Network Anonymity

Kenan Avdić* and Alexandra Sandström†
Email: {*kenav350, †alesa938}*@student.liu.se*
Supervisor: David Byers, {*davby@ida.liu.se*}
Project Report for Information Security Course
Linköpings universitet, Sweden

Abstract

In this paper we intend to examine to what degree it is currently possible to achieve anonymity, with special emphasis on the Internet. We also intend to examine the availability of such techniques for an average internet user. There is a historical and current perspective included, from mix-nets to today's systems such as Tor, Freenet and others. We discuss the efficiency of these systems and their vulnerabilities. Finally, we discuss the results from a technical and social point of view on anonymity and privacy, both on the Internet and in general.

1. Introduction

Anonymity and privacy have been a concern of humanity for a long time. The word *anonymity* is derived from Greek and literally means to be without a name, unnamed. Pseudonymity is another, closely related concept, where a person uses an alternate name or identity to his own. There exists one distinction; the pseudonym is usually reused and can be recognised by others. Other than recognition, pseudonymity also allows for example relationships such as trust or friendship, which are unattainable with full anonymity, to be imparted upon pseudonyms.

There are various reasons for which a person may wish to be anonymous or pseudonymous. Some of the example occasions are social situations, art, politics, commerce and crime. The reasons are varying; in a social situation, an anonymous (or pseudonymous) person can be certain to avoid preconceptions and receive a more objective and unbiased opinion in discussions [1]. In a repressive political regime, an anonymous person may freely express their political opinion without fear of repercussions. [2] Pseudonymity is widely used on the Internet, e.g. many people use several different aliases to explore different aspects of themselves, such as hobbies, or pretend to be someone else for the purpose of social experimentation. Naturally, anonymity can also be used in crime and for other offensive or disruptive purposes. A high degree of anonymity imparts an increased difficulty in accountability.

Remote means of communication and in particular the Internet have made the difficulty of obtaining anonymity significantly lower than it has been historically.

Nevertheless, obtaining anonymity on the internet can be a difficult matter, depending on what degree of anonymity is required. Invariably, the Internet protocol version 4 has a destination IP-address (and so it will in the future with IPv6) that can be traced to a physical interface, i.e. a computer and its location and possibly the owner or user. Currently, discovering someone's identity can be done via the internet service provider, and can in some countries be, due to recent increasingly harsher intellectual property legislation, a trivial matter. Due to obvious technology limitations it is not possible to achieve total anonymity, as the unique IP-address must be available for communication. Thus, in order to achieve a measurable degree of anonymity, one cannot rely on the secrecy of an IP-address, i.e. that an IP-address will not be tracked back to the user; other techniques must be employed. Many techniques used today are generally based on the principle of plausible deniability. This principle is often used by the military and is based on the avoidance of responsibility by not revealing a relationship between the originator and an effect. In networking, this is today most often achieved by using proxies and/or tunnelling, which hide the identity of the message originator. Many commercial anonymisation services today utilise this principle.

We intend to examine available anonymisation techniques today, some principles they are based upon and their measurement techniques. We are also interested in public availability of the implementations of these techniques, as well as commercial anonymisation services.

This paper is organised as follows. Section 2 provides background and overview of the techniques used for anonymisation on the Internet. In section 3 we describe the metrics of anonymity, i.e. to which degree a user can be anonymous on the Internet. Section 4 presents previous and similar works. Finally, we conclude in section 5.

2. Background

The seminal paper on untraceable electronic e-mail by David Chaum [3] introduced the concept of mixes. A mix is a relay that employs public key cryptography to anonymise the sender of an e-mail. A mix also uses batching and pads all blocks to a fixed size. A number of messages are batched together and sent out at the same time to prevent tracking them to a single sender.

Each outgoing block is padded to the same size so that an attacker is not able to distinguish between different messages. This property is called *bitwise unlinkability*. It was later shown that the original mix does not achieve this [4]. Chaum suggested two different methods of interconnecting mixes, one of which is the mix-cascade. Mix-cascades are today used in almost all practical anonymous channels. Another term introduced by Chaum is the *anonymity set*. The anonymity set and other metrics are covered in greater detail in section 3.

2.1. Remailers

An anonymous remailer is an electronic mail relay used to conceal the sender of an e-mail. There are several types of remailers, called Type 0, 1, 2 and 3.

