Surfacing Collaborated Networks in Dark Web to Find Illicit and Criminal Content

Ahmed T. Zulkarnine, Richard Frank, Bryan Monk, Julianna Mitchell, Garth Davies
International CyberCrime Research Center (ICCRC)
School of Criminology, Simon Fraser University
Burnaby, Canada
{tzulkarn, rfrank, bmm8, jdm17, garthd}@sfu.ca

***Abstract*— The Tor Network, a hidden part of the Internet, is becoming an ideal hosting ground for illegal activities and services, including large drug markets, financial frauds, espionage, child sexual abuse. Researchers and law enforcement rely on manual investigations, which are both time-consuming and ultimately inefficient. The first part of this paper explores illicit and criminal content identified by prominent researchers in the dark web. We previously developed a web crawler that automatically searched websites on the internet based on pre-defined keywords and followed the hyperlinks in order to create a map of the network. This crawler has demonstrated previous success in locating and extracting data on child exploitation images, videos, keywords and linkages on the public internet. However, as Tor functions differently at the TCP level, and uses socket connections, further technical challenges are faced when crawling Tor. Some of the other inherent challenges for advanced Tor crawling include scalability, content selection tradeoffs, and social obligation. We discuss these challenges and the measures taken to meet them. Our modified web crawler for Tor, termed the "Dark Crawler" has been able to access Tor while simultaneously accessing the public internet. We present initial findings regarding what extremist and terrorist contents are present in Tor and how this content is connected to each other in a mapped network that facilitates dark web crimes. Our results so far indicate the most popular websites in the dark web are acting as catalysts for dark web expansion by providing necessary knowledgebase, support and services to build Tor hidden services and onion websites.**

Keywords*— *Tor Network, Web Crawler, Criminal Network, Dark Web, Web Graph, Social Network Analysis.

I. INTRODUCTION

The dark web, an intentionally hidden part of the internet, helps to protect the privacy of internet users from traffic analysis attacks [1] [2]. This portion of the internet can only be accessed through specialized dark web browsers or technologies [3]. Mansfield [3] identified the central characteristics of this true dark web: 1) decentralization – most commonly peer-to-peer technology is used, where content resides on a group of distributed personal computers instead of a centralized server; 2) internet infrastructure utilization, where the dark web is built atop the public internet; and 3) usage of non-standard computer protocols and ports unreachable by those outside the network. Some of these characteristics occasionally align between the surface and the deep web. In the dark web, the user can only access the contents and resources through specific protocols and the network is built such that the user's identity can't be traced back. Various dark web software are widely available to the public, with Tor Network being the most popular [4]. Historically, the Tor Network has been recognized as an ideal hosting ground for many illegal activities and services, from large drug markets [5] to child sexual abuse [6]. Numerous studies have established that extremist and terrorist organizations have used, and continue to use, the public internet for the purposes of recruitment, private communication, networking and inciting violence [7]. However, very little is known regarding illicit and extremist activities on the Tor network. The first part of this paper discusses the current studies that have been undertaken by researchers to better understand the Tor topology. Consequently, we explored dark web studies that have assessed the presence of illicit and criminal contents in the dark web.

Manual investigations, which are commonly used by researchers and law enforcement, are both time-consuming and ultimately inefficient. Some research has utilized automated mechanisms to explore Tor but there have been no prior studies which have systematically explored or assessed the content available within the Tor hidden network [8] [9]. In addition, the most comparable research used manual means of uncovering Tor hidden services and failed to provide specific focus on extremist or terrorist activities [9].

