DARK WEB: THE DARK SIDE OF INTERNET TO SUPPORT SECURITY AND PRIVACY

Raja Mondal[1], Anirban Bhar[2], Soumya Bhattacharyya[3]

[1] *B. Tech student, Department of Information Technology, Narula Institute of Technology, Kolkata, India.*
[2,3] *Assistant Professor, Department of Information Technology, Narula Institute of Technology, Kolkata, India.*

ABSTRACT

A large architecture of various computer networks makes up the Internet as a whole. Open websites that use search engines like Google, Firefox, etc. make up the network. The "Deep Web" and the "Surface Web" are the two main Internet subsets. All websites that are indexed and accessible to the general public are considered to be part of the surface web, whereas websites that are not indexed are considered to be part of the deep web. The Dark web, however, is the most dangerous aspect of the Deep web. The World Wide Web's hidden content is what is known as the "Dark Web." The Deep Web includes the Dark Web. The Onion Router, or TOR, can be used to access it. Dark Web actors are hidden and anonymous. Specialized browsers like TOR and I2P offer three considerations: privacy, anonymity, and the likelihood of going undetected. We'll examine and draw conclusions about the impact of the Dark Web on several spheres of society in this post. For a while, the number of typical anonymous Dark Web users (using TOR) in Kosovo and across the globe is given. Search results from Ahimia and Onion City Dark Web are used to determine the impact of hidden resource websites. On the Dark Internet, anonymity cannot be guaranteed. TOR has committed to it and plans to execute covert operations.

Keyword: - *Dark Web, TOR, Deep web, Encryption, Privacy, Anonymity, I2P, Computer Network.*

1. INTRODUCTION

Everyone uses the internet on a daily basis. Surface Web and Deep Web are the two components that make up the internet. Websites like Amazon, Wikipedia, Facebook, YouTube, and others are already well-known to us. These websites fall within the category of the Surface Web, which search engines like Google, Bing, Yahoo, etc. can index. The "Visible Web" or Surface Web is well-known. But the internet is only one aspect of this. The 4% of the Internet's surface that is accessible to the general public is known as the Surface web. There is a tone of information on the surface web, and it is legal. Some websites are hidden and inaccessible to the general population, who makes up 96% of the Internet, far away from these websites. The Deep Web exists in this peculiar place. Since authentication is required to access data like private accounts or online banking, Users have access to confidential information on websites on the "Hidden Web". If these websites are indexed, then anyone can obtain the material by searching for your name, and the public will be made aware of your personal information. To protect privacy, pages that have been validated are not indexed. With Deep Web verification, one can visit these domains. A minor portion of the Deep Web is the Dark Web. "Invisible Web" is another name for the dark web. The public is unable to access the dark web. About 45,000 Dark web sites, or 0.01% of the Internet, exist.

The World Wide Web's (WWW) hidden content that lives in the Deep Web is known as the "Dark Web." Although it uses the Internet, accessing it involves additional software, authorization, setups, etc. Dark Web can only be accessed via specialized anonymous software installed in browsers like TOR (The Onion Router), Subgraph, Waterfox, I2P (Invisible Internet Project), etc. It ensures confidentiality and these are not indexed by the web search engines. It's a representation of the Surface Web. Both excellent and bad entertainers may operate in secrecy because to the encrypted internet. For the buyers and business, it is a fact. The simplest way to navigate the dark web is through TOR. It can be applied to both lawful and illicit operations. Legal uses of the dark web include the freedom of speech for journalists and informants, the

global flow of data with privacy, and other uses that are prohibited, such as trafficking in illegal substances, using stolen credit card information for gambling, etc.

A branch of the Surface web serves as the Dark web. It operates similarly to the Clearnet. Its websites are comparable to those of the Surface Web; for instance, it features commercial websites that resemble Amazon or Flipkart and offer the ability to buy or sell things, as well as Wikileaks, which is identical to the Wikipedia on the Surface Web. It uses the TCP/IP protocol standard to transport traffic (HTTP and FTP) within and across networks. HTML pages and their advantages make up its content. The dark web is notorious for unlawful activity and is seen as a dangerous location where the sale of illegal goods like drugs occurs. If privacy is the main concern, it might be a very secure way to browse the internet.

The remaining sections of the essay are structured as follows. The distinctions between the deep web and the dark web are discussed in Section 2. The Onion Router's overview is presented in Section 3. Sections 4 and 5 respectively explain how TOR and the Dark Web operate. Section 6 includes a list of the functions and uses of TOR and the Dark Web, followed by requirements for maintaining anonymity. The paper is finally ended along with a few predictions for the future of research.

