Comparison of Low-Latency Anonymous Communication Systems - Practical Usage and Performance

Thorsten Ries, Andriy Panchenko, Radu State and Thomas Engel

Interdisciplinary Centre for Security Reliability and Trust
University of Luxembourg
Email: {thorsten.ries, andriy.panchenko, radu.state, thomas.engel}@uni.lu

Abstract

The most popular system for providing practical low-latency anonymity on the Internet is Tor. However, many other tools besides Tor exist as both free and commercial solutions. In this paper, we consider five most popular low-latency anonymisation services that represent the current state of the art: single-hop proxies (Perfect Privacy and free proxies) and Onion Routing based solutions (Tor, I2P, and JonDonym). We assess their usability and rank them in regard to their anonymity. We also assess their efficiency and reliability. To this end, we define a set of metrics and present extensive measurements based on round-trip time, inter-packet delay variation and throughput. Apart from the technical realization, economic aspects are also crucial for anonymous communication systems. In order to attract more users, which is mandatory in order to improve anonymity per se, systems need to exhibit a certain payoff. We therefore define an economic model that takes all relevant aspects into consideration. In this paper, we describe the results obtained, lessons learned, and provide guidance for selecting the most appropriate system with respect to a set of requirements.

1 Introduction

For various reasons, people want to protect their identity when communicating over the Internet. Doing so, they protect their privacy. Freedom of expression may be one motivation, while another reason may be a company or customer with the need to stay anonymous[1] for certain business transactions.

Based on this need, the aim of this paper is to compare existing implementations of anonymising systems with respect to users' requirements such as performance and usability, also taking into account aspects of anonymity and security as well as the *real costs*, i.e., monetary costs the user faces. To this end, we assessed five tools that represent the different approaches and the current state-of-the-art in practical anonymisation: free proxies, Perfect Privacy [2], JonDonym [3], Tor [4], and I2P [5].

In recent years, research in anonymity has been very active, with many approaches developed. However, only a very few of these reached wide-scale deployment and are used in practice. The predominant system in use today is Tor, developed by Dingledine et al. (2004). Tor is considered to be a low-latency anonymisation tool, which means that data is supposed to be delivered within a reasonable time, allowing the usage of interactive applications such as web browsing.

In contrast, high-latency systems such as Mixmaster and Mixminion, developed by Moeller et al. (2003) and Danezis et al. (2003) respectively, provide a high degree of anonymity and should be considered for exchange of "more sensitive" information. As a drawback, communications like anonymous web browsing would not practically possible because of the long delays. Beside these, several other anonymisation tools exist as both free or commercial solutions following different design approaches; current low-latency approaches can basically be divided into single-hop proxies and Onion Routing approaches, initially invented by Reed et al. (1998).

The easiest solution to hiding the identity of a user is the use of a single proxy server. Traffic is routed through a node that strips off the origin IP address and presents its own instead. The main problem of single-host proxies is that they are a single point of failure in regard to availability and trust.

The current step in the evolution of anonymous networks is Onion Routing, where messages are encrypted in layers and sent from one node to the next. At each hop one layer of encryption is removed (or added, depending on the direction) and the result further forwarded.

Further, users need to distinguish between services where one entity operates both the anonymisation nodes, and the information service (e.g., Perfect Privacy) and services where nodes can be operated by independent third parties (e.g., Tor, I2P).

However, independent of the used anonymisation technique, users' identities may still be discovered using other techniques such as information leakage at the application layer. This can be accomplished through analysis of the HTTP headers or by intersection attacks, using language or font presets for instance as proved by Raymond (2000) and Wright et al. (2003). Therefore, either a service to alter HTTP header information should be provided by proxy service operators, or it is recommended to use filtering proxy on the user side before sending the data to the anonymisation network.

This superficial classification of anonymisation systems already shows the complexity a user faces deciding upon an appropriate solution. During this selection process, several aspects are usually considered. In addition to the most important aspect, the degree of anonymity and performance plays a large

[1] The term *anonymity* derives from the Greek word $\alpha\nu\omega\nu\upsilon\mu\iota\alpha$ and means "without a name" or "namelessness"

[2] http://www.perfect-privacy.com

[3] https://anonymous-proxy-servers.net/en/index.html

[4] http://torproject.org

[5] http://www.i2p2.de/

role, as do reliability, usability, and economic aspects (see Fig. 1).

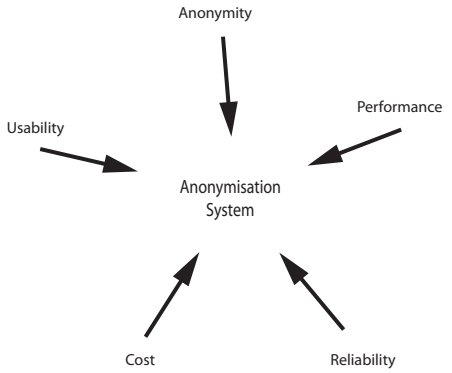


Figure 1: Aspects of systems selection

In this paper, we examine all relevant incentives and combine them to provide usage guidance on already-deployed anonymisation networks by classifying the systems and showing their strength and weaknesses. Applying this guidance, users can select the anonymisation service that best suits their needs in a concrete situation.

The remainder of this paper is organized as follows: firstly, we give an overview of the tools we compare. Section 2 describes the anonymisation systems we used for comparison, and is followed by an overview of related work (Section 3). In Section 4, we briefly examine usability in regard to its impact on acceptance of the tools. Further, we measure and evaluate the performance in terms of round trip time (RTT), Inter-Packet Delay Variation (IPDV), and throughput. We classify anonymisers in regard to their efficiency in Section 5 before addressing practical issues of anonymity and security in Section 6. In Section 7, we calculate and discuss the aspect of reliability, before all relevant aspects are combined to describe the economic impact on a user's decision process in Section 8. Finally, Section 9 concludes with the lessons learned and future work.

2 Anonymisation Systems

The simplest way of hiding someone's identity is to use of a proxy server. The receiver of the message only gets the IP address of an intermediate server, not of the sender. The main drawback is that adversaries can easily de-anonymise users by compromising a server or simply providing one. However, this service may still assure a basic level of anonymity. Due to the simple setup, proxy servers are very common, either as free or commercial solutions and can be easily found in the Internet. The providers of these proxy servers are mostly unknown, so one does not know how trustworthy they actually are. Commercial services exist too, such as Perfect Privacy, which currently provides 48 servers in 23 countries worldwide, allowing users to choose either their preferred proxy or a self-defined cascade of proxies, meaning that several proxy servers are combined into a chain. This may increase anonymity and security against an external adversary, but still has the drawback that the service as a whole is operated by a single entity. Perfect Privacy offers a variety of ways of connecting: users can simply use the servers as an HTTP- SOCKS proxy[6]. In addition, users can connect to the proxy server via OpenVPN, PPTP VPN, or SSH. As long as the user does not use their own layer of encryption, the traffic from the proxy server to the destination is not encrypted and consequently completely visible to the server provider. This is true for all anonymisation tools presented here.

