

The Deep Web

Radhika Bailurkar, Aniket Goswami

Abstract— Almost the entire world uses the internet these days and yet while browsing we come across short of information. This is due to the sole reason that the world wide web provides just a small portion of information. A lot more information is hidden in the so called Deep Web. The Deep Web has a lot of potential, such as it contains a lot of data useful for various activities. The Deep Web is as Bad as Good, allowing all illegal activities which should not be mentioned. Hence the Deep Web isn't easily accessible. Our general search engines cannot access the data in Deep Web because these engines make use of the indexing technique which the Deep Web does not support. Therefore special search engines such as Tor, Infomine, Deep web tech etc. All though browsing in Deep Web is risky but with a bit of skill and some knowledge, you can illuminate a lot of valuable information.

Index Terms— Surface Web, Deep Web, Dark Web, Ice Berg Concept, Tor , Silk Road.

1 INTRODUCTION

People today use internet for number of reasons like news, entertainment, communication and for several other reasons. This all accounts to only a small part of what we know as the World Wide Web. The part of internet which is easily accessible by regular search engines is called as "Surface Web". On the contrary "Deep Web" cannot be access using these search engines. The concept of Deep Web was introduced by the U.S. Military from the Department of Defense Advanced Research Project Agency(DARPA). Deep web is very massive in size, the traditional search engines provide us only small amount of information compared to Deep Web which is 500 times larger than surface web. The term Deep Web was discovered in 2001 by BrighthPlanet. One of the subset of Deep web is the Darknet which mainly deals with drugs, pornography, weapons, hitman, counterfeit, forgeries and hacking. This section is collectively known as the "Dark Web". However for accessing the Dark Web we require a secure environment in our computer using a suitable VPN. This problem can be overcome by using a Tor browser for searching the Dark Web websites.

2 SIZE OF DEEP WEB

Talking about the web, i.e. information that is accessible from a particular web browser, not the internet itself it is nearly impossible to scale the size of Deep Web. Estimates based on study done at the University of California, Berkeley in 2001 that deep web consists of 7.5 petabytes of information hidden information. All this hidden data can be summed up to what popularly known as the "IceBerg" Concept. This term was coined by Denis Shestakov. As explained in the diagram using traditional search engines we are accessible to 4% of the Surface Web. But by using the Tor browser our range of exploration is greatly enhanced thus allowing us over 90% of information which is not accessible to Surface Web crawlers. On further exploration we arrive at the bottom where the Dark Web resides. Here we find only illegal activities which are not at all involved with deep web stuff.



3 HOW TO ACCESS THE DEEP WEB

Normally all the search engines make use of web crawlers to find the data stored on the World Wide Web. This crawlers or spiders follow the hyperlinks to the domain. But the search engines cannot see through data stored in Deep Web. There are private websites that require login passwords before you can access the contents. Because these crawlers cannot think, they cannot enter passwords on a login screen or keywords in search box. Thus, this databases must be searched separately.

The Invisible Web refers to part of World Wide Web that's not indexed by the search engines. However this doesn't mean that we do not have any source to access it. Thanks to browsers like Tor, I2P and Freenet which give us the freedom to access the Deep Web links. These browsers show us every hidden and protected articles, websites and information regarding your desired search without getting tracked.

Lets get to know a few resources which will be our diving vessel for Deep Web. Some of these are invisible web search engines with specifically indexed information.

4 TOR



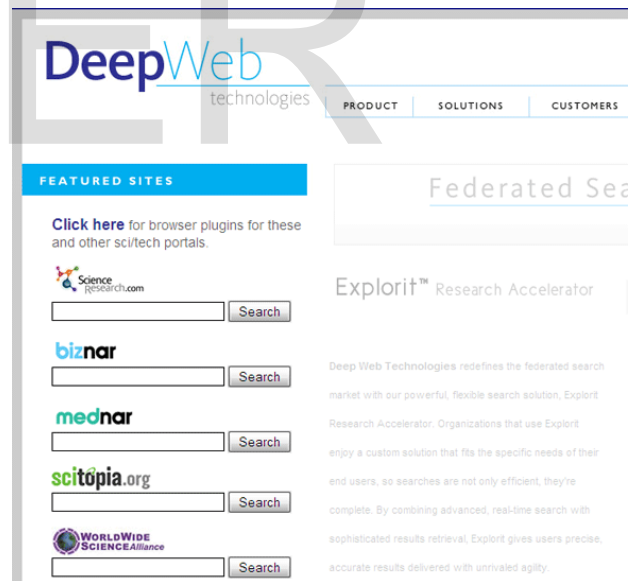
Tor known as the "Onion Router", is used on Windows, mac OS or Linux. Tor is a tool to protect the users privacy. Tor is as a gateway into the Deep web. It aims to encrypt its users identities and their locations from monitoring by using VPN. The VPN is designed to provide security to the users. The most famous sites of Tor is Silk Road sites for selling the drugs which was shut down by the FBI in 2013. With the help of Tor we can see the websites of Deep/Dark web otherwise it is invisible to user

5 INFOMINE



Infomine is a kind of search engine developed many online libraries in the United States. This search engine consists of several categories like databases, electronic journals, electronic books, bulletin boards, mailing lists, online library card catalogues, articles, directories of researches and many other resources. It is primarily used for university searches. Its valuable to use sites such as INFOMINE to search for hidden resources that don't usually appear in traditional search engines.

