Quantifying and Measuring Anonymity

Steven J. Murdoch

University of Cambridge Computer Laboratory
http://www.cl.cam.ac.uk/~sjm217/

Abstract. The design of anonymous communication systems is a relatively new field, but the desire to quantify the security these systems offer has been an important topic of research since its beginning. In recent years, anonymous communication systems have evolved from obscure tools used by specialists to mass-market software used by millions of people. In many cases the users of these tools are depending on the anonymity offered to protect their liberty, or more. As such, it is of critical importance that not only can we quantify the anonymity these tools offer, but that the metrics used represent realistic expectations, can be communicated clearly, and the implementations actually offer the anonymity they promise. This paper will discuss how metrics, and the techniques used to measure them, have been developed for anonymous communication tools including low-latency networks and high-latency email systems.

1 Introduction

Anonymous communication systems seek to hide patterns visible in communications to obscure relationships between people and the activities they carry out, typically over the Internet. Such systems have become increasingly popular as a result of the Internet developing into an important tool in the support and promotion of human rights. Examples of uses include the publication of videos showing human rights abuses, journalists soliciting information on government corruption, and law enforcement agencies monitoring websites operated by organized crime.

In all these examples there are motivated individuals who would want to discover the identity of the users of the anonymous communication system. Therefore it is of critical importance that the level of protection that the anonymous communication system provides is well understood. Overestimating the level might result in users putting themselves at unacceptable amounts of risk; underestimating the level might result in users avoiding using a system unnecessarily.

The task of measuring the level of anonymity offered by anonymous communication tools is challenging particularly because of the narrow safety margins which they necessarily offer. A system operating perfectly can only hide the real sender or receiver of a message within the ranks of the users of that system. An

DPM 2013 and SETOP 2013, LNCS 8247, pp. 3–13, 2014. DOI: 10.1007/978-3-642-54568-9 1

attacker who wants to de-anonymise a user can often also take into account auxiliary information collected through means other than monitoring the anonymous communication system.

For example, suppose a company discovers that a whistleblower has leaked documents, sent through an anonymous communication system, proving that management have authorised the bribing of government officials. If that anonymous communication system only had a million users that day, then there are at most a million candidates for who leaked the document. Intersecting the set of users of the system with the set of people who had access to the documents in question might leave only a handful of possibilities. Even a small amount of information disclosed by the anonymous communication system could leave the whistleblower singled out.

In contrast, encryption systems draw their strength from the large number of possible keys that could have been used to encrypt the information – far more than the number of users of the system. Adding to the key length imposes a linear cost to users of the system but increases the time needed to attack the system exponentially. As a result, modern encryption systems have a very large safety margin and so even serious weaknesses in encryption algorithms rarely have a practical effect on their security.

Therefore research on anonymous communication systems has focussed on improving security through increasing their number of users and decreasing the information disclosed to an observer. However, achieving either of these goals typically comes at a significant cost to users by reducing network capacity. As a result, it is not feasible to achieve the same safety margins that encryption systems offer and so it is important to develop ways to accurately measure the level of protection offered by anonymous communication systems. Then appropriate design choices can be made to provide the right trade-off between performance and security.

2 Email mixes

One of the early applications of anonymous communication technology was to email. In a scheme proposed by Chaum [2] a user selects one or more "mixes" as a path through which his message should be sent. Messages are encrypted by a sender under multiple layers of public-key encryption. Outside each layer of encryption is the address of the next mix, which allows messages to be routed. This mix can remove the next layer of encryption, and will find the address of the next mix in the path to which the message should be sent. Once the message reaches the last mix in the path, the plaintext of the message will be available along with the address of the ultimate destination of the message.

Each mix will see the immediate source of the message and the immediate destination. Therefore the first mix will know the sender's address but not the recipient's, and the last mix will know the recipient's address, but not the sender's. Similarly, someone observing messages flowing through the network will not be able to match incoming messages to outgoing messages based on the content

because a decryption operation is carried out at each step which only a specific mix has the private key necessary to perform, and message lengths are fixed. Messages are also delayed at each mix, for a random period of time or until a particular number of messages have been received by a mix (or some combination of these) so as to complicate matching based on the time messages are sent and received.

In this way, the email mix network provides "unlinkability" [6] to messages because the attacker should not be able to link which messages entering the mix network correspond to which messages leaving the mix network. The mix network can also be seen to offer anonymity to its users – for each message leaving the network it should not be possible to establish its sender and for each message entering the network it should not be possible to establish its recipient. An attacker does however know a list of possible candidate senders for each message which leaves the network – the "sender anonymity set". Similarly there is a "recipient anonymity set" for each message sent.