This paper highlights some technological challenges faced when attempting to survey illicit and extremist content using tools capable of shining a light onto this anonymous network. Tor functions differently in transport (TCP) level and uses socket connections. Some of the other inherent challenges for advanced Tor crawling include scalability, content selection tradeoffs, and social liability (being a good citizen by not interrupting or overburdening web servers). With these challenges in mind, we developed an enhanced web crawler, termed the "Dark Crawler". It automatically searches websites on the internet based on pre-defined keywords and/or image hash values, and follows hyperlinks within the webpages, seeking out content of interest. We will also discuss different strategies and components incorporated into analysis, which are based on the data extraction from the Tor network using the

978-1-5090-3865-7/16/$31.00 ©2016 IEEE

Dark Crawler. Through social network analysis, our initial results provide insight into how owners and administrators of dark websites are ensuring a better quality of product and support.

II. LITERATURE REVIEW

A. Tor Network Topology and Limitations

Despite the existence of other dark web software, the Tor network remains one of the most frequently used and well-known, and is favored by users who wish to hide their activities [10]. The Tor Network, also known as the Onion Router [11], was initially developed by the U.S Navy in the 1990s with the purpose of protecting U.S intelligence communications online [12]. This network was launched officially in 2002 to the public with the core objective of open source information gathering [11].

The Tor network provides anonymity by routing traffic through other users who are using the specialized web browser "Tor" and have declared themselves to be *nodes* within the network [13]. Whenever a Tor user, referred to as a *source*, joins the network though TorBrowser, a *virtual circuit* is constructed using a random selection of Tor nodes (i.e. a selection of usually 3 computers running the TorBrowser). This virtual circuit is used for approximately ten minutes, after which a new virtual circuit is created. This circuit contains three types of nodes: (1) entry nodes – the first node in the circuit which accepts incoming traffic; (2) intermediate nodes – which pass data along from a node to the next; and, (3) exit nodes – the last node in the circuit which delivers traffic to the open internet [14]. When a source requests access to a website, the "web request" is encrypted through multiple layers and sent to the entry node. From the entry node, the "web request" is passed to a virtual circuit which contains randomly chosen intermediate nodes placed around the world. At each hop, a single layer of encryption is taken off from the "web request" before passing it to the next intermediate node [5]. Once the "web request" reaches the exit node, all of the encrypted layers are taken off and the unencrypted "web request" is sent to the web server along the public internet [14]. In this process, the source information is lost and the user in the Tor network remains anonymous. Traffic can only be traced back to the previous link in the virtual circuit.

Despite Tor's impressive architecture, researchers have attempted to reveal and exploit its weaknesses. Passive traffic analysis and active watermarking are two prominent traffic analysis techniques to attack the Tor network [15]. In a passive traffic analysis attack, the traffic is recorded unreceptively and statistical measures are incorporated to identify the relationship between a sender's outbound traffic and the receiver's inbound traffic [15]. MIT researchers, in collaboration with the Qatar Computing Research Institute (QCRI) have devised a machine learning algorithm which examines traffic patterns of data packets of immediate nodes. This algorithm can determine 1) whether a virtual circuit is an introductory point circuit (circuit used by a location-hidden[1] service to advertise its existence in Tor network), a rendezvous-point circuit (circuit used by the Tor user to connect to these hidden services) or is an ordinary web-browsing circuit. They were able to classify the various nodes with 99% accuracy, nodes hosting hidden services with 88% accuracy, and the websites the user was browsing with 88% accuracy [16]. Watermarking is a more direct attack, where each transaction at each node of the network is recorded. However, this method requires the attacker to take control of the corresponding entry and exit nodes. The malicious exit node manipulates the cell (data packet) from a TCP stream and embeds a "secret signal" (series of bits) using Key Based Random Permutation (KBRP) [15]. The secret signal is transported along with target traffic and reaches the malicious entry onion router. Afterwards, an accomplice of the entry node detects the secret signal based on the received cell. This method effectively and accurately detects the anonymous communication relationship among Tor users [15]. Tor network developers officially acknowledged this attack method and have taken steps regarding this vulnerability. These short-term steps include the removal of the attacker from the system, software updates to prevent these relays and improving hidden service design [17].