Another low-latency anonymisation approach provides the possibility of active mixing[7] of the traffic together with Onion Routing. A popular example of using this approach is JonDonym. Started as an open source project at the TU Dresden, JonDonym (formerly known as JAP) became a popular tool to gain anonymity in the Internet. Users can choose between several fixed paths, known as cascades, with nodes provided by JonDonym operators and nodes operated by other organizations or individuals. Currently, there are 34 nodes in the network forming 16 cascades. The operators of JonDonym provide two kinds of service: a free service, usually having two nodes in a cascade with several hundreds users and a commercial service with usually three nodes in a cascade. Compared to the free service, the number of concurrent participants is relatively low (less than 100). Even though traffic mixing is supported in this approach, to the best of our knowledge, it is not activated because of performance issues.

Today's most widely used anonymisation system is Tor. Also based on onion routing, Tor tries to provide an acceptable degree of anonymity, while allowing the use of interactive web applications. Recently, Dingledine (2009) showed that Tor has about 300,000 users daily and about 2,000 relaying nodes. The main difference from JonDonym is its volunteer-based node operation. In order to achieve optimal system performance, Tor currently relies on directory servers, which gather all relevant information about the network and provide information about the performance of nodes to the clients.

I2P is a system similar to Tor and JonDonym. In contrast to JonDonym and Tor, the main objective of I2P is communication within its own network and not with external services. As a consequence, there is currently only one outbound HTTP gateway responsible for all outgoing web traffic. Another difference from Tor and JonDonym is its fully-distributed network, which has no centralised server for coordination and organisation. Hence, the network consists of a set of nodes that communicate with each other in order to achieve anonymity. All traffic is encrypted using *garlic encryption*, which combines multiple messages into one single message to make traffic analysis more difficult.

3 Related Work

In the many years since the establishment of the Internet, network performance has been an extensive field of research, showing different issues and optimizations in a large number of publications, e.g., by Keshav (1999). In recent years, logical networks, also known as overlay networks, were introduced to allow the easy creation of additional network services without modification of the underlying network. These have became a popular topic of interest in network research and shifted several network paradigms to the application layer. Peer-to-peer networks and other overlay network topologies were introduced to improve data exchange or to add additional functionality. Among these is anonymity, which elicited so much interest, that a special field of research, *anonymous communication*, was established.

Several surveys on anonymous communication systems exist, e.g., conducted by Kelly (2009) or Ren

[6] SOCKS is an Internet Protocol to facilitate routing of packets using a proxy server

[7] actively delaying and batching messages

et al. (2009). In the work of Pries et al. (2008), in which the authors describe the concepts of basic anonymous communication, as well as implemented systems, the need of low-latency anonymous communication systems is highlighted. However, most surveys focus on MixNet based schemes based on the approach of Chaum (1981) for anonymous remailers and Onion Routing (particularly on Tor); minor work has been conduction on other network routing-based techniques like Crowds and P2P networks such as Tarzan, which was developed by Freedman et al. (2002) and MorphMix, an approach by Rennhard et al. (2002). Due to its widespread usage with about 300,000 users daily, existing performance measurements in anonymous communication mainly concentrate on Tor. The main objective is the improvement of performance, for instance using alternative methods of path selection. Very often, authors of related publications concentrate on throughput improvements in Tor and either propose algorithms to achieve higher performance *or* higher anonymity as shown by Snader et al. (2008). In contrast, the importance of latency in anonymisation networks as performance metric is highlighted by Murdoch et al. (2008). However, both publications consider only a single property, while our study combines these with the variance to determine the overall performance.

Other relevant matters in choosing the appropriate anonymisation system are rarely considered. The optimal system needs to be reliable, and also has to be usable and cheap. Economic aspects are covered by Acquisti et al. (2003) to build a general model in order to describe the incentives for participation in anonymous networks. This approach was elaborated by Ngan et al. (2010), going one step further and describing incentives for relaying traffic within Tor with the aim of an overall performance improvement.

The work of Dingledine et al. (2006) emphasizes the usability and the network effect in anonymisation networks. The authors argue the importance of usability to increase the user base and, consequently, on the achievable anonymity. Related to both usability and performance is the time needed for sending and receiving messages. Even when just surfing in the Internet, users expect an appropriate performance. If these expectations are not met, users will most likely not use the service. Various studies have attempted to find out the maximum tolerable time for loading a website. Different numbers can be found in literature, depending on the culture, etc., but recent studies studies, e.g., by Kopsel (2006) and Wendolsky et al. (2007) conclude that about four seconds is a maximum tolerable delay for most users.

To the best of our knowledge, to date there has been no practical comparison of all relevant aspects (degree of anonymity, performance, usability, reliability, and cost) of already deployed low-latency anonymisation tools. This paper aims to close this gap.

4 Usability

As already mentioned, usability is a crucial aspect since it is essential to attract more users, which is a prerequisite for improving anonymity. The higher the number of participants, the better the theoretical anonymity due to the increased size of the anonymity set (as in the work of Pfitzmann et al. (2009)). Consequently, providers of anonymity services aim to have a high number of users, which, which incurs the cost of a degradation in the system's performance. However, even before evaluating the systems' features, the user informally evaluates the usability of the anonymisation system during installation and initial configuration. This is of particular importance, as she may already form a negative opinion of the system and may reject its further use.

To evaluate usability, we use the *cognitive walkthrough* (CW) method, developed by Wharton et al. (1992). Hereby users try to accomplish tasks with the aim of identifying usability issues. The particular evaluation was divided into three steps:

1. CW1: Installation of the anonymising software.
2. CW2: Configuration of the browser/other software.
3. CW3: Verification of the anonymised connection.

In the following, we describe these steps more in detail.

CW1: Installation of the anonymising software

Although some prerequisite software installations may be challenging to inexperienced users, all systems provide well documented websites to support users during the installation process. Very often, step-by-step instructions are given, which vary from a simple double-click (JonDonym) to some more advanced configuration being necessary (Tor and I2P).

