Attacking Unlinkability: The Importance of Context

Matthias Franz[1], Bernd Meyer[1], and Andreas Pashalidis[2]

[1] Siemens AG, Corporate Technology,
Otto-Hahn-Ring 6, 81739 Munich, Germany
{matthias.franz, bernd.meyer}@siemens.com
[2] NEC Europe Ltd, Network Laboratories
Kurfürsten-Anlage 36, 69115 Heidelberg, Germany
andreas.pashalidis@netlab.nec.de

Abstract. A system that protects the unlinkability of certain data items (e. g. identifiers of communication partners, messages, pseudonyms, transactions, votes) does not leak information that would enable an adversary to link these items. The adversary could, however, take advantage of hints from the context in which the system operates. In this paper, we introduce a new metric that enables one to quantify the (un)linkability of the data items and, based on this, we consider the effect of some simple contextual hints.

1 Introduction

A number of privacy-preserving systems, such as mix networks, anonymous credential systems, and secret voting schemes, protect the unlinkability of certain data items of interest. Mix networks, in particular, protect the unlinkability of the messages that enter the network with respect to their recipients, the messages that leave the network with respect to their senders, and, hence, the identifiers of communicating parties with respect to communication sessions. Since their introduction [9], a number of different mix network variants have been proposed (see, for example, [4, 19, 26, 33, 34]), some of which have also been implemented and deployed. Anonymous credentials, on the other hand, protect the unlinkability of the pseudonyms and the transactions with respect to the users they correspond to. Since their introduction into the digital world [10], a number of anonymous credential systems have been proposed (see, for example, [7, 8, 11–14, 29, 32, 38]). Secret voting schemes protect the unlinkability of votes with respect to the users that cast them. Such schemes have evolved from ostracism [24] to sophisticated cryptosystems; for an overview of the current state of the art the reader is referred to [1].

The problem of analysing how well the above types of system protect unlinkability has received some attention during recent years. The focus of most works is, however, on mix networks (see, for example, [2, 15, 16, 25, 27, 30]). This is not surprising since mix networks provide the basis for anonymous communication and are, as such, necessary for preserving privacy in a number of settings, including the setting of anonymous credentials [17] and, sometimes, the setting of voting systems (see, for example, [6]).

An adversary that wishes to link the protected items may use information that is leaked by the system during its operation, or hints from the environment in which the system operates. In contrast to existing literature, the focus of this paper is on the latter. That is, we study a number of simple contextual hints and their effect on unlinkability. Our results apply to *all* types of unlinkability-protecting system, including mix networks, anonymous credentials, and secret voting schemes. The rest of the paper is organised as follows. Section 2 introduces the metric for unlinkability that is used throughout the paper. Section 3 examines seven different types of hint and their effect on unlinkability. Finally, Section 4 concludes.

2 Measuring unlinkability

Consider a set of elements A and a partition $\pi \vdash A$ of that set. Note that we do not distinguish between π and the equivalence relation it defines. In the sequel, we write $a_1 \equiv_\pi a_2$ if the elements $a_1, a_2 \in A$ lie in the same equivalence class of π, and $a_1 \not\equiv_\pi a_2$ otherwise. Let $\tau \vdash A$ denote a 'target' partition, chosen uniformly at random. We use entropy as a metric for unlinkability. That is, the unlinkability of the elements in a set A against an adversary $\mathcal{A}$ is defined as

$$\mathcal{U}_A(\mathcal{A}) = -\sum_{\pi \in \Pi} \Pr(\pi = \tau) \log_2(\Pr(\pi = \tau)),$$

where $\Pi = \{P : P \vdash A\}$ denotes the set of partitions of A and $\Pr(\pi = \tau)$ denotes, in $\mathcal{A}$'s view, the probability that π is the target partition τ. We further define the *degree* of unlinkability of the elements in A against an adversary $\mathcal{A}_H$ with access to a hint H about τ as

$$\mathcal{D}_A(\mathcal{A}_H) = \frac{\mathcal{U}_A(\mathcal{A}_H)}{\mathcal{U}_A(\mathcal{A}_\emptyset)},$$

where $\mathcal{A}_\emptyset$ denotes the adversary without any hints. That is, $\mathcal{A}_\emptyset$ knows A but has no information about τ. The set of candidate partitions for $\mathcal{A}_\emptyset$ is therefore $\Pi_A(\mathcal{A}_\emptyset) = \Pi$, i. e. the set of all partitions of A. The number $|\Pi_A(\mathcal{A}_\emptyset)| = B_{|A|}$ of such partitions, a Bell number [3, 35], is given by the recursive formula

$$B_{n+1} = \sum_{k=0}^{n} \binom{n}{k} B_k \tag{1}$$

where $B_0 = 1$.[3] Since τ is chosen uniformly at random, the unlinkability of the elements in A is therefore at its maximum, i. e. $\mathcal{U}_A(\mathcal{A}_\emptyset) = \log_2(B_{|A|})$ bits. This

[3] The first few Bell numbers are 1, 1, 2, 5, 15, 52, 203, 877, 4140, 21147.

is the best case from a privacy point of view: all partitions of A are equally likely to be the target partition τ.

Remark 1: In the setting of unlinkability-protecting systems, the goal of the adversary is to identify a target partition from an 'anonymity set' of candidate partitions. The fact that the information-theoretic metric we use for unlinkability is identical to the metric introduced for anonymity in [18, 36], is therefore natural.

Remark 2: $\mathcal{U}_A$ is a measure of the information that is contained in the probability distribution that the adversary assigns to the set of all partitions of A. Since we assume that τ is selected uniformly at random, this distribution is, a priori, uniform. However, a hint may enable the adversary to change his view such that, a posteriori, some partitions are more likely than others. The hints we consider in this paper enable the adversary to exclude a number of candidate partitions (i.e. to reduce the size of the 'anonymity set') while the remaining partitions remain equally likely.