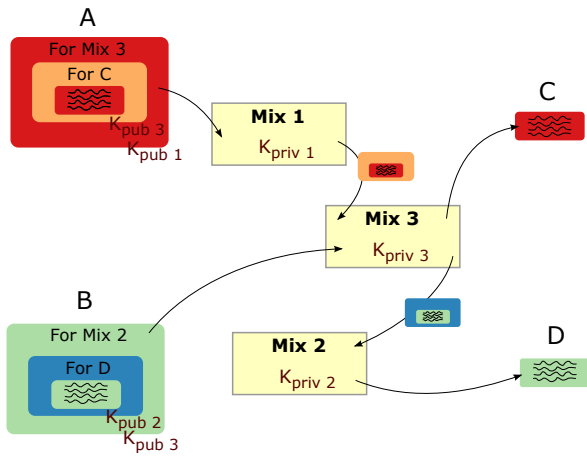

Fig. 1. A two-hop mix network. A is sending a message to C, via Mix 1 then Mix 3. B is sending a message to D via Mix 3 then Mix 2

2.1 Measuring anonymity

Much of the research on email mixes has focussed on how to quantify the anonymity provided. Berthold *et al.* [1] proposed to simply count the size ("cardinality") of the anonymity set: a larger list of candidates for the true sender or receiver corresponds to better anonymity. By taking the logarithm of the set size, base 2, this quantity can be expressed in bits. An ideal anonymous communication system will have an anonymity set size of the number of users and the probability of each user being the sender or recipient of a particular message will

be equal. Looking at the anonymity set as a probability distribution over possible senders/receivers of a message, the ideal anonymous communication system produces the uniform distribution.

However real anonymous communication systems will not achieve this ideal. It is typically possible to distinguish senders from recipients by observing the direction of flow of data. Also by taking into account that it will be unlikely (for usability reasons) that mixes will delay messages for a long period of time, not every possible sender/recipient will be equally likely the true sender/recipient. In an extreme case an attacker may know that a single user may almost certainly be the sender of a message yet based on cardinality this system is indistinguishable from an ideal one of the same size, as shown in Figure 2.

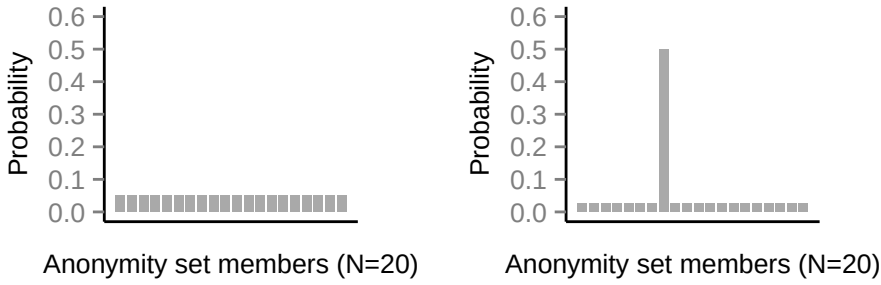

Fig. 2. Two possible distributions over a 20-element anonymity set. The left distribution is uniform (all elements at $\frac{1}{20}$); the right has one element at probability $\frac{1}{2}$ and the others at $\frac{1}{38}$

For this reason, other proposed metrics take into account the unevenness of the probability distribution. One such metric is the "degree of anonymity" proposed by Reiter *et al.* [7]. Although originally developed for analysing a system for anonymising web traffic it can equally be applied to email mixes. The 6 point scale is described in Table 1.

The degree of anonymity metric differentiates between the two anonymity set distributions of Figure 2. The left graph shows that users are beyond suspicion whereas the right is barely probable innocence. For all reasonable purposes, the left graph corresponds to a better system so taking into account the unevenness of the distribution has produced a better metric, but ignoring the cardinality of the set has a weakness too.

For example, an anonymity set probability distribution over 101 senders, with the most likely sender having probability 0.01 and others probability 0.0099 offers possible innocence. Whereas an uniform anonymity set probability distribution over 4 senders has each sender assigned a probability of 0.25. Although the latter system has a better degree of anonymity, the probability of an attacker successfully identifying a user is much higher than the former.