2.1.1. Type 0. The first widely used practical application of an anonymous system began with the anon.penet.fi remailer. This remailer had over 300,000 users at its peak. It was located in Finland and operated by Johan Helsingius between 1993 and 1996. It was a simple map of pseudonyms, where any user could obtain a pseudonym and connect it with their actual e-mail address. This principle showed vulnerable to legal threats and was compromised first in 1995, when the Church of Scientology accused an anon.penet.fi user of copyright crimes. Shortly thereafter, Helsingius was forced to shut down the service due to new litigation against it.

2.1.2. Type 1. The Cypherpunk or Type 1 remailer is similar to the Type 0 remailer. However, it stores no user data and offers only anonymity. It uses PGP [1] or GPG [2] encryption and can use a cascade for additional security. When employing a cascade, several relays are used to deliver the message. The message is then encrypted by each relays' encryption key in reverse sequence of the delivery path. The message is sent to the first relay, which decrypts the first layer to reveal the next relay (or recipient) and forward the message. The layered encryption provides unlinkability between the sender and the recipient.

2.1.3. Type 2. The Mixmaster or Type 2 remailer [5] is conceptually very similar to the Type 1. However, Mixmaster employs some additional principles to make attacks significantly more difficult, particularly some concepts already proposed by Chaum [3]. First, as in the original mix, the messages are split into equal sized blocks and padded as necessary. Second, the messages are pooled before being sent to avoid traffic analysis attacks. Third, the mixes send dummy packets to each other to further confuse the attacker.

1. Pretty Good Privacy; http://www.pgp.com
2. GNU Privacy Guard; http://www.gnupg.org

2.1.4. Type 3. Mixminion or Type 3 remailer adds to the capabilities of the previous remailers by introducing receiver anonymity. Moreover, simultaneous sender and receiver anonymity is supported. Another addition is message modification (i.e. tagging) [4], [6] protection in the form of an integrity check. The Type 3 remailer is one of the foci of current research.

2.1.5. Nym servers. Pseudonymous or *nym* servers [7] use existing mix network infrastructure (such as Mixmaster or Mixminion) combined with the private information retrieval protocol [8] to provide pseudonymity. This is achieved by using public key cryptography where each pseudonym is identified by a long term public key.

2.2. Onion Routing

Onion routing [9] is a technique similar to what Mixmaster employs. It is based on Chaum's concept of mix cascades [3]. Each message is encrypted in layers by the sender. Each respective layer is encrypted in the reverse routing order of the message, i.e. if a message M is to be routed through routers A, B and C, it is encrypted as:

$$\{\{\{M\}PK_A\}PK_B\}PK_C$$

Clearly, this routing order is one-way only. If two-way communication is required, two routes must be established, one in each direction. The onion decreases in size as it is propagated through the network and its layers are "peeled" off, so its payload is padded at each hop to avoid that the message changes in size. Only the final proxy can discern the real data from the padding and knows its combined size.

After a route has been established by an onion, each circuit stream is assigned an identifier at each proxy. This identifier is used in subsequent communication so that each proxy knows which other proxy follows it in the routing order.

2.2.1. Tor. The Tor architecture [10] is similar to the Onion routing concept, but changes some key design properties. It takes a more practical approach to solving the same issues. To begin with, circuits in Tor are not established using an onion, but use iterative circuit setup. A user's Onion Proxy (OP) establishes an encrypted connection using a symmetric key with the first Onion router (OR1) on her chosen path. Once this connection is established, the OP can send a special *create* message to OR1, asking it to extend the path to the next router in her path, OR2. OR1 then establishes a new connection to OR2 and the circuit is extended. It can then iteratively be further extended until a chosen path is achieved.

Another key difference is that Tor does not use message padding or mixing. Instead, it uses fixed size communication datagrams called *cells*. It also offers end-to-end integrity checking, as opposed to its predecessor that only uses random nonces.

Additionally, in Tor, each circuit uses multiplexing so that it is not necessary to establish new circuits for new connections. One circuit can be used for several streams. This is naturally claimed to increase efficiency and improve anonymity.

Finally, the original onion router concept suggested communicating network information to nodes by using flooding. Tor uses a set of trusted nodes instead, called *directory servers*. The directory servers are periodically polled by user proxies for the latest network state.