CW2: Configuration of the browser/other software

As already mentioned, we tested both free and commercial systems. The two commercial systems, Perfect Privacy and the premium service of JonDonym, have to be paid for in advance. This can by done by credit card or anonymously by using vouchers (see Section 6). Thus, the process of paying makes some additional effort necessary, but is relatively easy to handle.

The aim of Tor is to protect data transport. For web browsing, there are no specific measures to hide potentially unmasking information as such as browser type, language settings, and so forth, which is sent by default to the web server. Therefore the developers highly recommend the installation of a local proxy server that modifies or deletes this information before sending the data. After the installation of the local proxy server, the final step is the same for all other tested systems: the users have to configure the application (in this case, the browser) in order to use a proxy server. Depending on the browser, the step of proxy configuration may be difficult for a less sophisticated user the first time because of the sometimes not obvious location of these settings. Only Tor simplifies this process by installing an add-on (*Torbutton*) that that allows the proxy settings to be easily switched on and off.

CW3: Verification of the anonymised connection

Once the user has configured the browser or the additional software, she needs to verify whether the anonymisation service is running properly. On dedicated web sites that reveal the IP address of the connecting user, it is easy to check the system's functionality. Some of these web sites[8] provide additional information about security/anonymity issues, like the connecting IP address, HTTP header information and whether Java/Javascript is turned on in the browser. Except for finding an appropriate website, this step was found to be relatively easy to accomplish.

[8] E.g., http://test.anonymity.com

5 Performance

Probably the most important aspect for users on the Internet, even when acting anonymously, is performance. In particular, Round Trip Time (RTT), Inter-Packet Delay Variance (IPDV) and throughput have a significant influence on the overall performance as perceived by users. Because this has a direct impact on the user's satisfaction, we examine these parameters in detail and calculate the overall efficiency.

5.1 Testbed environment

All measurements described in this section were performed between a client (running Ubuntu 10.04, Intel Core2 Duo, 3GHz, connected at 100Mbit/s to the campus network and with 300Mbit/s to the outside world) located at the University of Luxembourg and two web servers, one located in Luxembourg and one on St. Vincent Island. The basic measurement setup is depicted in Figure 2. In order to allow the comparison of all tools under the same conditions, we used the HTTP protocol as the least common denominator supported by all tools.

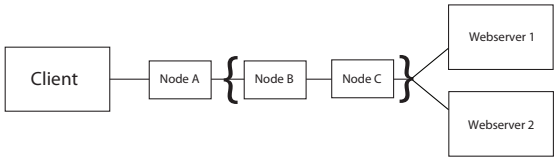


Figure 2: Testbed setup with either one proxy or a chain of intermediate nodes

For RTT measurements, we chose the Apache web server benchmarking tool[9]. It allows measurement of the time a request needs to get about 200 bytes from an HTTP server. Even though this approach involves a certain overhead, it allows a relative comparison of the systems. In order to consider time-shifts and varying network usage at different times of a day, we repeated the measurements over six days.

Measurements of IPDV were conducted every minute over a period of four days, using a dedicated client-server application. We measured the inter-arrival time between every sequence sent with a one second interval in between. The main motivation of the following measurements is the question of whether it is possible to use applications such as VoIP over the anonymisation systems.

Finally, we measured the throughput for three consecutive days using GNU *wget*[10]. We chose to download files of two sizes (100KB and 1MB) to examine the interaction between the amount of the transferred data and the TCP slow start algorithm. We used these file sizes to identify differences between small and large files based on a recent report that states that the average size of a web site is 320KB (Google (2010)). Thus, we cover cases of both smaller and larger files.

All measurements were performed using the already described anonymisation tools, applying the following settings:

- Free proxies (FP) were chosen from a web page listing free proxy servers ranked by their performance[11]. As these servers typically have a high fluctuation, we had to switch between servers during the test, causing significant downtimes (see Section 7).
- Perfect Privacy (PP) currently offers servers at 23 locations worldwide. Some locations provide only a single server, others up to eight for the purpose of load balancing. We used three randomly chosen nodes out of 48, located in Amsterdam, Moscow, and Chicago.
- JonDonym, using three different random premium service cascades (out of nine), having three nodes each. Measurements were not performed using free cascades because the user limit is often reached and, consequently, the service continually becomes unavailable.
- Tor with its default configuration, changing circuits at least every 10 minutes.
- I2P, which also changes internal paths every 10 minutes, but uses always the same single outbound server with estimated 1,000 concurrent users [12].

In addition, we performed the same measurements without any anonymisation tool. This information serves as the reference value to calculate the efficiency and performance losses of anonymisation tools.

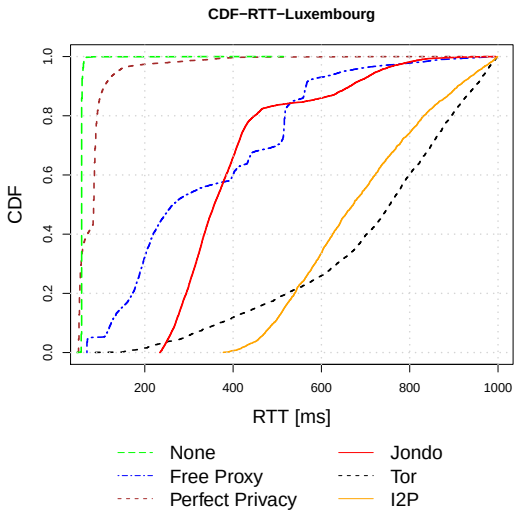


Figure 3: CDF Luxembourg

5.2 Round Trip Time

Network latency and RTT have a fundamental impact on end-to-end performance in computer networks. Voice over IP (VoIP) applications for instance require a RTT of less than 600ms[13] to provide adequate quality.