Example: Consider an anonymous help line where a clerk offers advice over the telephone. Suppose that, one day, the clerk receives four calls, denoted $A = \{\lambda_1, \lambda_2, \lambda_3, \lambda_4\}$. Without any additional information, all $B_4 = 15$ partitions of A constitute valid ways to link these calls. Since without any additional information all these partitions are equally likely, the unlinkability of the calls is, in this case, $\log_2(15) \simeq 3.9$ bits, and the degree of unlinkability is $\log_2(15)/\log_2(15) = 1$.

The clerk, however, has some additional information: he realised that the calls λ_1 and λ_2 were made by men, and that the calls λ_3 and λ_4 by women (however, the clerk does not know whether or not the same person called twice). This hint effectively rules out all partitions where λ_1 or λ_2 appears in the same equivalence class as λ_2 or λ_4. In particular, only four partitions remain valid, namely $\{(\lambda_1, \lambda_2), (\lambda_3, \lambda_4)\}$, $\{(\lambda_1, \lambda_2), (\lambda_3), (\lambda_4)\}$, $\{(\lambda_3, \lambda_4), (\lambda_1), (\lambda_2)\}$, and $\{(\lambda_1), (\lambda_2), (\lambda_3), (\lambda_4)\}$. Since these four partitions are equally likely, the unlinkability of the calls is, in this case, $\log_2(4) = 2$ bits, and the degree of unlinkability is $\log_2(2)/\log_2(15) \simeq 0.52$.

3 The importance of context

In this section, we examine seven types of hint that an adversary may obtain from the operational context of the system. In particular, we examine hints that reveal to the adversary (a) the number of equivalence classes in τ, (b) the cardinality of equivalence classes in τ, (c) the fact that all equivalence classes in τ have a given cardinality, (d) a 'reference partition' the equivalence classes of which have exactly one representative in each equivalence class in τ, (e) a set of element pairs that are equivalent in τ, (f) a set of element pairs that are not equivalent in τ, and (g) a combination of (e) and (f).

3.1 The number of equivalence classes

Consider an adversary $\mathcal{A}_{H_1}$ with a hint $H_1 = (\alpha)$, where $\alpha \in \mathbb{N}$ and $1 \le \alpha \le |A|$, that reveals how many equivalence classes τ has. $\mathcal{A}_{H_1}$ can restrict its attention to $\Pi_A(\mathcal{A}_{H_1}) = \{P : P \vdash A, |P| = \alpha\}$, i.e. the partitions that divide A into exactly α equivalence classes. The number of such partitions, which is a Stirling number of the second kind [22], is given by

$$|\Pi_A(\mathcal{A}_{H_1})| = \frac{1}{\alpha!} \sum_{k=0}^{\alpha} (-1)^k \binom{\alpha}{k} (\alpha - k)^{|A|}.$$

Since τ is chosen uniformly at random, the unlinkability of the elements in A is $\mathcal{U}_A(\mathcal{A}_{H_1}) = \log_2(|\Pi_A(\mathcal{A}_{H_1})|)$ bits. Figure 1 shows the degree of unlinkability $\mathcal{D}_A(\mathcal{A}_{H_1})$ as a function of $|A|$.

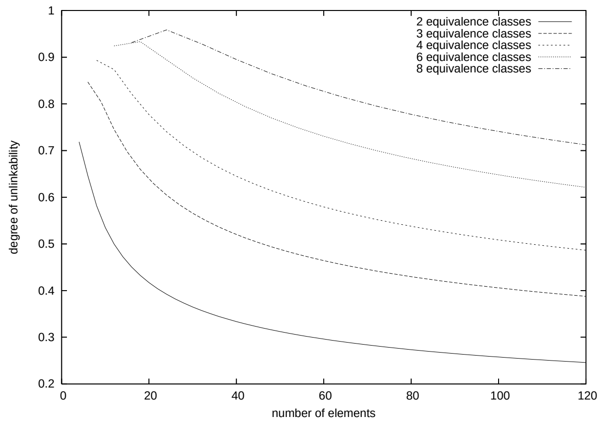


Fig. 1: Degree of unlinkability $\mathcal{D}_A(\mathcal{A}_{H_1})$ of elements in a set A as a function of $|A|$, if it is known that they must be divided into α equivalence classes.

How to obtain this hint: The number α typically is the number of users in a system. In the setting of mix networks, this number may be known to the operator of the network if users are required to register themselves or pay a fee. Otherwise, obtaining such a hint may be tricky due to the possibility of sybil attacks [20]. Whether or not it is straightforward to obtain this hint in the setting of anonymous credentials depends on the application. In the case of cash, for example, the financial institution is very likely to know how many users participate in the system. Similarly, in the case of demographic or personal credentials (such as age certificates or driving licences), the issuing authority is also likely to know the number of users in the system. In the setting of secret voting, there exist multiple ways to obtain the number of voters. The number of casted ballots, for example, may be conclusive about the number of voters.

3.2 The cardinality of equivalence classes

Consider an adversary $\mathcal{A}_{H_2}$ with a hint $H_2 = (\beta_1, \beta_2, \ldots, \beta_\alpha)$, where $\sum_{i=1}^{\alpha} \beta_i = |A|$ and $1 < \alpha < |A|$, that reveals the sizes of the equivalence classes in τ. That is, if $\tau = \{T_1, T_2, \ldots, T_\alpha\} \vdash A$, H_2 reveals that $|T_1| = \beta_1$, $|T_2| = \beta_2$, and so on. $\mathcal{A}_{H_2}$ can restrict its attention to $\Pi_A(\mathcal{A}_{H_2}) = \{P : P = \{T_1, T_2, \ldots, T_\alpha\} \vdash A, \forall 1 \leq i \leq \alpha, |T_i| = \beta_i\}$, i. e. the partitions that divide A into exactly α equivalence classes with cardinalities $\beta_1, \beta_2, \ldots, \beta_\alpha$. The number of such partitions is given by

$$|\Pi_A(\mathcal{A}_{H_2})| = \frac{|A|!}{\prod_{i=1}^{\alpha}(\beta_i!) \prod_{i=1}^{|A|}(\gamma_i!)} \tag{2}$$

where, for all $1 \leq i \leq |A|$, $\gamma_i = |\{\beta \in H_2 : \beta = i\}|$ (for a proof see Appendix B). That is, γ_i is the number of equivalence classes in τ that have cardinality i. Since τ is chosen uniformly at random, the unlinkability of the elements in A is $\mathcal{U}_A(\mathcal{A}_{H_2}) = \log_2(|\Pi_A(\mathcal{A}_{H_2})|)$ bits. It is perhaps worth noting that there exist hints of type H_2 which do not reveal any information as to whether any two given elements are equivalent or not. This is in contrast to what is claimed in [37] (see Appendix A).

