR exit routing traffic and presents the exit routing traffic by country and the geo-location of the IP addresses. This study also investigated where TOR clients and routers are located geo-politically since a client's IP address is visible to a router when that router is used as the entrance node on the client's circuit, through the TOR network. In the current TOR implementation, only particular routers, called entry guards, may be used for the first hop of a client's circuit. A router is labelled as an entry guard by the authoritative directory servers. All TOR router IP addresses are maintained by the directory servers, and the researcher kept track of the router IP addresses by simply polling the directory servers periodically.

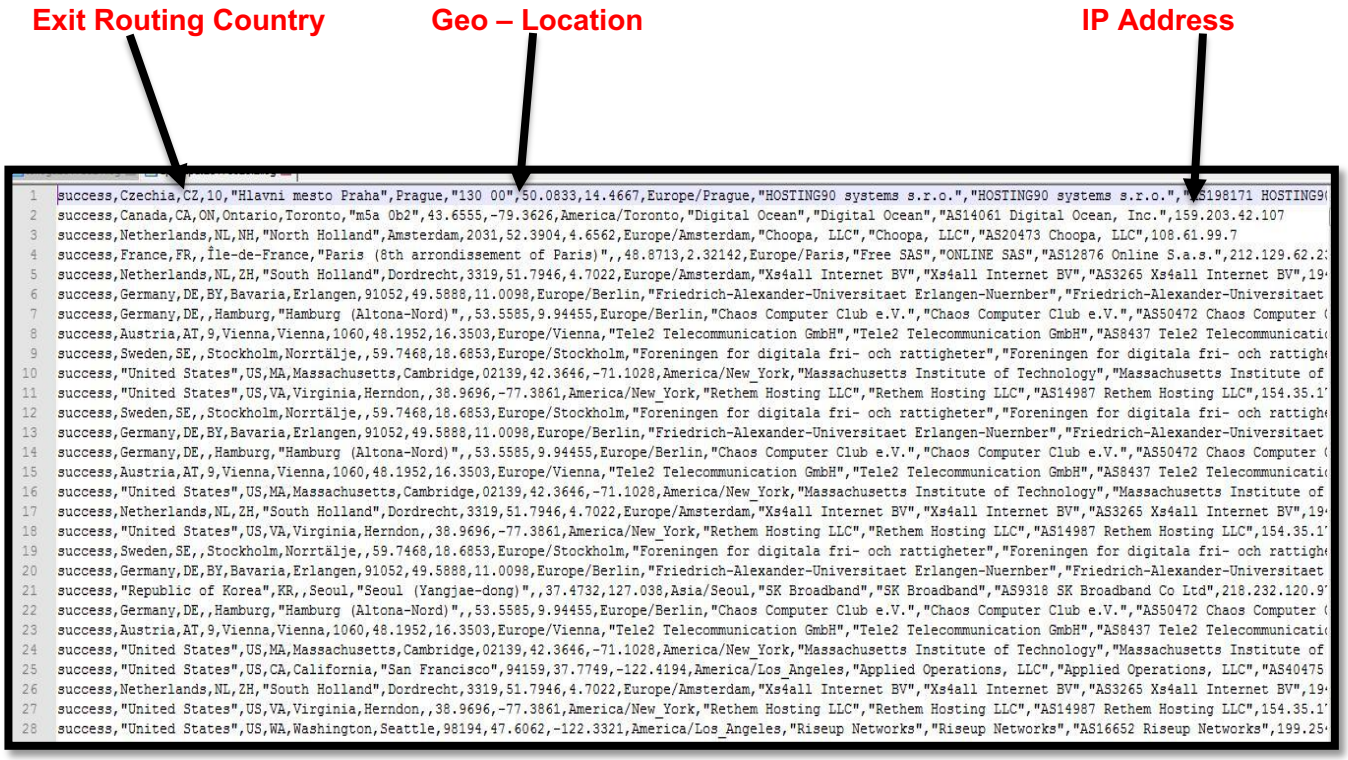


Figure 4. 10: TOR Exit Routing traffic

In order to put the raw geopolitical client distributions into perspective, Table 4.2 includes a ratio of the percentage of TOR users to the percentage of Internet users by country, by using data on the distribution of broadband Internet users by country. The percentages presented above were computed by dividing the total number of TOR clients located in each country by the total number of TOR clients observed in this study, which provides the percentage of TOR users located in each country. For example, the relative TOR usage for Germany is computed as follows: The percentage of the total Internet users who are from Germany is 3.9% and according to the client observations, Germany makes up 543 of the 2486 total TOR clients, which is 22%. Thus, the ratio of TOR users to Internet users in Germany is 5.64.

Table 4. 2: Exit Routing Traffic by Country

Client Distribution		**Router Distribution**		**Relative TOR Usage**	
Country	Total	Country	Total	Country	Ratio
Germany	543	Germany	374	Germany	5.64
United State	487	United States	326	United States	2.6
France	341	France	69	France	10
Netherlands	216	Netherlands	35	Netherlands	22.89
Moldova	116	Sweden	35	Sweden	25
Sweden	115	Moldova	24	Moldova	72
Canada	104	Russia	28	Canada	5.3
Russia	86	Canada	28	Russia	1.3

Table 4.3 shows the exit routing traffic by country. In total there were 45 countries from various continents and no connections made to any countries associated with extremist groups. There were also no connections made to any African countries even though there are a substantial amount of TOR relays located in Africa. More than half of TOR users are located in Europe, which is also the region with the highest penetration, as the service is used by an average of 80 per 100,000 European Internet users (Carvalho, 2017). However, there was quite a large portion of traffic directed towards countries where there are serious Internet censorship laws, and of those, China is ranked second. This finding is related to previous studies conducted by Clark (2017), Goldsmith (2016) and Dobson (2016) where the utilisation of TOR is prominent in Internet censorship countries, where they utilise TOR to remain anonymous and engage in freedom of speech forums and chats.

Table 4. 3: Exit Routing Traffic

Country	Count of success
Germany	22.01%
United States	19.60%
France	13.72%
Netherlands	8.69%
Republic of Moldova	4.67%
Sweden	4.63%
Canada	4.19%
Russia	3.46%
Austria	2.98%
United Kingdom	2.21%
Switzerland	1.61%
Republic of Korea	1.61%
Ukraine	1.01%
Romania	0.93%
Italy	0.93%
Brazil	0.89%
Denmark	0.68%
Norway	0.68%
Poland	0.64%
Bulgaria	0.56%
Finland	0.52%
Hungary	0.52%
Czechia	0.52%
Spain	0.36%
Republic of Lithuania	0.32%
China	0.28%
Ireland	0.28%
Belgium	0.20%
Czech Republic	0.20%
Luxembourg	0.16%
Portugal	0.12%
Japan	0.12%
Estonia	0.12%
Latvia	0.12%
Hong Kong	0.08%
Palau	0.04%
Turkey	0.04%
Panama	0.04%
Singapore	0.04%
Saint Kitts and Nevis	0.04%
Greece	0.04%
Croatia	0.04%
Australia	0.04%
India	0.04%
Grand Total	100.00%

The exit routing traffic by country can be graphically presented by utilising Gephi. The graph presented below is a Gephi FA2 graph with South Africa as the country of origin. Figure 4.11 is a graphical representation from the data that was entered into Gephi. The graph presents all exit routing traffic from the TOR node and reveals a total of 45 different countries that South Africans connected to from the TOR exit node. The connecting countries represent different continents, however there was no traffic directed towards any of the African countries. Carvalho (2017) noted that on average of 60 per 100,000 Internet users, the Middle East and North Africa have the second highest usage rate in relation to the number of TOR users as a percentage of the Internet population. This study however does not support that of Carvalho (2017), Dingeldine (2015) and Sancho (2015) as there was no exit routing traffic to any of the African countries nor to Italy. With the experiment utilising South Africa as the country of origin, there would be an expectancy to log traffic to one of the African countries as

TOR has a large presence in these countries. There was also no traffic directed towards any countries associated with extremist groups, however there was exit routing traffic to countries with serious Internet censorship laws. Some of these countries such as China prohibit the use of Google and any explicit web content (Sancho,2015).

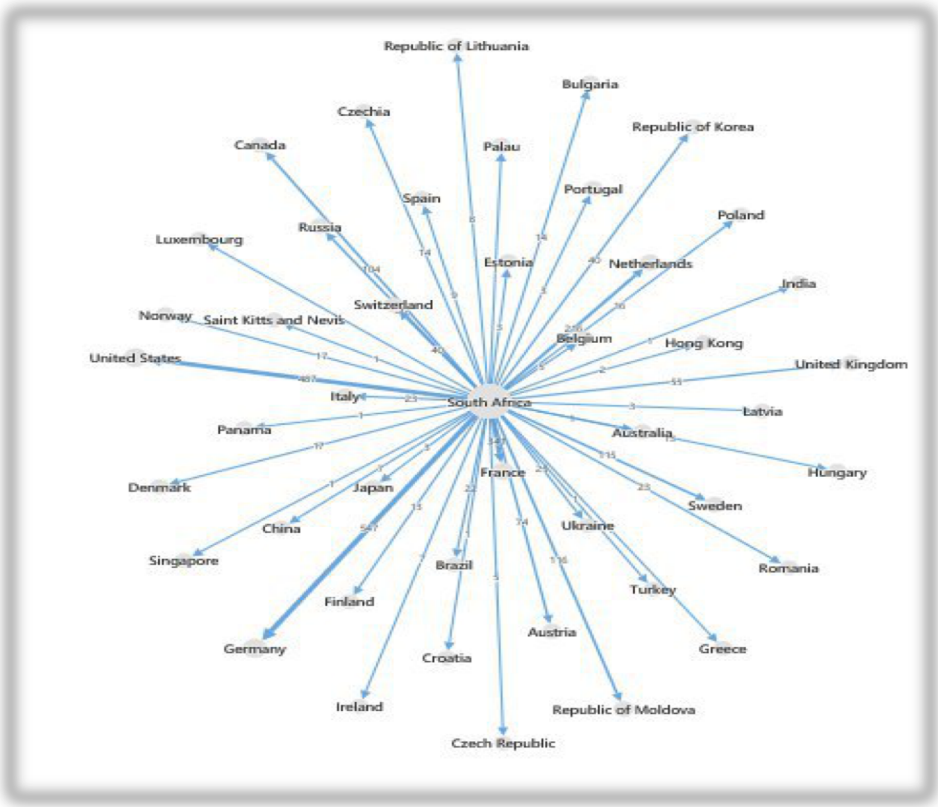


Figure 4. 11: Network Visualisation

There were 44 nodes connected by 45 edges of the complete network and Table 4.4 reports the network metrics. The measurements are calculated with directed edges, and self-loops being removed. The degree distribution represents the number of connections that a node has with other nodes. The node being a representation of a country and the connection representing a border. For a directional graph there are two kinds of degrees:

In Degree: The number of edges entering a vertex

Out Degree: The number of edges leaving a vertex.

