A Prototype Design for a Framework to Analyse the Traffic Flow in Darknet

K.Sowjanya Naidu, L.S.Chakravarthy

Abstract— The malicious activities in the darknet are an emerging threat to the cyber space. Darknet sites operate using TOR(The Onion Router) hidden services which provides the feature of disguising the users of the transaction in the darknet market place. Hence identifying and monitoring such illegal activities in the marketplace has become a tedious task for the cyber and law enforcement officials. This paper presents a prototype for a framework which analyse the traffic flow in the darknet as finding the exact sender and receiver is almost impossible as the TOR is increasing the layers of security to the maximum extent making it impossible to track the users in the transactions. Here we give a methodology using webcrawlers and extract the data from the darknet sites to find the domain of the traffic flow through which the broad area of traffic can be sorted out which would be beneficial for the cyber and law enforcement agencies to find the illicit trade in the darknet market places.

Keywords— Darknet, TOR, webcrawlers, marketplaces.

INTRODUCTION:

The internet we are searching called the World Wide Web(WWW) consists of 4-10% of the whole internet and the remaining 90-96% has the content which is not catalogued or indexed[1].

Internet can be broadly divided into Clear Web/ Surface Web and Deep Web as shown in Fig.1. The Clear Web or Surface Web mainly consists of Web content or web pages which can be accessed using the traditional search engines like google and yahoo or any other standard browser with a hassle free access. The content in the deepweb cannot be accessed by the search engines such as google etc. DeepWeb essentially consists of the databases and the credentials of corporate firms which are intentionally hidden due to many security reasons.

Embedded within the deepweb is the Darknet or Darkweb which consists of professional hackers who intentionally break the networks and create havoc, commit extortion, Paedophiles, arms, drugs, Human trafficking etc. Darknet sites are hosted with Domain Name System(DNS) root such as .BIT domains which are not controlled or managed by Internet Corporation for Assigned Names and Numbers ICANN and such sites hosted on limited-access network infrastructure requires special software - TOR to access it[2] . In a research conducted by the Kings College London, 57% of the darknet sites indulges in criminal activities including drugs, extreme pornography, finance etc. The darknet lacks ethics and trade is anonymous, unknown buyers and sellers transact with each other with the help of Bitcoin or other Crypto currency which provide high security to the transaction[3].

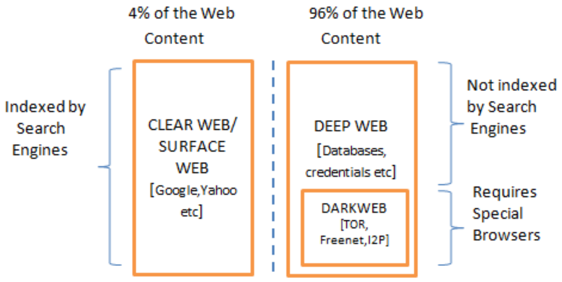
4% of the Web Content
96% of the Web Content
Indexed by Search Engines
CLEAR WEB/ SURFACE WEB [Google,Yahoo etc]
DEEP WEB [Databases, credentials etc]
DARKWEB [TOR, Freenet,I2P]
Not indexed by Search Engines
Requires Special Browsers

Figure 1: Layers of Internet

Individuals can access the dark web by using the special software such as TOR which can be called as "The Onion Router" in which the sender and receiver are kept anonymous. TOR is an open source software that assured the anonymity by camouflaging the actual IP address of sender and receiver. Initially , the US Naval Research Lab in 1990's developed this Onion Routing , which is the basis for TOR as a way to protect naval communication so that an enemy cannot trace computer messages and detect the ships location[4].

TOR disguises the actual IP and makes the transaction complete there by providing high anonymity for the people involved in the transaction. Because of the high anonymity provided by the TOR browser, it is getting difficult for the cyber and digital forensic officials to locate the genesis of the traffic or location of the person in the Darknet. The term Darknet became popular world wide with the shut down of the SilkRoad a popular darknet site by the Federal Bureau of Investigation(FBI) in October,2013. SilkRoad is an online site for selling drugs, fake passports, drivers licences and other illegal service providers like forgers, hackers etc, SilkRoad used this underground computer network TOR that sends the computer messages through at least 3 servers to disguise the users. As the concept of TOR is well known to US officials, the SilkRoad is brought down in Oct, 2013 and its CEO Ross Ulbricht was imprisoned[5].