Table 1. The 6-point degree of anonymity scale

	Degree	Attacker observation
Best anonymity	Absolute privacy	No evidence whether or not a sender sent any message
	Beyond suspicion	A sender sent a message, but all senders are equally likely to have sent any message
	Probable innocence	A sender is no more likely to have been the originator of a message than to not have been
	Possible innocence	A sender has a nontrivial probability of not being the originator of a message
	Exposed	The originator has been identified
Worst anonymity	Provably exposed	The originator has been identified and the identity can be proven to others

It therefore follows that both cardinality and unevenness of distribution should be taken into account, and so Shannon entropy was proposed as a metric by Serjantov and Danezis [8]. Here, if the probability that user i was the true sender is p_i, and there are N members of the anonymity set, then the entropy of the anonymity set S is:

$$\mathrm{H}(S) = -\sum_{i=1}^{N} p_i \log_2 (p_i)$$

For the probability distributions in Figure 2, the left distribution has entropy ≈ 4.32 bits (the same as the cardinality, in bits – $\log_2(20)$), but the right distribution only has entropy ≈ 3.12. The anonymity set discussed above, of 101 senders with one at probability 0.01 and others at 0.0099, gives entropy 6.66 bits (only $10^{-5}\%$ less than the entropy of the uniform distribution over 101 senders). Whereas the uniform distribution over 4 senders is 2. We can see that entropy takes into account both cardinality and unevenness, and also gives similar values to similar distributions, but it is still possible to find examples which raise the question of whether entropy is the best metric.

For example, in Figure 3 the two very different distributions have the same entropy. However, from the perspective of an attacker the anonymity might be very different. The de-anonymisation of communications is seldom used as an end in itself, but rather to guide further investigation. An attacker analysing the left distribution would need to investigate 10 senders before getting a 50% probability of having found the right sender. In contrast the attacker could achieve the same goal with the right distribution after trying only one user.

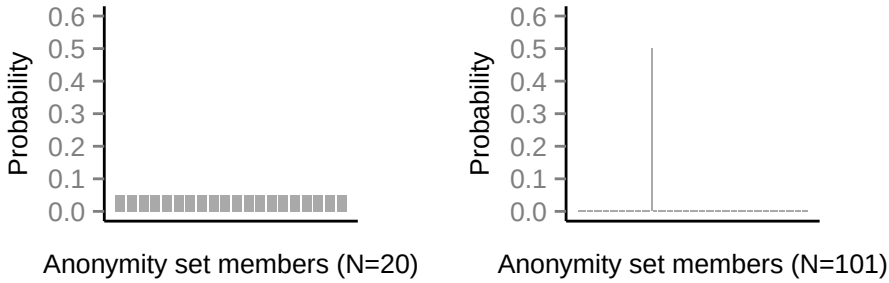

Fig. 3. Two possible distributions. The left graph is the same as in Figure 2 – the uniform distribution over 20 senders. The right diagram is a probability distribution over 101 senders, with one having probability $\frac{1}{2}$ and the others having probability $\frac{1}{200}$. Both have entropy $\log_2(20) \approx 4.32$

One way of differentiating between the two distributions is to note that the number of users is rarely under direct control of the system designer so a reasonable metric could examine the ratio between the security of the ideal system for a given user base to the actual security achieved for the same user base. This metric was proposed as the "degree of anonymity" by Diaz *et al.* [3], but to differentiate from the Crowds degree in Table 1 we will use the term "normalized entropy" to refer to the Diaz degree. Where $\mathrm{H}(S)$ is the entropy of the anonymity set S and N is the cardinality of the anonymity set, the normalized entropy is defined as:

$$\mathrm{d}(S) = \frac{\mathrm{H}(S)}{\log_2(N)}$$

However, even when the sizes and entropy of the anonymity sets are the same there may be questions as to which distribution is better. For example, the two distributions in Figure 4 have the same entropy (≈ 3.12) and cardinality (20) and therefore the same normalized entropy. They also have the same degree of anonymity – probable innocence. The left distribution has one sender at probability $\frac{1}{20}$ and the other 19 at $\frac{1}{38}$. The right distribution has 5 senders at probability $\frac{a}{5}$ and the other 15 at $\frac{1-a}{15}$ where $a \approx 0.86$ is the solution to the equation defined by Tóth *et al.* [9]:

$$a \log_2\left(\frac{a}{5}\right) + (1-a)\log_2\left(\frac{1-a}{15}\right) = \frac{\log_2\left(\frac{1}{2}\right)}{2} + \frac{19\,\log_2\left(\frac{1}{38}\right)}{38}$$