As a special case, consider an adversary $\mathcal{A}_{H_3}$ with a hint $H_3 = (\alpha)$, where $\alpha \in \mathbb{N}$ divides $|A|$, that reveals the fact that τ has α equivalence classes *of the same cardinality* $|A|/\alpha$. $\mathcal{A}_{H_3}$ can restrict its attention to $\Pi_A(\mathcal{A}_{H_3}) = \{P : P \vdash A, \forall p \in P, |p| = |A|/\alpha\}$, i. e. the partitions that divide A into exactly α equivalence classes of equal cardinality $|A|/\alpha$. The number of such partitions is given by

$$|\Pi_A(\mathcal{A}_{H_3})| = \frac{|A|!}{\alpha!((|A|/\alpha)!)^\alpha} \tag{3}$$

(for a proof see Appendix B). Since τ is chosen uniformly at random, the unlinkability of the elements in A is $\mathcal{U}_A(\mathcal{A}_{H_3}) = \log_2(|\Pi_A(\mathcal{A}_{H_3})|)$ bits. Figure 2 shows the degree of unlinkability $\mathcal{D}_A(\mathcal{A}_{H_3})$ as a function of $|A|$.

How to obtain this hint:[4] In the setting of mix networks, this hint may be obtained if it is known how many messages each user sends in each session. In the setting of anonymous credentials, it is possible to obtain this hint if it is known how many pseudonyms each user has. In the setting of secret voting, this hint may be obtained if it is known how many ballots each user casted.

3.3 A reference partition

Consider an adversary $\mathcal{A}_{H_4}$ with a hint $H_4 = (\rho)$, consisting of a 'reference partition' $\rho = \{R_1, R_2, \ldots, R_{|A|/\alpha}\} \vdash A$ such that, for all $1 \leq i \leq |A|/\alpha$, $|R_i| = \alpha$ (note that α divides $|A|$), and that reveals the fact that each of the equivalence classes of τ contains exactly one element from R_i. $\mathcal{A}_{H_4}$ can restrict its attention to $\Pi_A(\mathcal{A}_{H_4}) = \{P : P \vdash A, P \text{ is a transversal of } \rho\}$, i. e. the partitions that

[4] This paragraph refers to hints of both type H_2 and H_3.

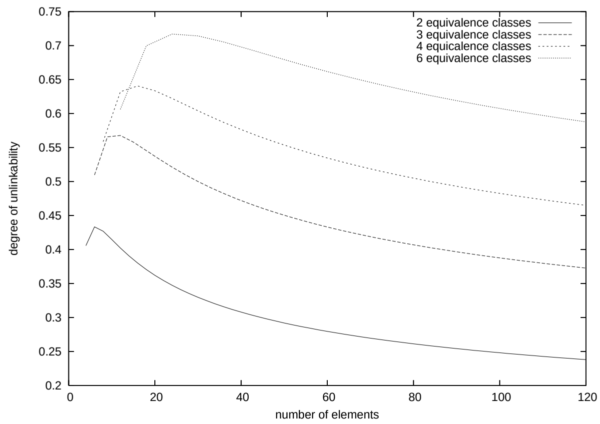


Fig. 2: Degree of unlinkability $\mathcal{D}_A(\mathcal{A}_{H_3})$ of elements in a set A as a function of $|A|$, if it is known that they must be divided into α equivalence classes of equal cardinality $|A|/\alpha$.

divide A into α equivalence classes of equal cardinality $|A|/\alpha$, where each class contains exactly one element from each of $R_1, R_2, \ldots, R_{|A|/\alpha}$. The number of such partitions is given by

$$|\Pi_A(\mathcal{A}_{H_4})| = (\alpha!)^{(|A|/\alpha)-1} \tag{4}$$

(for a proof see Appendix C). Since τ is chosen uniformly at random, the unlinkability of the elements in A is $\mathcal{U}_A(\mathcal{A}_{H_4}) = \log_2(|\Pi_A(\mathcal{A}_{H_4})|)$ bits. Figure 3 shows the degree of unlinkability $\mathcal{D}_A(\mathcal{A}_{H_4})$ as a function of $|A|$.

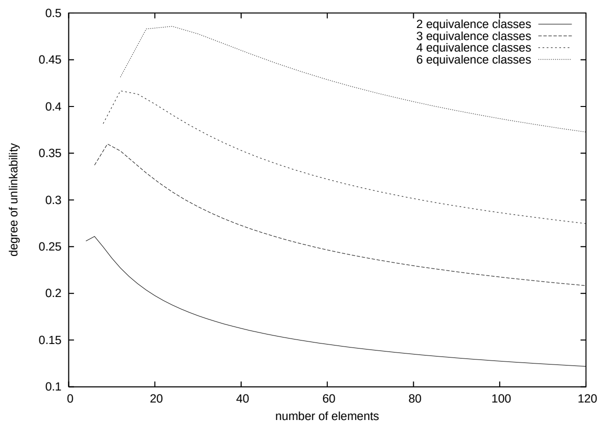


Fig. 3: Degree of unlinkability $\mathcal{D}_A(\mathcal{A}_{H_4})$ of elements in a set A as a function of $|A|$, if it is known that they must be divided into α equivalence classes of equal cardinality $|A|/\alpha$, such that each class contains exactly one element from each equivalence class of a given partition.