2.3. Crowds

Crowds [11] is a simple but elegant concept. It attempts to anonymise web clients by allowing users (which they call *jondos*) to forward each others' requests. Each pair of jondos share an encryption key for communicating that is established when a jondo joins the crowd.

Crowds introduced a new concept called *initiator anonymity*; any node receiving a request cannot know if the sender was the node that sent it or some other node that preceeds it.

2.4. Peer-to-peer storage & publication systems

A peer-to-peer architecture is a popular network architecture where nodes (called peers) communicate directly with each other without any central infrastructure. This architecture is often combined with an implementation of an application layer overlay network or a DHT [3] for indexing and peer discovery. Peer-to-peer is a robust scalable infrastructure, but it does not provide anonymity by design.

Peer-to-peer anonymous storage and publication systems are primarily concerned with providing or reinforcing freedom of speech and overcoming censorship. They are sometimes called *censorship resistant* systems. Generally, they are based on transparently distributing encrypted files or parts of files in such a way that hosts can be (and preferably are) oblivious to the type of information they hold.

2.4.1. Freenet. Freenet is the largest and most popular practical solution today. It is a "cooperative distributed filesystem incorporating location independence and transparent lazy replication" [12]. The files are stored in a local datastore, and are most often encrypted (as is recommended to avoid legal liability). Locating and storing of files is done using a hash of its filename. The communication is very similar to how IP [13] works; each node knows only its adjacent nodes and all requests have a hops-to-live value which is decremented for each hop.

Freenet allows for authorship verification by signing files cryptographically. This is a type of pseudonymity which in turn allows for building a reputation.

3. Distributed hash table (DHT) Table with pairs (key; value) created collectively by nodes that they use to retrieve values associated with a given key.

Freenet has received a large amount of media attention, unfortunately mostly from the copyright enforcement point of view.

2.4.2. GNUnet. GNUnet [14] is another practical solution for anonymous file-sharing and publication. One difference in GNUnet's protocol, GAP, compared to other systems, is that it does not attempt to conceal the connection between the initiator and the responder. Instead, the connection is direct and the principle of *plausible deniability* is used to decouple the initiator with the action. This is done (as with Crowds) by making an initiator indistinguishable from an intermediary.

GNUnet can be configured for either high anonymity or high performance. Each node can change a number of parameters to surrender some percentage of its anonymity in order to achieve better efficiency.

2.4.3. TAZ/Rewebbers. The TAZ/Rewebber system [15] focuses on publisher protection of HTTP [16] traffic only. It comprises of two components. The first is the rewebber network, which is a network of HTTP proxies providing an anonymisation service to HTTP servers using them. The proxies use public key encryption to mask the real address of the target, i.e. the address is encrypted with the proxies public key. To prevent search engines from indexing the content and revealing the service, every resource is encrypted with a symmetric key which is also distributed with the URI, e.g. given a proxy P and a server A:

$$http://P/\{K_P, http://A\}_{KP(P)}$$

As these URI's can become quite long due to the encryption, the second component, the TAZ (Temporary Autonomous Zone) server is employed to reduce their length. TAZ servers is thus a database that maps a long public-key-encrypted address of a target server to a much shorter .taz resource.

2.5. Peer-to-peer communication systems

Peer-to-peer anonymous communication systems focus on distributed anonymous communication. Compared to more centralised solutions, these systems have an additional desired quality, *relay homogeneity*. Each node within such a system both originates and relays traffic, which makes the node's own traffic indistinguishable from the cover traffic.

2.5.1. Tarzan. Tarzan is a peer-to-peer anonymous IP network overlay. Tarzan uses layered encryption and multi-hop routing (similar to the Chaum's mix) with best-effort delivery model. Compared to other anonymization systems Tarzan works on a lower level, the network layer. As such, it is more difficult to use practically (as it may require kernel patching etc.). On the other hand, it also makes it capable of transparently encrypting any type of communication, by intercepting it on a lower layer.

Each node in Tarzan has a unique identity (based on its IP-address and public key) for every node and creates virtual circuit chains (tunnels) that are used to exchange encryption keys and forward anonymized messages.

2.5.2. MorphMix. MorphMix [17] is a peer-to-peer circuit-based mix network designed to improve upon some weaknesses of Tarzan. For example, in Tarzan, each node is required to maintain a global view of the entire system, allowing scaling only up to about 10000 nodes. MorphMix improves upon this by letting each node maintain a view of only its neighbours.