Evaluating our measurements, significant differences were identified: the commercial approaches, Perfect Privacy and JonDonym, show the lowest average RTTs together with the free proxy, while Tor and I2P are significantly slower by a factor of three to four. The Cumulative Distribution Function (CDF) plots (Fig. 3 and 4) show the fraction of measurements of RTT that are below a certain value. Taking Tor and I2P as an example, the tests on the server in Luxembourg show that Tor can achieve lower RTTs, but between 550ms and 1s, I2P

[9] http://httpd.apache.org/docs/2.0/programs/ab.html
[10] http://www.gnu.org/software/wget/
[11] http://proxy.speedtest.at/proxybyPerformance.php?offset=0
[12] http://stats.i2p.to/
[13] http://www.itu.int/itudoc/itu-t/aap/sg12aap/history/g.114/g114.html

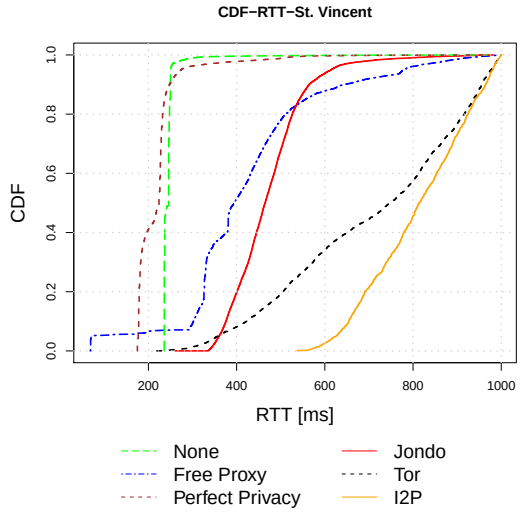
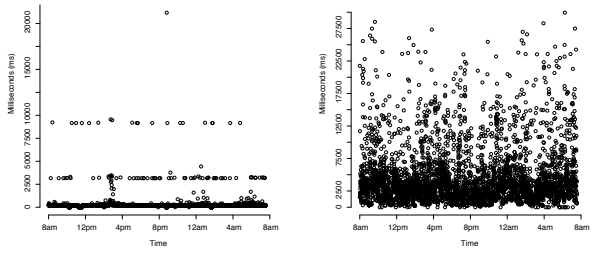


Figure 4: CDF St. Vincent



(a) Perfect Privacy (b) Tor

Figure 5: RTTs measured during one day using Perfect Privacy(a) and Tor(b).

performs better, meaning that, for instance in 60% of the measurements, Tor showed a RTT of about 800ms, while I2P achieved about 720ms. The results show also that VoIP is only possible with Perfect Privacy (Amsterdam), FP, and JonDonym with some restrictions, because their RTTs are less than 400ms for at least 80% of the measurements.

During our measurements, Perfect Privacy shows three distinct levels of RTTs with each level fluctuating in only a narrow band of a few milliseconds (Fig. 5(a)). Due to the usage of TCP packets, this pattern is most likely created by packet retransmits. The timeout of TCP packets on GNU/Linux is 3000ms and would explain the very constant additional delays. This suggests that there was congestion on the communication line or the proxy server. The same behaviour has been observed on other Perfect Privacy proxy servers as well. Tor instead shows a wider variance of RTT values (see Fig. 5(b)) due to the variety of possible circuits. Possible retransmits are not easily detectable in this plot.

5.3 Inter-Packet Delay Variation

Interactive real-time applications such as VoIP depend heavily on a constant IPDV. While multimedia streaming applications can compensate differing IPDV by the use of buffers, this is not possible in VoIP. In the sense of Quality of Service of VoIP, IPDV should be <100ms to avoid distortion[14].

Figure 6(a) shows the IPDV observed at both

[14] http://www.gig-ip.com/help/voip_and_qos_sensors.htm

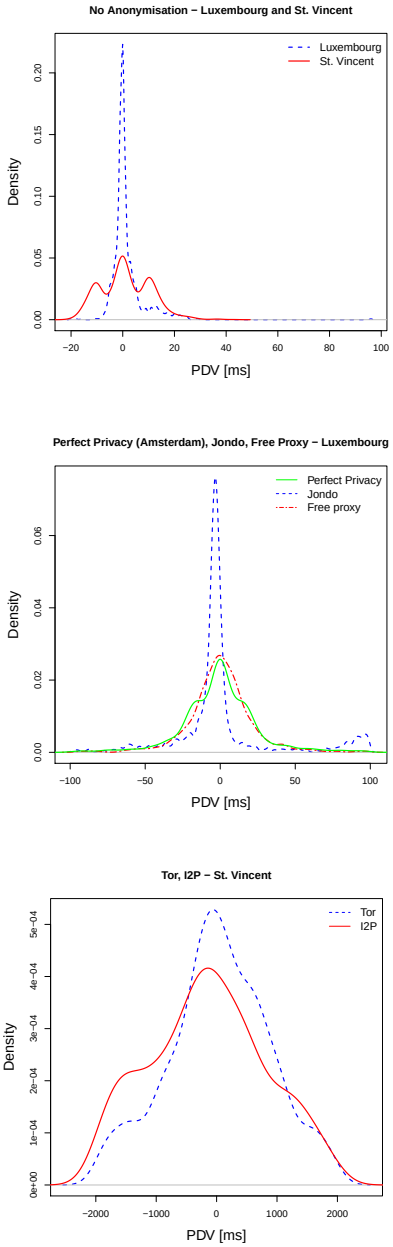


Figure 6: Inter-Packet Delay Variations

servers without anonymisation service. The server in St. Vincent has a smoother and wider distribution, probably caused by the longer distance between client and server, compared to the server in Luxembourg. However, Figure 6 suggests that apart from Perfect Privacy and JonDonym, no other anonymisation service would be able to comply with the recommended value of IPDV for VoIP applications. The values for Tor and I2P are far too high for this kind of communication (Fig. 6(c)), most likely due to congestion. Perfect Privacy, the free proxy, and JonDonym in particular provide a IPDV of less than 50ms (cf. Fig. 6(b)), and satisfy the requirements for carrying VoIP traffic.

5.4 Throughput

In order to evaluate application performance within different scenarios, we measured throughput while transferring files with the sizes of 100KB and 1MB. Due to similarities of the results, only the throughput results to St. Vincent server are shown here.

The first CDF graph (Figure 7(b)) shows the throughput of an anonymisation system while trans-