Table 4.4 outlines the degree distribution as being 0.978, which means that on average, each country has in the region of one neighbour. The Average Weighted Degree is the average sum of weights of the edges of the node.

A graph density represents how close the network is to being complete. For a directed graph it is calculated by e/ (v*(v-1)), where:

e- Represents the edge and,

v- The number of vertices.

If a node is connected to every other node, then the density of the graph = 1. The graph density distribution for the data set presented above is at 0.022, which is therefore low. Therefore, it can be concluded that only one node is connected to all other nodes in other words South Africa is the country of origin and all TOR traffic was distributed to other countries from the country of origin.

Table 4. 4: Gephi Statistics

Degree Distribution	0.978
Average Weighted Distribution	55.24
Graph Density Distribution	0.022

Figure 4.12 is a graphical representation of the average weighted degree. The average weighted degree is 55.244, which weight represents the border that two countries share. From the distribution it can be deduced that on average each country shares on average a border length of 55.244.

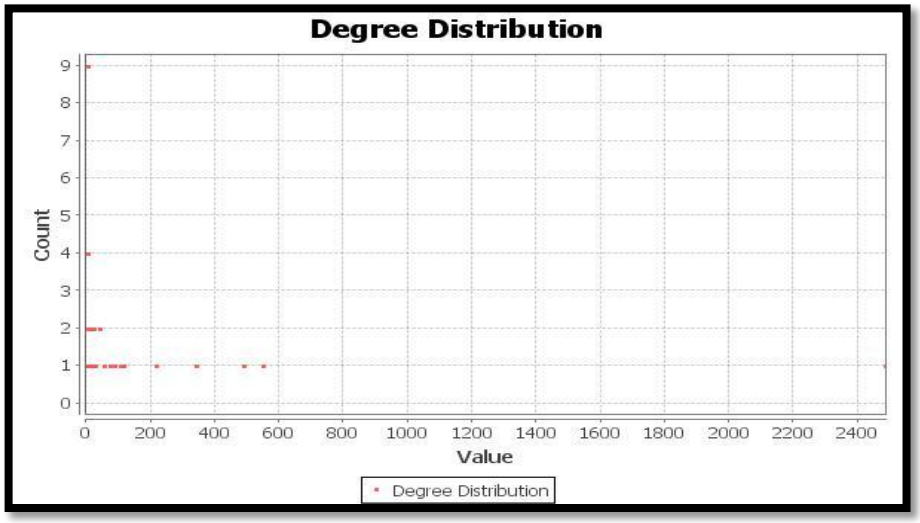


Figure 4. 12: Degree Distribution

The average weighted graph distribution obtained from Gephi is presented in figure 4.13 and aims to present the largest exit routing traffic to a country. The blue arrows gives an indication that the largest exit routing traffic was directed to Germany and closely followed by the U.S.A. From the average weighted degree graph above it can be concluded that the USA is closely followed by Germany as the greatest contributor to the largest amount of exit routing traffic. This finding is directly correlated with the statistics that are provided by TOR (2016) and Dingeldine (2015) on the greatest number of nodes that are found on the TOR network. Carvalho (2017), Katyal (2016) and Crampton (2016) noted over 50% of TOR users being located in Europe with Internet user statistics reaching on average 100 000 Internet users per day. Carvalho (2017) and Katyal (2016) further noted that Italy is second only to the United States, in terms of average number of users, as over 126,000 people access the Internet through TOR every day from the United States. This is in accordance with a study conducted by Dingeldine (2015), McCoy (2008) and Katyal (2016) found that majority of TOR traffic was directed to the U.S.A and European countries such as Germany.

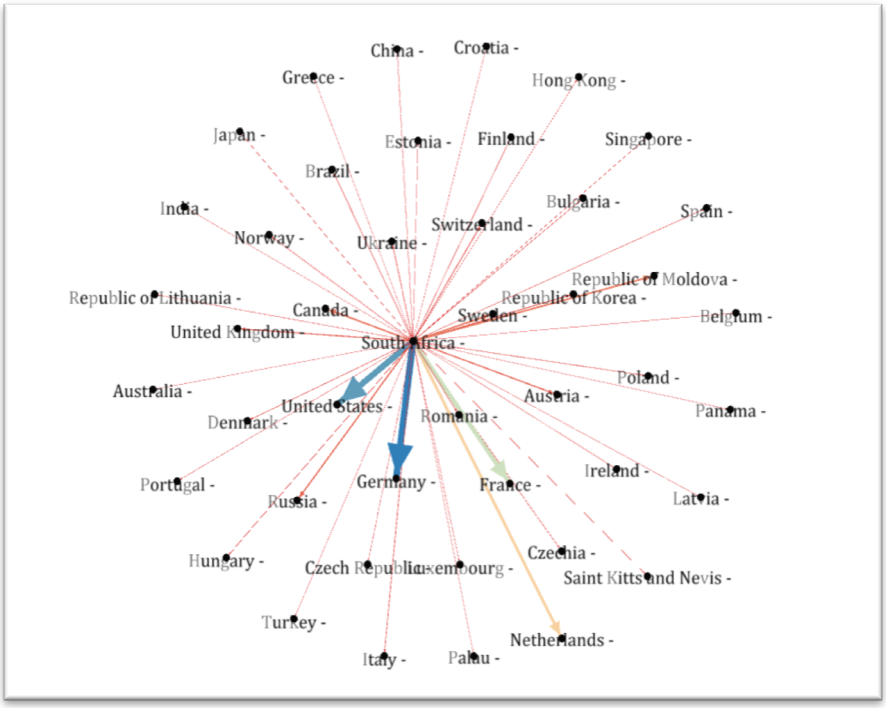


Figure 4. 13: Average Weighted Degree

4.7 Summary

The primary objective of the study was to understand TOR usage by South Africans in accessing the Dark Web. In particular, the study provided observations that helped understand how the Dark Web is being used, how the Dark Web is being misused and for what purpose is the Dark Web being used in South Africa. The study further analysed exit routing traffic by traffic and the attainment of the geo-location of the IP addresses. This chapter presented the data from the experiment conducted over a five-week period and the web content from the experiment was further analysed utilising a variety of open source and proprietary declassification tools. A summary of the findings is presented below:

1. The primary usage of TOR is to access social media websites on the Dark Web possibly at work, as the use of TOR will allow to bypass a proxy server whilst maintaining anonymity.
2. The predominant exit routing traffic was to Germany closely followed by the U.S.A. This finding was in line with previous studies as the largest number of TOR relays are to be found here.

3. There are connections to countries with serious Internet censorship laws such as China and the Republic of Korea which will be of concern to law enforcement agencies.
4. South Africans are utilising the Dark Web to access educational resources.
5. South Africans are engaging in illicit activities on the Dark Web as the web content analysed presented traffic to onion addresses associated to child pornography, sale of drugs, sale of firearms and the visitation to sites where spyware and ransomware can be obtained.

The next chapter will aim to further analyse the data and present new findings that will add to the findings of previous studies undertaken. The chapter will further analyse what social media websites were visited and the illicit trade that individuals were engaging in. The methodology of this study is broken into three parts as outlined in chapter one and three. The final step in the in the experiment process is the validation of the results that was presented at the SADC cybersecurity conference. The questions posed at that conference together with the responses will be presented in the next chapter

CHAPTER FIVE

PRESENTATION OF RESULTS

5.1 Introduction

TOR has become synonymous with Internet censorship countries as it allows individuals Internet freedom, however the utilisation of TOR in accessing the Dark Web has received a lot of attention from nation states (Katyal, 2016). The utilisation of the Dark Web as a means of communication and recruiting by ISIS has become a serious concern for all countries fighting the war on terrorism (Weimann, 2016). The previous chapter presented the experimental results of this study and this chapter will aim to comprehensively analyse the results to obtain an overview of the actual usage of the Dark Web individuals are engaging in. This chapter will further analyse the actual social media websites visited during the experiment, the countries with serious Internet censorship laws and what actual illicit trade individuals are engaging in. The chapter will also present new findings that contradict previous studies thus adding to the literature on this research topic.

5.2 Focus Group Response

The results of the experiment were presented to the South African Development Community (SADC) cyber security group which comprised delegates from 14 countries who were debated topics on the cyber security landscape in various groups. The South African contingent comprised members from the South African State Security Agency, Chief Director of Information Security Services, Council for Scientific and Industrial Research and the General of the South African Army. For the purpose of this study all respondents will remain anonymous and the responses from the focus group is presented in table 5.1

Table 5. 1: Focus Group Response

No	Question	Response
1.	When was the experiment conducted? And for how long?	End of September for a duration of five weeks. This was in accordance with previous research conducted.
2.	Do you have the IP addresses of the clients? **Response**: The government websites came under attack during this period and we assume the attack originated from the Dark Web.	Yes. Why?
3.	Were clients accessing TOR through a VPN, logged?	Yes. The results were logged. A VPN by nature in conjunction with TOR as a browser should allow for the anonymous browsing on the Dark Web. However, due to high level methodology utilised in the experiment, IP addresses of the clients were obtained. We also display stats of the number of VPN's accessing TOR during the conducting the experiment.
4.	There is quite a bit of malicious traffic that were visited during the experiment. Can these websites be further declassified?	Yes. Although majority of the traffic is associated to that of social media, there is a still a concern on the large amount of malicious traffic that was visited during the two-week period, these vary from Pornography to drug purchases. This is just an initial analysis of the results. A further analysis will be taken. The aim is to declassify the websites into actual URL's visited.
5.	Approximately how much of traffic was logged during the experiment?	Over 11 million websites were visited during the experiment. There is not much TOR usage from a South African context and on average, there was around 7000 users of TOR daily. This is small, in comparison countries such as Germany and USA. The small number of users are directly in line with the amount of traffic logged.

6.	Will you make the results available to the State Security Agency?	Yes. The results will be made available upon completion of the study.
7.	How reliable are the results?	The results and methodology of the experiment was validated by an independent security consultant that undertook this before on behalf of the government. The methodology of setting up a TOR exit node can also be found on the TOR website. The methodology was also aligned to previous research conducted.
8.	Would you make the methodology of the experiment available?	Yes. This will be made available upon request. The TOR exit node will be placed at a university and allow other universities to connect to it for further research purposes.
9.	The exit routing traffic by country is very interesting. Was there any traffic to Muslim states?	There was no exit routing traffic to any of the Islamic states, however there was traffic that was routed to countries where there are serious Internet censorship laws. Some of these countries such as China, even ban the use of Google. It also interesting to note the routing traffic to North Korea.

