ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)
Vol 4, Issue 3, March 2017

Deep Web and Dark Web

[1]Chandra Shekar N, [2]Sugali Vijay, [3]M Balachandra Reddy
[1] Assistant Professor, [2] Engineering 1st Year, [3] Engineering 1st Year

Abstract— the Deep Web refers to any Internet content that, for various reasons, can't be or isn't indexed by search engines like Google. Which includes dynamic web pages, blocked sites (like those that ask you to answer a CAPTCHA to access), unlinked sites, private sites (like those that require login credentials), non HTML/-contextual/-scripted content, and limited-access networks

***Index Terms*—deep web, credentials, Private sites, captcha, blocked sites.**

I. INTRODUCTION

Limited-access networks cover all those resources and services that wouldn't be normally accessible with a standard network configuration and so offer interesting possibilities for malicious actors to act partially or totally undetected by law enforcers. These include sites with domain names that have been registered on Domain Name System (DNS) roots that aren't managed by the Internet Corporation for Assigned Names and Numbers (ICANN) and, hence, feature URLs with nonstandard top-level domains (TLDs) that generally require a specific DNS server to properly resolve. Other examples are sites that registered their domain name on a completely different system from the standard DNS, like the BIT domains. These systems not only escape the domain name regulations imposed by the ICANN; the decentralized nature of alternative DNSs also makes it very hard to sinkhole these domains. Also under limited-access networks are darknets or sites hosted on infrastructures that require the use of specific software like TOR to access. Much of the public interest in the Deep Web[1] lies in the activities that happen inside darknets. Unlike other Deep Web content, limited-access networks are not crawled by search engines though not because of technical limitations. In fact, gateway services like tor2web offer a domain that allows users to access content hosted on hidden services. While the popular imagery for the Deep Web is an iceberg, we prefer to compare it to a subterranean mining operation in terms of scale, volatility, and access. If anything above ground is part of the ―searchable Internet,‖ then anything below it is part of the Deep Web—inherently hidden, harder to get to, and not readily visible..

II. USES OF DARK WEB

A smart person buying recreational drugs online wouldn't want to type related keywords into a regular browser. He/She will need to anonymously go online using an infrastructure that will never lead interested parties to his/her IP address or physical location. Drug sellers wouldn't want to set up shop in an online location whose registrant law enforcement can easily determine or where the site's IP address exist in the real world, too. There are many other reasons, apart from buying drugs, why people would want to remain anonymous or set up sites that can't be traced back to a physical location or entity. People who want to shield their communications from government surveillance may require the cover of darknets. Whistleblowers may want to share vast amounts of insider information to journalists without leaving a paper trail. Dissidents in restrictive regimes may need anonymity in order to safely let the world know what's happening in their country. On the flip side, people who want to plot the assassination of a high-profile target will want a guaranteed but untraceable means. Other illegal services like selling documents such as passports and credit cards also require an infrastructure that guarantees anonymity. The same can be said for people who leak other people's personal information like addresses and contact details.

III. THE DARK WEB VERSUS DEEP WEB

Much confusion lies between these two, with some outlets and researchers freely interchanging them. But the Dark Web is not the Deep Web; it's only part of the Deep Web. The Dark Web relies on darknets or networks where connections are made between trusted peers. Examples of Dark Web systems include TOR, Freenet, or the Invisible Internet Project (I2P). Taking on the mining tunnel metaphor[2], the Dark Web would be the deeper portions of the Deep Web that require highly specialized tools or equipment to access. It lies deeper underground and site owners have more reason to keep their content hidden.

ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)
Vol 4, Issue 3, March 2017

IV. NUMBER SITES ON DEEP WEB BASED ON LANGUAGE

In terms of raw number of domains, English was the main language of choice by at least 2,154 sites out of the 3,454 successfully scouted domains. That roughly makes up 62% of the total number of sites. This was followed by Russian (228 domains) then French (which may include French and Canadian-French sites, 189 domains)[3].