6 DEEP WEB TECH



Deep web Tech gives you five search engines for specific topic. These search engines cover science, medicine, business. Mednar provides us results generated from over 60 medical collections through federated search. In this type of search, here the user enters a single search query and the engine submits the query to several other search engines. These results are then filtered to create one single authorized result page.

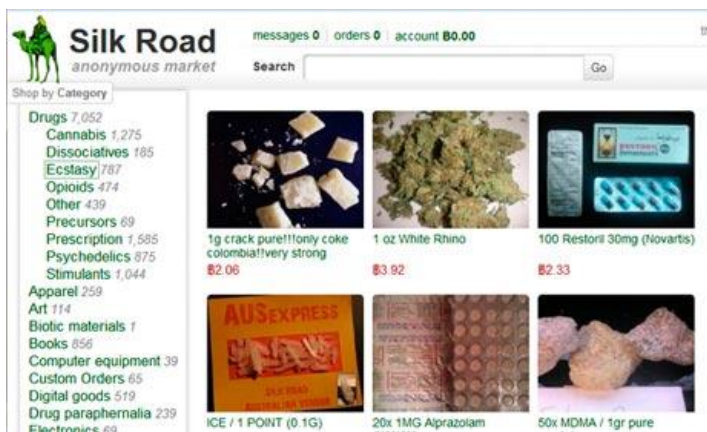
Using this specific search engines, you can query the underlying databases in Deep Web.

7 THE WWW VIRTUAL LIBRARY



Like a regular library instead of books the WWW Virtual Library consists of pages of links to sites about a specific topic. This was started by Tim Berners-Lee, the creator of web. This site is again subdivided into several categories like agriculture, law, computer science and international affairs. It is site where anyone can anonymously edit the content. Since it is run by the volunteers there are many number of links which haven't yet edited since they were last posted

8 THE SILK ROAD



Silk Road was an online black market and the first modern "Dark Net" market, best known as a platform for selling illegal drugs. It was launched on February 2011. Being a part of Deep Web, it was accessed using the Tor browser thus providing anonymity and security to the user. Initially there were a limited number of new seller accounts available which the sellers had to purchase in an auction. Later, the new seller got

an account on a basis of a fixed fee. Items were available through bitcoins.

Soon it emerged to be one of the popular marketplace in the Dark Web leading to its shutdown by the FBI in October 2013 and arrest of its father founder "Ross William Ulbricht" also pseudonymously known as "Dread Pirate Roberts".

A month later on 6 November 2013, Silk Road 2.0 came online, run by former administrators of Silk Road. It too was shutdown the following year by arrest of its alleged operator in the so-called "Operation Onymous".

Ulbricht was convicted of eight charges related to Silk Road in U.S. Federal Court in Manhattan and was sentenced to life prison without possibility of parole.

9 THE BRIGHTER SIDE OF DEEP WEB

The Dark Web is a home to alternate search engines, email services, file storage, file sharing, social media, chat sites, whistle blowing sites as well as sites that provide a safer meeting ground for political dissidents and anyone else who may find themselves on the fringes of the society.

In an age where the government is constantly monitoring the network, the Dark Web offers some relief to people who prize their anonymity, so that to avoid political, economic or social harassment. Secondly the right to express our own thoughts on any topic without fearing persecution is like a boon in certain parts of the world.

For citizens living under oppressive governments, the Dark Web offers more secure way to communicate with like minded individuals. Unlike Facebook or Twitter which can be easily monitored.

Sadly it is also the honey that attracts all the criminal activity which gives deep web such a bad name.

10 THE DARK SIDE OF DEEP WEB

The ability to anonymously access the content makes Deep Web very attractive to criminals. Networks that provide anonymity such as Tor, represent as a valuable instrument for cyber criminals to access sensitive personal information from normally restricted databases, illicit drugs, stolen credit card numbers, human trafficking, weapons, exotic animals and everything else which is banned by the government.

The range of Deep Web searching greatly increased in October 2013 when the FBI shut down the underground site "Silk Road". Most of the transactions on the Deep Web accept BitCoin currency for payments allowing the purchase of any items by preserving the anonymity of the transaction, encouraging the development of trade in respect to any kind of illegal activities.

11 CONCLUSION

Data in the Deep Web is very hard to search. Actually, it is because of the vastness of Deep Web which is nearly impossible to measure. This doesn't mean that the unseen should be ignored. Browsers like Tor help us to establish a connection with the it, thus providing us a secure environment to browse through the web pages without being monitored. Unlike Surface Web, over the years the Deep Web is constantly growing deeper and darker and thus getting more complicated. This can be understood by the IceBerg concept coined by Denis Shestakov. But its ever increasing size is creating a breeding grounds for all sorts of illegal activities like selling drugs, weapons, child porn and other confidential information. But as we know that every dark cloud has a silver lining, the brighter side of it should not be forgotten at all. Its anonymity has been a boon in the countries with oppressive leaders, by giving them the freedom of speech to express their thoughts. Deep Web resides in a different world of information and by having some skill and knowledge we can extract valuable information from it.

REFERENCES

- [1] http://en.wikipedia.org/wiki/Deep_Web
- [2] <http://computer.howstuffworks.com/internet/basics/how-the-deep-web-works.htm>
- [3] www.press.umich.edu/jep/07-01/bergman.htm
- [4] <http://www.makeuseof.com/tag/deep-web-important-think/>
- [5] <https://brightplanet.com/2014/03/clearing-confusion-deep-web-vs-dark-web/>