How to obtain this hint: In the setting of mix networks this hint may be obtained if each of the α users sends exactly one message through the network in β communication sessions. An adversary that wishes to divide the set of messages that leave the network (there are $\alpha \cdot \beta$ of them) into α subsets of equal cardinality β, such that each subset contains the messages sent by a single user, can construct a reference partition $R_1, R_2, \ldots, R_\beta$ by grouping the messages that leave the network according to communication sessions (i. e. such that, for all $1 \leq i \leq \beta$, R_i contains the messages that leave the network in session i). In the setting of anonymous credential systems, this hint may be obtained if each user has established exactly one pseudonym with each organisation in the system; an adversary that controls the organisations knows the reference partition as a side effect of normal operation. In the setting of secret voting, this hint may be obtained in special cases, such as the case of a combined election where each of the α voters is asked to answer β different questions on separate ballots. An adversary that wishes to divide the set of casted ballots (there are $\alpha \cdot \beta$ of them) into α subsets of equal cardinality β, such that each subset contains the ballots casted by a single user, can construct a reference partition by grouping the ballots according to the question they correspond to.

3.4 Breach of privacy: linking case

Consider an adversary $\mathcal{A}_{H_5}$ with a hint $H_5 = (L)$, where the set L consists of distinct pairs $\{a_1, a_2\} \subseteq A$, and that reveals the fact that, for all $\{a_1, a_2\} \in L$, $a_1 \equiv_\tau a_2$. Note that $|L| \leq |A|(|A|-1)/2$. $\mathcal{A}_{H_5}$ can restrict its attention to $\Pi_A(\mathcal{A}_{H_5}) = \{P : P \vdash A, \forall \{a_1, a_2\} \in L, a_1 \equiv_P a_2\}$. That is, the adversary can restrict its attention to those partitions that divide A such that, for all $\{a_1, a_2\} \in L$, a_1 and a_2 are equivalent. The number of such partitions is given by

$$|\Pi_A(\mathcal{A}_{H_5})| = B_{\Phi(A,L)} \tag{5}$$

where $\Phi(A, L)$ denotes the number of connected components in the graph (A, L) with vertices the elements in A and edges the elements in L. For a fixed L, and since τ is chosen uniformly at random, the unlinkability of the elements in A is $\mathcal{U}_A(\mathcal{A}_{H_5}) = \log_2(|\Pi_A(\mathcal{A}_{H_5})|)$ bits. If, on the other hand, L is chosen at random, then the expected value of (5) is given by

$$\mathrm{E}(|\Pi_A(\mathcal{A}_{H_5})|) = \mathrm{E}(B_{\Phi(A,L)}) = \sum_{k=1}^{|A|} B_k \Pr(\Phi(A, L) = k)$$

where $\Pr(\Phi(A, L) = k)$ denotes the probability that the graph (A, L) consists of exactly k connected components. Figure 4 shows the expected degree of unlinkability $\mathrm{E}(\mathcal{D}_A(\mathcal{A}_{H_5})) = \log_2(\mathrm{E}(B_{\Phi(A,L)})/\log_2(B_{|A|})$ as a function of $|A|$ and $|L|$, for the case where the elements in L are selected uniformly at random. Note that, in this case, the graph (A, L) is a random graph with $|L|$ edges,[5] and the

[5] See, for example, [5, 23] for a treatment of such graphs.

probability $\Pr(\Phi(A, L) = k)$ depends only on $|A|$ and $|L|$. Due to lack of an exact formula for $\Pr(\Phi(A, L) = k)$ (but see [21, 28]), the values shown in the figure are based on simulation. It is, of course, by no means necessary that the elements in L are selected uniformly at random; depending on the context and the power of the adversary, these elements may be selected by some other process that may lead to a faster or slower decrease in unlinkability.

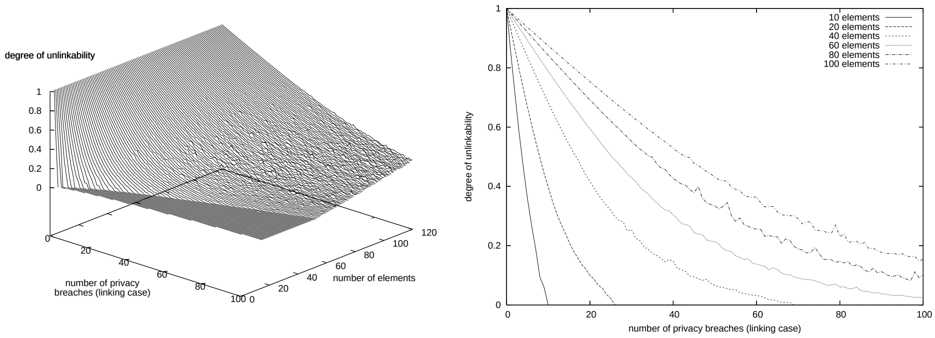


Fig. 4: Expected degree of unlinkability $\mathrm{E}(\mathcal{D}_A(\mathcal{A}_{H_5}))$ as a function of the number of elements $|A|$ and the number of privacy breaches (linking case) $|L|$. The elements in L are selected uniformly at random.

How to obtain this hint: Each element $\{a_1, a_2\} \in L$ can be seen as a privacy breach that tells the adversary that a_1 and a_2 are linked. In the setting of mix networks, a_1 and a_2 could be messages that leave the network; an adversary can link them based on e. g. their content or recipient. In the setting of anonymous credential systems, a_1 and a_2 could be transactions; an adversary can link them based on contextual information such as credential type [31], timing, location, or an identical piece of information that is attached to both transactions, e. g. a telephone number or an email address. In the setting of a combined election, a_1 and a_2 could be ballots; an adversary can link them based, for example, on the handwriting they may contain.

