An Overview of Some Popular Network Anonymity Systems

James Fell, james.fell@alumni.york.ac.uk

10th March 2015

When accessing the Internet under normal circumstances a user's identity is typically not very difficult for other parties to determine. Every time a web site or other network resource is connected to that server can log the user's IP address. At the very least, the user's IP address can easily be traced back to their Internet Service Provider. This has led to a demand for services that attempt to provide network anonymity by obscuring a user's real IP address. Many, although not all, also attempt to provide some degree of privacy by encrypting the contents of network communications between the user and the anonymity service.

It is worth stressing here the fact that anonymity and privacy are not the same thing. If your identity cannot be ascertained then you have anonymity, but not necessarily privacy. If in addition to this the content of your communications cannot be read by adversaries then you have privacy. Depending on the threat model different users will have different requirements and this should be kept in mind for the following discussion.

There are many reasons why a person may seek to use an online anonymity service, some legitimate and some not. A few examples of legitimate users include confidential sources communicating with journalists, law enforcement personnel who need to remain anonymous whilst connecting to an illegal web site they are investigating, people living under oppressive regimes who wish to access censored information without retaliation from the state, and business owners doing research on their competitors [1]. Examples of bad users would include people involved in distributing child pornography and spammers. The unfortunate reality is that it is impossible to deny privacy enhancing technologies to the bad users without also denying them to the good users. The rest of this essay focuses solely on the technical issues and not the politics.

There are currently numerous technical solutions available aimed at providing users with network anonymity whilst using the Internet. By far the most sophisticated and popular is Tor, which we will ultimately focus much of our attention on. It is however useful to also look at the main alternatives in order to give a broad overview of the subject. We will therefore introduce the following seven technologies in order; VPN services, anonymous proxies, anonymous remailers, JonDonym (aka JAP), Tor, Freenet and I2P.

Virtual Private Network technology was originally designed to allow distant entities to connect their LANs together over the Internet in a secure manner. This is still a widely used application for VPNs although more recently a large number of privacy services have appeared that attempt to use VPN technology for network anonymity purposes. While there is plenty of academic literature concerning the traditional corporate scenario for VPN usage, meaning connecting remote offices together securely, there has not been very much work regarding the more recent anonymity centred model. In 2012 Jacob Appelbaum et al conducted research [2] at the University of Washington to investigate whether or not the design of VPN technology is really suitable for this new usage model which it was not originally intended for.

The basic architecture of a VPN based privacy service is such that the user creates an encrypted tunnel from their computer to the provider's network and then routes all their Internet traffic through it. In theory, this makes traffic between the user and the VPN provider unreadable to third parties and also hides the user's true IP address from servers they connect to on the Internet. However, there

are several weaknesses.

As explained in [2], when a user connects to a VPN service the new tunnel is typically made available to applications on the workstation as a virtual network adapter and it is up to the local operating system's routing table whether a particular network connection goes through the VPN tunnel or goes directly out to the Internet. The routing table must therefore be reconfigured both on connection to and disconnection from the VPN service. In many cases applications such as web browsers must also be restarted to ensure that their traffic takes the correct route. In the case of the VPN's original intended use, this situation is perfectly fine. If a connection that was meant to go over the corporate VPN link accidentally gets sent to a different network interface, most likely the connection will simply fail as it will be addressed to a non-routable, private RFC 1918 IP address. In the case of a VPN privacy service however, getting this right all of the time is critically important as even a single DNS request going directly to the Internet when it was meant to go through the VPN tunnel can be enough to break anonymity.

Even if the routing table is always handled correctly, there are issues relating to client software such as web browsers being used sometimes with the VPN and sometimes without. For example, browser fingerprinting can often be used to identify a user with a very high certainty [11]. This is because a user's particular combination of browser version, plugins, screen resolution and other configuration details often combine to give a unique fingerprint. If this fingerprint is linked to the user's real identity from their unprotected browsing, then when they use the same browser through the VPN service they can be identified once again. There is also an issue of cookies that are acquired through normal usage being sent to web sites when later using the VPN and again revealing the user's true identity. While these issues as covered in [2] are certainly true, I feel it is important to note here that this is a potential problem with all anonymity systems and not just VPN services. The authors of [2] did not make this fact clear and may leave some readers with the impression that only VPNs have this risk. A separate web browser installation should always be reserved only for anonymous usage, regardless of the anonymity system.