From the engaging dialogues on the results presented, it became very apparent that the Dark Web is off a serious concern to SADC. African states have come under serious attacks over the last few years with many stemming from the Dark Web. The question posed on when the experiment was undertaken and if the addresses from the experiment had been acquired shows the seriousness on the utilisation of TOR to conduct illegal activities. It was during this period that the South African government websites came under serious attacks with the no trace of the individuals that executed this attack. There is a really good possibility that this attack was generated from the Dark Web.

5.3 Dark Web usage in South Africa

South African TOR usage in comparison to Europe and America is relatively small. Figure 5.1 provides an indication of the number of daily connected TOR users from South Africa which ranges between 4000 – 6000 users daily. From the data presented in this study the total number of websites that were visited during the experiment was eleven million seven hundred and sixty-three thousand, four hundred and eighty hits on various categories of websites. The study can therefore conclude that the average TOR user from South African visits seventy-eight websites on the Dark Web daily.

Average number of websites visited daily on Dark Web by South Africans:

Average daily users TOR = (Sum of TOR users per day ÷ 30) = 5000
Total TOR users for the month = 5000 x 30 = 150 000
Total Number of hits on websites = 11 763 480
Average number of websites visited by South Africans = 11 763 480 ÷ 150 000 = 78

There has yet to be a study that presents data on the average number of Dark Web websites that TOR users visit per country. There is no data from a South African context on the usage of TOR and the average number of Dark Web websites that South Africans access on a daily basis. Studies by Flores (2016), Dingeldine (2015), Sancho (2015) and Weimann (2016) provided sample data on websites accessed on the Dark Web by utilising TOR. A flaw in their study was that they did not present the total number of websites or the average number of websites that were accessed during the period of the experiment.

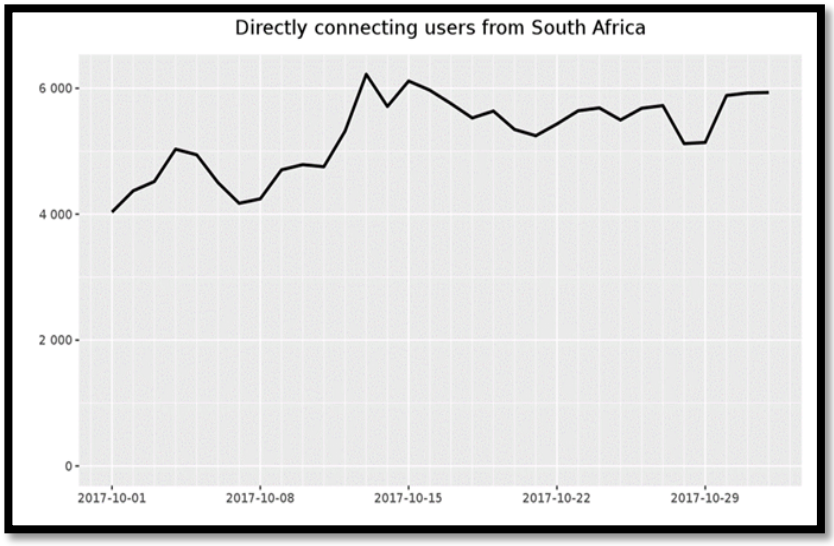


Figure 5. 1: South African TOR Usage

5.4 Dark Web Misuse in South Africa?

With the anonymity that TOR provides, individuals tend to have more confidence in engaging in illegal activities on the Dark Web. From a government monitoring perspective many nation states do not conduct any form of traffic analysis on the Dark Web, thereby making it easier for individuals to engage in these illicit activities. The Dark Web provides anonymity to individuals, which encourages them to engage in illegal activities as noted by Chertoff (2015), Dingeldine (2015) and Goldsmith (2016). The transactions and the purchase of prominent goods and services traded would provide insight into the kind of activities individual would engage in if their identity was hidden.

Majority of the activities on the Dark Web have an impact on the real world, this is unlike the cybercriminal underground (Chertoff, 2015). Many of the malicious tools and services sold in the cybercriminal underground can be used to gain profit (Carvalho, 2017). Those peddled in the Dark Web assassination services, for example, obviously serve a different, more sinister purpose (Greenberg, 2015). Pornography is freely available on the surface as it is on the Dark Web, however much of the pornography found on the Dark Web is that of child pornography as noted by Norton (2016) and Trendlabs (2017).

Table 5.2 presents a breakdown of pornographic activities from a South African perspective. Dingeldine (2015), Norton (2016) and Trendlabs (2017) stated that 80% of pornographic content on the Dark Web is child pornography. This is further entrenched by Vitare (2014) where he stated that 60% of child pornographic content is found on European servers whilst another 37% is found on servers found in the USA. It is clearly evident that the majority of exit routing traffic from South Africa is from Europe and the USA, which is also in accordance to Norton (2016) who stated that one can probably find questionable pornography on the Surface Web, or perhaps the Dark Web such as incest, bestiality or (fake) rape videos. A study conducted by the University of Portsmouth in 2017 revealed that over 80% of TOR network visits is related to pedo sites. A report by Trend Micro Research in 2017 also identified 8,707 “suspicious” pages. A further analysis of the “Surface Web” sites that those sites linked to reveal that they can be classified into various categories ranging from child exploitation (26%) to Disease vector sites (33.7%),

From the statistics presented in the table, it can be concluded that the visitation of child pornographic websites is a serious issue, however one that is not as bad as that presented by Vitare (2014), Dingeldine (2015) and Norton (2016). These studies results are therefore not in accordance with that of Vitare (2014), Dingeldine (2015) and Norton (2015) as child pornography does not make up the majority of pornographic material that was accessed. A flaw in Dingeldine (2015), Vitare (2014) and Norton (2016) was the monitoring of traffic to a limited number of countries only. These countries according to Norton (2016) is where the majority of pornographic material can be found on the Dark Web.

Table 5. 2: Pornography

Website Classification	Percentage
Child Pornography	33%
Other	67%

One of the primary aims of this study was to ascertain the illegal activities that individuals in South Africa were engaging in. Figure 5.2 provides an in-depth analysis of the illicit websites that were declassified during the experiment. These are not classified to a specific theme only as presented in studies by Chen (2012) and

Westlake (2017). Data from this study concluded that there is no exit routing traffic to countries that associate themselves with terrorism.

The data from this study also provides an overview of some of the malicious websites that were visited for religious and political reasons. Flores (2016) who noted that the usage of TOR by terrorist groups for anonymous communication is also supported by Weimann (2016), Ghaffar (2016) who stated that terrorism and terrorist organisations on the Dark Web are a leading concern for law enforcement because they are a threat to national security. Extremists have utilised the benefits of the Dark Web, with organisations such as ISIS using it to radicalise and recruit young and vulnerable people from across the globe. Weimann (2016) and Moore (2016) also noted that terrorist organisations, such as ISIS, use the Dark Web to communicate with each other and organise their activities. It also can be used to provide information to terrorists or criminal groups, including “maps, photographs, directions, codes and technical details of how to use explosives, poisons, weapons, chemicals, and so on” (Weimann, 2016). The use of the Dark Web by South Africans to visit these religious websites could be the result of anonymous communication to religious extremist groups. There is a possibility that South Africans are engaging in religious and political chats and forums on extremist websites such as ISIS.

The illicit traffic represented in figure 5.2 further provides evidence that this study will contribute to the findings by Flores (2016), Warf (2016) and Weimann (2016) as one of the limitations of the researcher’s studies was the monitoring of the drug trade on the Dark Web only in countries such as the USA and Europe. The findings listed below will further contribute to the studies by Flores (2016), Warf (2016) and Weimann (2016):

1. The visitation of drug websites on the Dark Web is not associated with South Africa. The Dark Web is famous for drug websites such as Silk Road and Grams as noted by Vinto (2015), Christin (2015) and Van Buskirk (2015). Silk Road predominant trade routes were USA, China, India and other parts of Europe as noted by Flores (2016), Dolliver (2015) and Christen (2015). Vinto (2015) and Dolliver (2015) noted the emergence of Grams as an identical TOR hidden service such as Silk Road, operating the same routes. There has yet to be research that shows illicit drug trade usage occurring in South Africa. The traffic directed to these websites clearly indicate the presence of drug sales occurring

online in South Africa, however the exact nature of these transactions still remains unknown.

2. There were large amounts of traffic to government websites during the experiment. The experiment was undertaken during the same period that three South African government websites came under attack by "H4ksniper". This concern was raised when the results were presented at the SADC cyber security conference where the general of the South African army noted the hacking of these websites during the same period as the experiment was conducted with the hackers still remaining at large as there was a belief that the attack was conducted on the Dark Web. Studies by Sancho (2015). McCoy (2008) and Dingeldine (2015) presented no web traffic directed towards any governmental websites during the conducting of their experiment. This presents an interesting finding which should lead to future research. Further monitoring on the usage of TOR by South Africans need to be undertaken particularly looking at Dark Web traffic to governmental websites.

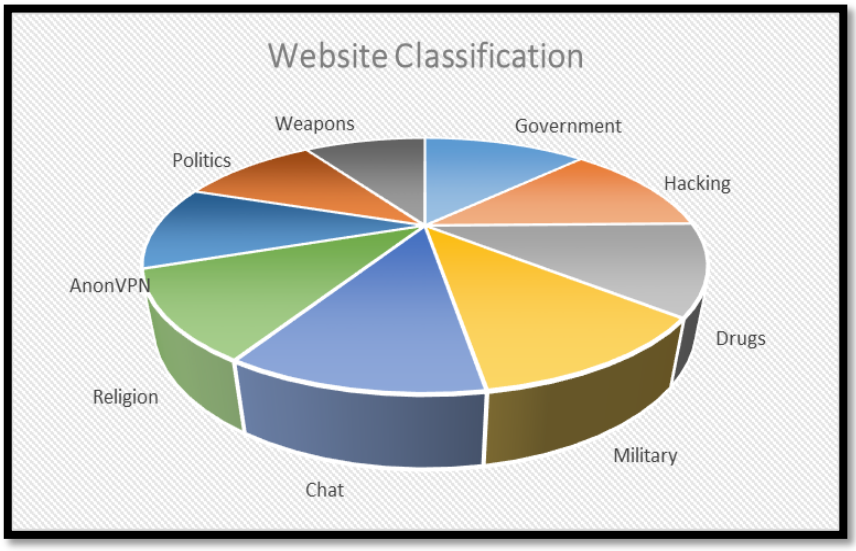


Figure 5. 2: Illicit Traffic

5.5 Uses of the Dark Web in South Africa

The Dark Web is closely associated with the engaging of illegal activities due to its anonymous nature. Previous research by King (2016), Goldsmith (2016), Norton (2016) and Yaneza (2014) reaffirm this. The engagement of illicit usage in South Africa

was never analysed before and the illicit usage that occurred during the experiment is shown in figure 5.3.