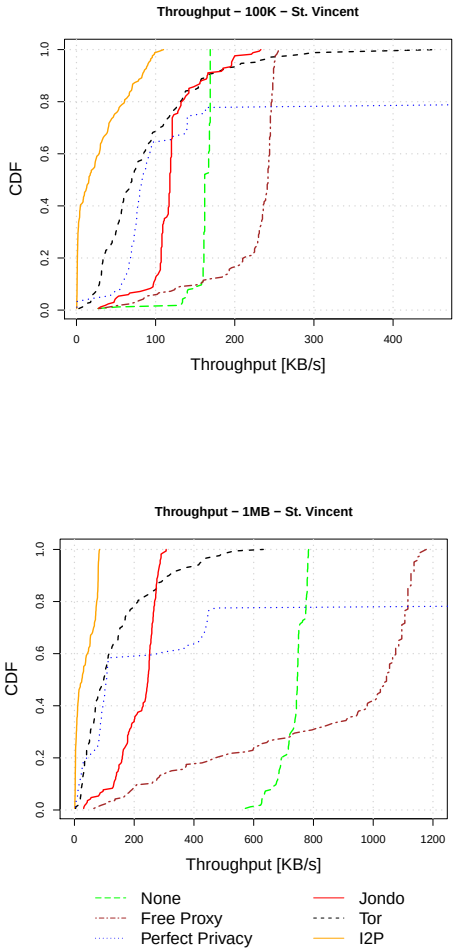


Figure 7: CDF Throughput

ferring 1MB of data. I2P shows generally the lowest throughput, while the maximum throughput was achieved by the Perfect Privacy proxy in Amsterdam (1044KB/s). This value was even higher than the throughput measured without anonymisation (746KB/s). This is an astonishing result, because our tests were conducted utilising the HTTP header option *no-cache*, so that there should be no caching on the proxy server. The only explanation could be the usage of a more powerful connection via Amsterdam compared to the native connection.

As can be seen in Figure 7(a), the values for transmission of 100KB files were significantly lower, for Perfect Privacy and without anonymisation tool, by a factor of 4 - 4.5. This effect may result from TCP slow start, when the hosts involved try to achieve the highest throughput for this particular connection by adapting TCP window size. In Tor and I2P, the throughput is almost constant for the different file sizes. The reason for this behaviour, again, is presumed to be congestion within the network.

5.5 Systems efficiency

Even though low-latency anonymisation systems in general provide the possibility is being used for interactive applications such as surfing on the Internet, the question remains how they compare to the behaviour and performance without any anonymisation. In this section, we calculate the efficiency of RTT and throughput. We did not include IPDV because as long as the required threshold value is not exceeded (e.g., 100ms for VoIP), the actual value is of no particular interest to the average user.

For the throughput, we calculate the efficiency (E_T) for the transfer of 100KB and 1MB data respectively as ratio of the mean throughput ($T_{mean}(AS_i)$) of the corresponding anonymisation system AS_i to the throughput ($T_{no\text{-}anon}$) measured without anonymisation tool (161KB/s and 734KB/s) using eq. 1.

$$E_T(AS_i) = \frac{T_{mean}(AS_i)}{T_{no-anon}} \quad (1)$$

RTT efficiency (E_{RTT}) is calculated similarly to E_T: it is the ratio of the RTT without anonymization (RTT_{no_anon}) and the mean RTT of AS_i ($RTT_{mean}(AS_i)$).

$$E_{RTT}(AS_i) = \frac{RTT_{no-anon}}{RTT_{mean}(AS_i)} \quad (2)$$

To further guide users' decisions, Table 1 lists the efficiencies of throughput and RTT. The value of 1 is the reference value, accomplished without anonymisation and the higher the value, the more efficient is the system.

Apart from I2P, all tools show an acceptable efficiency for 100KB files, but the rates decrease tremendously for the large files. Here, JonDonym and Tor show throughput efficiencies of less than 30%, I2P only 5%.

RTT total efficiency values are calculated as the mean of Luxembourg and St. Vincent. Compared to bandwidth efficiencies, these values are even worse. While again the Perfect Privacy server in Amsterdam performed well, all other E_{RTT} values are below 40%. Altogether, these figures show that current anonymisation systems still suffer from poor performance.

Table 1: Throughput and RTT Efficiency

Anonymisation System	E_T (100KB)	E_T (1MB)	E_{RTT}
Free Proxy	1.19	0.88	0.27
PP Amsterdam	1.37	1.17	0.89
PP Moscow	0.88	0.71	0.37
PP Chicago	0.85	0.37	0.40
JonDonym	0.75	0.30	0.36
Tor	0.55	0.19	0.07
I2P	0.18	0.05	0.08

6 Anonymity and Security

Anonymity may be quantified using different metrics, as a survey of Kelly et al. (2008) shows, but none of which is comprehensive. In this section, we establish a thorough classification of anonymity for all tested anonymisation services. Because a quantitative comparison of all services is difficult up to impossible (as there is no existing metric that would consider all possible attacks) using existing approaches such as entropy, which is described by Diaz et al. (2002), we performed an educated anonymity/security appraisal and ranked attackers in regard to their ability and costs to de-anonymise users. This ranking is based on our subjective assessment and may differ from other classifications.

The idea is simple: we identified the different roles of adversaries against systems for anonymous communication. We then ranked these adversaries with respect to their power. In order to quantify the anonymity, we ranked the power of an adversary on a scale between 0 and 1. The value of 1 means that the adversary can de-anonymise the involved entities with a high probability, whereas the value of 0 means that the adversary is generally harmless with respect

to the considered anonymisation technique. For instance, while a web service provider has limited power to identify a user coming from an anonymisation network, an Internet exchange (IX) and a Government has much greater power and abilities. Figure 8 shows the results of our appraisal.

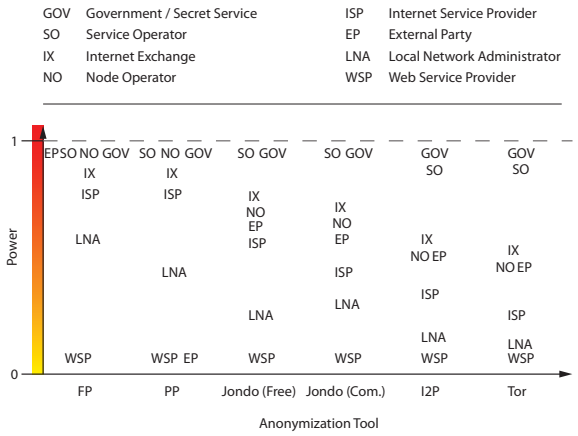


Figure 8: Classification of anonymisation systems

No single low-latency anonymisation technique can provide an adequate protection against an attacker having a government or anonymisation service operator status. Therefore, users of anonymisation systems are required to trust the service operator. Using for instance Tor, users get good protection against the Web Service Provider, the Local Network Administrator (LNA), as well as the ISP. This is due to the encryption used between the sender and the first Tor node. A node operator and the External Party (EP) have some more power, as they can add as many nodes to the network as they have resources. Here, an External Party is defined as an entity outside the anonymisation system, that is trying to become a part of it. Hence, every other entity we consider in our categorization can be seen as an EP too.