2. LITERATURE REVIEW

With the help of the Dark Web, it was easier to share weapons and spread child pornography. The TOR network is used to provide network information, and customers can easily afford to use an encrypted way of anonymity. As a result, innumerable literary works enable the advancement of study, and as a result, the TOR routing with the other concepts is provided with the aid of the various US intelligence systems. 2007 (Navara & Nelson) It makes it possible for the Dark Network system to be utilized for both unlawful and legal purposes. The privacy of the programme is conveniently illustrated for the purpose of analyzing the data, and study is also continued with the help of the ISI testing frameworks. This is done with an adequate review of the network trackers. The behavior of the literature review is based on a careful examination of the various facets of the Dark Web, which are satisfactorily explained. The work also assists to explain the key components of the researcher's study [1].

In a different study by Barnett et al., the role of spiders—defined as computer programmes used to browse information on the World Wide Web—and the simplicity of access that can be gained through the registration process are studied, making it possible to easily gather the precise and necessary information on the various forms. Interest in social network analysis (SNA) has led to the development of graph-based methods that allow for the evaluation of network structure through the representation of population power. 2007 (Navara & Nelson) Using social networks to illustrate the impact of social ties makes it simpler for real-world networks to do so. For the analysis of forum posting and website connections, specific SNA techniques have been created.

Understanding "remote networks" and their unique properties is the main goal. To identify militant websites and content supporting terrorism, intricate coding algorithms have been created [2].

The identification of violent and extremist websites that pose serious concerns is made possible by sentiment and impact analysis. Terrorism In order to save, incorporate, handle, and understand the variety of intelligence relevant to terrorism for international / national security goals, informatics is defined as the application of specialist knowledge processing, research methods, and methodologies. The methodology draws from a variety of disciplines, including computer science, mathematics, astronomy, economics, and social sciences.

The term "anonymity" in the Dark Web comes from the Greek word "anonymia," which describes concealing one's identity to others. Our digital footprints are stored online as data if we perform any action on the site. We can infer that anonymity is guaranteed if the Internet Protocol address cannot be registered. The TOR client distributes Internet traffic throughout the globe using volunteer server networks [3].

This makes it simpler to conceal information from customers and reduce the possibility of tracking their activity. Dark Web also has negative effects because it motivates criminals to engage in cybercrime and cover their tracks. T is said to be a useful tool for governments to exchange sensitive information, for journalists to get around censorship, and for activists to "shelter" from oppressive regimes. A network of computers may communicate securely thanks to onion technology1. Asymmetric encryption is used to encrypt and deliver messages to all network nodes. 2014; Jonason et al. We have several usage examples in this part that are pertinent to the data we gathered and extracted from our software. The first analysis of the information obtained over the previous two years looks at the distribution of languages across all active Deep Web sites.

Two distinct techniques are used to identify the language, including a Python programme called guess language that use an offline trigram-based approach. Google Translation is (a); (b). In order to address each system's weaknesses, specific findings are contrasted. For instance, Google Translate has a huge bias in the data since it has no concept of "hidden

language" (for example, where there is no data on a page). Instead, it falls back on English in ambiguous situations. The importance of the language is displayed in the following table as a percentage of [4] Domains with pages in that language. We eliminated all pages identified as "unknown" and pages smaller than 1 kb from the statistics computation because they lacked sufficient data for accurate detection.

The goods and services we discover on the Deep Web pretty well represent the kinds of purchases people wish to make provided their anonymity is guaranteed. While the lack of formal identification puts them in grave risk, it also gives them a hazy sense of security that permits them to sell mainly illegal items and services. Some of the things we've seen on the Deep Web frequently had more serious effects on the "real world," as opposed to covert crimes. Except for the fact that the pages selling them really exist, we are unable to guarantee the legitimacy of the products and services featured here [5]. Although we were unable to include all products and services, we did include some of the major groups that would help us get a better sense of the country.

3. STRUCTURAL VIEW POINT

There are three parts that consists of world wide web.

Surface Web:

The surface web, often referred to as the visible web, indexable web, or clearnet, is the first component. It contains content on the internet. Users only have access to the surface web during normal daily activities. It may be found using common search engines and is accessible with common web browsers like Mozilla Firefox, Microsoft's Internet Explorer or Edge, and Google Chrome that don't need any extra configuration. Google, Yahoo, Facebook, YouTube, Wikipedia, regular blogging websites, and essentially anything we can see on any search engine's result page are some instances of surface web.