Another major Tor limitation has been identified by McCoy and his colleagues [1]. Malicious exit nodes, if exploited to maximum capability, can make the Tor network vulnerable to some extent. Usage of insecure protocols by nodes are prominent in the Tor network, hence malicious routers can capture certain login information such as usernames and passwords [1] by logging traffic.

B. Illicit and Criminal Content

Tor network facilitates dark web markets (referred as darknet) that provide an environment conductive for illegal transactions. Buyers and vendors using these platforms have the opportunity to proceed with illegal transactions while reducing concerns over criminal sanctions. Bitcoin, a virtual trading currency, is the most commonly used currency in illegal virtual marketplaces [5].

A variety of illicit and criminal commodities are found in these dark virtual spaces, ranging from hiring assassins, acquiring and sharing child pornography, sharing criminal ideologies and hacking information, buying and selling illegal drugs, stolen social security numbers, as well as other fraudulent identity information [18].

The Tor network lacks full scale search engine functions such as those provided by Google to the surface web. Illicit offerings therefore are listed in numerous websites and databases that users have to access. For example, despite the closure of SilkRoad and its 13,648 different drug deals, thousands of new dealers have found dozens of new drug internet marketplaces [19] through the various dark web indexes.

According to Dredge [11], 24% of the dark web's content is accounted for by drug-related sites, which attract about 5% of the dark web traffic. Power [19] provided a new list of a dozen

[1] Location hidden services allow Tor users to provide a TCP service, such as web server without revealing its IP address [30]

dark web drug stores that dominated the market: Blue Sky, Hydra, Agora, Evolution, Pandora, Pirate, BlackBank, Tor Bazaar, Cannabis Garden, Cloud Nine, Andromeda, Outlaw, and Alpaca. Purchasing drugs online is becoming more attractive because of several factors: including convenience, choices, prices and user ratings [19]. Because numerous drug e-commerce platforms are available, the competition has increased with websites trying to constantly find novel ways to be more popular and scalable [19]. Dredge [11] claims 80% of the dark web traffic is generated by requests to access child exploitation and abuse material. According to this study, 45,000 hidden services are estimated to be present at Tor network at any given time. Only 2% of the Tor network content contains child exploitation materials. The cause of such high demand of child abuse content is still a question to researchers and many critics have questioned the accuracy of such data.

Terrorist organizations also use markets to finance their activities through the buying and selling of illegal weapons and drugs. O'Neil [20] identified the "Armory", which was assumed to be the most well-known weapon marketplace in the Dark Web after Silk Road, containing around 400 items for sale ranging from bombs to bullets. The minimum purchase was $1050 and the price of the majority of offered items was many times above legal market price. An AK-47 automatic assault rifle, for example, sold for $2800. The legal price of AK-47 starts from $499 and according to Oxford University research, the average global price of this lethal weapon is $534 [31]. The legal purchase of this weapon requires the buyer to be registered at governmental agencies in countries such as the United States. This e-commerce platform used a private shipping system with bases in 15 countries, allowing buyers and sellers to avoid the postal system while using a large fleet of trucks [21]. Hence, it made it more difficult for law enforcement to track the dark web criminals.

Among the variety of criminal recruitment services, hiring hitmen and hackers top the list. According to Daily Mail Online in the UK, dozens of hitman are available for recruitment through the Tor network [22]. Generally, the price of hitmen for an American or Canadian operation is ten thousand dollars, whereas the same job in Europe is approximately twelve thousand dollars.