Another important finding of [2] is that very few VPN privacy service providers actually make their users aware of the risks. Instead their marketing mostly makes inaccurate claims about the strength of the anonymity and privacy being provided without defining a clear threat model or explaining the many ways in which it can all go horribly wrong.

In 2011 Li et al carried out research [3] at the University of Nevada to investigate and compare patterns of usage for anonymous proxies, anonymous remailers, I2P, JonDonym/JAP and Tor. As part of this work they “identified 1,441 anonymity proxy servers, 15 remailers, 11 JAP mixers, 483 I2P relays, and 10,510 Tor relays .” They performed geolocation lookups on all of these IP addresses to analyse how each of the systems are spread amongst different countries. They also checked for the presence of all the IP addresses they had identified in a blacklist of 195,919 spammer IP addresses. Finally, Li et al used a modified P2P client to gather 114,593 IP addresses of file sharers and again checked the various anonymity system IP addresses they had gathered against it.

Anonymous proxies are a very simple method of anonymity. A user configures an application such as a web browser to proxy its connections through a specially configured server on the Internet. The other end of the connection sees the IP address of the proxy server instead of that of the user. Of course the proxy server can log the real IP address of the user as well as the sites he visits just like a VPN provider can and this information may later be disclosed. The connection between the user and the proxy is also often not encrypted which means this method provides even less privacy than a VPN. Nevertheless it is popular amongst some types of users as it is very simple to set up and there is an almost endless supply of free proxy servers. The research in [3] found that for the 1441

identified anonymous proxies the USA, China and then Brazil were the top three locations. It was also found that they were by far the most abused anonymity system in regards to the sending of spam and illegal P2P file sharing. In total 1140 of the 1441 proxy IP addresses collected were found to be on the spam blacklist and 36 of them were found to be on the P2P file sharing list.

Remailers are services which allow the sending of emails anonymously. Here we see an increase in sophistication compared to VPNs and anonymous proxies. This is because remailer architectures are based on the idea of a mix network, a concept first suggested by David Chaum in 1981 [4]. In a mix network messages are relayed through multiple hops instead of a single hop and each mix (or node) in the network can only identify the hop before it and the hop after it, rather than knowing the initial origin and final destination. Traffic passing through a mix network is encrypted using each node's public key. In the case of a remailer, where high latency is usually acceptable, it is also the case that multiple emails are bundled together by each mix and then sent out to the next mix together in one batch. This makes traffic analysis far more difficult. The research in [3] found only 15 active remailers and concluded that due to past abuse by spammers this is now quite a rare service. They found that the USA, Germany and then Netherlands were the top three locations of anonymous remailers.

JonDo (also called JAP or Java Anon Proxy) is software that is used to route traffic through the commercial JonDonym service (the free version is called AN.ON or Anonymity Online) [5]. Like the anonymous remailers, JAP is also based on the concept of a mix network although this time the focus is on web browsing rather than email and so low latency is a requirement. Traffic is passed through mix cascades encrypted to the public key of each mix and is kept at a constant rate to make traffic analysis more difficult. The research presented in [3] found JAP, like remailers, to have minimal usage at that time. They found that for the 11 identified JAP mixes Germany, the USA and then Netherlands were the top three locations. None of the 11 JAP mixes showed up on the spam blacklist or on the P2P file sharing list.

The Tor system (originally an acronym standing for The Onion Router) was developed in the late 1990s by the US Naval Research Laboratory. A 'second generation' version based on this was then released to the public in 2002. A detailed explanation of how Tor works is given by its creators in [6]. Although this paper is from 2004 and many improvements have been made to the Tor network since then, it is still a good starting point for understanding the technical details of how the network functions.

The network consists of thousands of volunteer run relays combined with nine servers called directory authorities that are run by the Tor Project itself. The directory authorities maintain a digitally signed public list of IP addresses of all the relays in the network. A connection through the Tor network is called a circuit and typically consists of a path of three relays. For the first relay in the circuit the Tor client selects a pool of several relays to use repeatedly long term. These are known as entry guards and make profiling attacks much harder than they would be if users selected a new entry relay at random for every connection. The next relay in the circuit is termed a middle relay and the final relay is termed an exit relay. Packets of data each 512 bytes long, known as cells, are sent through a Tor circuit having layers of encryption with one layer being removed by each relay as per the original onion routing concept. Unlike in a traditional mix network, in onion routing public key cryptography is only used to set up circuits and negotiate symmetric keys, with symmetric encryption then used to encrypt the actual messages being relayed.