Deep Web:

Similar to how it expresses itself—below the surface and not as dark—is the deep web. Deep Web lists all the websites that are not indexed, which means that when we conduct a search, the search engines will not return them to us [8]. Because these are personalized for the users and there is no need to index every URL, the Deep Web also includes a particular section of the mainstream web that is legal, such as Amazon, IMDB, and Netflix. Most people utilize the deep web without even realizing it. Because they may be accessed through an application programme interface, some websites, like Facebook and Instagram, are also referred to as being on the Deep Web. The info that can't be found using standard search engines is all present on the deep web. Although its true extent is unknown, it is thought to be 400–500 times larger than the Surface web. Google and other regular search engines cannot access data stored on the Deep Web. Due to data inconsistencies and nominal problems, indexing websites becomes complicated. To access the data, these anonymous websites require login information. These websites have time limits and become unavailable after a specific amount of time. The depths of the internet are a limitless reservoir of knowledge. Deep web refers to all the websites and pages that do not immediately appear when we perform an Internet search. Everything that needs a login is included, such as personal accounts and online banking. Because it protects privacy and personal information, deep web is both extremely safe and crucial. It is essential for day-to-day living. The users are given security and privacy [6].

Dark Web:

Because the Dark Web is a small portion of the Deep Web, it is primarily used for illegal material and activities. The Dark Web is a website protected by TOR that is used for discreet communication [7]. It is located in the internet ravine's bottommost layer. By utilizing an anonymous browser that needs validation and verification, one can visit it. You are dealing with strange servers as you use the dark web. Everything stays inside the TOR network, which similarly provides users with safety and privacy. While all Deep Webs are the Dark Web, not all Dark Webs are Deep Webs.

4. WORKING WITH TOR AS A MEANS OF ENTERING THE DARK WEB

Websites that are located on unindexed or hidden portions of the Internet are referred to as the "dark web." It can only be found on a specific network and cannot be found on a regular web.

The major entry point to the Dark Web is the TOR (The Onion Router) Browser, which is how we can access it. TOR was primarily created for a variety of unique uses. The Naval Research Laboratory in the United States created TOR in the twenty-first century [8]. The purpose of creating TOR was to give American military soldiers fighting overseas secrecy. Tor uses multiple layers of encryption. Its name, Onion Router, refers to its numerous layers, much like an onion. TOR utilizes a multi-layer architecture that makes it possible to identify the current and prior devices that it has passed through, but not the origin and destination devices. Before transferring the data packets to the next device in the chain, each device decrypts the layer of the onion that it is now in. The information stored on the next device is decrypted by the current device. It is aware of the data's route, from where it has received the data to where it transmits the data packet, but it is unaware of the origin and destination of the data packets. This is Onion Routing's primary standard. Data packets cannot be traced as a result without knowing the network's origin or destination. By employing addressing systems made up of keys generated at random and denoted by the address ending with. onion, TOR keeps track of the "hidden sites." The simplest way to navigate Dark Web sites is using TOR. In order to prevent exposing and DDoS (Distributed Denial of Services) attacks, TOR URLs are typically complex and challenging to remember [9, 10]. You can visit the Dark web sites if you have any of these browsers set up on your PC. It is important to note that the majority of evil and tainted dark web sites require passwords and only permit access by invitation.

The "Onion Routing" method is the basis for how TOR works, in which user data is first encrypted before being sent through a number of relays that are part of the TOR network. The user is protected and their identity is concealed via layered encryption. By omitting connections between the user's computer and the target through the network of intermediary relays, it operates. These can be located all over the world and are entirely operated by volunteers who are willing to give up some data transfer capacity for the cause. Since each relay has such a large volume of data transfer to offer, having more relays is beneficial when anonymity and network speed are the primary concerns [11]. The difficulty of tracking a user increases with the number of relays.

This suggests that we have used a multi-layered encryption scheme to cover our unique data, much like an onion. Each relay contains information that it needs to know, including where the data came from and where it goes when it is transmitted.