3.5 Breach of privacy: unlinking case

Consider an adversary $\mathcal{A}_{H_6}$ with a hint $H_6 = (U)$, where the set U consists of distinct pairs $\{a_1, a_2\} \subseteq A$, and that reveals the fact that, for all $\{a_1, a_2\} \in U$, $a_1 \not\equiv_\tau a_2$. Note that $|U| \leq |A| \cdot (|A| - 1)/2$. $\mathcal{A}_{H_6}$ can restrict its attention to $\Pi_A(\mathcal{A}_{H_6}) = \{P : P \vdash A, \forall\{a_1, a_2\} \in U, a_1 \not\equiv_P a_2\}$. That is, the adversary can restrict its attention to those partitions that divide A such that, for all $\{a_1, a_2\} \in U$, a_1 and a_2 are in different equivalence classes. The number of such partitions is given by

$$|\Pi_A(\mathcal{A}_{H_6})| = \sum_{U' \subseteq U} (-1)^{|U'|} B_{\Phi(A, U')} \tag{6}$$

where $\Phi(A, U')$ denotes the number of connected components in the graph (A, U') with vertices the elements in A and edges the elements in U' (for a proof see Appendix D). For a fixed U, and since τ is chosen uniformly at random, the unlinkability of the elements in A is $\mathcal{U}_A(\mathcal{A}_{H_6}) = \log_2(|\Pi_A(\mathcal{A}_{H_6})|)$ bits. If, on the other hand, U is selected at random, the expected value of (6), for a given number n of elements in U, is given by

$$\mathrm{E}(|\Pi_A(\mathcal{A}_{H_6})|) = \sum_{\substack{U \subseteq Z \\ |U|=n}} \Pr(U) \sum_{U' \subseteq U} (-1)^{|U'|} B_{\Phi(A,U')} \tag{7}$$

where Z denotes the set of all distinct pairs $\{a_1, a_2\} \subseteq A$ and $\Pr(U)$ denotes the probability that U is selected. Figure 5 shows the expected degree of unlinkability $\mathrm{E}(\mathcal{D}_A(\mathcal{A}_{H_6})) = \log_2 \mathrm{E}(|\Pi_A(\mathcal{A}_{H_6})|)/\log_2(B_{|A|})$ as a function of $|A|$ and $|U|$, for the case where the elements in U are selected uniformly at random.[6] It is, of course, by no means necessary that the elements in U are selected uniformly at random; depending on the context and the power of the adversary, these elements may be selected by some other process that may lead to a faster or slower decrease in unlinkability.

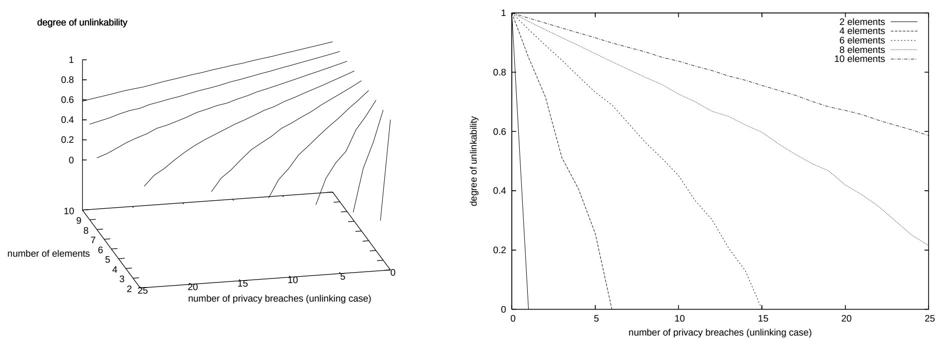


Fig. 5: Expected degree of unlinkability $\mathrm{E}(\mathcal{D}_A(\mathcal{A}_{H_6}))$ as a function of the number of elements $|A|$ and the number of privacy breaches (unlinking case) $|U|$. The elements in U are selected uniformly at random.

How to obtain this hint: Each element $\{a_1, a_2\} \in U$ can be seen as a privacy breach that tells the adversary that a_1 and a_2 are not linked. In the setting of mix networks, a_1 and a_2 could be messages that enter the network; an adversary can unlink them based on e. g. their content or sender. In the setting of anonymous credential systems, a_1 and a_2 could be transactions; an adversary can unlink them based on contextual information such as credential type,timing, location, or a piece of information that is attached to both transactions, e. g. two differing telephone numbers or email addresses. In the setting of a combined election, a_1

[6] Since evaluating (7) takes time exponential in $|U|$, the results shown in Figure 5 were obtained by simulation.

and a_2 could be ballots; an adversary can unlink them based, for example, on the handwriting they may contain.

Example: Let us briefly revisit the example from Section 2 at this point. Since the clerk knows that the calls λ_1 and λ_2 were made by men, and the calls λ_3 and λ_4 by women, he can effectively unlink λ_1 and λ_2 from λ_3 and λ_4. That is, he has a hint $H_6 = (U) = (\{(\lambda_1, \lambda_3), (\lambda_1, \lambda_4), (\lambda_2, \lambda_3), (\lambda_2, \lambda_4)\})$. In order to evaluate (6) the value of $\Phi(A, U')$ must be determined for each subset $U' \subset U$. In this example, we have

- the case where $U' = U$ and $\Phi(A, U') = 1$,
- four cases where $|U'| = 3$ and $\Phi(A, U') = 1$,
- six cases where $|U'| = 2$ and $\Phi(A, U') = 2$,
- four cases where $|U'| = 1$ and $\Phi(A, U') = 3$, and
- the case where $U' = \emptyset$ and $\Phi(A, \emptyset) = 4$.

That is, (6) evaluates to $B_1 - 4B_1 + 6B_2 - 4B_3 + B_4 = 1 - 4 + 12 - 20 + 15 = 4$, which coincides with the result from the trivial approach in Section 2.