MorphMix has a collusion detection mechanism to detect malicious nodes in the network. The initiator stores previous path selections and compares them to new selections to try and avoid malicious nodes when planning a path for the message. Furthermore, an already connected node, a *witness node* is polled to verify neighbour information. Upon detecting a malicious node, a circuit is torn down and no longer used.

2.6. Commercial and other systems

The commercially popular systems today are mostly based on virtual private networks. They generally employ the L2TP [18] and PPTP [19] protocols or SSH to tunnel data to a proxy server. This proxy is the endpoint of the tunnel. The anonymity here is thus based on trust imparted on the proxy not to disclose the identity of the user. These centralised solutions are very weak to some attacks, specifically the proxies can easily be blocked, and traffic analysis is trivial if it is possible to eavesdrop on the proxy itself (more on this in section 3).

2.6.1. Anonymizer. The Anonymizer is the earliest of such proxy systems. The company was founded in 1995 by the author of the Mixmaster remailer. In addition to providing a VPN tunneling solution, Anonymizer uses IP-address rotation to further obfuscate the originator of a message.

Unfortunately, due to the fact that the Anonymizer states that they may reveal personal information to a third party [20], it becomes clear that the service stores user information, even if it may not store their activities. This is highly detrimental to the apparent degree of anonymity that the Anonymizer provides. Furthermore, the company itself was in 2008 acquired by a company that delivers solutions to the United States national security community.

2.7. Data collection methodology

A literature study was done providing an overview of the history of techniques used. Several papers concerning both the theoretical and the practical approach were consulted and used as a basis for this work. As there is quite a large body of work available on this subject, some anonymisation techniques may be underrepresented or even omitted. Furthermore, we do not verify previous studies, but accept them as stated.

3. Solution and Analysis

An *anonymity set* first proposed by Chaum [3] is the traditional method of quantifying anonymity. The larger the size of this set the greater is the anonymity provided by the system. The modern accepted definition of anonymity follows:

Anonymity of a subject means that the subject is not identifiable within a set of subjects, the *anonymity set*.

Other significant terminology [21] includes:

Unlinkability of two or more items of interest (IOIs) from an attacker's perspective means that within the system, the attacker cannot sufficiently distinguish whether these IOIs are related or not.

Unobservability is the state of IOIs being completely indistinguishable from any other IOI of the same type. [22]

A *pseudonym* is an identifier of a subject other than one of the subject's real names; A subject is pseudonymous if a pseudonym is used as identifier instead of one of its real names.

3.1. Anonymity metrics

A basis to measuring anonymity can be said to be the *anonymity set*. The anonymity set is the set of all the possible subjects who are considered capable of causing an effect. The larger this set is, and the more evenly distributed the communication within it (sending and receiving), the more anonymity it yields.

One suggested practical metric for anonymity is the *degree of anonymity* (fig. 1) [23], [24]. It was first proposed as a spectrum ranging from *absolute privacy* [4] to *provably exposed* [5]. Since then, it has been expanded with a metric based on Shannon entropy [25]. This information theoretic based metric allows us to calculate the measure of anonymity given by a system as a function of its maximum achievable anonymity. The key properties used in the anonymity quantification are the anonymity set (the larger it is, the higher the anonymity provided) and the number of attackers or colluding nodes. In one particular case, crowds, the anonymity provided yields an almost linear relationship between these two [26].

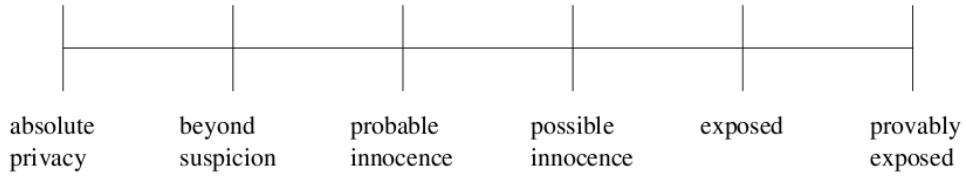
3.2. Threats and weaknesses of anonymity systems

The simplest anonymity solution is naturally also the easiest one to attack. Commercial and other systems based on a single proxy are highly susceptible to traffic

4. Absolute privacy – adversary can in no way distinguish the situations in which a user performed an action and in which he did not

5. Provably exposed – the adversary can prove that a user performed a certain specific action

Figure 1: Initially proposed degrees of anonymity



analysis. The proxy host is simply eavesdropped on, and correlations are easily found in the incoming and outgoing traffic. Additionally, in the case of commercial services, there is a need for an amount of trust in the proxy that users may not be prepared to bestow. The Type 0 anon.penet.fi remailer's legal issues and subsequent closure also show that a centralised solution or one for which a single individual is responsible can be weak to pressure to threats outside the network, in this case, civil law. The same applies to commercial solutions.

