Invisible Web

Sumeet Raghunath Koyande, Yache Vishal Shivraj Reddy, Mayuri Dendge

Bharati Vidyapeeth Institute of Management and Information Technology, CBD Belapur,
Navi Mumbai, Maharashtra, India

ABSTRACT

In this research paper our topic is to analyze and dig deeper into deep web and what can we find when we go dive deep into the internet world.

Web means internet as we all know, today almost everyone usesthe internet. We use internet for sharing information or to get information. The web which we use is called as Surface web, which is huge. This surface web contains billions of websites which are indexed on the goggle or other web search engines. Search engines can easily search for this indexed website or even the data on that website with the help of crawlers and modern search algorithm.

Another part of the web is the deep web, which is almost 95% of total web i.e. although the surface web is huge but it's only 5% of the entire web. Deep web websites are not visible to everyone, because these websites are not indexed on any search engine. To search these links, we need a TOR browser, a special browser which hides the IP address of the computer by bouncing the connection around a distributed network [1].

The deep web also contains large databases such as iCloud, Google's databases So that this information should not visible to anyone from the surface web. The deep web provides a good amount of security for such systems because this links only be accessible for specific users and with specific credential and specific address. All the data we store on the cloud is stored on physical location but the only way to access this data is thru deep web links only. These links and databases controlled by the companies which provide cloud services. Because of this all the data is stored on servers also and not visible to anyone and users can

only access their data with their login credential and data of other users remain safe [3] [4].

I. INTRODUCTION

The deep web is the hidden part of the World Wide Web who contains is not indexed on any search engine. It is used for mail services and online banking which need another level of protection.

We can simply imagine deep web like an iceberg in the ocean. Only the 10 percent of the iceberg is on the surface of the water and the rest of the 90 percent of the iceberg is under the water, the same way the surface web or visible web is only 4% which we can search or access. And rest of the 96% web is occupied by the deep web which we don't have access.

This deep web contains the information and databases of the companies which they want to hide from normal users for security purpose. The deep web is not only used by the companies but also used by the individuals like journalists to share the secret information.

To access the deep web, we need a browser called as TOR browser, which jumps the IP from one location to another location to ensure the anonymity on the web. The study says that accessing the deep web is not illegal but hiding the identity makes it illegal to use the deep web.

Dark web is the part of the deep web where all the criminal activities happen. This is the place full of cyber criminals and peoples who do any illegal

activities. There isa dark market in dark web where we can find all illegal materials which we cannot purchase anywhere like drugs, guns from the dark market website like Silk Road, etc.

Use of the dark web is not illegal, but the activities happening on the dark net is illegal. It is because anonymity needs to maintain while accessing deep web.

Fig.1.1 Surface Web vs Deep Web

II. SURFACE WEB:
Regardless of what we search on the internet using our normal browser, this all data belongs to the surface web. All the activities we do on the surface web are easily traceable by any authorized organization or by any individual who is having knowledge about computer networking.

This is the part of the web which is accessible and visible to us using the common browser and search engine. The surface web has become the part of the web at the time the first web browser is created.

Surface web is the part of the web whose contents are indexed on the web search engines like goggle, yahoo, etc. which is visible for everyone to access. On search engine giant goggles website it has mentioned that there are 130 Trillian individual web pages are indexed on the goggle search engine, it is a huge number but still its only 4% of the total web [1].
➢ On average almost 30,000 websites get hacked every day.

III. DEEP WEB:
The deep web is another part of the web it is also called as invisible web, we can assume it as a part of the iceberg under the water. Contains dynamic web pages which are not indexed by common search engines like Google. The name deep web is given because such content is beyond the reach of our regular search engines. The deep web is a huge part of the total World Wide Web. Its huge in size and can't be measured because the majority of the content on the deep web is locked.

Deep web contains a big amount of structured data on the web, accessing such a large amount of data is not an easy task for data analysts. Websites in a deep web are also contained HTML documents with hidden code attached behind to it. The anonymity which deep web offers is far superior than surface web, in comparison anonymity offered by the surface web is not even comes close to the anonymity offered by the deep web.

IV. WHY IT IS IMPORTANT
➢ It is so massive that, 96% of the total web is situated in the deep web.
➢ It is estimated that deep web is almost 400 to 500 times larger than the surface webbecause no can actually measure the size of the deep web as the majority of the content is not accessible[3].

Also, there is a huge number of websites yet to discover on the deep web, because nobody has access to this website without secret links or invitations. Big databases of companies are stored on the deep web, also the cloud services or the data which we store on the cloud uses the deep web for data storage for security purpose because no one can get access to these data without special links. All the cloud services like Gmail, iCloud, Amazon use the deep web link.

V. DARK WEB:
This is the wildest area on the internet where there are no rules or restrictions are on any user. This is facilitated by the hackers and computer users from the entire world who thought there should be no limitation, law enforcement agencies or supervision on what we search or what people do on the internet.

Dark web is the deepest part of the deep web where all the criminal activities happen. peoples often get confused with dark web and deep, where the dark web is the part of the deep web. These dark net websites

are only accessible with the tor browser only and these sites use .onion domain, .onion domain mainly focused on providing anonymous access to the internet. The anonymity of the dark web users stays anonymous because of an extra layer of encryption.