As well as using Tor to connect anonymously to resources on the Internet such as web sites, the network makes it possible for users to anonymously run their own TCP based services. These are known as Hidden Services and are accessed within the network using the .onion top level domain [6]. To create a hidden service, after first setting up a web server or other service, a Tor user

generates a new private/public key pair for it. A hash of the public key serves as the .onion domain name. The user then builds circuits to several relays picked at random known as Introduction Points. These circuits are left open and for each .onion domain on the network, the directory authorities maintain a public list of these introduction points along with the public key. For another user to then connect to the hidden service, he would first build a circuit to a randomly picked relay termed a Rendezvous Point as well as to one of the hidden service's published introduction points. The user would send the identity of his rendezvous point to the hidden service through the introduction point. The hidden service would then itself build a circuit to the rendezvous point. After this, the client and the hidden service can both communicate through the rendezvous point without either knowing the other's identity.

The research in [3] identified a total of 10,510 unique IP addresses acting as Tor relays, although this was over a period of five days and only about 1500 Tor relays were ever active at any one time. This was still by far the most popular of the systems examined. The relays were distributed over 95 countries with Germany, the USA and France being the top three countries. In total 164 of the 10,510 relay IP addresses collected were found to be on the spam blacklist and 17 of them were found on the P2P file sharers list.

Freenet is a system which “operates as a self-organizing P2P network that pools unused disk space across potentially hundreds of thousands of desktop computers to create a collaborative virtual file system” [10]. Freenet is not a network anonymity service in the typical sense. It does not provide anonymous web surfing or anonymous access to the wider Internet but rather a censorship resistant way of anonymously publishing and hosting files within the Freenet network itself. Each file that is stored on Freenet is referenced using a unique GUID key based on a cryptographic hash. It is theoretically impossible to work out which node(s) on the network are storing any specific file and therefore it is not possible to suppress a particular piece of content. It is also impossible for the operator of a node to determine what files it is itself storing.

Freenet supports many interesting features such as Signed Subspace Keys which allow users to digitally sign content that they insert into the network. Beyond inserting individual files into the network it is also possible to create entire web sites known as Freesites and there are also emailing and instant messaging services that can be used within the network. More information can be found on the project's own web site [12]. Various attacks on Freenet have been discussed in the academic literature [13, 14].

I2P (Invisible Internet Project) is a P2P service which is also based on the concept of a mix network [7]. In this case the network specifically employs a method known as garlic routing. This is very similar to the onion routing used by Tor except with garlic routing the innermost layer of a packet can contain multiple separate messages. I2P allows users to connect to each other by routing traffic through multiple intermediate peers. Such a connection is known as a tunnel in I2P parlance and each peer in the network is known as a router. Unlike a Tor circuit, each tunnel can only work in one direction and hence a typical connection requires two different tunnels to be created, one inbound and one outbound.

Unlike Tor, I2P does not have centralised directory authorities. Instead it relies on a distributed database feature called the NetDB (Net Database) which stores a list of inbound tunnel endpoints (known as inbound gateways) for each user. For one user to send a message to another, he sends it down his own outbound tunnel addressed to the inbound gateway of the second user, as retrieved from the NetDB. Neither of the users is able to determine the other's true identity.

Similar to Tor hidden services it is possible for users to host web sites and other services anonymously within the I2P network. Indeed, unlike Tor, this is actually the main focus of the

project rather than using it to connect anonymously to Internet resources. Such hidden services are known as Eep sites and use the .i2p top level domain. Although not very popular, it is also possible to use I2P to connect out to Internet resources by using the small number of outproxy routers on the network. These are the I2P equivalent of Tor's exit relays.

The research in [3] identified over a seven day period only 483 I2P routers in 29 different countries with France, Germany and then the USA being the top three locations. In total 29 of the 483 router IP addresses collected were found to be on the spam blacklist and none were on the file sharing list.