Considering an attacker able to investigate one possible sender, the left distribution is worse for privacy, with a 50% chance that the attacker will succeed compared to 17.2% for the right. On the other hand an attacker able to investigate 5 senders will succeed in the left distribution with probability 61% but 86% with the right distribution. Tóth *et al.* [9] proposed using min-entropy – $-\log_2(\max_i p_i)$ to quantify the minimum security achieved by any user, which

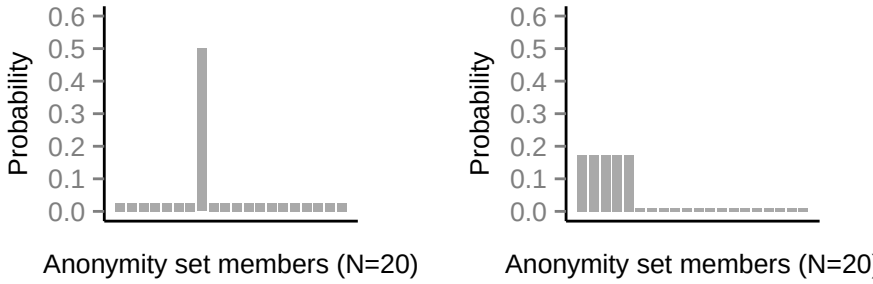

Fig. 4. Two possible distributions, with identical cardinality (≈ 4.32 bits), entropy (≈ 3.12 bits) and normalized entropy (≈ 0.72)

matches the effective security in the case of an attacker able to investigate one sender. Under this metric the left distribution gives 1 bit and the right gives ≈ 2.54 bits.

In fact, cardinality, entropy and min-entropy are all special cases of Rényi entropy, for $\alpha = 0$, $\alpha \to 1$ and $\alpha \to \infty$ respectively where:

$$\mathrm{H}_\alpha(S) = \frac{1}{1-\alpha} \log_2 \left(\sum_{i=1}^{N} {p_i}^\alpha \right)$$

Figure 5 shows $\mathrm{H}_\alpha(S)$ over a range of α for the anonymity set distributions in Figures 3 and 4. As expected, cardinality depends only on the number of senders and min-entropy depends only on the probability of the most-likely sender. For many values of α there can be a conceivable model of the attacker in which the $\mathrm{H}_\alpha(S)$ will make sense as a metric for anonymity. For example H_0 will be the number of questions that the attacker needs to ask if a question can eliminate half of the candidates; $\mathrm{H}_{\to 1}$ will be the same if the attacker can choose questions which will eliminate an arbitrary subset of the anonymity set. $\mathrm{H}_{\to \infty}$ represents the security if the attacker can investigate one user.

3 Low-latency anonymous communication systems

As we have seen above, metrics implicitly define a threat model in terms of the attacker's strategy. Measuring the security of a network according to a metric also requires defining the attacker's capabilities. For email mixes the attacker capability commonly assumed is "global-passive" – the attacker can monitor the entire network but cannot interfere with network traffic nor view the internal processing of any mix. There is debate as to whether this is appropriate as few attackers can monitor the entire Internet, and computer security is not good enough to ensure that email mixes are not compromised. However where the global-passive model fails is the analysis of low-latency general-purpose anonymous communication system.

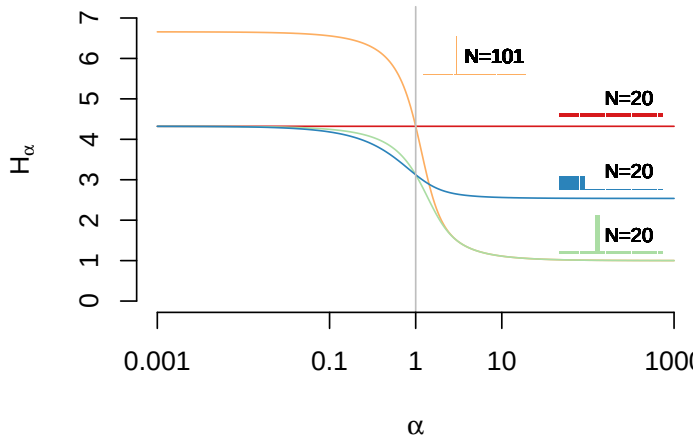

Fig. 5. Rényi entropy for four anonymity set distributions for a range of α

The leading low-latency anonymous communication system is Tor [4]. It is capable of anonymising any TCP-based protocol, and it introduces very low levels of delay compared to email mixes (milliseconds rather than days). It therefore has much wider applicability and so has far more users than any email mix. Users send traffic through the network by building a "circuit" through 1 or more (usually 3) Tor nodes, with nested layers of symmetric encryption, to fulfil a similar purpose as the nested public-key encryption of email mixes. Symmetric keys are negotiated with authenticated Diffie-Hellman key agreement. On top of the circuit-level encryption, TLS tunnels are maintained between each pair of Tor nodes which are exchanging traffic.