The largest proportion of websites visited during the experiment was classified as social media sites. According to Statistics South Africa (2017) 70% of South Africans weekly activities are spent online, visiting social media sites. As a large proportion of traffic passing through the exit node was during working hours one could ascertain that employers were utilising TOR at work to gain access to social media sites and Job Search sites on the Dark Web. The utilisation of TOR to bypass the proxy server will result in application software being updated as shown in figure 5.3 where a portion of the traffic is directed to software updates. This however places a huge risk not only to the companies' network, but also for individuals' accessing social networks on the Dark Web. Studies by Flores (2016), Django (2018), King (2016), Norton (2016) and Dingeldine (2015), whilst monitoring traffic on the Dark Web found the predominant usage of the Dark Web was to access illicit websites and engage in unlawful trade. The findings presented in figure 5.3 do not support the findings from these studies. A potential flaw in their studies was that they only monitored exit routing traffic to countries such as Germany, USA and parts of Europe, where Vitare (2014), Crampton (2016) and Dobson (2016) noted that the majority of pornographic websites and drug trade could be found in countries such as the USA, Germany and Europe.

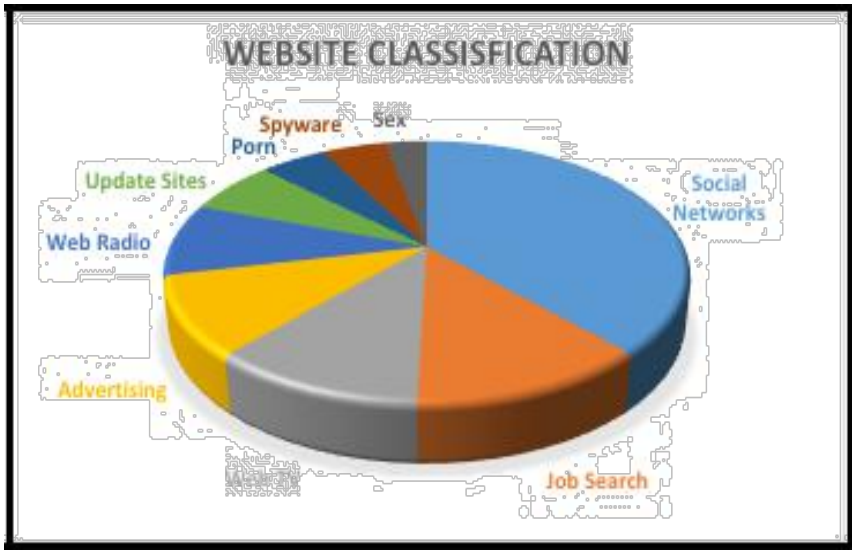


Figure 5. 3: Website Classification

5.6 Purpose of using the Dark Web in South Africa

Social media pervades our everyday lives, whilst some use it as a means of communication others utilise it to conduct scam activities. Facebook currently is the largest social media platform in the world, and it is no surprise that it made up the most amount of social media web traffic as shown in figure 5.4. There is however no way to ascertain how it was being used or misused by South Africans in this study. A latest report released by PlainSite (2019) claimed that 50% of the accounts on Facebook are fake and individuals utilise these in scamming others. Flores (2016) and Paganini (2017) findings in terms of social media activities on the Dark Web were related to scam emails and chat groups, with these scammers utilising applications such as Facebook and Facebook Messenger to undertake these illicit activities.

Studies by Django (2018), Hacquebord (2015), Paganini (2017) and McCoy (2008) on observations of Dark Web traffic, found the dominant amount of traffic passing through the exit node was that off malicious activities. The primary use of TOR in studies by Django (2018), Dingeldine (2015) and King (2016), was the usage of the Dark Web to purchase drugs and to visit pornographic websites. A further study by Flores (2016) and Paganini (2017) only noted the use of emails by West Africans when targeting individuals and businesses. The findings from this study do not correlate with that off Django (2018), Hacquebord (2015), Paganini (2017) and McCoy (2008) as they noted the major use of the Dark Web was to access pornographic websites, illicit trade and the use of TOR to distribute scam emails on the Dark Web. The studies did not log Facebook or any other social media platform as the primary activity on the Dark Web.

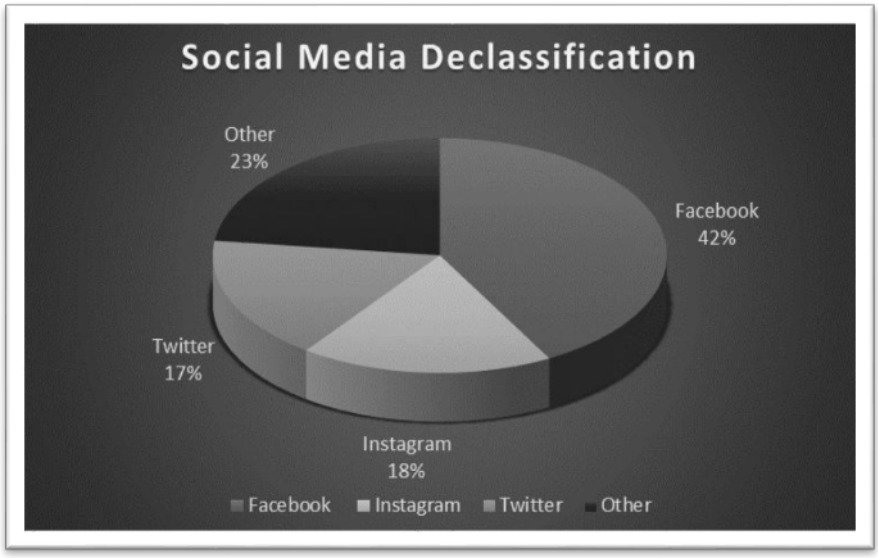


Figure 5. 4: Social Media

Online dating has become increasingly popular with the likes of applications such as Tinder. As with other social media applications it poses a serious risk to individuals as many have gotten scammed with fake profiles. Figure 5.5 presents a classification of the chat activity associated with the social media activities. There was a total of one hundred and thirty-six thousand three hundred and nighty eight visits to dating websites. Online dating has become extremely popular with a large number of dating applications and websites been developed. Findings from Flores (2016) and Weimann (2016) noted the use of the Dark Web by West Africans when targeting individuals and businesses with scam activities. Flores (2016) further noted the use of the Dark Web by West Africans in developing scam emails and posing fraudulent profiles on social media websites. This study however did not log any traffic to any off the West African countries. This study therefore proves the use of TOR by South Africans on these Dark Web dating websites to be associated with scam activities and remaining anonymous in doing so. This finding will therefore add to the findings to that of Flores (2016) and Weimann (2016) where the issue of the utilisation of the Dark Web in scamming individuals and businesses is not only confined to West African countries but also associated to South Africa as well. This will also tie in with theoretical framework of the study that is based on the Space Transition Theory as proposed by Jaishanksr (2007), where he explains the behaviours of the persons who bring out their conforming and non-conforming behaviours in the physical space and virtual space.

Jaishanksr (2007) also noted that in the Space Transition Theory people behave differently when they move from one space to another. Individuals that repress crime in a physical space, would commit such crimes in the virtual space. Jaishanksr (2007) also noted concepts such as Cyber stalking and Cyber as defamation and where offenders make use of online space because of its anonymity and widespread approach.

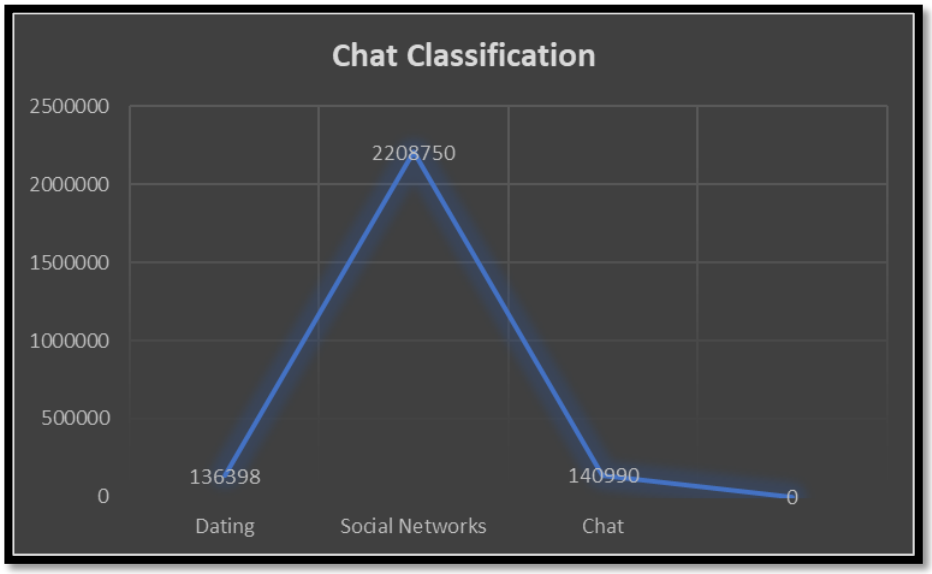


Figure 5. 5: Dating Websites

Chat groups and forums on the Dark Web have drastically increased in number and many of the extremist groups like ISIS have moved their websites onto the Dark Web so as to provide their members with some sort of anonymity (Weimann, 2016). These do not only serve as a means of communication but also as a means of recruitment. ISIS utilises the Dark Web as a means of communication to carry out coordinated attacks. Figure 5.6 shows the further declassification of the 23% other found in figure 5.4. The study logged no traffic to any Islamic states or countries associated with extremist groups. The further declassification of the other 23% of the social media activity represents chat groups of a religious and political nature. This does not correlate to the findings presented by Carvalho (2017), Weimann (2016), Ghaffar (2016) and Moore (2016) where it was noted religious chats and forums are found in Islamic and extremist countries. Weimann (2016) and King (2016) also

associated religious chat groups with extremist groups in different countries. The results presented in figure 5.6 could be interpreted in one of the following ways:

1. There are religious extremist groups in SA utilising the Dark Web to communicate with each other within the country.
2. The use of the Dark Web by South Africans is being utilised to engage in political and religious chat with extremist groups in other countries. This further means the presence of these groups in multiple countries other than the most commonly known ones.

Future studies should aim to ascertain the exact nature of these chat groups. A further declassification and penetration into these forums are essential as they potentially pose a major threat to S.A.