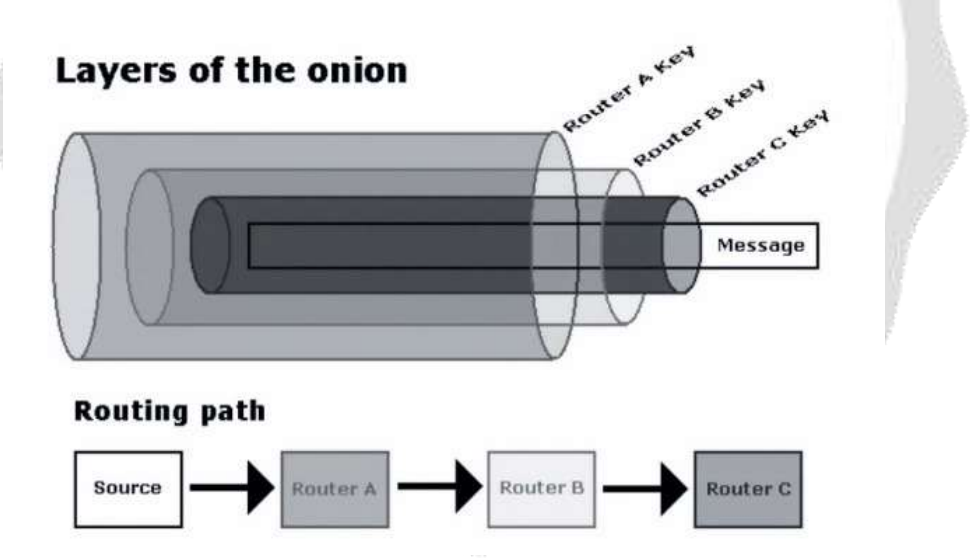

Fig -1: Onion Routing

5. SECURITY ISSUES

Virus: A Virus is a program that is loaded to your machine without your knowledge and runs against your wishes. They are computer programs that bind themselves to or corrupt the machine or files which appear to circulate to other machines on the network by clicking on them, by fax, by mobile devices, etc. They interrupt the operation of the machine and affect the data stored either by changing it or by deleting it entirely. Definition of viruses: (1) Melissa, (2) Sasser, (3) Zeus, (4) Conficker, (5) Stuxnet, (6) Mydoom, (7) Red Code.

Warms: Worms, unlike viruses, do not require a host to cling to. They're only replicating until they've used up all the resources left on the machine. The word "worm" is often used to mean "self-replicating" malware (MALicious softWARE). It has some free memory of drives or other computers. Example of heat: (1) Badtrans, (2) Bagle, (3) Gun, (4) ExploreZip, (5) Kak worm, (6) Netsky, (7) SQL Slamme.

Hacker: A rising hacker is a person who breaks into computers, usually by obtaining access to administrative controls.

White hat hacker: A white hat hacker is a information security expert who hacks into secure systems and networks to check and determine their protection. White hat hackers use their expertise to boost security by revealing bugs to malicious hackers (known as black hat hackers) who can identify and manipulate them. While the methods used are similar, if not equivalent, to those used, Malicious hackers, white hat hackers, have permission to recruit them against the company that recruited them.

Grey Hat Hacker: The word “white hat” or “blue hat” applies to a computer hacker or information security specialist who can often break the law or traditional ethical norms, but who has no criminal intent characteristic of a black hat hacker.

Black Hat Hacker: A black hat hacker is a person with advanced computer skills whose aim is to break or circumvent Internet security. Black hat hackers are also known as crackers or dark-sided hackers. The general opinion is that while hackers are constructing stuff, crackers are smashing things.

6. METHODS USED IN DARK WEB

If the proper safety measures are not implemented, the dark web may be a very deadly place. There are important measures we must follow if we want to learn how to surf the dark web anonymously and safely. But we must remember that things change quickly and that cybercriminals become more skilled every day.

i) Always Use a VPN to Access the Dark Web: Even if we use the Tor browser, anyone with enough time, money, and knowledge may still track our traffic and identify us. In fact, a flaw in the Tor browser that occasionally allowed the leak of genuine IP addresses was discovered in 2018. So, if we wish to utilize Tor anonymously, we can connect to the black web via either a VPN or Tor Bridges, which are Tor nodes that are not publicly indexed. When using a VPN for the dark web, only an encrypted tunnel to a VPN server will be visible to our ISP, not that we are connecting to a Tor node.

ii) Get the official Tor Browser download page: Even though Tor has had security issues in the past, it remains the most widely used method of accessing the black web. Hackers and governmental organizations find the Tor browser to be an intriguing target. To spy on users while they are on the dark web or to breach users before they even access it, fake Tor browser versions have been developed. It should not be surprising to learn that there are many bad actors out there given its dominant position in the industry and the type of stuff we may access when utilizing it. These individuals attempt to spoof the app and trick you into downloading a corrupted version in its place. As a result, we must only ever download the Tor browser from the official website. To find it, go to torproject.org. It costs nothing to download and use the browser.

iii) Exercise Security Caution: The dark web is a favorite hangout for hackers, cybercriminals, virus developers, and other sketchy characters who we really don't want anywhere near our computer.