There is one attack that has long been, and still is, a subject of active research, the *(n–1) attack*. As mentioned earlier, mixes pool messages into batches before sending them out, to avoid traffic analysis. However, an attacker can flood the mix with (n–1) messages (where n is the message count when the batch is transmitted) and discover the path of the last message in the batch. Generally, this type of attack is ineffective against more advanced mixes.

If an attacker knows that there is some repeated and persistent communication between node A and a set of other nodes, given enough time, he can employ an advanced form of traffic analysis called *intersection attack* to deduce which nodes node A is talking to.

If a global passive attacker can examine the size of data, e.g. web pages, he can deduce a connection from file sizes. Usually, padding is employed to make the data fragments the same size. Such an attacker can also examine the timings of messages moving through the system to find correlations. This is known as the *timing attack*.

A *Sybil attack* [27] is specific to peer-to-peer systems. If the identity of nodes is not verified, a single node may present several identities and pose as many, and so significantly influence the peer-to-peer network topology. It may even partition the entire overlay network to its benefit.

3.2.1. Threat models. The main threat models to consider when designing an anonymous systems are the following [28]:

- **A Global passive attacker** is the most commonly used attacker type in the litterature. This type of attacker can observe all communication links but remains passsive, i.e. unable to alter any traffic.
- **A Global passive attacker with many compromised mixes** is similar to the global passive, but in addition assumes that the attacker has access to all but one mix in the mix-network, i.e. can access parameters such as private keys and other sensitive data. This is an even stronger type of attacker.
- **A Global active attacker** is the same as the passive one, but is able to control the communication medium such as injecting arbitrary amounts of data or delaying existing traffic.
- **A Global active attacker with many compromised mixes** is an active type of the global passive attacker with compromised mixes, i.e. the attacker has full control over the compromised mixes and can change their data.
- **A sub-global attacker** is an aggregate type of attacker that comprises of most of the practical real-world attackers. The attacker in this model controls some known or unknown passive communication links and nodes.

3.3. Evaluation and Comparison

Anonymity schemes can be either centralised, where only few nodes act as servers for the rest, or peer-to-peer, where every node may act as a server or client. Of these, peer-to-peer systems are those considered to be more vulnerable to attacks on anonymity [29]. Due to the architecture, the peer-to-peer setting introduces many more opportunities for attack. Additionally, current peer-to-peer technology does not provide an adequate level of anonymity. On the other hand, centralised approaches have the obvious drawback in difficulty of scaling for a large population of users. Especially in the later years, as the interest for anonymity services is increasing so does a need for a scalable solution. For example, one of the most popular systems currently in use, Tor, is running at near maximum capacity. Nevertheless, only centralised scheme solutions are widely deployed today, while research is on-going for a more scalable solution. One example of this is implementation of an anonymity layer on top of existing technologies and applications such as bittorrent.

Due to its architecture, a centralised systems infrastructure is easier to attack. Network targets are easily identifiable, and the infrastructure may also be associated with a physical person vulnerable to e.g. legal pressure (e.g. Tor exit nodes).

The censorship resistant systems, however, are naturally almost all based on peer-to-peer technology. It would be difficult, if not impossible, to conceive a solution to the problem of censorship resistance based on a centralised solution.

Remailers provide the highest anonymity today. Type 0 is not used, as it is completely defenseless to attacks. Types 1 and 2 are somewhat vulnerable, but still considered adequate by many users. Type 3 provides state-of-the-art anonymity. Remailers can only be used for electronic mail, generally use passive global attacker type as the attacker model, are difficult to use and are not targeted towards the casual user.

Tor is the most popular implementation for low-latency anonymous communication. It provides limited anonymity, as it does not have a strong attacker model. On the other hand, it can be used for all kinds of communication (compared to remailers) is simpler to use and more user-friendly.

4. Related Work

There has been several papers examining an overall view of internet anonymity. Of these perhaps the most comprehensive is the work done by Serjantov [28] and Danezis & Diaz [22].