3.6 Breach of privacy: combined case

Consider an oracle which answers questions of the form 'are the elements (a_1, a_2) linked?' by either 'yes' or 'no', depending on whether $a_1 \equiv_\tau a_2$ or $a_1 \not\equiv_\tau a_2$. An adversary $\mathcal{A}_{H_7}$ with access to such an oracle obtains, in effect, a hint $H_7 = (L, U)$, where L and U are as described above. Note that $L \cap U = \emptyset$ and $|L| + |U| \leq |A| \cdot (|A| - 1)/2$. $\mathcal{A}_{H_7}$ can restrict its attention to $\Pi_A(\mathcal{A}_{H_7}) = \{P : P \vdash A, \forall\{a_1, a_2\} \in L, a_1 \equiv_P a_2 \wedge \forall\{a_1, a_2\} \in U, a_1 \not\equiv_P a_2\}$, i.e. to those partitions that divide A such that, for all $\{a_1, a_2\} \in L$, a_1 and a_2 are equivalent and, for all $\{a_1, a_2\} \in U$, a_1 and a_2 are not equivalent. The number of such partitions is given by

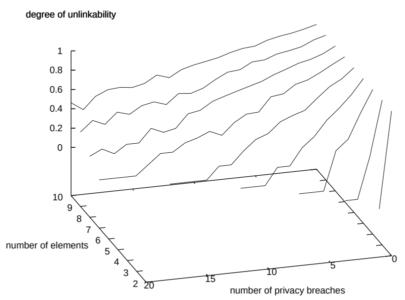
$$|\Pi_A(\mathcal{A}_{H_7})| = \sum_{U' \subseteq \tilde{U}} (-1)^{|U'|} B_{\Phi(\tilde{A}, U')} \tag{8}$$

where $\tilde{A}$ denotes the set of components of the graph (A, L), the set of edges $\tilde{U}$ contains the edge $\{c_1, c_2\}$, where $c_1, c_2 \in \tilde{A}$ and $c_1 \neq c_2$, if and only if U contains a pair $\{a_1, a_2\}$ such that either ($a_1 \in c_1$ and $a_2 \in c_2$), or ($a_1 \in c_2$ and $a_2 \in c_1$), and $\Phi(\tilde{A}, U')$ denotes the number of components in the the graph $(\tilde{A}, U')$ with vertices the elements in $\tilde{A}$ and edges the elements in U'. In effect, the difference between equations (6) and (8) lies in the fact that the latter operates on a quotient graph — induced by L — of the graph on which the former operates.

For a fixed set of oracle calls, i.e. a fixed L and U, and since τ is chosen uniformly at random, the unlinkability of the elements in A is $\mathcal{U}_A(\mathcal{A}_{H_7}) = \log_2(|\Pi_A(\mathcal{A}_{H_7})|)$ bits. If, on the other hand, τ and the oracle calls are selected at random, the expected value of (8), if exactly $n = |L| + |U|$ oracle calls are made, is given by

$$\mathrm{E}(|\Pi_A(\mathcal{A}_{H_7})|) = \sum_{\substack{L, U \subseteq Z \\ |L| + |U| = n}} \Pr(L \wedge U) \sum_{U' \subseteq \tilde{U}} (-1)^{|U'|} B_{\Phi(\tilde{A}, U')} \tag{9}$$

where Z denotes the set of all distinct pairs $\{a_1, a_2\} \subseteq A$ and $\Pr(L \wedge U)$ denotes the probability of selecting τ and oracle calls such that L and U are the results of the oracle's answers. Figure 6 shows the expected degree of unlinkability $\mathrm{E}(\mathcal{D}_A(\mathcal{A}_{H_7})) = \log_2 \mathrm{E}(|\Pi_A(\mathcal{A}_{H_7})|)/\log_2(B_{|A|})$ as a function of $|A|$ and $|L \cup U|$, for the case where τ and the elements in $L \cup U$ are selected uniformly at random.[7] It is, of course, by no means necessary that τ and the elements in $L \cup U$ are selected uniformly at random; depending on the context and the power of the adversary, these elements may be selected by some other process that may lead to a faster or slower decrease in unlinkability.



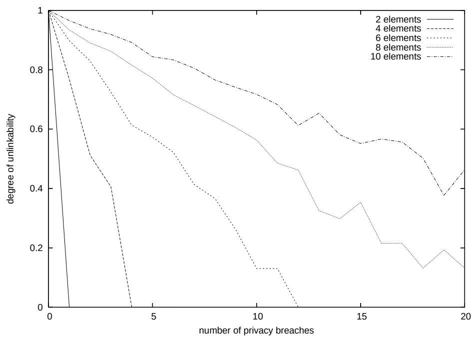


Fig. 6: Expected degree of unlinkability $\mathrm{E}(\mathcal{D}_A(\mathcal{A}_{H_7}))$ as a function of the number of elements $|A|$ and the number of privacy breaches $|L \cup U|$. The target partition τ and the elements in $L \cup U$ are selected uniformly at random.

How to obtain this hint: See sections 3.4 and 3.5.

4 Conclusion

In this paper, we considered the setting of a system that protects the unlinkability of certain elements of interest, and an adversary with the goal to nevertheless link these elements. We studied how a number of contextual hints, if disclosed to the adversary, affect its ability to link the elements. We conclude that, an adversary that knows only the number or the cardinality of the equivalence classes that the elements must be divided into (or a 'reference partition' as described in Section 3.3), is, in most cases, unable to link the elements with certainty. However, as Figures 1, 2, and 3 demonstrate, such knowledge nevertheless reduces the degree of unlinkability of the elements to a significant extent.

By contrast, an adversary that breaches privacy by linking and/or by unlinking pairs of elements, is able to identify the target partition (i.e. uniquely link all elements) after a certain number of breaches have occurred. However, if

[7] Since evaluating (9) takes time exponential in $|U|$, the results shown in Figure 6 were obtained by simulation.

the adversary is limited to linking (resp. unlinking), then this required number of privacy breaches can occur only in the extreme case where all elements are equivalent (resp. if each element constitutes a separate equivalence class) in the target partition. Figures 4, 5, and 6 demonstrate the significance of such breaches in an 'average' case, i.e. in the case where randomly selected pairs are linked or unlinked. Note that linking (Figure 4) has a significantly more dramatic effect on unlinkability compared to unlinking (Figure 5). This however, is not surprising, since 'belonging to the same equivalence class' is a transitive relation, while 'belonging to different equivalence classes' is not.