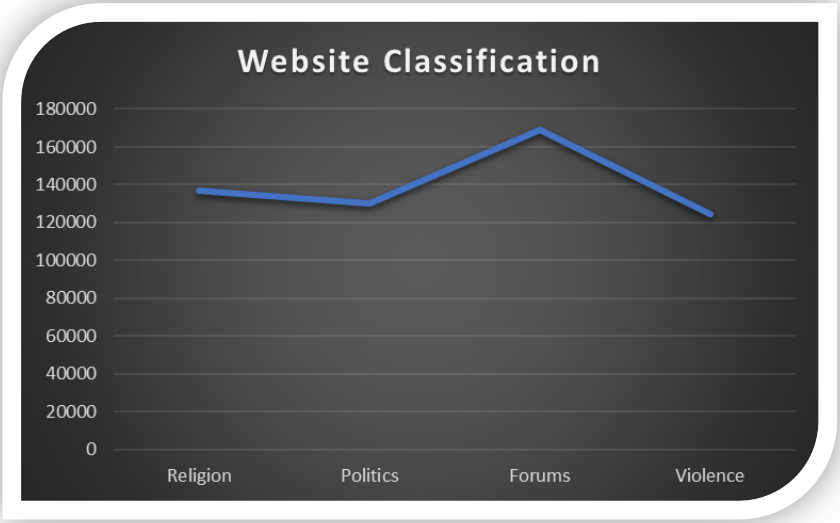


Figure 5. 6: Social Media Classification

5.7 Extent of Using the Dark Web in South Africa

Internet censorship countries utilise TOR to access the Dark Web and remain anonymous. There has been increased pressure by nation states to enforce governance of the Internet and curb the usage of TOR in accessing the Dark Web (Dobson, 2016). Countries such as Korea have serious Internet censorship laws that even ban the usage of Google. Figure 5.7 provides an illustration of the Internet censorship countries South Africans connected to. According to the TOR logs, seven

of the forty-five countries have serious Internet censorship laws against citizens with varying degrees of censorship. Carvalho (2017), Clark(2017) and Katyal (2016) notes that when looking at the number of TOR users as a percentage of the larger Internet population, the Middle East and North Africa has the second highest rate of usage, with an average of over 60 per 100,000 Internet users utilising the service. Carvalho (2017) further states that TOR is particularly popular in Israel, which accounts for more TOR users than India, while having less than 4% of its Internet users. The findings from the experiment are in contradiction to Carvalho (2017), Clark (2017) and Katyal (2016) who noted the number of TOR users as a percentage of the Internet population, the results presented gave an indication of the popularity of TOR by the Middle East and North Africa as they had the second highest rate of usage with these countries averaging over 60 per 100,000 Internet users utilising the service. Carvalho (2017) further noted TOR as being popular is Israel and the usage in this country accounts for more than that in India. One would expect the data from this study would correlate to that of Carvalho (2017) Clark(2017) and Katyal (2016) where there would be exit routing traffic to North African countries. There is however a serious concern as according to Clark (2017), the Russian government blocks websites critical of the government and websites associated with militant or extremist organisations. There was substantial exit routing traffic to Russia and the Republic of Korea.

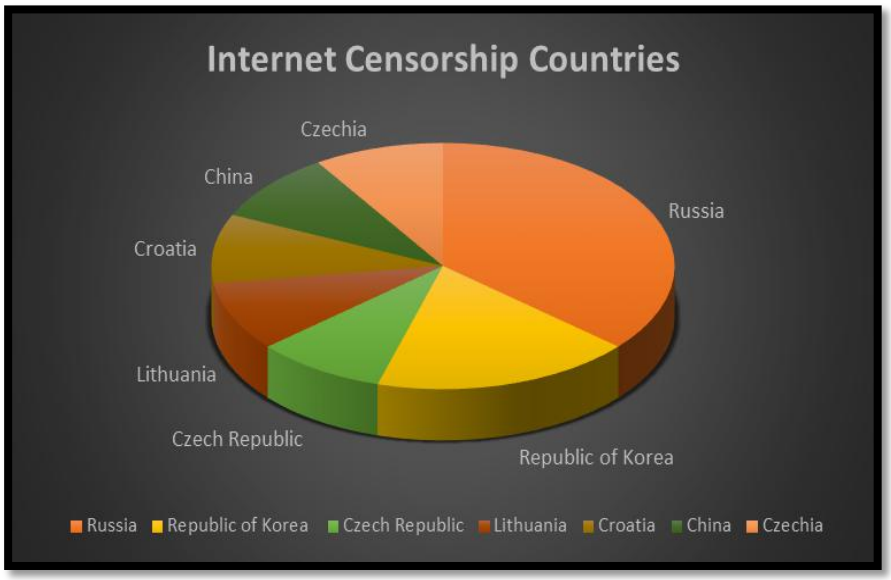


Figure 5. 7: Censored Countries

5.8 Summary

A further analysis on the presentation of results as presented in chapter four was illustrated in this chapter. A further analysis into the social media websites that were accessed was presented and it was found that Facebook was the primary contributor. It still leads to an unanswered question as to how Facebook was being utilised. As majority of the visits to the social media websites were during working hours, one would assume it is being utilised to bypass a proxy server at work as many organisations block social media websites. There is reason however to believe that the TOR version of Facebook is also being utilised to create fake accounts and to scam people, whilst remaining totally anonymous in South Africa. A serious concern is the visitation of child pornographic websites that are being accessed by South Africans and the illicit trading of drugs occurring on the Dark Web in South Africa. Previous studies have not shown any drug sales on the African continent with the drug trade on the Dark Web primarily occurring in the U.S.A and European countries. The analysis presented in this chapter provides justification that further research needs to be conducted aligned to this study. Further studies need to be dwelled into the actual social media activities that individuals are engaging in and the kind of drug trade that is occurring. The next chapter will provide a summary on the findings of the study, discuss the limitations of this study and the need for future work that needs to be undertaken in this research area.

CHAPTER SIX

SUMMARY AND FURTHER CONTRIBUTIONS

6.1 Introduction

The background to this study was provided by discussing events in the recent months pertaining to the increase in cyber-attacks and the problems facing nation states. As a result of censorship and monitoring of Internet usage, people have resorted to censorship resistance and anonymity preserving systems, like TOR. With the increase in the usage of anonymous networks, cybercrime has increased dramatically over the past few years and identifying the source of these attacks pose a huge problem to nation states. Identifying the actual source of anonymous traffic and anonymous communication systems has become a priority for many national security state agencies. This research aimed to identify TOR usage utilising South Africa as the country of origin in accessing the Dark Web. The study aimed to takes steps to fill the gap that solely relied on network measurements, to aid the traffic analysis that helped to verify the identification of anonymous connections. The strategy can be classified in the same category as those presented by Murdoch (2015) and Mittal (2011). Unlike the previous researchers, this study did not confine the declassification of websites into specific content types, nor not logging exit routing traffic by country.

This study presented the architecture of a system which was designed and deployed to detect eavesdropping on TOR. The experiment was conducted by running the exit node between the end of September and the first week of November 2017, to detect websites being accessed by South Africans on the Dark Web utilising a technique of eavesdropping. In Chapter four, the results of the experiment were presented, and it was reported that the majority of the network traffic was associated to that of the social media and there was also a large portion of traffic assigned to illicit activities such as drug sales and pornography. In total there were forty-five countries that the TOR exit node connected to and seven of these are associated with serious Internet censorship laws, some of which ban the usage of google and confine Internet usage to only certain content types. In Chapter five, the study provided an analysis of the results and it was reported that 42% of the social network traffic was assigned to

Facebook. It could thus be concluded that the primary purpose of TOR in a South African context, is to bypass a proxy server at work in order to gain access to social media websites such as Facebook on the Dark Web. The study provided a further declassification of the pornographic websites that were visited by South Africans and found that there was an alarming amount of traffic was directed towards child pornography. This finding was in in line with that reported by Dingeldine (2015) who stated that 80% of pornographic content on the Dark Web is child pornography. This is further entrenched by Vitare (2014) where he stated that 60% of child pornographic content is found on European servers whilst another 37% is found on servers found in the USA.

6.2 Summary

The experiment architecture designed and implemented in this study revealed South Africa's Dark Web usage. There was no previous study that was undertaken whereby Dark Web traffic was monitored in South Africa. The results obtained from the experiment will provide law enforcement agencies with a conceptual overview on the use and misuse of the Dark Web by South Africans. Previous studies undertaken by Dingeldine (2014), Chen (2014) and Westlake (2017) revealed the primary use of the Dark Web by individuals was to engage in illicit activities, however in a South African context the primary usage of the Dark Web was to engage in social media activities. The primary use of TOR in South Africa is therefore to bypass a proxy server at work in order to engage in social media activities. An overview on the findings of this study is presented below:

1. There was illicit trade that was logged during this study and in particular an area of concern is the online drug markets that individuals were engaging in. Dark Web drug stores were predominantly located in the USA and European countries as noted by Vinto (2015). There was never any drug trade occurring on the Dark Web from a South African perspective. Pornographic material is easily available all over the Internet, however child pornography is largely contained on the Dark Web. Dingeldine (2015) and Vitare (2014) revealed that 80% of the pornographic material found on the Dark Web is aligned to that of child pornography. There was

traffic that was directed to a large amount of child pornographic websites. This should be of concern to law enforcement agencies.

2. The exit routing traffic to countries revealed a substantial amount of traffic to countries such as the Republic of Korea, Russia and China. These countries have serious Internet censorship laws and as stated by Carvalho (2017) many even block search engines such as Google.
3. A critical finding of this study was the large amounts of traffic directed towards governmental websites during the experiment. The experiment was undertaken during the same period that three South African government websites came under attack by "H4ksniper", a hacker from Morocco. No previous studies logged any web traffic to any government websites. Attacks to governmental and organisational websites generally originate from the Dark Web (Vitare, 2014).
4. There was no exit routing traffic to any countries associated with extremist groups, however there was traffic directed to political and religious chat groups. From the social media category, 23% of the portion of the traffic was directed to political and religious social media platforms. Extremist groups such as ISIS have moved their websites onto the Dark Web and social media activities are found in extremist countries and other countries (Carvalho, 2017). This provides an indication that South Africans are engaging on these extremist social media platforms.
5. There was a substantial portion of traffic towards dating websites. Online dating has become extremely popular with a large number of dating applications and websites been developed and with the increase in these applications, there has been a rapid increase in the number of scamming activities (Flores, 2016). Flores (2016) further noted the use of TOR by West Africans in developing scam emails and posing fraudulent profiles on social media websites. Dating websites are not associated with illicit traffic and can be easily accessed on the Internet. The utilisation of TOR to maintain anonymity gives an indication that South Africans are utilising these websites in order to engage in scam activities.
6. TOR provides statistics on the number of daily TOR users by country. There is no information provided on the amount of Dark Web traffic. The data accumulated during the study, allowed for the calculation on the average number of Dark Websites that South Africans visited. There is no study that presents data on the average number of Dark Web websites that TOR users visit per country. There is no data from a South African context on the usage of TOR and the average number

of Dark Web websites that South Africans access on a daily basis. Studies by Flores (2016), Dingeldine (2015), Sancho (2015) and Weimann (2016) provided sample data on websites accessed on the Dark Web by utilising TOR.