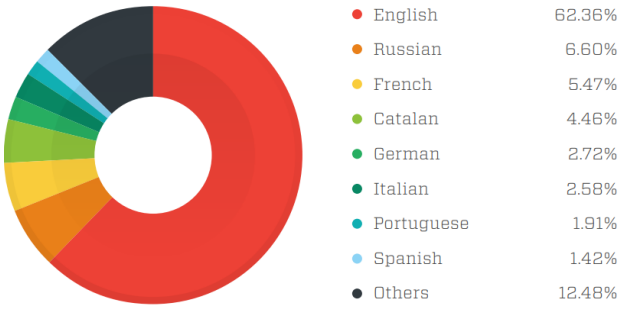

Fig 1 – Most popular languages based on number of domains containing pages that use them.

Looking at the language distribution based on number of URLs, Russian beat English because, despite having fewer sites, the number of sites that used Russian was bigger. At present, there's a particularly huge Russian forum not directly linked to malicious activities but mirrored in both TOR and I2P that on its own added to the total number of Russian pages.

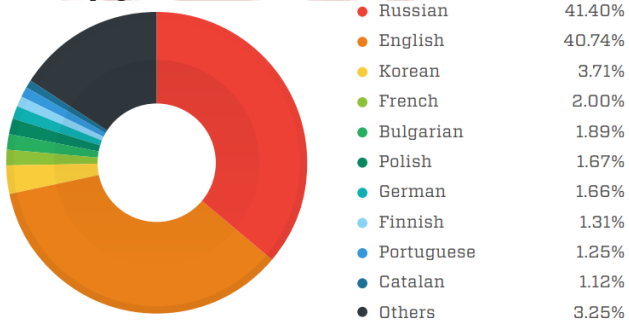

Fig 2 – Most popular languages based on number of URLs with content using them

V. COMMON USER PROFILE IN DEEP WEB

Profiling Deep Web users is challenging. More than language, the closest and most reliable estimate can probably come from looking at the different marketplace vendors. This gave us an idea as to what makes networks like TOR and I2P appealing. To get reliable information, we referred to the data available in https://dnstats.net/—a site that specializes in tracking activities in all darknet markets. An analysis of the top 15 vendors across all marketplaces[3] showed that light drugs were the most-exchanged goods in the Deep Web. This was followed by pharmaceutical products like Ritalin and Xanax, hard drugs, and even pirated games and online accounts. This data backed up the idea that a majority of Deep Web users—at least those who frequent the top marketplaces—go there to purchase illicit drugs.

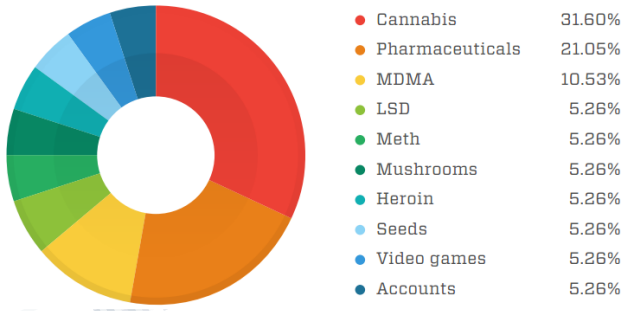

Fig 2 – Vendor breakdown

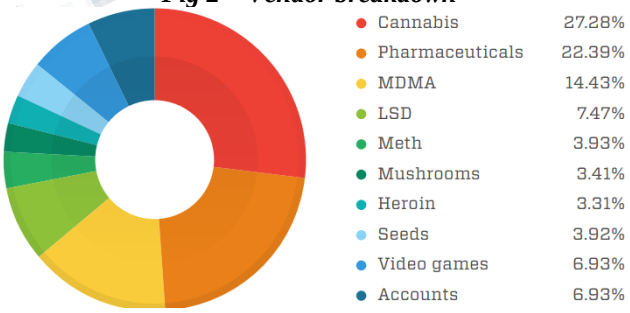

Fig 4 – Buyer breakdown

VI. WHAT MAKES DEEP WEB

Many users may not be aware that there are more than just standard sites in the Deep Web. Fig 5 tries to explain the type of protocols used in deep web. Fig

ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)
Vol 4, Issue 3, March 2017

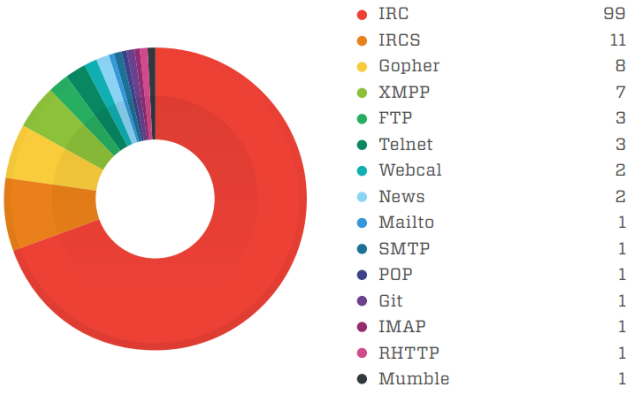

Fig V – Protocols found in Deep Web apart from HTTP/HTTPS More than 100 domains use either the IRC or IRCS protocol—normally chat servers that can be used as a rendezvous for malicious actors to meet and exchange goods or as a communication channel for botnets. The same concept applies to seven XMPP domains and one Mumble domain—protocols for chat servers that run in TOR.

VII. THE BAD STUFF THAT GOES ON IN DARK WEB

The Dark Web offers a certain level of anonymity that makes people in it more inclined to engage in illegal activities. The various transactions in it, including the makeup of prominent goods and services traded, very well paint a picture of what people would do if the secrecy of their identities was guaranteed. Unlike in the cybercriminal underground[4], most types of activities in the Dark Web have more apparent, if not drastic, effects on the real world. Many of the malicious tools and services sold in the cybercriminal underground can be used to gain profit. Those peddled in the Deep Web—assassination services, for example—obviously serve a different, more sinister purpose.

Few of the activities at deep web include,

1. The Malware Trade
2. Bitcoin and Money laundering services
3. Stolen accounts for sale
4. Passports and citizenships for sale
5. Leaked details: Government, law enforcement, and celebrities
6. Assassination services

VIII. TOR

Services that rely on rogue TLDs or Namecoins offer too high an adoption barrier to common users and very few advantages for average users to be considered publicly significant. But we've witnessed more and more usability improvements in systems like I2P and especially TOR that have slowly become viable solutions for everyone to go anonymous with hardly any setup requirements.TOR applications with anonymous browsers integrated are readily available for all major desktop and mobile platforms, allowing everyone to go anonymous and to access hidden sites in just a few touches. For journalists and people requiring an even higher level of privacy, a fully tailored OS integrating TOR[5] and anonymous browsers can be downloaded and installed on a USB key for use on any available machine. In practice, traffic between two TOR nodes is not traceable, but that to and from entrance and exit TOR gateways are. If an organization operates enough TOR gateways, there is a possibility that traffic using the TOR network can be tracked. Using TOR in countries that don't have enough budgets to operate a critical mass of gateway nodes can be considered safe. But in other countries with high intelligence service budgets like the United States or China, using TOR may not be as safe. In addition, any anonymizing system is only as effective as its user. As advanced as an anonymizing system may be, even those like TOR and I2P only cover the transport layer of communication but remain powerless toward the content of communication. Simply put, no anonymizing system can hide a user who posts his/her home address and details in the open. We can, therefore, list two major types of risk linked to anonymity in the Deep Web: Environmental vulnerabilities, Social vulnerabilities Environmental vulnerabilities refer to every possible flaw that can be linked to other software used together with TOR. For example, a notorious bug affecting the Adobe® Flash® version embedded in the browser that comes with a version of TOR once put the whole system in jeopardy since it was possible to exploit the bug to leak sensitive data despite the use of TOR. Social vulnerabilities are related to user behavior and the precautions users may take other than simply using TOR. Dread Pirate Roberts who was recently convicted to life in prison due to his Deep Web marketplace was caught by the FBI due to his use of a private email address in a public forum. Correlating the

ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)
Vol 4, Issue 3, March 2017

identities of Deep Web users with their Surface Web alter egos is an interesting research field that involves disciplines like social network analysis and stylometry.