In 2009 Abou-Tair et al carried out research [8] at the University of Siegen to investigate and compare the relative usability of mixmaster remailers, I2P, JonDonym and Tor. The motivation behind this research was the fact that anonymity systems generally become more secure as their user base grows (hence the size of the anonymity set grows) and the more user friendly a system is the more people are likely to use it. The study was based on the researchers' experience of installing, configuring, testing and then deactivating each of the four systems. Their experience of those four steps was measured against a list of eight usability guidelines including how easy the process is to understand, how easy it is to tell that a step has been completed successfully, how clear the user interface and documentation is and how obvious it is whether the system is active or not at any given time.

Tor was found by the researchers to be generally user friendly and easy to work with. Installation is performed using a wizard process and use of the network is activated and deactivated using a button that is added to the user's web browser. Clear warnings are also given to the user that no system can provide perfect anonymity and that there are risks. This finding tends to correlate with what was seen in [3] with the large size and popularity of the Tor network.

I2P performed far less well. The whole process of installation and use was found to be very confusing. For example, Java is required to have already been installed and yet if it isn't installed it is not auto-installed by the setup application and no directions as to how to manually get it are provided to the user either. The documentation was also found to be aimed at highly technical users and would be very unclear to novice users. The user interface and terminology was found to be unclear. This also tends to correlate with what was seen in [3] with the relative unpopularity of I2P.

JonDonym was found to share one of the problems that [2] criticised in VPN services, namely that it's web site gives the impression that this is a totally secure system without honestly explaining to users the many risks that are present in reality. It also suffers from a confusing naming convention. The software client is sometimes known as JonDo and sometimes known as JAP. The network itself is known either as AN.ON for the free version or JonDonym for the paid version which includes more functionality. This can be confusing to users. On the positive side, the installation process was found to be quite simple much like Tor with a wizard guiding the user through it. Once installed the system was also very simple to use and simple to deactivate when necessary.

Mixmaster was found to have many of the usability problems of I2P but to an even worse extent, such as being aimed only at highly technical users and having very complex documentation. The researchers also looked at a client for Mixmaster called Quicksilver however even using this the whole process was found to be confusing and far too complex to be accessible to novice users. Remailers were found to be both hard to use in [8] and unpopular in [3].

It is interesting to see that the findings of [8] regarding usability do not line up completely with the findings of [3] regarding popularity of network. JonDonym/JAP was found in [8] to be very easy to use and yet was found in [3] to be very unpopular. It seems likely that the superiority of the anonymity provided by Tor combined with the fact that it is completely free means that even with

high usability JAP cannot compete effectively for users.

As well as usability, it is obviously very important to users of anonymity systems to know that the system has been thoroughly tested from a security point of view and is resistant to attack. Unlike most of the alternatives, Tor has had a huge amount of attention from the academic and security communities over the years with many novel forms of attack being first discovered and then mitigated through improvement of the system's design. This is a huge field with hundreds of papers in the literature so we will limit our discussion here to one key issue; malicious exit relays.

In 2014 Winter et al published some research [9] into how the Tor Project could detect malicious exit relays. This is an important issue for the safety of the network as traffic from an exit relay to its ultimate destination cannot be protected by Tor's own encryption. If it is using an unencrypted connection, for example a Tor user is viewing a web site using HTTP instead of HTTPS, then the exit relay can spy on and also modify the traffic. Even if the connection is using end-to-end encryption such as with HTTPS a malicious exit relay can still attempt a man in the middle attack by replacing the legitimate SSL certificate with its own self-signed version. If the user ignores the resulting browser warning and proceeds then the connection can be monitored as if it was plain HTTP. It is therefore important to be able to continuously detect exit relays carrying out such attacks and have the directory authorities ban them from the network.

To this end Winter et al developed two software tools. The first tool is called exitmap and it tests an exit relay on the Tor network to see if it is carrying out any common man in the middle active attacks such as DNS poisoning, injecting javascript into web pages, rewriting HTTPS links within web pages as HTTP, or presenting the attacker's own self-signed SSL certificate instead of the legitimate site's. The exitmap tool checks for all of these in essentially the same way. It connects to a resource on the Internet by making a direct connection outside of Tor and also by using a Tor circuit ending with the particular exit relay to be checked. The two results are compared and if they do not match then there is a problem. For example, if the SSL certificate retrieved from Facebook by connecting to it through a particular exit relay is not the same as the SSL certificate retrieved when connecting directly to Facebook, then it is clear that this exit relay should be flagged as malicious.