Unfortunately, since internet anonymisation research is a highly active research field, (similarly to network security) both the anonymisation and disclosure technology is very much in a state of flux. Due to this, comprehensive studies on this technology quickly become outdated.

Wright et al. [30] provide an overview of then-current state of anonymity research. Furthermore, Palme and Berglund [1] have led an informal discussion about the anonymity use and social, legal and political aspects of anonymity and its impact on society. Camp and Osorio [31] take a look at current methods of staying anonymous from the e-commerce perspective.

5. Conclusions

To what degree is it then possible to be anonymous on the Internet today? In the literature, the definition of anonymity often seen as sufficient is that a user can complete a transaction without the possibility of it being distinguished from other users. An adversary can thus always determine some information about a user's activity, no matter how small.

From this point of view, none of the presented technologies provide true anonymity, in the sense that they invariably leave a trace in the system, especially when intentionally observed. Moreover, there is no common way to evaluate the techniques, on how resistant they are against various types of attacks and how good they are in preserving privacy and protecting the identity of the user. The degree of anonymity measurement is a good start but it is not a single comprehensive answer.

Anonymisation technology is still very much an interesting and active field of research. The initial mix-system paper by Chaum that sparked off the network anonymity research was barely 30 years ago, and the demand for privacy and anonymity is only increasing. We have looked at some anonymisation techniques available today, their strengths and weaknesses, and ways to measure anonymity.

References

[1] J. Palme and M. Berglund, "Anonymity on the internet," http://people.dsv.su.se/~jpalme/society/anonymity.html, 2002.

[2] J. D. Wallace, "Nameless in cyberspace: Anonymity on the internet," in *CATO Institute Briefing Papers*, no. 54, dec 1999.

[3] D. L. Chaum, "Untraceable electronic mail, return addresses, and digital pseudonyms," *Communications of the ACM*, vol. 24, no. 2, pp. 84–88, Feb. 1981.

[4] B. Pfitzmann and A. Pfitzmann, "How to break the direct RSA-implementation of mixes," in *Advances in Cryptology (Eurocrypt '89), LNCS*, vol. 434. Springer-Verlag, 1990, pp. 373–381.

[5] U. Möller, L. Cottrell, P. Palfrader, and L. Sassaman, "Mixmaster Protocol — Version 2," IETF Internet Draft, Jul. 2003.

[6] B. Pfitzmann, "Breaking and efficient anonymous channel," in *Advances in Cryptology (Eurocrypt '94), LNCS*, vol. 950. Springer-Verlag, 1994, pp. 332–340.

[7] L. Sassaman, B. Cohen, and N. Mathewson, "The pynchon gate: A secure method of pseudonymous mail retrieval," in *Proceedings of the Workshop on Privacy in the Electronic Society (WPES 2005)*, Arlington, VA, USA, Nov. 2005.

[8] B. Chor, O. Goldreich, E. Kushilevitz, and M. Sudan, "Private information retrieval," in *Proceedings of the IEEE Symposium on Foundations of Computer Science*, 1995, pp. 41–50.

[9] D. M. Goldschlag, M. G. Reed, and P. F. Syverson, "Hiding Routing Information," in *Proceedings of Information Hiding: First International Workshop*, R. Anderson, Ed. Springer-Verlag, LNCS 1174, May 1996, pp. 137–150.

[10] R. Dingledine, N. Mathewson, and P. Syverson, "Tor: The second-generation onion router," in *In Proceedings of the 13th USENIX Security Symposium*, Aug. 2004, pp. 303–320.

[11] M. Reiter and A. Rubin, "Crowds: Anonymity for web transactions," *ACM Transactions on Information and System Security*, vol. 1, no. 1, Jun. 1998.

[12] I. Clarke, O. Sandberg, B. Wiley, and T. W. Hong, "Freenet: A distributed anonymous information storage and retrieval system," in *Proceedings of Designing Privacy Enhancing Technologies: Workshop on Design Issues in Anonymity and Unobservability*, Jul. 2000, pp. 46–66.

[13] *RFC 791 Internet Protocol - DARPA Inernet Program, Protocol Specification*, Internet Engineering Task Force, Sep. 1981. [Online]. Available: http://tools.ietf.org/html/rfc791

[14] K. Bennett and C. Grothoff, "GAP – practical anonymous networking," in *Proceedings of Privacy Enhancing Technologies workshop (PET 2003)*, R. Dingledine, Ed. Springer-Verlag, LNCS 2760, Mar. 2003, pp. 141–160.