III. METHODS

A. *Web Crawling*

We developed the "Dark Crawler", a modified crawler based on a custom build web crawler, Child Exploitation Network Extractor (CENE). CENE was created by a computer scientist in collaboration with criminologists [23] and initially focused on finding child exploitation material. CENE previously has demonstrated success in locating and extracting data on child exploitation images, videos, keywords and links on the public internet. Our enhanced "Dark Crawler" has been able to access Tor while simultaneously accessing the public Internet. It automatically searches websites within the Tor network based on pre-defined keywords, then stores this raw data in a database. In order to infiltrate the Tor network, we used a third-party local HTTP proxy software called Privoxy, which seamlessly connects the Dark Crawler to the Tor network. Privoxy is a non-caching web proxy that provides advanced

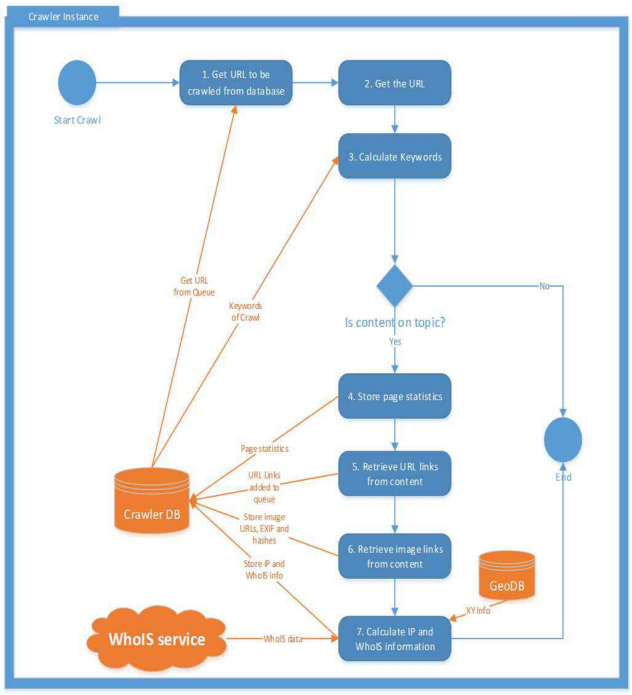

Fig. 1. *Workflow associated with a thread of Dark Crawler*

filtering options for controlling HTTP headers and data [24]. The "Dark Crawler" operates by starting the crawling process at user-specified websites, referred to as seeding sites. It retrieves HTML content of the pages from onion websites (websites in the Dark Web), analyzes them and recursively follows the outgoing links on the pages. In the process, the page statistics and content are stored in the database.

The images found on the webpages are hashed using MD5, SHA1 and PhotoDNA and subsequently also stored in the database. These hash values will be checked in the future against a hash-database which contains categorized hash values of child exploitation and terrorism images identified by law enforcement authorities. Finally, for surface websites, the WhoIS information is retrieved and the IP address is geo-located using an internal Geodatabase. The resulting XY coordinates are stored in the database associated to that webpage. See Figure 1 for a summary.

Christin [18], for example, used a web-crawler on Tor specifically to study the Tor-based Silk Road. However, for that project a very simple web-crawler was used to blindly scan a specific website and collect information about it. Extending a single website to a broad scan within Tor poses a significant challenge, as Tor websites are often not linked with one another, and many Tor websites are only found through searches of the public internet. Advanced Tor web crawling presents some inherent challenges:

1) Scalability: The dark web is very large and evolving every day. Seeking a broad coverage and freshness, the web crawler is required to achieve extremely high throughput, which poses many difficult engineering problems [25]. To address the scalability issue we used "distributed crawling"; that is, measures taken to ensure the crawling is distributed over

multiple machines to increase the collective throughput. The distribution has been done by partitioning the URL space, so that each crawler machine or node is responsible for a subset of the URLs on the web.

2) Content selection tradeoffs: Even the best crawlers do not crawl the whole web, or capture all of the changes [25]. Rather, web crawling is done selectively and in a carefully controlled order. The core goals include acquiring high-value content quickly, ensure eventual coverage of all reasonable content, and to get rid of low-quality, irrelevant and redundant content. In order to tackle content selection tradeoff, scope crawling has been used. Scope crawling limits crawling activities to pages that fall within a particular category or scope, and, as a result, acquires in-scope content much faster and cheaper. In order to achieve scope crawling, the database behind the crawler has been enhanced so that it can list the "don't explore domains" to speed up crawling. Previous versions of the crawler stored the webpages to be crawled in a single "queue," but this was inconvenient in terms of prioritizing important webites and data retreival, and, more importantly, was expensive in randomizing domains so the crawler did not get stuck in a certain domain. For the "Dark Crawler" we segregated the queue to into two, one for domains and another for pages (while still linking them to the queue) to help prioritize the important items and their order.