IX. THE FUTURE OF DEEP WEB

While public awareness may lead to the increased use of or interest in the Dark Web and other similarly intended sites in the Deep Web, users currently don't have enough reason to migrate their Internet browsing to specialized anonymizing software in the near future. Meanwhile, it's much more likely for technological developments related to the Dark Web to improve the stealthiness of darknets. Right now, there seems to be a race between ―extreme libertarians‖ and law enforcement agencies, with the former trying to find new ways to become even more anonymous and untraceable by the latter. But since the commerce of illicit goods is one of the most predominant activities taking place in the Deep Web, it has become essential in the context of high anonymity, to be able to guarantee trust and reputation among sellers and buyers without having to rely on an external authority like a banking institution as in ―canonical‖ ecommerce. We are foreseeing the rise of new, completely decentralized marketplaces that rely on the blockchain technology that Bitcoins and other cryptocurrencies [3] already exploit for transport and storage. As such, the technology will be used to implement full-blown marketplaces without a single point of failure and that rely on particular aspects of game theory to guarantee safe transactions, escrow mechanisms, and trust between actors who may be shady in nature to begin with. Cryptocurrencies go hand in hand with Deep Web marketplaces. In that regard, we'll see new, advanced ways to make Bitcoins even less traceable than they are now. There's also a huge possibility for malware to start exploiting the blockchain technology to embed them and be carried around in a new, fully decentralized way.

CONCLUSION

The most heavily traded goods by top 15 vendors in the Deep Web are light drugs, followed by prescription medicine like Ritalin and Xanax and synthetic, prohibited drugs. The Deep Web heavily uses protocols outside the standard HTTP/HTTPS, most commonly IRC, IRCS, Gopher, XMPP, and FTP. Certain parts of the Deep Web have become a safe haven for different cybercriminals and criminal activities. Prevalent malware families like VAWTRAK and CryptoLocker are using TOR as part of their configuration. Takedowns of criminal marketplaces are not particularly lasting or impactful solutions against the drug trade, as there continue to be dedicated online shops and forums that serve the demand for illicit drugs. The Deep Web is also rife with Bitcoin-laundering services like EasyCoin to further increase the anonymity of moving money through the Bitcoin system. A cybercriminal underground definitely operates in the Deep Web as well. Stolen accounts, passports, and identities of high-profile personalities are sold in professional-looking forums with complete pricing information and descriptions. Assassination services are also advertised and offered in the Deep Web.

Deep Web will continue to raise a lot of issues and be a point of interest for both law enforcers and Internet users who want to circumvent government surveillance and intervention. As such, researchers, scholars and students like us need to continue keeping tabs on the Deep Web as its role in the Internet grows.

REFERENCES

[1] The Tor Project, Inc. Tor Project. Last accessed on 11 June 2015, https://www.torproject.org/.
[2] Vincenzo Ciancaglini, Marco Balduzzi, Max Goncharov, and Robert McArdle. Trend Micro Security Intelligence. "Deep Web and Cybercrime: It's Not All About Tor." Last accessed on 10 June 2015, http://www.trendmicro.com/cloud-content/us/pdfs/securityintelligence/white-papers/wp-deepweb-and-cybercrime.pdf.
[3] Robert McArdle. (4 October 2013). TrendLabs Security Intelligence Blog. "Cybercrime in the Deep Web." Last accessed on 10 June 2015, http://blog.trendmicro.com/trendlabs-security-intelligence/cybercrime-in-the-deepweb/
[4] Vincenzo Ciancaglini. (8 November 2013). TrendLabs Security Intelligence Blog. "The Boys Are Back in Town: Deep Web Marketplaces Back Online." Last accessed on 10 June 2015, http://blog.trendmicro.com/trendlabs-security-intelligence/the-boysare-back-in-town-deep-web-marketplaces-back-online/