An even more powerful attacker is the Internet Exchange, as it can observe a considerable amount of traffic between the Tor nodes. Recent studies, e.g., by Edman et al. (2009) show that there is a certain risk that provider of large Autonomous Systems (AS) can control a significant number of entry and exit nodes, hence this is also true for the corresponding Internet Exchange. Service Operator and Government or Secret Service are the most powerful players. They may have enough power to bias path selection, analyse all network traffic, break the encryption, or even apply non-technical means to achieve their goal (e.g., rubber-hose cryptanalysis[15]). We also differentiate between the the two available versions of JonDonym. While the free version provides a path-length of two nodes, the premium services always use three nodes. However, we rank the LNA higher for the premium service due to simplified fingerprinting, as proved by Panchenko et al. (2009) because of a smaller number of users.

Calculating the degree of anonymity using this classification, single values are weighted, summed and normalized:

$$A = 1 - \frac{\sum_{i=1}^{n} (w_i * a_i)}{\sum_{i=1}^{n} w_i}, \qquad (3)$$

where a_i is the power of attacker i, w_i the weight an user puts on the attacker a_i and n is the number of attackers considered (here: eight). Table 2 shows the degree of anonymity of the tested anonymisation services for two particular cases. Case 1 (C1) shows the values without consideration of any user-based weighting; C2 could be an employee using services like eBay at the workplace. Here, we mainly consider the LNA and the WSP as critical, i.e., giving them a higher weight (LNA: 10, WSP: 3). The other identities are weighted 1. Overall, Tor achieves the highest degree of anonymity and the free proxy the worst. The degree of anonymity of Tor is even higher than in I2P, mainly given because of the single outbound node I2P provides. Due to the higher ranking of a LNA, the free version of JonDonym may be more appropriate in C2.

Table 2: Degree of anonymity

Anonymisation System	A (C1)	A (C2)
Free Proxy	0.18	0.32
Perfect Privacy	0.26	0.42
JonDonym	0.33	0.54
JonDonym	0.36	0.53
I2P	0.44	0.62
Tor	0.47	0.66

In addition to the degree of anonymity, other aspects, such as anonymous payment for the use of commercial anonymisation tools, are of relevance too, as they may directly influence the anonymity. For example, providing the real name and/or the bank account number would reveal the identity of the user to the company offering an anonymisation service.

The commercial service providers in this comparison, Perfect Privacy and JonDonym, offer an alternative by also accepting payments by anonymous payment schemes such as PaySafeCard[16] or UKash[17]. Users can anonymously buy a code in an ordinary shop and to pay for the anonymisation service with this code as with pre-paid telephone cards, but without any personal registration being required. Another possibility for ensuring anonymity during the payment process is the usage of anonymous credit cards, which work either like pre-paid cards or like gift cards. Pre-paid cards need to be charged before usage, gift cards can be bought already containing a certain balance.

Considering the difference between free and commercial service operators, we cannot preempt the user's decision to which service is more trustworthy; users have to trust the operator in both cases. Only the operator's intention may vary, and range from commercially-driven to belief in expression of freedom or the hope of creating a trap to harvest sensitive information.

7 Network Reliability

The next essential aspect, which is particularly important for user satisfaction, is the reliability of the network. We assess it in terms of the *failure rate*. To

[15] Torture of a person to extract cryptographic secrets, e.g., a password

[16] http://www.paysafecard.com/

[17] http://www.ukash.com

calculate the failure rate, all *unanswered* RTT benchmarking requests were counted during the period of experiment execution.

A common parameter to describe the failure rate is MTBF, which expresses the Mean Time Between Failures of a system. In this context, MTBF is calculated as the sum of the uptime periods divided by the number of downtimes:

$$MTBF = \frac{\Sigma(t_{down} - t_{up})}{n_{down}} \quad (4)$$

where n_{down} is the number of failures.

We also calculated the Mean Time To Recovery (MTTR), which is computed in the same way as MTBF (Equation 5). In order to evaluate reliability, both factors need to be considered.

$$MTTR = \frac{\Sigma(t_{up} - t_{down})}{n_{up}} \quad (5)$$

Tables 3 and 4 show that the loss of RTT connections occurred by the free proxy, which is the result of proxy servers going off-line periodically from time to time. During our experiment, we twice had to switch to a new proxy server. Another issue is related to the connection to the webserver in St. Vincent, which showed problems for two hours when some of the packets did not get through. The relatively high packet loss of JonDonym was caused by a service interruption of more than two hours. This interruption only affected JonDonym traffic to the server in Luxembourg while all other services were working well, including the connections to St. Vincent. This result suggests that there was a problem of connectivity between the final JonDonym relay and the server. However, the numbers only present a snapshot and may not necessarily reflect the long-term behaviour.

Table 3: Snapshot of MTBF and MTTR - St. Vincent (SV)

AS	PL	MTBF (SV)	MTTR (SV)
None	0.33%	72:54:35	0:00:39
Free Proxy	2.22%	0:10:05	0:04:14
PPA	0.29%	1:20:56	0:01:04
PPM	0.37%	1:05:35	0:01:12
PPC	0.29%	1:21:14	0:01:12
JonDonym	0.36%	1:04:51	0:01:07
Tor	0.36%	0:49:49	0:01:00
I2P	0.56%	0:19:22	0:01:23

Table 4: Snapshot of MTBF and MTTR - Luxembourg (Lux)

AS	PL	MTBF (Lux)	MTTR (Lux)
None	0%	144:23:04	0:00:00
Free Proxy	2.06%	0:38:54	0:00:59
PPA	0.05%	1:02:06	0:01:00
PPM	0.06%	0:05:04	0:01:00
PPC	0.02%	2:45:55	0:01:10
JonDonym	0.76%	3:44:16	1:05:04
Tor	0.06%	1:30:04	0:04:44
I2P	0.22%	0:02:04	0:45:00

Fig. 9 illustrates the number of unanswered RTT requests during the measuring period. A high number of lost messages without the use of any anonymisation system is a sign of a general network problem. However, the figure also shows the influence of path selection on reliability. It is again possible to see a high

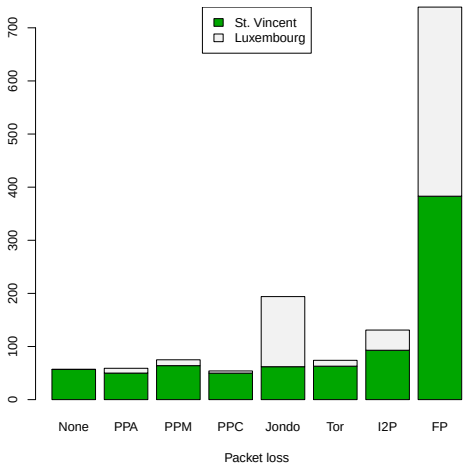


Figure 9: Lost packets during 6 days

number of lost packets for FP. This result confirms our observations during previous tests.