The summary of the findings presented above will add to previous studies undertaken pertaining to this topic. These findings revealed the current usage of the Dark Web in South Africa and will provide law enforcement agencies in South Africa with a comprehensive overview on the illicit trade that South Africans are engaging in on the Dark Web.

6.3 Future Studies

The aim of the study was to ascertain the usage of the Dark Web in South Africa and provide an analysis of any illicit activities' individuals may be engaging in on the Dark Web. The study satisfied answering all research questions, however there were limitations to the study:

1. **The study was conducted over period of five weeks with one week being utilised as a test period.**

The experiment should have been conducted over a longer period of time. Although the primary use of TOR is more aligned to engaging in social media, there were instances where malicious traffic was recorded. The utilisation of TOR in the visitation of dating websites implies South Africans are engaging in scam activities on these social media platforms. Penetration of these website and a further declassification of the websites need to be undertaken. There was never any evidence indicating that South Africans are engaging in the sale of drugs online. Malicious traffic of this nature needs to be further analysed and the sources of these identified.

2. **The experiment was undertaken over five consecutive weeks.**

The experiment should have been split into increments of one or two weeks spanning over greater months. For the deanonymisation of anonymous communication networks the experimental design of this study was aligned to that of McCoy and Bauer (2014) whose study was conducted by running the exit node for a period of four weeks. For the TOR crawling the experimental design was aligned to that of Chen (2014) and Westlake (2017). Additional experiments need to be undertaken in order to

continuously observe website traffic on the Dark Web, in order to obtain a better understanding on the usage of TOR.

3. **The study provided a general overview of Dark Web usage and the illicit activities that individuals engage in on the Dark Web.**

There was no in-depth analysis into the kind of chat topics individuals were engaging in on the religious chat groups and forums. An analysis on the actual usage of social media platforms such as Facebook by individuals needs to be determined. For future studies, the experiment needs to be conducted over a longer period of time and the website classification presented in this study needs to be utilised as a benchmark to further track any changes in TOR usage. The TOR traffic directed towards government websites during the cyber-attacks poses a serious issue and challenge to the South African State Security Agency. There needs to be constant monitoring of traffic directed towards government websites through TOR. A future study should concentrate purely on the amount of traffic directed towards government websites via TOR.

These limitations will provide a foundation for future research to be conducted. There is an urgent need for additional research to be undertaken in this field of study. There are certain illicit activities which are of serious concern and pose a possible threat to nations.

6.4 Conclusion

This study was aimed at understanding TOR usage utilising South Africa as the country of origin. In particular, the study provided observations that helped us understand how TOR is being used, how TOR is being misused, and which countries form the largest portion of exit routing traffic. Through the observations, the study has made several observations in showing the majority of exit routing traffic is directly aligned to the statistics that are provided by TOR. The study also provided a detailed analysis of website classification of TOR usage by South Africans. Although the primary usage of TOR in a South African context is to bypass a proxy server at work in order to access social media websites, there are instances where South Africans are utilising TOR to engage in malicious activities. This illicit traffic needs to be further analysed and should be of a serious concern to the security agencies in South Africa. Future experiments

conducted on TOR usage should be conducted over longer period's time and the results obtained analysed against this studies results. A further recommendation would be to compare the results obtained in this study to future studies and observe if there are any changes in TOR usage over the period of time.

BIBLIOGRAPHY

Ablon, L. 2014. Markets for Cybercrime Tools and stolen Data*. IEEE Transactions on Internet Computing,* 6 (2): 3-5

Adler, M. 2012. An Analysis of the Degradation of Anonymous Protocols*. ACM Transactions on Privacy and Security*, 15(3): 5-13.

Aldridge, J. & De´cary-He´ tu, D. 2015. A response to Dolliver's ''Evaluating drug trafficking on the Tor Network: Silk Road 2, the sequel''. *The International Journal of Drug Policy*, 26(11): 1124–1125.

Alsabah, M. 2014. Enhancing TOR's performance using real time traffic classification. *ACM Transactions on the Web,* 12(4): 73-84.

Anderson, C. 2016. Practical Anonymous communication on the mobile Internet using TOR*. IEEE Transactions on Knowledge and Data Engineering, 13(5):* 39-48.

Anon, 2012. Bright Planet. *ACM Transactions on Internet Technology,* 13(2):3-9

Anon, 2013. Number of Running TOR routers. *ACM Transactions on Internet Technology,* 15(1): 4-9

Bai, X. 2014. Traffic identification of TOR and web-mix. *IEEE Transactions on Knowledge and Data Engineering,* 1(2): 548-551.

Barbera, V. 2013. Attacking TOR onion routers on the cheap. *Springer Journal on Information Security,* 15(3): 664-681.

Barker, J. 2014. Using traffic analysis to identify the second-generation onion router*. IEEE Transactions on Services Computing*, 14(6): 72-78.

Barker, D., and Barker, M. 2016. Internet Research Illustrated. *Cengage Journal of Internet Privacy*, 8(6): 4.

Bartlett, J. 2015. The Darknet: Inside the Digital Underworld. *Springer Journal of Information Security*, 12(7): 13-23.

Baudrillard, J. 1984. Simulations. Semiotex, *New York Press*, 2(1):12-22

Bauer, K. 2014. Low resource routing attacks against TOR. *ACM Transactions on Internet Technology,17(5)*: 11-20.

Biryukov, A., Pustogarov, I., Thill, F. and Weinmann, R. 2014. Content and popularity analysis of TOR hidden services. *Springer Journal of Information Security*, 12(6): 9-15.

Biryukov, A.; Pustogarov, I.; and Weinmann, R. 2013. Trawling for tor hidden services: Detection, measurement, deanonymisation. *IEEE Transactions on Services Computing*, 6(5): 35-42.

Blond, L. 2014. One bad apple spoils the bunch: exploiting P2P applications to trace and profile TOR users. *IEEE Transactions on Knowledge and Data Engineering,* 1 (2): 121-142.

Bowen, B. 2014. Automating the injection of believable decoys to protect snooping. Association for Computing Machinery*. ACM Transactions on Storage, 6(7):* 81-86.

Branwen, G. 2015. Darknet Market Mortality Risks. *IEEE Transactions on Knowledge and Data Engineering*, 6(3): 3-6

Briere, M., K., Oosterlinck, and A. Szafarz (2015), 'Virtual Currency, Tangible Return: Portfolio Diversification with Bitcoins', ULB – Universite Libre de Bruxelles, Working Papers CEB: 13–031. Retrieved from http://search.proquest.com/docview/1438547719?accountid=14541

Burch, B. 2014. Tracing anonymous packets to their approximate source. *Journal of Management Information Systems,* 8(5): 319-328.

Carvalho, D. 2017. Darknet usage by Country- The anonymous Internet. *IEEE Transactions on Knowledge and Data Engineering, 6(3):* 2-6.

Caverlee, D. 2015. The Deep Web and Dark the Darknet. *IEEE Transactions on Services Computing,* 2(4): 5-8.

Chaabane A. 2012. Digging into anonymous traffic: A deep analysis of the TOR anonymising network. I*EEE Journal of Internet Computing, 5(8):* 167-174.

Chakravarty, C. 2012. Identifying proxy nodes in a TOR anonymisation circuit. *IEEE Transactions on Knowledge and Data Engineering,8 (2):* 633-639.

Charavarty, S. 2014. Detecting eavesdropping in tor using decoys. *ACM Transactions on Internet Technology,* 11 (2): 221-241.

Chaum, D. 2015. Untraceable Electronic Email, Return Addresses and digital Pseudonyms. *IEEE Transactions on Knowledge and Data Engineering,* 14(6): 84-90.

Chen, F. 2014. Toward improving path selection in TOR. *IEEE Transactions on Knowledge and Data Engineering,* 2(1): 1-6.

Chen, H. 2014. Dark Web: Exploring and Data Mining the Dark Side of the Web. *Springer Journal of Science and Business Media*. 6(2): 134-143

Chertoff, M. 2015. The Impact of the Dark Web on Internet Governance and Cyber Security. *Global Commission on Internet Governance*, 1(2): 101-105.

Christin, N. 2013. Traveling the silk road: A measurement analysis of a large anonymous online marketplace. *Springer Journal of Information Security and Privacy*, 9 (3): 3-11.

Christin, N. 2015. Silk Road: An analysis of a Large Anonymous Online Marketplace. *Springer Journal of Information Security and Privacy*, 10 (6): 2-8.

Ciancaglin, V. 2015. Below the surface: Exploring the Deep Web. *Trendlabs,* 1(6): 19-48.

Clark, A. 2014. TOR. The anonymous Internet, and if it's right for you. *Accessed in Gizmodo,* 1(2): 112- 113.

Clark, J. 2017. The shifting landscape of global Internet censorship. *Harvard Journal of Internet privacy and censorship, 2 (3)*: 18-22.

Clemmit, M. 2016. The Dark Web. *IEEE Transactions on Services Computing, 5 (3):* 12-15.

Comte, A. 1865. *A General View of Positivism*; Trubner and Co, *Cambridge University Press*, 12-19.

Crampton, J. 2016. The biopolitical mapping of cyberspace. *Journal of Management Information Systems, 4 (2):* 389-483.

Davies, N. 2014. Social Networking Analysis. *IEEE Journal of Internet Computing, 8(3):* 5-17.

Dhungel, P. 2016. Waiting for anonymity: Understanding delays in the TOR overlay. *IEEE Journal of Internet Computing, 3(2):* 1-4.

Dingeldine, R. 2015. The second-generation Onion Router. *Springer Journal of Information Security and Privacy,* 4(1)*:* 4-8.

Dingeldine, R. 2014. TOR: The second-generation Onion Router. *Journal of Management Information Systems*, 18(8): 303-319.

Dobson, J. 2016. The panopticons changing geography. *Geographical Review, 1(1):* 307-323.

Dolliver, D. 2015a. A rejoinder to authors: Data collection on TOR. *International Journal of Drug Policy*, 26(11): 1128–1129.

Dolliver, D. 2015b. Evaluating drug trafficking on the Tor Network: Silk Road 2: The sequel*. International Journal of Drug Policy*, 26(11): 1113–1123.