iv) Know Where You're Going: Getting around on the dark web might be challenging. We won't have the luxury of Google neatly indexing search results for us to browse while accessing the dark web. It can be challenging to find what we're seeking for as a result; it's easy to end up somewhere we don't intend to be. There is a tone of dark web site directories on the dark web itself. It is not advisable to access websites at random. There are several risks on the dark web that we must surely avoid. We may use a few directory sites to route us on the dark web in order to acquire a general feel of where we are. "The Hidden Wiki" is one of the sites that many new users frequently visit.

v) Make All of Your Transactions in Cryptocurrency: The dark web places a high value on anonymity. Everyone uses cryptocurrencies for online payments, which is another example of this. Companies and governments will be aware if we make a purchase using our bank account, credit card, PayPal, or any common payment method. This isn't the case with cryptocurrency because the buyer and seller can stay much more anonymous. We are likely dealing with a scammer, hacker, or spy if we come across someone on the dark web who wants to set up a transaction through a traditional bank.

vi) Shut Everything Down: It is necessary to properly close all open browser windows and any other associated material. Close the entire Tor browser. Close the operating system and restart in our default interface if we used TAILS.

7. CONCLUSIONS

Because of qualities like anonymity, the black web is unable to distinguish between malevolent and legitimate users. The enforcement agencies must meet this issue by putting in place strategies that protect user privacy while also apprehending criminals. Investigating fraudulent websites rather than fraudulent users is an effective way to do this. The Dark website is not illegal or bad to browse, but engaging in criminal activity is wrong. It provides a method of meeting others who share your interests and makes it possible to communicate further. Users' PCs may be infected by deanonymizing software installed by ethical hackers from government organizations. Enforcers pressing charges against the creators of a malicious site can stop the proliferation other sites once one is taken down. Another strategy for the government would try to break Tor, which would entail identifying each and every user. This would probably result in a more substantial type of service being formed, similar to the previous Silk Road trend. Furthermore, it would eliminate tools that genuine people could utilize, like dissidents.

Online users' anonymity is a double-edged sword that makes managing the system much more difficult. As policymakers advance, it is necessary to keep a closer eye on the growth of the Dark Web, and enforcement agencies must make sure they have the resources and legal backing they need to effectively control the Dark web.

8. REFERENCES

[1]. Harrison, J. R., Roberts, D. L., & Hernandez-Castro, J. (2016). Assessing the extent and nature of wildlife trade on the dark web. Conservation Biology.

[2]. Hurlburt, G. (2017). Shining Light on the Dark Web. Computer.

[3]. Jonason, P. K., Lyons, M., Baughman, H. M., & Vernon, P. A. (2014). What a tangled web we weave: The dark triad traits and deception. Personality and Individual Differences.

[4]. Navara, K. J., & Nelson, R. J. (2007). The dark side of light at night: Physiological, epidemiological, and ecological consequences. In Journal of Pineal Research.

[5]. Nilsson, R. H., Larsson, K. H., Taylor, A. F. S., Bengtsson-Palme, J., Jeppesen, T. S., Schigel, D., Kennedy, P., Picard, K., Glöckner, F. O., Tedersoo, L., Saar, I., Kõljalg, U., & Abarenkov, K. (2019). The UNITE database for molecular identification of fungi: Handling dark taxa and parallel taxonomic classifications. Nucleic Acids Research.

[6]. The Deep Web vs. The Dark Web. (2019, December 23). Retrieved from https://www.dictionary.com/e/dark-web/

[7]. Sinha, P., Jha, V. K., Rai, A. K., & Bhushan, B. (2017). Security vulnerabilities, attacks and countermeasures in wireless sensor networks at various layers of OSI reference model: A survey. 2017 International Conference on Signal Processing and Communication (ICSPC).

[8] Dalins, J., Wilson, C., & Carman, M. (2018). Criminal motivation on the dark web: A categorisation model for law enforcement. Digital Investigation, 24, 62–71.

[9] A public policy perspective of the Dark Web. (n.d.). Retrieved from https://www.tandfonline.com/doi/full/10.1080/23738871.2017.1298643

[10] Sharma, M., Tandon, A., Narayan, S., & Bhushan, B. (2017). Classification and analysis of security attacks in WSNs and IEEE 802.15.4 standards : A survey. 2017 3rd International Conference on Advances in Computing,Communication & Automation (ICACCA) (Fall).

[11] Biswas, R., Fidalgo, E., & Alegre, E. (2017). Recognition of service domains on TOR dark net using perceptual hashing and image classification techniques. 8th International Conference on Imaging for Crime Detection and Prevention (ICDP 2017).