3) Social obligations: The Crawler should not impose too much of a burden on the web sites they crawl [25]. In fact, without proper ordering of the crawler activities, the crawler might unintentionally carry out a denial-of-service attack. Thus the crawler was modified to only create two requests/connections to each website, and thus distribute the load across multiple websites.

B. Analytic Tools

A web application was developed to visualize and analyze the crawled data. This web application helped to manage the queue of the crawler, view crawled data (domain, pages and content) and initiate new crawls to be executed. In addition, it provided hyperlink analysis and graph construction based on the crawled data. Dealing with big data creates performance issues. Response latency time to visualize the crawled data was enormous. Initially the majority of HTTP requests to the web application were timing out due to database bottlenecking. Specifically, keyword search was extremely expensive. Utilizing SQL operator such as "LIKE" for keyword searching was exceptionally slow when database tables were enormous. In order to address this issue, we incorporated Sphinx into the web application. Sphinx is a standalone full text search engine, providing faster and efficient keyword searches (full-text) by creating indexes based on full-text database content. Utilizing Sphinx provided approximately 90% faster retrieval time for the keyword searches. The indexes of Sphinx are updated every day using batch processes to ensure fresh data is available for content analysis.

C. Web Graph, Degree Distribution and Centrality

Real-world complex biological, technological or social phenomena interpretation requires the understanding of structure and function of that complex network [26]. Social network analysis provides the core methods and models to analyze data generated from the internet [27]. Criminologists have used these methods and models to effectively gain more insight of crime and deviants. To gain insight into how illicit network functions, the ties between individuals in criminal enterprises need to be uncovered [28]. Malm and Bichler [28] conducted social network analysis on 1,998 individuals associated or involved with drug trafficking using police intelligence report from 2004 to 2006 as a data source. According to their result, individuals who are involved in financing, supply, smuggling and other niches as well have the highest fragmentation proportion [28]. Social network analysis is one of the core components in this dark web research. One of the core focuses of this research will be to examine the popularity of the Tor websites, as measured through degree distribution and network centrality. These analyses will occur within a graph constructed from crawled website data.

The degree of a website (vertex) v is the number of edges incident to that website v in a graph $G = (V, E)$. Degree of a website i denoted by k_i is the number of vertices adjacent to website i in an undirected graph. For directed graphs, in-degree of website i denoted by $k_{in}(i)$ is the number of directed links (edges) pointing to website i and out-degree $k_{out}(i)$ is the number of directed links (edges) pointing away from websites i to other websites.

In a collaborated online criminal world, network centrality answers the question: "What characterizes the most popular and important crime websites?" This measure plays an important role in understanding the importance of actors in a social network. In addition, to understand how an actor (i.e. a website or deviant individual) exerts its influence on other actors, we need to use an index that measures the centrality of an actor. In the dark web research, this measure will help us identify the most popular websites. In order to find out the central websites of a graph containing N websites (vertices), we will use degree distribution of the websites. The normalized in-degree score of a website i denoted by $< k_{in}(i) >$ is calculated by

$$< k_{in}(i) > = \frac{k_{in}(i)}{N-1} \quad (1)$$

Similarly, the normalized out-degree score of website i denoted by $< k_{out}(i) >$ is calculated by

$$< k_{out}(i) > = \frac{k_{out}(i)}{N-1} \quad (2)$$

Based on the normalized in-degree and out-degree score comparison, we will find the most central, as well as popular websites.