Durkheim, E. 1893/1933. The division of labour in society. *New York: The Free Press*. 21-35.

Edman, E. 2015. AS-awareness in TOR path selection. *Journal of Information Management,* 2(1): 380-389.

Edmundson, A. 2013. Security audit of safe-plug "TOR in a box. *Journal of Management Information Systems, 8(6):* 1-12.

Ehlert, M. 2012. I2P usability vs TOR usability. A bandwidth and latency comparison. *IEEE Transactions on Knowledge and Data Engineering, 8(7)*: 1-7.

European Central Bank (ECB). (2014), Virtual Currency Schemes, http://www.ecb.int/pub/pdf/other/virtualcurrencyschemes201210en.pdf (accessed Mar. 2, 2015).

Evans, N. 2011. A practical congestion at-tack on tor using long paths. *Journal of Management Information Systems, 6*(6*):* 33-50.

Feamster, N. 2015. Location diversity in Anonymity Networks*. ACM Transactions of Internet Technology*, 6(7): 66-76.

Finklea, K. 2016. Dark Web, special report for Congressional Research Service. *ResearchGate Journal on Cybercrime. 1(2):* 2-12.

Flores, R. 2016. Sextortion in the Far East. *Trend Micro Security News. 1 (1)*: 4-8.

Freeman, L. 2015. The development of Social Network Analysis. *Empirical Press, 1(1):* 5-8.

Freeman, D., and Hwa, T. 2015. Detecting clusters of fake accounts in online social networks. *ScienceDirect Journal of Internet Computing*, 5(6): 135-143.

Freeman, D. 2015. Using naive Bayes to detect spammy names in social networks. *IEEE Transactions on Services Computing*, 8(7): 11-14

Fu, X. 2012. One cell is strong enough to break TOR's anonymity. CA, *Journal of Management Information Systems Quarterly*, 3(5): 578-589.

Giddens, A. 1991. Modernity and self-identity. Self and society in the late modern age. Polity Press. pp. 214

Gilad, Y. 2012. Spying in the Dark: TCP and TOR traffic analysis. *Springer Journal of Internet Security and Privacy,* 6(8): 100-119.

Goldberg, I. 2014. On the security of the TOR authentication protocol. Cambridge, *Springer Journal on Information Security and Privacy,* 7(4): 5-9

Goldschlag, P. 2015. Hiding Routing Traffic through TOR and I2P. *Springer Journal on Information Security and Privacy,* 7(4): 137-150.

Goldsmith, J. 2016. Who controls the Internet? *University of Chicago Law review, 1(1):* 1217-1222.

Goncharov, M. 2015. Russian Underground 101. Trend Micro Security Intelligence. *ACM Transactions on Storage,* 7(6): 87-95.

Gottfredson, M. 1990. A general Theory of Crime. California: Stanford University Press, 3(3): 2-4.

Grabosky. 2014. “Virtual Criminality: Old Wine in New Bottles?” Social & Legal Studies. 10(1): 243–49.

Greenberg, A. 2015. Read the Transcript of Silk Road’s Boss Ordering 5 Assassinations. *IEEE Transactions on Services Computing,* 5 (4): *34-45.*

Grinberg, R. 2015. ‘Bitcoin: An Innovative Alternative Digital Currency’. *Hastings Science & Technology Law Journal*, 5(1): 160–207.

Gros, S. 2015. Protecting TOR exit nodes from abuse. *Springer Journal of Information Security and Privacy,* 6(7): 1246-1249.

Gu, L. 2014. The Chinese Underground in 2013. *Trend Micro Security Intelligence,* 4(5): 15-22.

Hacquebord, F. 2015. The Mysterious MEVADE Malware. *ACM Transactions on the Web,* 8 (3): *3-8.*

Hardy, R. 2016. Reputation in the Internet black market: an empirical and theoretical analysis of the Deep Web*. Journal of Information Management*, 6(6): 2-11

Hussain*, G. 2016. "Jihad Trending: A Comprehensive Analysis of Online Extremism and How to Counter It". A special report by Quilliam. 12-19.*

Hevner, A. 2015. Design Research in Information Systems Research. *Journal of Management Information Systems, 3(6):* 75-105.

Hirschi, T. 1969. *Causes of delinquency.* Los Angeles: University of California Press, 4(6): 3-5

Hogan, P. 2015. The Development of Social Network Analysis. A Study in the Sociology of Science. *Vancouver Empirical Press,* 1(1): 20-32.

HoneyNet, 2011. *The HoneyNet Project.* [Online] Available at: http:/www.honeynet.org/. [Accessed 15 October 2017].

Hopper, N. 2011. How Much Anonymity Does Network Latency Leak? *ACM Transactions on the Web*, 5(2): 82-91.

Houmansadr, A. 2014. The parrot is dead: Observing unobservable network communications. *ACM Transactions on Internet Technology, 15(7):* 65-79.

Huber, M. 2012. TOR HTTP usage and information leakage. *Springer Journal of Information Security and Privacy,* 6(7): 245-255.

Incorporated, T. 2015. The Deep Web: Anonymising Technology for the Good. *Trend Micro Security News, 4(3):* 4-12

Isdals, M. 2015. Privacy-preserving P2P data sharing with Oneswarm. *Journal of Management Information Systems* Quarterly, 4(6): 111-122.

Isdals, T. 2014. Privacy-preserving P2P data sharing with Oneswarm. *Springer Journal of Information Security and Privacy,* 5(7): 111-122.

Jaishanksr, K. 2007. Establishing a Theory of Cyber Crime. *International Journal of Cyber Criminology,* 1(4): 17-9.

Jameson, F. 1991. Postmodernism, Or the Cultural Logic of Late Capitalism*, Duke University Press.* 2-22

Jin, J. 2015. On the effectiveness of low latency anonymous network in the presence of timing attack. *Journal of Management Information Systems*, 7(12): 429-438.

Johnson, A. 2014. Users get routed: Traffic correlation on TOR by realistic adversaries. *Journal of Information Management,* 8(1): 121-131.

Katyal, N. 2016. Criminal law in cyberspace. *IEEE Journal of Internet Computing, 3(2):* 1003-1010.

Kruithof, K., Aldridge, J., H´etu, M. and Sim, E. 2016. An Analysis of the Size, Scope, and the Role of the Netherlands. Internet-facilitated Drugs. *ACM Transactions on Privacy and Security*, 6(9): 87-92

King, G. 2015. How censorship in China allows government criticism but silences collective expression. *Springer Journal of Information Security and Privacy,* 7(7): 23-34.

Lavrinc, D. 2015. Someone Bought a Tesla Model S with Bitcoins. *Journal of Management Information Systems, 8(7): 34-43*

Lenhard, J. 2011. Performance measurements of TOR hidden services in low bandwidth access networks. *Springer Journal of Information Security and Privacy,* 2(1): 324-341.

Levine, B. 2012. Timing attacks in low-latency mix-based systems*. Springer Journal of Information Security and Privacy*, 2 (1): 251-265.

Li, C. 2012. A new tunable mechanism of TOR based on the path length. *IEEE Transactions on Services Computing,* 2(16): 661-665.

Loesing, K. 2015. Performance measurements and statistics of TOR hidden services. *IEEE Internet Computing,* 10(2): 1-7.

Marxs, P. 1961. Theories of Society; Foundations of Modern Sociological Theory. *The Free Press of Glencoe*, 1061-1062.

McCanne, S. 2011. [Online] Available at: http://www.tcpdump.org/ [Accessed 11 November 2017].

McCoy, K. 2014. *Shining light in dark places: Understanding the TOR Network. IEEE Journal of Internet Computing*, 9 (4): 23-76.

Merton, K. 1938. Social Structure and Anomie, *American Sociological Review.* 3(1): 672-682.

Mittal, P. 2011. Stealthy traffic analysis of low latency anonymous communication using throughput fingerprinting. *ACM Transactions on Privacy and Security,* 3(2): 215-226.

Moghaddam, B. 2012. Protocol obfuscation for TOR bridges. *ACM Transactions on Privacy and Security. 2(3):* 97-108.

Moore, D., and Rid, T. 2016. Cryptopolitik and the Darknet. *IEEE Transactions on Knowledge and Data Engineering*, 123-135.

Mulazzani, M. 2015. Anonymity and monitoring: how to monitor the infrastructure of an anonymity system. *IEEE Transactions on Services Computing*, 9(4): 539-565.

Murdoch, S. 2015. Low-Cost Traffic Analysis of TOR. *ACM Transactions on Internet Technology*, 8(6): 183-195.

Murdoch, S. 2012. Hot or Not: Revealing hidden Services by their clock skew. *IEEE Journal of Internet Computing*. 27-36.

Norgaard, J., Walbert, H. and Hardy, R. 2017. Shadow Markets and Hierarchies. *IEEE Transactions on Knowledge and Data Engineering,* 13(4): 76-87

Norton, Y. 2016. Sex Addiction as affect dysregulation. *Journal of Clinical Investigation,* 1(1): 1444-1451.

Oh, S. 2017. Fingerprinting keywords in search queries over TOR. *Journal of Management Information Systems,* 19(7): 251-270.

Ovelier, P. 2006. Locating hidden servers. *Journal of Management Information Systems.* 12(1):161-174.

Owen. G. 2017. Tor: Hidden Services and Deanonymisation. *Springer Journal of Information Security and Privacy*, 4(2): 4-7.

Paganini. P, 2017. Digging into the Dark Web. *Journal of Management Information Systems*, 17(2): 10-12.

Pancheko, A. 2015. Performance analysis of anonymous communication channels provided by TOR. *Springer Journal of Information Security and Privacy,* 17(5): 221-228.

Pappas, V. 2010. Compromising Anonymity using packet spinning. *Journal of Information Management,* 2(3): 161-174.

Pries, R. 2015. On performance bottleneck of anonymous communication networks. *Springer Journal on Information Security and Privacy*, 16(5): 1-11.

Provos, N. 2011. *A virtual honeypot framework. IEEE Transactions on Services Computing,* 3(1): 1-14.

Pseudonym, 2014. *Official I2P site on Garlic Routing. ACM Transaction on Privacy and Security,* 4(8): 222-243

Reed, M. 2014. Anonymous connections and Onion Routing. *IEEE Journal of Internet Computing,* 4(6): 482-494.

Reiter, K. 2015. Crowds: Anonymity for web transactions. *ACM Transactions on Privacy and Security,* 12(3): 66-92.

Ritzer, G. 2015. Sociological Theory. New Delhi: *McGraw Hill*. ISBN No.13: 978-0078027017: 12-23.

Ritzer, G. 2008. The McDonaldization of Society*, Pine Forge Press, Los Angeles*. 351-384.

