Analysis of the Tor Browser and its Security Vulnerabilities

A. Kapadia

December 9, 2014

Abstract

Due to its notoriety of ensuring privacy on the internet, the Tor browser has made the use of onion routing more and more common amongst the public and everyday people. Tor is used for both licit and illicit purposes. While it is sometimes labeled as the *dark corner* of the web, referring to its past relation with Bitcoin and the virtual drug marketplace, Silk Road, it is also used by those living in internet censored countries, and even news organizations to ensure the protection and privacy of whistleblowers. This paper will discuss the strengths, and more importantly the weaknesses and security exploits of using the Tor browser in order to conceal information and user identity.

Contents

1 Introduction

Since its creation in the 1960s, the world-wide system of interconnected networks, otherwise known as the “internet,” has significantly grown to become the dominant tool used by millions of people for communication, entertainment, and social networking. With a substantial amount of data and personal information being shared every second, the question of anonymity is frequently brought up. Due to the recent privacy violations and unwarranted searches performed by the NSA, society’s increasing desire for internet privacy has resulted in its growing curiosity with the software notorious for anonymity and criminal activity known as Tor. Originally an acronym for “The Onion Router,” Tor is a free software that relays internet traffic in order to conceal a user’s IP address and location. The way Tor works its *magic* is through the process of onion routing, which is similar to an advanced method of proxy routing. When using an onion routing client, such as Tor, instead of non-securely sending data packets, Tor takes the user’s non-secure data packets and sends it to a node, which then encrypts the data. Then this encrypted information is relayed over and over to more nodes, each of which further encrypts the data, until the packets finally arrive at an “exit” node. Only once the data packets have reached this exit node, can the information be decrypted. Through these extra layers of encryption, onion routing improves the internet privacy of a user. But how much more secure is onion routing? What happens if someone knows where these exit nodes are and intercepts the information? In this paper, we will look at specific techniques, such as traffic-analysis and web fingerprinting, that exploit the vulnerabilities of onion routing.

2 To the Community

In today's society, due to the rise in information-based advertising, government monitoring, bulk storage of raw personal information and data, and the increased usage of social media websites, such as Facebook and Twitter, it has become significantly harder and maybe even impossible to ensure complete privacy. As more of society's personal information is being shared, sometimes without it even knowing so, more people have become aware of the issues at hand and are turning towards methods of preventing the handing out of their personal information. A major issue that concerns everyone is how large corporations use personal information they are sold from data brokers in order to profit off of the public through marketing and sales. For example, corporations heavily target people with specific advertisements based on their web searches and previous purchases on websites, such as Amazon and eBay. There have even been issues regarding insurance coverage, where insurance companies have started charging certain policyholders that eat unhealthy foods higher premiums. In addition, those who live in countries where information is censored by the government may be at risk if for browsing certain blocked websites or trying to access censored information. Because today's society is heavily dominated by technology, privacy and security are strongly connected to all aspects of life. Therefore, using alternative methods of browsing the web, such as Tor, which improve privacy and reduce the amount of personal information being shared to third-parties should be encouraged to the public.

Although it is a positive step forward that more everyday people use Tor to improve their privacy online, it is not correct if those who use it believe that it is one-hundred percent safe. In this paper, I will address some of the security concerns regarding the Tor browser in order to inform the public of the potential weaknesses that may lead to the leakage of user information and data. Alternative methods of ensuring data encryption and web browsing are significantly more secure and less prone to government monitoring and information-based advertising, however, is still very important for those who use such services to be aware of the possible security vulnerabilities and exploits.

3 Action Items

When a client browses the internet using web browsers, such as Google Chrome or Mozilla Firefox, they reveal specific information about themselves such as their location and packet data to intermediate routers. Because these routers are controlled by internet service providers (IPSs) that may be susceptible to “malicious attackers, eavesdroppers, and legal pressure,” a user’s information can become vulnerable [1]. In order to protect personal information, a user would have to use specific proxy browsers, such as Tor in order to encrypt their communication traffic and packet information. The following section will describe various malicious attacks that can be performed on the Tor browser. The purpose of this information is to inform users of the browser of possible vulnerabilities in order to spread awareness and caution.

3.1 Web Fingerprinting

Website fingerprinting refers to the process of attempting to access a client’s browsing behavior, or more specifically, determining the web pages they have previously visited by passively eavesdropping and analyzing their communication traffic. The communication traffic of a client will include certain information, such as “packet lengths, order, and timing information” [1]. Observing this encrypted information can allow a local attacker to extract information and draw certain conclusions leading to the identification of the web page. In general, an attacker’s strategy when performing a fingerprinting attack is to observe packet traces from his target client’s communication traffic and then compare these to the packet traces of specific web sites he believes the client may have visited.

When attackers observe information extracted from packet traces, they analyze many aspects of the data, such as HTML markers. When a browser accesses a web page, the HTML document is accessed first. Because of this, by “counting the size of incoming packets between the first outgoing packet (request for the HTML document) and follow-up outgoing packets [an attacker] can extract the HTML document’s size” [2]. Like

HTML markers, number and size markers are also used to compare packet traces as they "indicate direction changes in traffic flow" and "[reflect] how many packets were previously sent into the respective direction"[2]. Attackers also make use of the individual packets, and analyze the number of packets being sent in addition to the actual size of each specific packet. By analyzing many properties of packet streams, an attacker can potentially utilize it to identify web pages a client have browsed.