IV. RESULTS, ANALYSIS AND DISCUSSION

We present initial findings regarding extremist and terrorist content present in Tor and how this content is interconnected in a web graph, facilitating dark web crimes. The database used to analyze content contained 10,163 distinct Tor domains (onion

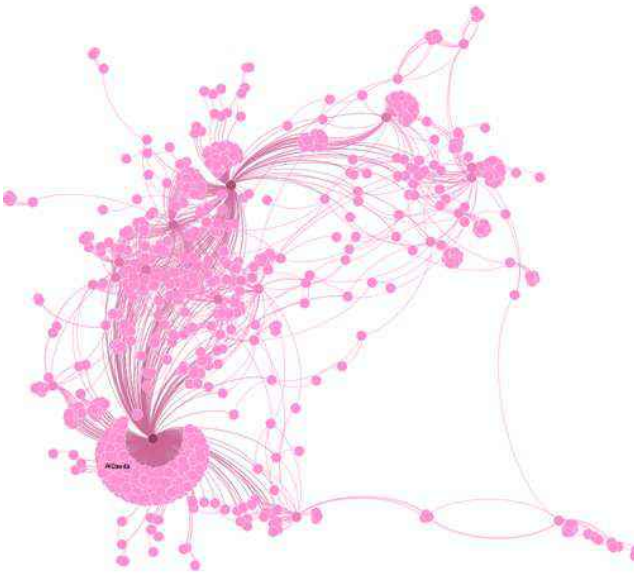

Fig. 2. *Collaborated network with extracted Tor websites.*

addresses) and 54,141 Tor web pages. This required approximately 260GB of data to be stored in the database.

The Tor network analysis suggests the networking topology is sparse, with an overall 2% of the edges connecting the websites with each other. A directed graph has been constructed using 766 Tor webpages where each webpage is considered as a vertex and the hyperlinks as edges. Figure 2 represents a directed graph that has been constructed to visualize the collaborated network retrieved.

In-degree scores provide a good overview of the most popular websites. These scores indicate the websites that people are willing to link to with their website. According to our finding, it can be inferred that Tor users are interested in diverse services offered in the dark web. Interestingly, our list tops with a web hosting and design website within the Tor network. This site offers customers a fully customized e-commerce onion website. To ensure anonymity, this onion site has integrated its payment processing system with the (mainly) anonymous bitcoin currency. Reliability can be a challenge for dark websites, who attempt to enhance their services by providing daily backup to the clients.

TABLE I. Top 5 popular Tor website based in-degree centrality

#	In-degree Distribution of Tor Websites			
	Website	***URL***	$< k_{in} >$	***Category***
1	TorShops	http://shopsat2dotfotbs.onion:80	0.022149	Hosting
2	Hidden Service	http://nfokjgfj3hxs4nwu.onion	0.019688	Infrastructure Management
3	Russian Anonymous MarketPlace	http://ramp2bombkadwvgz.onion:80	0.014766	Marketplace
4	Torch: Tor Search Engine	http://xmh57jrzrnw6insl.onion:80	0.013946	Search Engine
5	Galaxy2	http://w363zoq3ylux5rf5.onion:80	0.013946	Social Networking

The second most popular site indicates why the dark web has expanded rapidly. This site provides the necessary documentation for setting up hidden services in the Tor network. More specifically, it provides technical directions to system administrator how to setup Apache webserver in Ubuntu using an onion domain.

The third most popular website is a Russian marketplace which sells a good range of "light" drugs [29]. Website reviewers in certain blogs have revealed [29] that the inefficiency of Russian law enforcement has lead this Tor website to be popular and a stable marketplace. The Torch search engine was revealed as the 4th popular tor website in our centrality analysis. Finally, a social media website named Galaxy 2 find its place the top 5 list.