Saleh, S. 2017. Shedding light on the dark corners of the Internet: A survey of TOR research. *IEEE Transactions on Knowledge and Data Engineering,* 18(6): 1-35.

Sancho, D. 2015. Steganography and Malware: Concealing Code and C&C. *Trendlabs Security Intelligence*. 2(2): 13-23.

Stevens, G. 2016. The truth about the Deep Web. *Accessed in KernelMag,* 1(1): 8-13.

Stobing, C. 2018. Using deep web search engines for academic and scholarly research. *Journal of Management Information Systems*, 19(2):121-134

Tang, C. 2012. An improved algorithm for TOR circuit scheduling. *ACM Transactions on Privacy and Security,* 4(22): 329-339.

Tarde, G. (1903). The Laws of Imitation, pp 195-199, H. Holt and Company, New York.

Tonnies, *P.* 1859. Library of Liberal Arts. *Journal of new Librarianship*, 1(2):12-25

Tor Project. 2015. Ethical Tor Research: Guidelines. https: //blog.torproject.org/blog/ethical-tor-research-guidelines.

TOR, 2016. *TOR Metrics Portal.* [Online] available at: http://www.torproject.org/ [Accessed 22 September 2017].

Trendlabs, 2017. The many faces of cybercrime. *Trend Micro Security News*. 3(3): 65-76

Ulrich, B. 1992. Risk Society*:* Towards a New Modernity. *Sage Journal of Risks in Society and Societal Behaviour*, 3(2): 2-8.

Van Buskirk, J., Roxburgh, A., Naicker, S., & Burns, L. (2015). A response to Dolliver's Evaluating drug trafficking on the Tor network. *International Journal of Drug Policy*, 26(11): 1126–1127.

Veblen, L. 1991. Gemeinschaft und Gesellschaft, *Journal of Opladen: Leske Budrich.* 3(2). 12-17

Vinto, K. 2015. Silk Road CreaTORRoss. *International Journal of Drug Policy.* 6(2): 112-134

Vitaris, B. 2016. Russian is collecting encryption keys as Änti- Terrorism legislation goes into effect. *ACM Transactions on Privacy and Security*, 12(3): 12-32

Warf, B. 2016. Geographies of global Internet Censorship. Accessed in *GeoJournal,* 2(1): 12-20.

Weber, M. 1991. The Nature of Social Action in Runciman. *Journal of Selections in Translation*. 2(3): 35-43

Weimann, G. 2016. Terrorist Migration to the Dark Web. *Journal of Perspectives on Terrorism,*1(1): 1-4.

Weimann, G. 2016. Going Dark: Terrorism on the Dark Web, *Studies in Conflict & Terrorism*, 1(1): 195-206.

Wendolsky, R., 2015. Performance comparison of low latency anonymisation services from a user perspective. *Journal of Management Information Systems,* 13(3): 31-35.

Westlake, B. 2017. Assessing the validity of automated WebCrawler's as data collection tools to investigate online child sexual exploitation. *Journal of Sexual Abuse,* 29 (7): 2-10

Winter, P. 2014. How the great firewall of China is blocking TOR. *Springer Journal of Information Security and Privacy, 7(6):* 1-4.

Wright, M. 2015. An analysis of the degradation of anonymous protocols. *Journal of Management Information Systems, 11 (6):* 116-122.

Wright. P. 2017. Pentagon Hunts for ISIS on the Secret Internet. *IEEE Transactions on Services Computing*, 3(1): 4-8

Yaneza, J. 2014. Defending Against TOR-Using Malware, Part 1. *Trendlabs Security Intelligence Blog,* 6 (1): 12-18.

APPENDIX A

TOR CONFIGURATION FILE WITH SAMPLE RESULTS

```
1##TORConfigurationFile
2
3
4 ## Section 1: Basic Settings ##############################################
5
6.  ## Send all messages of level 'notice' or higher to /var/log/TOR/notices.log
7.  #Log notice file /var/log/TOR/notices.log
8.  ## Send every possible message to /var/log/TOR/debug.log
9.  Log debug file /var/log/TOR/debug.log
10
11. ## If set to 1, TOR will scrub any personally identifying information
12. ## from the log files
13. SafeLogging 0
14
15. ## Start the process in the background
16. RunAsDaemon 1
17
18. ## The directory for keeping all the keys/etc.
19. DataDirectory /var/lib/TOR
20
21
22 ## Section 2: Client Settings ###############################################
23
24. ## Define on which port the SOCKS client will listen
25. #SocksPort 0 # don't function as a client
26. SocksPort 9050 # what port to open for local application connections
27
28. ## Define on which IP address the client will listen
29. #SocksListenAddress # accept connections only from localhost
30. SocksListenAddress 0.0.0.0 # accept connections from anyone
```

```
31
32
33 ## Section 3: Relay Settings ############################################
34
35. ## A unique name for the server
36. Nickname TORPhD
37
38. 38 ## The IP or FQDN of the server
39. Address……..
40
41. ## Contact info to be published in the directory
42. ContactInfo Nobody <nobody AT example dot com>
43
44. ## Port to advertise for TOR connections
45. ORPort 9001
46
69
47. ## Publish server descriptors of certain versions
48. PublishServerDescripTOR v2, v3
49
50. ## A comma-separated list of exit policies
51. #ExitPolicy accept *: * # any exiting ports/protocols allowed
52. ExitPolicy reject *: * # no exiting ports/protocols allowed
53
54
55
56 ## Section 4: Authoritative Directory Settings ########################
57
58. ## What port to advertise for incoming directory connections
59. DirPort 9030
60
61. ## Become an authoritative directory
62. AuthoritativeDirectory 1
63
```

```
64. ## Specify which versions to publish
65. V2AuthoritativeDirecTORy 1
66. V3AuthoritativeDirecTORy 1
67
68. ## To make routers show up as "named" in the directory
69. NamingAuthoritativeDirectory 1
70
71. # Display this HTML page at the root of the directory server's port
72. DirPortFrontPage /var/lib/TOR/webpage.html
73
74. ## Entrance policy for this directory server
75. DirPolicy accept *: * # any ports/protocols allowed
76
77
78 ## Section 5: Settings for running TOR on a private network ############ 79
80. ## List of authoritative routers
84
85. ## Makes running on a private TOR network possible by speeding up the
86. ## voting process and disabling some security restrictions
87. TestingTorNetwork 1
```

Success	Czechia	50.0833	14.4667	Europe/Prague	HOSTING90 systems s.r.o.2	130.193.15.186
success	Republic of Korea	37.4732	127.038	Asia/Seoul	SK Broadband	218.232.120.97
success	Republic of Korea	37.4732	127.038	Asia/Seoul	SK Broadband	218.232.120.97
success	Republic of Korea	37.4732	127.038	Asia/Seoul	SK Broadband	218.232.120.97
success	Republic of Korea	37.4732	127.038	Asia/Seoul	SK Broadband	218.232.120.97
success	Republic of Korea	37.4732	127.038	Asia/Seoul	SK Broadband	218.232.120.97
success	Republic of Korea	37.4732	127.038	Asia/Seoul	SK Broadband	218.232.120.97
success	Republic of Korea	37.4732	127.038	Asia/Seoul	SK Broadband	218.232.120.97
success	Republic of Korea	37.4732	127.038	Asia/Seoul	SK Broadband	218.232.120.97
success	Republic of Korea	37.4732	127.038	Asia/Seoul	SK Broadband	218.232.120.97
success	Republic of Korea	37.4732	127.038	Asia/Seoul	SK Broadband	218.232.120.97
success	Republic of Korea	37.4732	127.038	Asia/Seoul	SK Broadband	218.232.120.97
success	Republic of Korea	37.4732	127.038	Asia/Seoul	SK Broadband	218.232.120.97
success	Republic of Korea	37.4732	127.038	Asia/Seoul	SK Broadband	218.232.120.97
success	Republic of Korea	37.4732	127.038	Asia/Seoul	SK Broadband	218.232.120.97
success	Republic of Korea	37.4732	127.038	Asia/Seoul	SK Broadband	218.232.120.97
success	Republic of Korea	37.4732	127.038	Asia/Seoul	SK Broadband	218.232.120.97
success	Republic of Korea	37.4732	127.038	Asia/Seoul	SK Broadband	218.232.120.97
success	Republic of Korea	37.4732	127.038	Asia/Seoul	SK Broadband	218.232.120.97
success	Republic of Korea	37.4732	127.038	Asia/Seoul	SK Broadband	218.232.120.97
success	Republic of Korea	37.4732	127.038	Asia/Seoul	SK Broadband	218.232.120.97
success	Republic of Korea	37.4732	127.038	Asia/Seoul	SK Broadband	218.232.120.97
success	Republic of Korea	37.4732	127.038	Asia/Seoul	SK Broadband	218.232.120.97
success	Republic of Korea	37.4732	127.038	Asia/Seoul	SK Broadband	218.232.120.97
success	Republic of Korea	37.4732	127.038	Asia/Seoul	SK Broadband	218.232.120.97
success	Republic of Korea	37.4732	127.038	Asia/Seoul	SK Broadband	218.232.120.97
success	Republic of Korea	37.4732	127.038	Asia/Seoul	SK Broadband	218.232.120.97
success	Republic of Korea	37.4732	127.038	Asia/Seoul	SK Broadband	218.232.120.97
success	Republic of Korea	37.4732	127.038	Asia/Seoul	SK Broadband	218.232.120.97
success	Republic of Korea	37.4732	127.038	Asia/Seoul	SK Broadband	218.232.120.97
success	Republic of Korea	37.4732	127.038	Asia/Seoul	SK Broadband	218.232.120.97
success	Republic of Korea	37.4732	127.038	Asia/Seoul	SK Broadband	218.232.120.97
success	Republic of Korea	37.4732	127.038	Asia/Seoul	SK Broadband	218.232.120.97
success	Republic of Korea	37.4732	127.038	Asia/Seoul	SK Broadband	218.232.120.97
success	Republic of Korea	37.4732	127.038	Asia/Seoul	SK Broadband	218.232.120.97
success	Republic of Korea	37.4732	127.038	Asia/Seoul	SK Broadband	218.232.120.97
success	Republic of Korea	37.4732	127.038	Asia/Seoul	SK Broadband	218.232.120.97

success	Republic of Korea	37.4732	127.038	Asia/Seoul	SK Broadband	218.232.120.97
success	Republic of Korea	37.4732	127.038	Asia/Seoul	SK Broadband	218.232.120.97
success	Republic of Korea	37.4732	127.038	Asia/Seoul	SK Broadband	218.232.120.97
success	Republic of Korea	37.4732	127.038	Asia/Seoul	SK Broadband	218.232.120.97
success	Republic of Korea	37.4732	127.038	Asia/Seoul	SK Broadband	218.232.120.97