Finally, note that the list of hints studied in this paper is by no means exhaustive and that some types of hint may be of more practical relevance than others. Identifying other, practical types of hint that help an adversary to link otherwise unlinkable elements, and studying their effect on unlinkability, is a direction for further research.

Acknowledgements

The authors would like to thank Michael Braun for his insightful suggestions, and Svante Janson for his pointers to some of the literature on random graphs.

A Counterexample to the theorem in [37]

Theorem 1 in [37] claims that it cannot be reached that, for all arbitrarily chosen pairs $\{a_1, a_2\} \subseteq A$, $\Pr(a_1 \equiv_\tau a_2) = \Pr(a_1 \not\equiv_\tau a_2) = 1/2$, from the point of view of $\mathcal{A}_{H_2}$.[8] This is wrong as the claim does not hold, for example, if $|A| = 4$ and $H_2 = (1, 3)$. We remark that, more generally, the claim does not hold for all solutions of the system of equations

$$\sum_{\beta \in H_2} \beta = |A|$$
$$\sum_{\beta \in H_2} \beta^2 = (|A|^2 + |A|)/2.$$

B Proof of (2) and (3)

Consider the task of dividing the elements in a set A into α subsets such that, for all $1 \leq i \leq \alpha$, the ith subset contains exactly β_i elements. One can perform this task by first ordering the elements in A, and then putting the first β_1 elements into the first subset, the next β_2 elements into the second subset, and so on. If one performs this task for all $|A|!$ orderings of A, one ends up with only

[8] The claim has been rephrased in order to fit our notation.

$|A|!/(\beta_1! \cdot \beta_2! \cdots \beta_\alpha!)$ different outcomes, because permuting the elements in each subset does not make a difference. Moreover, since the equivalence classes of a partition are *not* ordered, i.e. one can permute the equivalence classes of the same size without changing the partition, the number of *distinct* partitions that divide A into α subsets of cardinality $\beta_1, \beta_2, \ldots, \beta_\alpha$, is given by (2). Equation (3) follows as a special case. $\square$

C Proof of (4)

Consider a set A, a partition $\{R_1, R_2, \ldots, R_\beta\} \vdash A$ that divides A into $\beta = |A|/\alpha$ subsets of equal cardinality α, and the task of dividing A into α subsets of equal cardinality β, such that each subset contains exactly one element from $R_1, R_2, \ldots, R_\beta$. For ease of presentation, assume that, for all $1 \leq i \leq \beta$, there exists an ordering on the elements in R_i. Then one can perform this task by grouping the first element in each of $R_1, R_2, \ldots, R_\beta$ into Q_1, the second element in each of $R_1, R_2, \ldots, R_\beta$ into Q_2, and so on. By doing this, one ends up with a partition $\{Q_1, Q_2, \ldots, Q_\alpha\} \vdash A$ that meets the requirements.

It is possible to construct another partition $\{Q_1, Q_2, \ldots, Q_\alpha\} \vdash A$ that meets the requirements by permuting the elements in $R_1, R_2, \ldots, R_\beta$ and then repeating the above procedure. Indeed, one can construct all partitions that meet the requirements by repeating the above procedure for all combinations of permutations of the elements in $R_1, R_2, \ldots, R_\beta$. If one does this for all such combinations, of which there exist $\kappa = |R_1|! \cdot |R_2|! \cdots |R_\beta|! = (\alpha!)^\beta$, each of the resulting κ partitions will appear exactly $\alpha!$ times, namely once for each permutation of the sets $Q_1, Q_2, \ldots, Q_\alpha$. Thus, the number of *distinct* partitions that divide A into α subsets of equal cardinality β, such that each subset contains exactly one element from $R_1, R_2, \ldots, R_\beta$, is given by (4). $\square$

D Proof of (6)

Let (A, U) denote an undirected graph without loops, $\Phi(A, U)$ the number of connected components of (A, U), B_n the number of partitions of a set with n elements (see (1)), and $\Psi(A, U)$ the number of partitions of A which are such that no edge $e \in U$ connects two vertices in the same equivalence class. That is, $\Psi(A, U) \stackrel{def}{=} |\{P : P \vdash A, \forall\{a_1, a_2\} \in U, a_1 \not\equiv_P a_2\}|$. We prove (6), i.e.

$$\Psi(A, U) = \sum_{U' \subseteq U} (-1)^{|U'|} B_{\Phi(A, U')},$$

by induction over $|U|$. We actually prove a stronger result, namely that the above equation holds not only if (A, U) is a simple graph, but also if it is a multigraph without loops.

Proof. If $U = \emptyset$, then $\Psi(A, U) = B_{|A|}$ in accordance with (6). For $U \neq \emptyset$, let e denote an edge in U, and $Y = U \setminus \{e\}$. We distinguish between the following two cases.

Case 1. There exists an edge $e' \in Y$ connecting the same pair of nodes as e. In this case,

$$\Psi(A, U) = \Psi(A, Y)$$

by definition of Ψ,

$$= \sum_{Y' \subseteq Y} (-1)^{|Y'|} B_{\Phi(A,Y')}$$

by induction since $|Y| = |U| - 1$,

$$\begin{aligned} &= \sum_{Y' \subseteq Y} (-1)^{|Y'|} B_{\Phi(A,Y')} \\ &+ \sum_{\substack{U' \subseteq U \setminus \{e'\} \\ e \in U'}} \left((-1)^{|U'|} B_{\Phi(A,U')} + (-1)^{|U' \cup \{e'\}|} B_{\Phi(A,U' \cup \{e'\})}\right) \end{aligned}$$

because $(-1)^{|U' \cup \{e'\}|} B_{\Phi(A,U' \cup \{e'\})} = -(-1)^{|U'|} B_{\Phi(A,U')}$, and, finally,

$$= \sum_{U' \in U} (-1)^{|U'|} B_{\Phi(A,U')}$$

since for all $U' \subseteq U$ it holds that either $e \notin U'$, $e \in U'$ and $e' \notin U'$, or $e \in U'$ and $e' \in U'$.