Figure 6 shows how encryption is performed in Tor. Unlike email mixes however, Tor does not attempt to make all messages the same length as others. This is because TCP network traffic is highly variable and maintaining a constant rate of traffic would either dramatically reduce the maximum bandwidth of the network or add a massive amount of overhead. Therefore, like any other proposed low-latency general-purpose anonymous communication system, an attacker monitoring the entire Internet would be able to correlate network connections entering and leaving the network.

Since analysing a low-latency anonymous communication network under the global-passive model leads to the conclusion that any system is insecure, metrics assuming the global-passive model do not help guide the design of low-latency anonymous communication systems intended to defend against other threat models. It is therefore advantageous to directly measure the probability that the user's security will be compromised, under the assumptions of the actual attackers' capabilities, rather than using proxies for this probability such as the various types of entropy.

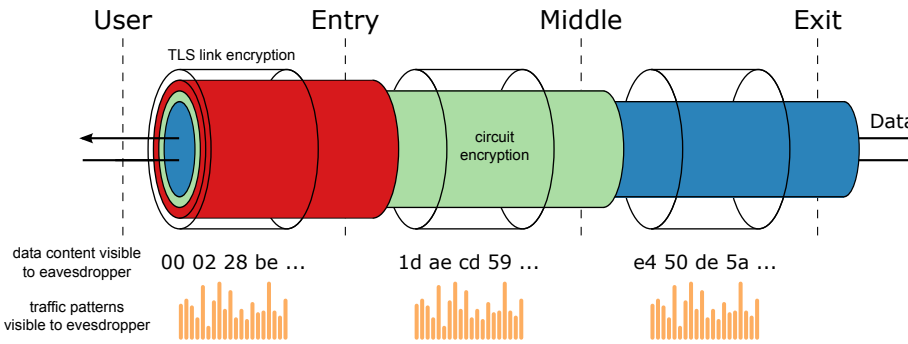

Fig. 6. Use of encryption in Tor. Telescoping circuit encryption shares keys between the user and each of the nodes in the circuit. On top of the circuit encryption, TLS authenticated encryption is performed on a link basis. Together, the circuit and link encryption ensure that incoming and outgoing traffic on a node cannot be linked based on content, but it may be linkable based on timing patterns

This is the approach used by Murdoch and Watson [5] to analyse the security of Tor directly, against an adversary who wishes to insert malicious nodes into the network so as to de-anonymize users. The goal of this analysis was to discover which of proposed schemes, for selecting Tor nodes in a circuit, is more secure. One candidate scheme was to select from nodes uniformly at random and another is that a user would select nodes with a probability weighted proportional to how much bandwidth that node has available. While the uniform selection had a higher entropy, the bandwidth-weighted scheme had better security.

The result can be seen in Figure 7. In these figures, each point shows a particular attacker strategy in terms of the number of malicious nodes added and the bandwidth given to each node. The colour shows what proportion of circuits will be compromised as a result. The attacker capabilities are represented as a line on the graph, showing what is the maximum investment possible: in this case the attacker has a bandwidth budget of 100 MB/s and can distribute this over a small number of high-bandwidth nodes or a large number of low-bandwidth nodes. The left graph shows the uniform-selection scenario, where the optimum attacker strategy will be to have a large number of low-bandwidth nodes resulting in 80% of circuits compromised. In contrast, the right shows the bandwidth weighted scenario where no matter how the attacker allocates resources, no more than 20% of circuits will be compromised.