Hence due to many reasons, the Darkweb is getting harder to crack. The layers of the Onion router indicate the security it is providing and TOR is adding additional layers to make it more sophisticated to crack the network. The security of the TOR is being increased to a level where the darknet sites cannot be directly seen and can be accessed only by invitation. The crypto currency Bitcoin, which was reportedly being used by the darknet sites for transactions is

Revised Manuscript Received on April 12, 2019.

K.Sowjanya Naidu, Faculty, Department of Computer Science and Engineering, GVP College of Engineering(A), Visakhapatnam, AP, India. (E-mail: sowjanya.k31@gmail.com)

Dr.L.S.Chakravarthy, Faculty, Department of Computer Science and Engineering, GVP College of Engineering(A), Visakhapatnam, AP, India. (E-mail: chakri.ls@gmail.com)

being replaced by a new crypto currency “Monero”, which offers stealth mechanisms and prevents the tracing of transactions which was observed to be a vulnerability of Bitcoin[6]. Due to these additional security reasons the people who are using darknet are exponentially increased and posing a serious threat to the digital forensic officials.

Darknet is a pivot of botnet activities in addition to malicious activities like drugs, arms etc. A bot is considered as one of the most complicated tool of cyber crimes. A bot net is a large number of infected machines spread across various parts of the globe. Botnets are installed on victim machines either knowingly or unknowingly. The central hacker controls all the other botnets called the “bot master”. The victim will not even know that his/her machine is infected by bot. Bots are installed through spam emails, spyware, phishing personal information, DDoS, ads etc. Bots use the data retrieved from victims system and sends to the Botmasters who inturn use the information for various purposes ranging from personal to financial gains. Botnet masters use darknets and TOR browser as it provides high security and guises the people in transaction.

The main advantage of the Darknet is the increased security and privacy of the transactions. Due to the increased security and anonymity of the transactions, the dark net is being used by the people who want to avoid internet surveillance for illegitimate and illegal purposes[7]. The increased anonymity of the user is leading to more transactions in the secret market places which include drugs, arms etc which effects the economic growth of the country and hence it has to be kept in check by the digital and forensic officials.

As the other side of the coin, there exists disadvantages too. The darknet is getting more darker which indicates that the security is being increased to guise the users in the darknet. Many number of crypto currencies are coming up in competence with Bitcoin as anonymous payment system. The darknet customer base is increasing as the security of the transactions is increasing. As a result of this increased customer base, the darknet marketplace is expanding exponentially to meet the customer demands and needs. The private delivery services which ensure untraceable, secure and anonymous delivery is cropping up. Hence despite of the disadvantages there is an increased need for the framework which analyses the network traffic flow.

In this paper, we propose a framework for analysing the traffic of the data flow in the darknet sites which can be useful for the cyber and digital forensic officials for tracing the domain of the traffic flow.

METHODOLOGY & RESULTS

The goal of the research is to analyse how the TOR layers influence the darknet traffic flow based on the data and how to reduce the processing complexity and thereby simplifying the data analysis. TOR provides the privacy and the anonymity to the sender and the receiver while transacting in the dark web pages. The ability of the user to search the dark web for illegal means has become easier and many resources are available to find the dark websites of interest. As the users of darknet increases, the data grows in velocity and volume , we need more flexible , efficient and scalable frame work for analysis of data traffic in the dark net.

Hence, in this regard we here by propose a framework to analyse and research on unsolicited information and it is designed in two phases, and they are:

Phase1:

The phase 1 consists of installing the Tor browser and connecting to the darknet websites. Then the scripts are written in high level language like python and scripts are enabled and webcrawlers are made to run on different darkweb sites and information is gathered. The information is gathered from the dark sites and fed to the databases and databases are populated for the further processing of data.

From the figure below the internet is connected to the router and the router is connected to the Darknet router. The dashed arrow consists of the darknet traffic flow and. The darknet router is installed with TOR and is used to access the darkweb sites. The scripts are run on different darknet sites webcrawlers are enabled which extract the data and the data is fed to the databases. The databases which are populated with the data is processed for the analysis in Phase 2.