The researchers ran the exitmap tool for a period of seven months to carry out multiple checks per week of all of the approximately 950 exit relays that were present at the time. Over this period they identified 40 exit relays that were misbehaving in some way. Some of these turned out to be misconfigured rather than deliberately malicious. For example, two of the relays turned out to be running anti-virus software which was breaking into IMAPS sessions to scan for viruses. However, the majority were blatantly malicious attempts to attack users.

The second tool presented in [9] is called HoneyConnector and is designed to expose exit relays that are engaged in passive sniffing of traffic. To do this the tool again establishes a Tor circuit ending with the particular exit relay to be checked. The tool then makes unencrypted connections to servers controlled by the researchers and logs into FTP and IMAP accounts that were set up specifically as bait for the test. Each exit relay has the credentials for a different account passing through it. The researchers were then able to log any unauthorised attempts to use those credentials over the coming weeks and therefore see which specific exit relays had been sniffing traffic. The HoneyConnector tool was deployed for a period of four months and passed unencrypted login credentials through all of the exit relays on multiple occasions. During this period they “registered a total of 255 login attempts with 128 sniffed plaintext credentials, tracing back to 27 sniffing exit relays ” [9].

The research performed in [9] was very interesting and useful, and clearly provided the Tor Project

with an effective technical solution for continuously banning malicious exit relays from the network. However, I believe that as well as this more emphasis should be placed on user education. If users of the Tor network always use end-to-end TLS encryption and do not proceed with connections when they get a warning about the certificate then there is little risk from malicious exit relays. All of the attacks rely on the user either failing to use TLS or accepting a self-signed certificate even though they are visiting a site like Facebook that clearly would not be using a self-signed certificate.

In conclusion, there are many different services aimed at providing anonymity and privacy for Internet users and none of them are perfect. There are usability issues to consider as well as a range of pitfalls especially for less experienced users that can lead to deanonymisation through configuration or other user errors. Each service is also vulnerable to attacks by well funded adversaries.

References

[1] – The Tor Project, “Who uses Tor?”. [Online] Available: https://www.torproject.org/about/torusers.html.en [Accessed: 10th March 2015]

[2] – Appelbaum et al, “vpwns: Virtual Pwned Networks”, The 2nd USENIX Workshop on Free and Open Communications on the Internet, August 2012.

[3] – Li et al, “An Analysis of Anonymity Technology Usage ” in Traffic Monitoring and Analysis, Lecture Notes in Computer Science Volume 6613, 2011, pp 108-121.

[4] – David Chaum, “Untraceable electronic mail, return addresses, and digital pseudonyms”. Communications of the ACM, Vol 24, No 2, February 1981.

[5] – JAP Anonymity and Privacy, “Architecture of the Anonymization Service”. [Online] Available: https://anon.inf.tu-dresden.de/desc/desc_anon_en.html [Accessed: 10th March 2015]

[6] – Roger Dingledine, Nick Mathewson, and Paul Syverson. “Tor: The Second-Generation Onion Router” in USENIX Security. USENIX, 2004.

[7] – Invisible Internet Project, “Introducing I2P”. [Online] Available: https://geti2p.net/en/docs/how/tech-intro [Accessed: 10th March 2015]

[8] – Abou-Tair et al, “Usability Inspection of Anonymity Networks ” in the Proceedings of the World Congress on Privacy, Security, Trust and the Management of e-Business (CONGRESS '09), August 2009, pp 100-109.

[9] – Winter et al, “Spoiled Onions: Exposing Malicious Tor Exit Relays ”, The 14th Privacy Enhancing Technologies Symposium, July 2014.

[10] – Ian Clarke et al, “Protecting Free Expression Online with Freenet” in IEEE Internet Computing, January 2002.

[11] – Electronic Frontier Foundation, Panopticlick Browser Fingerprinting Tool. [Online] Available: https://panopticlick.eff.org/ [Accessed: 10th March 2015]

[12] – Freenet Project, Official Website. [Online] Available: https://freenetproject.org/ [Accessed:

10th March 2015]

[13] – Guanyu Tan et al, “A Traceback Attack on Freenet” in Proceedings of IEEE INFOCOM 2013. Turin, Italy, April 14-19, 2013.

[14] – Todd Baumeister et al, “A Routing Table Insertion (RTI) Attack on Freenet” in Proceedings of the 2012 International Conference on Cyber Security, pp 8-15.