Another observation from reliability tests is the influence of the server location. Excluding the free proxy due its outages, connections to the server in St. Vincent summed up over all services show a higher loss in total (822) compared to Luxembourg (561). The reason for this issue may be a general network problem, not related to any anonymisation service. Normalising these results, the number of unanswered RTTs for all services, except I2P and the free proxy server, is quite low.

8 Economic aspects

Apart from the technical aspects of finding the appropriate anonymisation tool, users have also to decide on the economic value/cost. Some services rely on an active participation where users pay indirectly (C_i) by providing, e.g., computational resources. Using I2P for instance requires provision of bandwidth and computational power in order to use the network, while in Tor, users have the choice of either donating resources by acting as a relay node or as a client only. Paying indirect costs may be negligible in most cases, but may, on the other hand, limit the maximum achievable performance. We calculate C_i as the sum of C_r, C_e and C_b, where C_r is the relaying/routing costs, C_e costs of de- or encryption (computational effort) and C_b the costs of providing additional bandwidth.

$$C_i = \Sigma(C_r, C_e, C_b) \quad (6)$$

Using commercial anonymisation systems, users have to pay fees. These direct costs C_d are based on usage time or data volume. For instance the business model of Perfect Privacy is based on a monthly fee, offering a data flat rate. JonDonym instead bills according to the amount of data. Consequently, the overall costs are calculated as a sum of the two costs:

$$C = (C_i + C_d) \quad (7)$$

The payoff costs for every tool are then calculated as follows:

$$P_i = \frac{w_a A * w_e E * w_u U}{w_r R} - w_c C \quad (8)$$

where A is the degree of anonymity, E the efficiency, R the reliability and U the usability; $w_{\{a,e,u,r,c\}}$

are the different weights the user defines depending on her particular needs. Consequently, the calculation strongly depends on user requirements and has to take into account the actual situation.

At first glance, users main goal may be a high degree of anonymity together with a high efficiency and reliability of the system at low cost. This does not mean that users will not pay for such a service; statistics for Perfect Privacy show about 25,000 concurrent connections. Even though this might be not the accurate statistic, it still shows that a fast and reliable service can attract many users willing to pay a certain amount of money. Another example is Tor, with about 300,000 daily users. Tor is also known for its high anonymity and reliability, even at no cost, but with the drawback of poor performance. For the launch of future anonymisation systems, especially for commercial products, the operators need to take all these aspects into consideration.

9 Conclusion and future work

In this paper, we have defined a set of metrics in order to compare and evaluate five already deployed state-of-the-art anonymous communication systems in regard to their anonymity, performance, reliability, usability, and cost.

Besides the installation, which is relatively simple for all systems, the usability of the tools is generally good and should not be a reason for non-use. Usability does not vary not much from the users' point of view; they always have to configure their application, i.e., web browser, to use a proxy server, a process which is practically the same for all tools.

In order to provide a comprehensive comparison of the anonymisation tools, we ranked them in regard to the power of possible adversaries. Our classification is subjective and may vary from other opinions, but gives users an indication of strength and weaknesses of corresponding anonymisers. A future goal might be to simplify the presentation of the results and present them to users in a more appropriate way. This will be subject of further work. In addition, we measured throughput, RTT, IPDV and failure rate and calculated efficiencies. The results show that the proxy-based anonymisation systems outperform the Onion Routing approaches in throughput and RTT, but provide less anonymity. This trade-off applies to all systems and in the end, the user must decide which system best fits her requirements. However, web browsing is possible using all the tested tools, even though I2P in particular has long response times. Latency-critical applications like VoIP, which rely on highly responsive networks, are only usable to a certain extent with the the systems we examined.

An important finding is the efficiency of the throughput performance of single proxy solutions. They perform as well the *native* communication, sometimes even better.

We also observed that the selected anonymisation path and the recipient's location have a strong influence on performance and reliability. The Perfect Privacy proxy in Amsterdam, which outperformed the communication without anonymisation, demonstrates this. In general, all Perfect Privacy proxies we tested, as well as Tor, showed particularly reliability in terms of successful connections. While JonDonym and I2P were slightly less reliable, the most unreliable service was the free proxy service due to a high fluctuation of nodes. This demonstrates the main problem with single-proxy solutions but makes them applicable to high-performance short term downloads.

Economics in anonymity is still an under-investigated field of research, with only a few publications. In this paper, we show that besides real costs in terms of money, all relevant aspects such as performance, anonymity, reliability, usability, and cost, need to be evaluated in order to calculate system's payoff. However, as already mentioned, this calculation strongly depends on individual users' preferences.

To summarise our results, we established a comparison table, containing all examined anonymisation systems. We classified the systems in five groups, expressed on a scale of very good (*++*), good (*+*), average(*0*), bad(-) and very bad(- -). Table 5 shows this classification.

Table 5: Evaluation of anonymous communication systems

Anonymisation system	**V**	**A**	**E**	**R**	**C**
Free Proxy	0	- -	+	- -	++
Perfect Privacy	0	0	+	+	-
JonDonym	0	0	-	+	-
Tor	0	++	- -	++	++
I2P	0	+	- -	+	++

U = Usability, A = Anonymity, E = Efficiency, R = Reliability and C = Cost.

Overall, Tor shows the best results, followed by I2P. They score well in all disciplines except performance, which is their main weak point. Here, single proxy solutions score with the best performance. Unfortunately, their degree of anonymity is poor and additionally reliability leaves much to be desired. JonDonym performs averagely, showing no particular strength or weakness. However, it is very difficult to consider all users' requirements and it is finally up to them to evaluate the results in order to find the most appropriate solution.

To conclude, future work will be necessary in the following areas:

- Extending the usability evaluations by also involving less sophisticated users,
- Further investigation of the very high throughput via certain anonymisation paths,
- Economic aspects need to be evaluated in more detail, especially in regard to business purposes,
- It may be worthwhile to include social aspects into the proposed payoff function, e.g., particular group behaviour in an anonymous network.

Overall, this comparison shows the need and motivation to spend further effort on the improvement of existing anonymisation services or to work on alternative solutions.

Acknowledgement. This work has been partially supported by EC FP7 EFIPSANS project (INFSO-ICT-215549). Furthermore, we would like to thank Dominic for his extensive proof reading.