VI. DARK WEB SERVICES
➤ Silk road is the most famous online market where we can purchase almost anything which we can imagine from drugs to guns.
➤ We can also hire a hitman to kill real humans, the prices are different for normal persons and govt. officials or politician all the way from 20,000 euro to 1,00,000 euro.
➤ Anybody can purchase guns without a license.

Dark web mainly used for illegal activities such as black market (dark net market), hacking, child pornography, frauds, red rooms, etc. The websites in the dark web are mainly broken, inaccessible or abandoned using standard web browsing techniques. The term inaccessible is only used for our understanding, these websites are only accessible to those who know what they are looking for and also these websites strive to be private.

The only this which is more appealing to the users is anonymity associated with the dark web, and many people support the dark web for anonymity only. Anonymity is also the major point of argument, criminal activity happens on the dark web because of this anonymity factor only.

VII. TOR BROWSER:
Tor is a specialized browser used to access the deep web, the tor can be described as a software which helps you to bypass all the surveillance and the monitoring system. In technical terms, tor is an onion routing system, which is a technique for anonymous communication over the network. Once the browser establishesa connection to the tor network, then we are ready to dive into the dark web. All the data goes from tor is repeatedly encrypted and sent through several other onion routers, every router through which data is passing puts one more layer of encryption because of which anonymity gets maintained. The tor or onion routing was originally developed by the U.S. Naval Research Laboratory, where the primary purpose of protecting the government communication [5].

VIII. MAIN HEADLIGHTS OF TOR
➤ Provides complete security by connecting to other tor nodes while connecting to the internet or searching for the data on the internet or deep web.
➤ Continuously keeps on jumping the IP address from one point to another so that nobody will be able to identify the current location of the user. This provides the complete anonymity on the internet.
➤ The anonymity and the encryption maintain by the tor is exceptional as compared to the regular browsers, which does not provide any kind of secure channel other than secure HTML channel.

We can surf dark net sites and regular websites with TOR. Accessing the dark net is not illegal because it contains verity of information and different services, although there are some services and activities on the dark web is illegal.

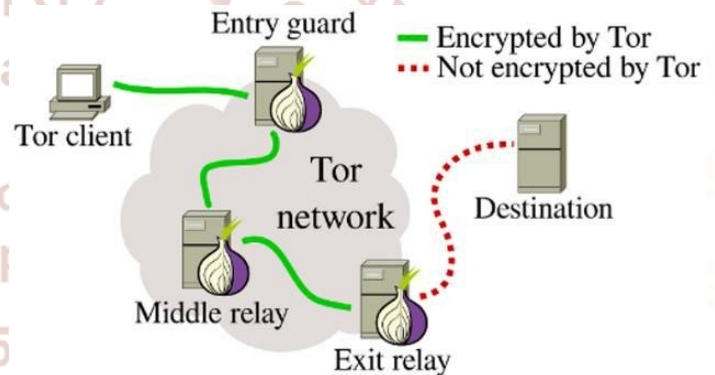Also using tor browser is not illegal because it's just the tool.

Fig.1.2 Tor network

IX. CRYPTO CURRENCY:
Crypto currencies like bit coin are started getting everyone's attention nowadays, because of its rapid growth. These currencies are decentralized currencies, which mean there is no central authority to have control of these currencies which leads to extreme fluctuation and unpredictable behavior. These currencies work through block chain system which is a distributed ledger and public transaction database so anyone can see all the transaction which are happening worldwide. All the transactions are visible to the public, still who is the sender and to whom he is sending the money.

The Mining Algorithm is as follows:
➤ Step 0 - Retrieve the hash of the previous block from the network.

➤ Step 1 - Gather a list of potential transactions known as a "block". This list of transactions comes from the peer-to-peer bitcoin network.

➤ Step 2 - Calculate a hash for a block of potential transactions along with a random number.

➤ Step 3 - If the hash is more than the currently set difficulty level, then you have mined that block. If not, start over from Step 1. Any additions to the list of transactions from step 1 along with a change in the random number from Step 2 mean that there's a chance that the criterion will be met in the next go around [6].

There is no mechanism to trace these transactions because of which the transactions made with crypto currencies are not taxable which is a huge problem for all the governments and central banks in the world. These currencies are very important for deep web transaction because all transactions on the dark web are done with crypto-currencies only. These currencies are used not only because anonymity, but these currencies are also decentralized and provide very secure transaction than regular payment gateways.

The only way to get bit coin is thru mining or purchasing the bit coin in a change of real money. Bit coin uses SHA-256 hashing algorithm for encryption.

X. CONCLUSION:

Surfing or using deep web or dark web is not illegal by any means, usage of tor browser is also not illegal. But to access deep web we need anonymity which only tor browser can provide and hide our identity while we are online is illegal by laws and all these things makes assessing deep web illegal in some countries. Also, the services which dark web provides are all illegal, that's why these services are not available on the surface web so it is also illegal to use these services.

REFERENCES

1. https://www.google.com/search/howsearchworks/crawling-indexing/

2. http://www.sciencefocus.com

3. http://www.sciencefocus.com/qa/how-many-terabytes-data-are-internet

4. http://www.worldwidewebsize.com/

5. https://davidenewmedia.wordpress.com/accessing-the-dark-web/tor/

6. https://bitcoin.stackexchange.com/questions/12603/the-bitcoin-mining-algorithm-from-a-programmers-viewpoint