TABLE II. Top 5 popular website based on out-degree centrality

#	Out-degree Distribution of Tor Websites			
	Website	***URL***	$< k_{out} >$	***Category***
1	PunPun	http://htzdaj24brekerl2.onion:80	0.420016	Directory
2	Hidden Links V0.1.1	http://hlinkhign4obv3a3.onion.link:80	0.171452	Directory
3	TorVPS	http://torvps7kzis5ujfz.onion.link:80	0.14274	Hosting
4	Onion Soup	http://s7kgnncq3zbe3yza.onion:80	0.113208	Political Blog
5	Onion Soup	http://soupksx6vqh3ydda.onion:80	0.112387	Political Blog

However, our main interest lies in illicit and criminal contents. Hence we have analyzed the content of the 4th and 5th top ranked outgoing websites. Onion Soup is a political blog that provides anti-US government and anarchist perspectives. This site contains political forums that mostly cover on "criticizing the break-down of rationality in US politics". Some of its content also focuses on US government activities abroad.

TABLE III. Prominent websites for keyword Searches

Keyword	**Extremist and Terrorism promoting website**		
	Website	***URL***	***Category***
Terrorist	TorLinks	http://torlinkbgs6aabns.onion/#political	Discussion Board
	Freenet	http://freenet7cul5qsz6.onion	Discussion Board
	FuckOff-And-Die.Com's Onion portal	http://3il6wiev2pnk7dat.onion	Discussion Board
	"name unavailable"	http://uudllt7casd3cykd.onion	Discussion Board
Extremism	Contranumenism Manifestation	http://contra6am7tdml6h.onion	Organization's Website
	Hack Canada	http://hackcanl2o4lvmnv.onion/	Organization's Website
National Secuirity	Freenet	http://freenet7cul5qsz6.onion	Discussion Board

Finally, based on keyword searches we tried to identify the prominent websites that promote extremist and terrorism discussions and perspectives. The first 20 Tor webpages of each keyword search were analyzed to gain a better understanding of the content. The websites which contain discussion or

commentaries that focus on government actions and events related to extremist and terrorist activities. Few prominent anti-religious organization's websites contain content which explicitly supports and advocates the use of violence to achieve religious goals.

V. CONCLUSION

In this paper, we discussed about the technical challenges and measures taken for handling broad scale web crawl specific to the Tor network. Analyzing big data introduced significant challenges and thus full-text search engine integration helped to provide a far-reaching dark web content analysis from an architectural point of view. We performed social network analysis on the extracted data which indicated the most popular websites are facilitating the expansion of dark web. Our analysis also suggests interest of Tor users ranges on variety of services offered in the Tor network. In addition, we can infer the most popular terrorist and extremist discussions focus on anti-government rationality and advocacy of religious violence.

In addition, to perform more detailed content analysis, there are a number of extensions to this work we envision in future. First, it will be interesting to incorporate keyword and automated sentiment analysis in the dark crawler. This will help us to classify and analyze a) violent extremist and terrorist contents b) anti-extremist websites c) neural sources and c) un-related content. Secondly, to obtain a better understanding of the architecture of the Tor network, it will be helpful to develop and evaluate more social network analysis metrics such as characteristic path lengths, clustering co-efficient, etc. Finally, a longitudinal study of these metrics will provide powerful insight regarding how the attributes of these popular websites helps them to scale, survive or decays in a long run.

ACKNOWLEDGEMENT

This research is supported by Defense Research and Development Canada (DRDC). The authors wish to thank the anonymous referees for many helpful suggestions that improved the presentation of the manuscript.

REFERENCES

[1] D. McCoy, K. Bauer, D. Grunwald, T. Kohno and D. Sicker, "Shining light in dark places: Understanding the Tor network," in *Privacy Enhancing Technologies*, 2008.

[2] D. Lacey and P. M. Salmon, "It's Dark in There: Using Systems Analysis to Investigate Trust and Engagement in Dark Web Forums," in *Engineering Psychology and Cognitive Ergonomics*, Springer, 2015, pp. 117-128.

[3] S. Mansfield-Devine, "Darknets," *Computer Fraud \& Security,* vol. 2009, no. 12, pp. 4-6, 2009.