References

Acquisti, A., Dingledine, R. & Syverson, P. (2003), On the Economics of Anonymity, *in* R.N. Wright, eds, 'Financial Cryptography', Springer Lecture Notes in Computer Science, pp. 84–102.

Chaum, D.L. (1981), Untraceable electronic mail, return addresses, and digital pseudonyms , *in* 'ACM Journal of Communication', ACM, pp. 84–90.

Danezis, G., Dingledine, R. & Mathewson, N. (2003), Mixminion: Design of a Type III Anonymous Remailer Protocol , *in* 'Proceedings of the 2003 IEEE Symposium on Security and Privacy', IEEE 'Computer Society, pp. 2-15.

Díaz, C., Seys, S., Claessens, J. & Preneel, B. (2002), Towards Measuring Anonymity *in* ''R. Dingledine & P. Syverson, eds, 'Privacy Enhancing Technologies', Springer Lecture Notes in Computer Science, pp. 54—68.

Dingledine, R., Mathewson, N. & Syverson, P. (2004), Tor: the second-generation onion router *in* 'Proceedings of the 13th conference on USENIX Security Symposium - Volume 13', USENIX Association, San Diego, CA, p. 21.

Dingledine, R. & Mathewson, N. (2006), Anonymity Loves Company: Usability and the Network Effect *in* 'Proceedings of the Fifth Workshop on the Economics of Information Security (Weis 2006)'.

Dingledine, R. (2009), Tor and circumvention: Lessons learned *at* '26th Chaos Communication Congress', Berlin, Germany.

Edman, M. & Syverson, P. (2009), As-awareness in Tor path selection , *in* 'Proceedings of the 16th ACM conference on Computer and communications securit (CSS '09)', ACM, pp. 380–389.

Freedman, M.J. & Morris, R. (2002), Tarzan: a peer-to-peer anonymizing network layer , *in* 'Proceedings of the 9th ACM conference on computer and communications security (CCS '02)', ACM, pp. 193–206.

Google (2010), Web metrics: Size and number of resources, http://code.google.com/speed/articles/web-metrics.html,[Online; last accessed 2010-10-01].

Kelly, D.J., Raines, R.A., Grimaila, M.R., Baldwin, R.O. & Mullins, B.E. (2008), A survey of state-of-the-art in anonymity metrics , *in* 'Proceedings of the 1st ACM workshop on network data anonymization - NDA '08', ACM Press, p. 31.

Kelly, D. (2009), A taxonomy for and analysis of anonymous communications networks, Ph.D., Air Force Institute of Technology.

Keshav, S. (1999), On individual and aggregate TCP performance , *in* 'Proceedings of Seventh International Conference on Network Protocols', IEEE Computer Society Press, pp. 203–212.

Köpsell, S. (2006), Low Latency Anonymous Communication How Long Are Users Willing to Wait? , *in* 'Journal of Emerging Trends in Information and Communication', Volume 3995 Springer, Lecture Notes in Computer Science, Berlin / Heidelberg, pp. 221–237.

Möller, U., Cottrell, L., Palfrader, P. & Sassamanm L. (2003), Mixmaster Protocol — Version 2, *in* 'IETF Internet Draft', IETF.

Murdoch S.J. & Watson, R.N.M. (2008), Metrics for Security and Performance in Low-Latency Anonymity Systems, *in* 'N. Borisov & I. Goldberg, eds, 'Privacy Enhancing Technologies', Springer Lecture Notes in Computer Science, pp. 115–132.

Ngan, T.-W., Dingledine, R. & Wallach, D.S. (2010), Building Incentives into To, *in* R. Sion, ed, 'Financial Cryptography', Springer Lecture Notes in Computer Science, pp. 238–256.

Panchenko, A., Herrmann, D., Wendolsky, R. & Federrath, H. (2009), Website fingerprinting: attacking popular privacy enhancing technologies with the multinomial naïve-bayes classifier, *in* 'Proceedings of the 2009 ACM workshop on Cloud computing security', ACM, Chicago, Il, USA, pp. 31–42.

Pfitzmann, A. & Hansen, M. (2009), A terminology for talking about privacy by data minimization: Anonymity, Unlinkability, Undetectability, Unobservability, Pseudonymity, and Identity Management , *Online:* 'http://dud.inf.tu-dresden.de/Anon_Terminology.shtml'.

Pries, R., Yu, W., Graham, S. & Fu, X. (2008), On performance bottleneck of anonymous communication networks , *in* '22nd IEEE International Symposium on Parallel and Distributed Processing, IPDPS 2008', IEEE Computer Society Press, Miami, Florida USA, pp. 1–11.

Raymond, J.-F. (2000), Traffic Analysis: Protocols, Attacks, Design Issues, and Open Problems , *in* 'Proceedings of Designing Privacy Enhancing Technologies: Workshop on Design Issues in Anonymity and Unobservability', Springer Verlag, pp. 10-29.

Reed, M.G., Syverson, P.F. & Goldschlag, D.M. (1998), Anonymous connections and onion routing , *in* 'IEEE Journal on Selected Areas in Communications', IEEE Computer Society, pp. 482–494.

Ren J. & Wu, J. (2009), Survey on anonymous communications in computer networks, *in* 'Computer Communications' Volume 33, Issue 4, pp. 420-431.

Rennhard, M. & Plattner, B. (2002), Introducing MorphMix: peer-to-peer based anonymous Internet usage with collusion detection, *in* 'Proceedings of the 2002 ACM workshop on Privacy in the Electronic Society (WPES '02)', ACM, pp. 91–102.

Snader, R. & Borisov, N. (2008), A Tune-up for Tor: Improving Security and Performance in the Tor Network , *in* 'Proceedings of the Network and Distributed Security Symposium - NDSS '08', Internet Society.

Wendolsky, R., Herrmann, D. & Federrath, H. (2007), Performance Comparison of Low-Latency Anonymisation Services from a User Perspective, *in* '', N. Borisov & P. Golle, eds, 'Privacy Enhancing Technologies', Springer Lecture Notes in Computer Science, pp. 233–253.

Wharton, C., Bradford, J., Jeffries, R. & Franzke, M. (1992), Applying cognitive walkthroughs to more complex user interfaces: experiences, issues, and recommendations, *in* 'Proceedings of the SIGCHI conference on Human factors in computing systems', ACM Press, Monterey, CA, pp. 381–388.

Wright, M., Adler, M, Levine, B. N. & Shields, C. (2003), Defending Anonymous Communication Against Passive Logging Attacks , *in* 'Proceedings of the 2003 IEEE Symposium on Security and Privacy', IEEE Computer Society, pp. 28–.