[15] I. Goldberg and D. Wagner, "TAZ servers and the rewebber network: Enabling anonymous publishing on the world wide web," *First Monday*, vol. 3, no. 4, Aug. 1998.

[16] R. Fielding, J. Gettys, J. Mogul, H. Frystyk, L. Masinter, P. Leach, and T. Berners-Lee, *RFC2616 Hypertext Transfer Protocol – HTTP/1.1*, Internet Engineering Task Force, 1999. [Online]. Available: http://tools.ietf.org/html/rfc2616

[17] M. Rennhard and B. Plattner, "Introducing morphmix: Peer-to-peer based anonymous internet usage with collusion detection," in *In Proceedings of the Workshop on Privacy in the Electronic Society*, 2002, pp. 91–102.

[18] W. Townsley, A. Valencia, A. Rubens, G. Pall, G. Zorn, and B. Palter, *RFC2661 Layer Two Tunneling Protocol "L2TP"*, 1999. [Online]. Available: http://tools.ietf.org/html/rfc2661

[19] K. Hamzeh, G. Pall, W. Verthein, J. Taarud, W. Little, and G. Zorn, *RFC2637 Point-to-Point Tunneling Protocol*, 1999. [Online]. Available: http://tools.ietf.org/html/rfc2637

[20] "Anonymizer privacy policy," Oct. 2008. [Online]. Available: http://www.anonymizer.com/company/legal/privacy_policy.html

[21] A. Pfitzmann and M. Hansen, "A terminology for talking about privacy by data minimization: Anonymity, unlinkability, undetectability, unobservability, pseudonymity, and identity management, v0.33," http://dud.inf.tu-dresden.de/Anon_Terminology.shtml, Apr. 2010.

[22] G. Danezis and C. Díaz, "A survey of anonymous communication channels," Microsoft Research, Tech. Rep. MSR-TR-2008-35, Jan. 2008.

[23] A. Serjantov and G. Danezis, "Towards an information theoretic metric for anonymity," in *Privacy Enhancing Technologies*, Apr. 2002, pp. 41–53.

[24] C. Díaz, S. Seys, J. Claessens, and B. Preneel, "Towards measuring anonymity," in *Privacy Enhancing Technologies*, ser. Lecture Notes in Computer Science, R. Dingledine and P. F. Syverson, Eds., vol. 2482. Springer, 2002, pp. 54–68. [Online]. Available: http://dblp.uni-trier.de/db/conf/pet/pet2002.html#DiazSCP02

[25] C. E. Shannon, "A mathematical theory of communication," *Bell System Technical Journal*, vol. 27, pp. 379–423, 625–56, Jul, Oct 1948.

[26] C. Díaz, "Anonymity metrics revisited," in *Anonymous Communication and its Applications*, ser. Dagstuhl Seminar Proceedings, S. Dolev, R. Ostrovsky, and A. Pfitzmann, Eds., no. 05411, 2006.

[27] J. Douceur, "The Sybil Attack," in *Proceedings of the 1st International Peer To Peer Systems Workshop (IPTPS 2002)*, Mar. 2002, pp. 251–260.

[28] A. Serjantov, "On the anonymity of anonymity systems," Ph.D. dissertation, University of Cambridge, Jun. 2004.

[29] P. Mittal and N. Borisov, "Shadowwalker: peer-to-peer anonymous communication using redundant structured topologies." in *ACM Conference on Computer and Communications Security*, E. Al-Shaer, S. Jha, and A. D. Keromytis, Eds. ACM, 2009, pp. 161–172. [Online]. Available: http://dblp.uni-trier.de/db/conf/ccs/ccs2009.html#MittalB09

[30] J. Wright, S. Stepney, J. A. Clark, and J. Jacob, "Formalizing anonymity a review." in *University of York Technical Report*, no. YCS 389, Jun. 2005.

[31] J. Camp and C. Osorio, "Privacy-enhancing technologies for internet commerce," Harvard University, John F. Kennedy School of Government, Working Paper Series rwp02-033, Aug. 2002. [Online]. Available: http://ideas.repec.org/p/ecl/harjfk/rwp02-033.html