[4] K. Misata, "the tor project: An inside view," *XRDS: Crossroads, The ACM Magazine for Students,* vol. 20, no. 1, pp. 45-47, 2013.

[5] A. GREENBERG, *Hacker Lexicon: What Is the Dark Web? | WIRED,* 2014.

[6] M. Ward, *Tor's most visited hidden sites host child abuse images - BBC News,* 2014.

[7] G. Weimann, Terror on the Internet: The new arena, the new challenges, US Institute of Peace Press, 2006.

[8] T. Fu, A. Abbasi and H. Chen, "A focused crawler for Dark Web forums," *Journal of the American Society for Information Science and Technology,* vol. 61, no. 6, pp. 1213-1231, 2010.

[9] C. Guitton, "A review of the available content on Tor hidden services: The case against further development," *Computers in Human Behavior,* vol. 29, no. 6, pp. 2805-2815, 2013.

[10] S. Lu, *What is the dark web and who uses it? - The Globe and Mail,* 2015.

[11] S. Dredge, *What is Tor? A beginner's guide to the privacy tool | Technology | The Guardian,* 2013.

[12] J. B. Fagoyinbo, The Armed Forces: Instrument of Peace, Strength, Development and Prosperity, AuthorHouse, 2013.

[13] TorProject, *Tor Project: Anonymity Online,* 2015.

[14] B. Conrad and F. Shirazi, "A Survey on Tor and I2P," *Proc. 9th ICIMP,* p. 22, 2014.

[15] S. Deepa, "Masquerading Attack To Break Tor's Anonymity," *International Journal of Computer Trends and Technology (IJCTT) ,* vol. 4, no. 6, pp. 1722 - 1726, 2013.

[16] L. Hardesty, *Shoring up Tor | MIT News,* 2015.

[17] Arma, *arma's blog | The Tor Blog,* 2014.

[18] N. Christin, "Traveling the Silk Road: A measurement analysis of a large anonymous online marketplace," in *Proceedings of the 22nd international conference on World Wide Web*, 2013.

[19] M. Power, *Life after Silk Road: how the darknet drugs market is booming | Technology | The Guardian,* 2014.

[20] P. H. O'Neil, *Enter the Armory, the Dark Web's Walmart of weaponry,* 2014.

[21] P. H. O'Neill, *How an alleged Dark Net drug dealer got a 'second chance' from a judge,* 2014.

[22] D. M. Reporter, *The disturbing world of the Deep Web, where contract killers and drug dealers ply their trade on the internet | Daily Mail Online,* 2013.

[23] M. Bouchard, K. Joffres and R. Frank, "Preliminary analytical considerations in designing a terrorism and extremism online network extractor," in *Computational Models of Complex Systems*, Springer, 2014, pp. 171-184.

[24] S. Romanosky and C. Kuo, "FoxTor: Anonymous web browsing," *Tor GUI Competition,* 2006.

[25] C. Olston and M. Najork, "Web crawling," *Foundations and Trends in Information Retrieval,* vol. 4, no. 3, pp. 175-246, 2010.

[26] R. Albert and A.-L. Barabasi, "Statistical mechanics of complex networks," *Reviews of modern physics,* vol. 74, no. 1, p. 47, 2002.

[27] P. J. Carrington, J. Scott and S. Wasserman, Models and methods in social network analysis, vol. 28, Cambridge university press, 2005.

[28] A. Malm and G. Bichler, "Networks of collaborating criminals: Assessing the structural vulnerability of drug markets," *Journal of Research in Crime and Delinquency,* vol. 48, no. 2, pp. 271-297, 2011.

[29] Deep.Dot.Web, *Ramp (Russian Anonymous Marketplace) | Deep Dot Web,* 10.

[30] R. Dingledine, N. Mathewson and P. Syverson, "Tor: The second-generation onion router," 2004.

[31] P. Killicoat, Weaponomics: the global market for assault rifles, vol. 4202, World Bank Publications, 2007.