4 Conclusions

The above examples have shown a few examples from the wide variety of metrics for anonymous communication networks. These range from the discrete levels of the anonymity degree through various types of entropy then to directly quantifying the probability of user compromise. Each has their own advantages in

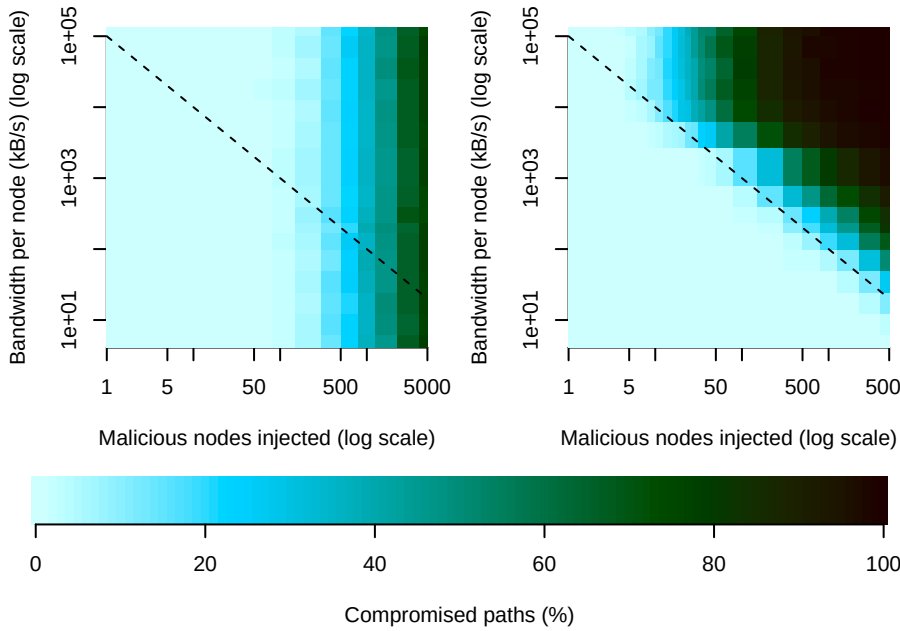

Fig. 7. Probability of path compromise for two different circuit selection algorithms – uniform on left and bandwidth weighted on right

terms of how easy they are to calculate for a given system, how simple a summary they provide, how versatile they are in terms of possible attacker models, and how representative the measurement is of the practical security provided. The narrow safety margins which necessarily follow from the problem-space that anonymous communication systems exist in poses challenges not only for system design but also quantification of security. Some of the lessons learnt by the anonymous communication community may be more widely applicable, and it is likely that much knowledge from other fields can contribute to the development of better metrics for anonymous communication systems.

References

1. Berthold, O., Pfitzmann, A., Standtke, R.: The disadvantages of free MIX routes and how to overcome them. In: Federrath, H. (ed.) Proceedings of Designing Privacy Enhancing Technologies: Workshop on Design Issues in Anonymity and Unobservability. pp. 30–45. Springer-Verlag, LNCS 2009 (July 2000)
2. Chaum, D.: Untraceable electronic mail, return addresses, and digital pseudonyms. Communications of the ACM 24(2) (February 1981)
3. Diaz, C., Seys, S., Claessens, J., Preneel, B.: Towards measuring anonymity. In: Dingledine, R., Syverson, P. (eds.) Proceedings of Privacy Enhancing Technologies Workshop (PET 2002). Springer-Verlag, LNCS 2482 (April 2002)

4. Dingledine, R., Mathewson, N., Syverson, P.: Tor: The second-generation onion router. In: Proceedings of the 13th USENIX Security Symposium (August 2004)
5. Murdoch, S.J., Watson, R.N.M.: Metrics for security and performance in low-latency anonymity networks. In: Borisov, N., Goldberg, I. (eds.) Proceedings of the Eighth International Symposium on Privacy Enhancing Technologies (PETS 2008). pp. 115–132. Springer (July 2008)
6. Pfitzmann, A., Hansen, M.: A terminology for talking about privacy by data minimization: Anonymity, unlinkability, undetectability, unobservability, pseudonymity, and identity management (Aug 2010), `http://dud.inf.tu-dresden.de/literatur/Anon_Terminology_v0.34.pdf`, v0.34
7. Reiter, M., Rubin, A.: Crowds: Anonymity for web transactions. ACM Transactions on Information and System Security 1(1) (June 1998)
8. Serjantov, A., Danezis, G.: Towards an information theoretic metric for anonymity. In: Dingledine, R., Syverson, P. (eds.) Proceedings of Privacy Enhancing Technologies Workshop (PET 2002). Springer-Verlag, LNCS 2482 (April 2002)
9. Tóth, G., Hornák, Z., Vajda, F.: Measuring anonymity revisited. In: Liimatainen, S., Virtanen, T. (eds.) Proceedings of the Ninth Nordic Workshop on Secure IT Systems. pp. 85–90 (November 2004)