While the technique of fingerprinting does not actually break any aspect of Tor's cryptography, "even if an attacker does not understand a message's semantic, he can try to match observed patterns to known patterns" and if his attempt at web fingerprinting is successful, he has in a sense " [destroyed] the whole protection and anonymity promised by" Tor [2].

3.2 Man-in-the-Middle Attacks

Another type of attack that can be performed on the Tor browser is a man-in-the-middle attack. While web fingerprinting can be used to identify specific web pages that a victim has visited, a MITM attack, if performed successfully, can be much more harmful as it allows for the access of sensitive unencrypted information.

All of the relay nodes that Tor uses to add further layers of encryption to data packets are highly encrypted; however, when these packets leave the final node, commonly referred to as the *exit* node, all of this information becomes unencrypted, and left vulnerable to attack. While this type of attack can be extremely malicious, the only way someone can perform it is by setting up an exit node, and proceeding to sniff for packets using an open-source packet analyzer, known as Wireshark.

The way an attacker would go about performing such an attack is by first setting up an exit node, which can be done with Vidalia*, a GUI that allows for more control when using Tor. The next step would be to set unencrypted HTTP and retrieve mail settings to be allowed. Finally, by using a tool, such as TCPflow or Wireshark, an attacker can analyze Tor traffic that is leaving his exit node. For example, the command "tcpflow -i eth0 port 80" allows an attacker to sniff port 80 for packets and capture "HTTP requests,

POP3 emails, and IMAP emails" in order to find sensitive information like login names and passwords. By setting up an exit node and sniffing live data packets, an attacker can potentially " read and intercept the traffic of people using TOR, before it reaches the final destination" [3].

* Unfortunately for OS X users, the Tor Bundle package no longer comes equipped with Vidalia. Therefore, since there are no easy relay packages for OS X, an attacker must find a different way to get the Tor binary onto his system, making this process more difficult.

3.3 Malicious JavaScript

One vulnerability of the Tor browser is that it can potentially allow an attacker to gain information about his victim through malicious scripts. Luckily, Tor comes equipped with NoScript, an extension that allows the client to disable JavaScript and all executable web content on the browser, at the click of a button. While this extension is a strong preventative method against potential threats, an unaware user who doesn't enable NoScript could be very prone to harmful script attacks.

For those who don't enable NoScript while browsing with Tor, there is a strong potential for disaster. When browsing non-secure websites, users are susceptible to many forms of malicious JavaScript, such as cross-site scripting (XSS). XSS is a type of security vulnerability that is found in thousands of websites. Attackers perform cross-site scripting by injecting client side scripts into web pages. By injecting scripts into web pages, users may be susceptible to small issues like pop-up messages or re-direct links. However, in more severe cases, attackers may inject scripts that when are run, manipulate a user's cookies and other personal information in order to gain authorization to user accounts, without the user ever noticing.

According to OWASP, the Open Web Application Security Project, a significant percentage of all websites are prone to cross-site scripting. Because of the vast number of security flaws within web applications, when a user browses the internet with Tor, his privacy and identity may remain vulnerable if NoScript is not enabled.

4 Conclusion

Through the process of layering encryption and “routing data through several overlay nodes” to improve security, anonymization networks like Tor provide users with a more secure browser that improves privacy [2]. Because it has become significantly more difficult for a user to maintain privacy and guarantee that none of his personal information is being leaked, more people have started turning to these anonymization networks.

In this paper, several possible methods of breaching security and attacking client’s using the Tor browser were described. A Tor user is prone to many vulnerabilities depending on their environment. By comparing and analyzing packet traces, a local attacker can target specific individuals by passively eavesdropping on their network traffic in order to determine which web pages they have been accessing. Based on whether or not the NoScript extension is enabled on the browser, Tor users who do not enable it remain susceptible to many forms of malicious JavaScript attacks, including cross-site scripting. In addition, if an attacker sets up an exit node using the Vidalia GUI, he can extract personal information from the unencrypted data being relayed, such as login usernames and passwords. Because of these techniques, even if someone uses Tor to access illegal drug marketplaces or censored websites, if their packet traces are being analyzed or they are the target of a fingerprinting attack, there remains a chance of being identified.

Although Tor does provide its users much more sophisticated security and anonymization compared to commonly used web browsers, it does not ensure absolute protection from all types of security attacks. Because Tor promises to “strengthen a user’s civil rights, to protect the privacy, or even to give a user the opportunity to evade the censorship,” it is clearly “vital to know the networks’ level of protection and to enforce their promised protection” [2].

5 References

[1] Effective Attacks and Provable Defenses for Website Fingerprinting, 2003. https://www.usenix.org/system/files/conference/usenixsecurity14/sec14-paper-wang-tao.pdf.

[2] Website Fingerprinting in Onion Routing Based Anonymization Networks. http://lorre.uni.lu/~andriy/papers/acmccs-wpes11-fingerprinting.pdf.

[3] Intercepting TOR traffic to Sniff Passwords or other Data, 2011. http://www.ubertechblog.com/2011/03/intercepting-tor-traffic-to-sniff.html.

[4] Practical Vulnerabilities of the Tor Anonymity Network, .http://www.syverson.org/tor-vulnerabilities-iccs.pdf.

[5] Tor: The Second-Generation Onion Router, 2004. http://freehaven.net/anonbib/cache/tor-design.pdf.

[6] A Formal Treatment of Onion Routing. http://cs.brown.edu/~anna/papers/cl05-full.pdf.