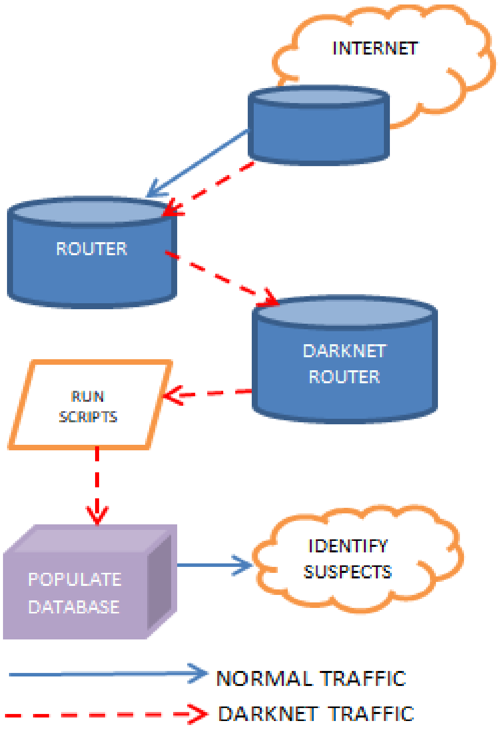

Figure 2:Darknet Traffic Flow

Phase 2:

This phase consists of the analysis of the data which is gathered from the darknet sites. Figure 2 shows the procedure of the proposed in-depth analysis method of how the data is collected and analysed in the dark net traffic. The procedure is composed of six main steps: collection, extraction, analysis, classification, identification and tracing. The each phase of the procedure is as follows:

1. Collection: During the first step, the data is collected from running the scripts on the dark websites and the data regarding the traffic is retrieved.
2. Extraction: The data which is collected from running the scripts is extracted and populated in the data bases.
3. Analysis: In this step the data which is populated in the data bases is fed to the traffic analysing tools like maltego to find the patterns in the traffic flow.
4. Classification: In this step the data patterns which are collected from the maltego analyser tools is analysed based on the traffic flow of the data.
5. Identification: Here, based on the patterns of the data collected after the processing of data, the traffic flow is calculated based on the patterns identified from the previous step.
6. Tracing: The final step comprises of the finding the patterns of maximum traffic flow thus giving data to the forensic officials which may be considered for further processing.

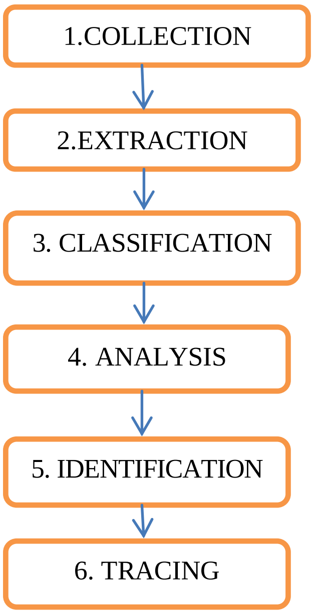

Figure 2: Phase2: Analysis of Data

The aim of this research was to empirically demonstrate the efficacy of the aforementioned steps of the analytical framework given in the above figure, with the goal of making the patterns of the data complexity in the dark net traffic flow to reduce and simplify the data analysis which can be used for identifying the dark net traffic domains.

RELATED WORK:

The aforementioned problems studied in this work have been explored in the literature to an extent. In the section, we will describe the related work done and existing techniques in the darknet. Darknet is an area of enormous research based of different fields.In [8] the authors focuss on police crackdown on cryptomarkets and the Opearation Onymous. Operation Onymous targeted many cryptomarkets and many dealers retired post operation from the crypto-markets. It focusses on pricing before and after the operation. The authors of [9] discuss about the market place Agora, and analysis of data is done with an algorithm developed in combination of Feature extraction and SVM for classification and compared with existing models. Author compares the data collected using web crawlers from Silkroad 2.0 and Agora[10]. Analysis is done comparing the price ranges and finally concludes that the drug trafficking is prevalent in market places. The author [11] in paper compares the market places Evolution and Silk Road 2.0 and analyses trade. In [12], focuses on types of theft and Australian Virtual Markets. It is concluded that Australian sellers fix more prices when compared to other drug trafficking sellers. This paper [13] discusses on the transactions that take place in darknet which are not genuine. The fraud that occurs in cryptomarkets with the data collected from various sources. It deals with how cryptomarkets work and how to improve crime prevention efforts. Buying in cryptomarkets mainly deal with drugs. The purity of the drug also plays a vital role. The purity of the drugs is concentrated and analysed. Author identification is the ultimate goal in the dark web market transactions. Due the increased security author identification is almost impossible but there are several ways to go the nearest point. One such method for identification is Stylometry[14]. The information is gathered from the individuals who maintain multiple accounts. By gathering and linking the information available from the multiple accounts the details can be known and can assist in investigations in exposing the anonymous identities. The increased security in the darkweb is due to the TOR layers. The TOR layers provide security by disguising the users of darknet. Hence the darkweb is a large repository of data, categorization of the data is needed which can be done by ATOL (Automated Tool for Onion Labelling) which is an assessment for the content in large scale repositories[15]. A study is also being conducted on different patterns in the trafficking of the products based on the shipping country in illicit markets[16]. The transactions in the secret marketplace is being carried out in Bitcoin crypto currency. This paper [17] specifies that the black market is being transformed into black e- commerce as the bitcoin activity is reduced as more opaque currency comes into markets with provide more stealth mechanisms to guise the users in the transaction. As the crime increases in the darkweb, the criminals interact with each other through forums[18]. Hence analysing this patterns enables us to have a better understanding of the behavioural patterns of the people involved in darknet.

CONCLUSION AND FUTURE WORK:

Darknet is comprised of huge amounts of confidential data. Given the scarcity of data of the individuals in the darknet, our paper gives a framework for analysing the traffic flow in the data transaction in the darknet, by running scripts and enabling webcrawlers for extracting the data and feeding the data to the data analysis tools for analysis of data. It gives the domain in which the transactions are more prevelant in the darknet transactions which would be beneficial for the cyber and digital forensic officials.

The paper we present here gives a prototype for analysing the traffic flow. It can be extended empirically and results can be established which in turn can be useful for cyber and digital forensic officials. As the darknet is getting more darker there is a need for the webcrawling algorithms. The webcrawlers used in traversing the darknet can be implemented further using webcrawling algorithms either by implementing the crawling methods or keyword selection methods.

REFERENCES:

1. Shining Light on the Dark Web ,George Hurlburt, STEMCorp, Cyber Trust
2. https://www.icann.org/news/blog/the-dark-web-the-land-of-hidden-services
3. Dark Web by Kristin Finklea, Specialist in Domestic security,Congressional research service
4. Personal use, social supply or redistribution? cryptomarket demand on Silk Road 2 and Agora, Jakob Demant & Rasmus Munksgaard &Esben Houborg
5. https://www.informationsecuritybuzz.com/news/secret-history-tor/
6. A. Greenberg, “Monero, the Drug Dealer’s Cryptocurrency of Choice, Is on Fire,” Wired, 25 Jan. 2017; wwwwired.com/2017/01/monero-drug-dealers-cryptocurrency-choice-fire
7. Complex Network Analysis of Darknet Black Market Forum Structure, toms reksna, Leiden, The Netherlands
8. Décary-Hétu & Giommoni 2016, Do police crackdowns disrupt drug cryptomarkets? A longitudinal analysis of the effects of Operation Onymous
9. Graczyk & Kinningham 2015, Automatic Product Categorization for Anonymous Marketplaces
10. Demant, Munksgaard, & Houborg 2016, Personal use, social supply or redistribution? cryptomarket demand on Silk Road 2 and Agora
11. Broséus et al 2017b, Forensic drug intelligence and the rise of cryptomarkets. Part I: Studying the Australian virtual market
12. Moeller et al 2017, Flow My FE the Vendor Said: Exploring Violent and Fraudulent Resource Exchanges on Cryptomarkets for Illicit Drugs
13. Damien et al 2016, Buying drugs on a Darknet market: a better deal? Studying the online illicit drug market through the analysis of digital, physical and chemical data
14. Ho & Ng 2016, Application of Stylometry to Dark Web Forum User Identification
15. Ghosh et al 2017, Automated Categorization of Onion Sites for Analyzing the Darkweb Ecosystem
16. Reksna et al 2017, Complex Network Analysis of Darknet Black Market Forum Structure
17. Foley et al 2018, Sex, Drugs, and Bitcoin: How Much Illegal Activity Is Financed Through Cryptocurrencies?
18. Pastrana et al 2018, CrimeBB: Enabling Cybercrime Research on Underground Forums at Scale