Case 2. There exists no edge in Y connecting the same pair of nodes as e. In this case, by definition of Ψ,

$$\Psi(A, U) = \Psi(A, Y) - X,$$

where X denotes the number of partitions of A such that the nodes connected by e are equivalent, but no edge in Y connects equivalent nodes. That is, $X = |\{P : P \vdash A, \forall \{a_1, a_2\} \in Y, a_1 \not\equiv_P a_2 \wedge a_e \equiv_P a'_e\}|$, where a_e and a'_e denote the nodes connected by e.

We now define $(\tilde{A}, \tilde{U})$ as the graph obtained from (A, U) by merging the nodes connected by e. Note that the edges in $\tilde{U}$ are in one-to-one correspondence with those in $U \setminus \{e\} = Y$, in particular $|\tilde{U}| = |U| - 1$. Also note that due to the merging, even if (A, U) is a simple graph, $(\tilde{A}, \tilde{U})$ may be a multigraph (although, since e is removed, without any loops). By construction of $(\tilde{A}, \tilde{U})$, we have $X = \Psi(\tilde{A}, \tilde{U})$ and, therefore,

$$\begin{aligned} \Psi(A, U) &= \Psi(A, Y) - \Psi(\tilde{A}, \tilde{U}) \\ &= \sum_{Y' \subseteq Y} (-1)^{|Y'|} B_{\Phi(A,Y')} - \sum_{\tilde{U}' \subseteq \tilde{U}} (-1)^{|\tilde{U}'|} B_{\Phi(\tilde{A},\tilde{U}')} \end{aligned}$$

by induction, since $|Y| = |U| - 1$ and $|\tilde{U}| = |U| - 1$,

$$= \sum_{Y' \subseteq Y} (-1)^{|Y'|} B_{\Phi(A,Y')} - \sum_{\substack{U' \subseteq U \\ e \in U'}} (-1)^{|U'|-1} B_{\Phi(A,U')}$$

because $\Phi(\tilde{A}, \tilde{U}') = \Phi(A, U')$ for the subset $U' \subset U$ containing the nodes corresponding to those in $\tilde{U}'$ and additionally e,

$$= \sum_{U' \in U} (-1)^{|U|} B_{\Phi(A,U')}.$$

This completes the proof. $\square$

References

1. B. Adida. Advances in cryptographic voting systems. PhD thesis, Massachusetts Institute of Technology, 2006.
2. D. Agrawal, D. Kesdogan, and S. Penz. Probabilistic treatment of mixes to hamper traffic analysis. In *2003 IEEE Symposium on Security and Privacy (S&P 2003), 11–14 May 2003, Berkeley, CA, USA*, pages 16–27. IEEE Computer Society, 2003.
3. E. T. Bell. Exponential numbers. *American Mathematical Monthly*, 41:411–419, 1934.
4. O. Berthold, H. Federrath, and S. Köpsell. Web mixes: A system for anonymous and unobservable internet access. In H. Federrath, editor, *Designing Privacy Enhancing Technologies, International Workshop on Design Issues in Anonymity and Unobservability, Berkeley, CA, USA, July 25–26, 2000, Proceedings*, volume 2009 of *Lecture Notes in Computer Science*, pages 115–129. Springer Verlag, Berlin, 2001.
5. B. Bollobás. *Random Graphs*. Number 73 in Cambridge Studies in Advanced Mathematics. Cambridge University Press, second edition, 2001.
6. D. Boneh and P. Golle. Almost entirely correct mixing with applications to voting. In V. Atluri, editor, *Proc. of the 9th ACM Conference on Computer and Communications Security*, pages 68–77. ACM Press, 2002.
7. J. Camenisch and A. Lysyanskaya. An efficient system for non-transferable anonymous credentials with optional anonymity revocation. In B. Pfitzmann, editor, *Advances in Cryptology — EUROCRYPT 2001, International Conference on the Theory and Application of Cryptographic Techniques, Innsbruck, Austria, May 6–10, 2001, Proceedings*, volume 2045 of *Lecture Notes in Computer Science*, pages 93–118. Springer Verlag, Berlin, 2001.
8. J. Camenisch and A. Lysyanskaya. Signature schemes and anonymous credentials from bilinear maps. In M. Franklin, editor, *Proceedings of the 24th Annual International Cryptology Conference, Santa Barbara, California, USA, August 15-19 — CRYPTO 2004*, volume 3152 of *Lecture Notes in Computer Science*, pages 56–72. Springer Verlag, Berlin, 2004.
9. D. Chaum. Untraceable electronic mail, return addresses, and digital pseudonyms. *Communications of the ACM*, 24(2):84–90, 1981.
10. D. Chaum. Security without identification: transaction systems to make big brother obsolete. *Communications of the ACM*, 28(10):1030–1044, 1985.

11. D. Chaum. Showing credentials without identification: transferring signatures between unconditionally unlinkable pseudonyms. In J. Seberry and J. Pieprzyk, editors, *Advances in Cryptology – AUSCRYPT 90*, volume 453 of *Lecture Notes in Computer Science*, pages 246–264. Springer Verlag, Berlin, 1990.
12. L. Chen. Access with pseudonyms. In E. Dawson and J. D. Golic, editors, *Cryptography: Policy and Algorithms, International Conference, Brisbane, Queensland, Australia, July 3-5, 1995, Proceedings*, number 1029 in Lecture Notes in in Computer Science, pages 232–243. Springer Verlag, Berlin, 1995.
13. L. Chen, M. Enzmann, A.-R. Sadeghi, M. Schneider, and M. Steiner. A privacy-protecting coupon system. In A. S. Patrick and M. Yung, editors, *Financial Cryptography and Data Security, 9th International Conference, FC 2005, Roseau, The Commonwealth of Dominica, February 28–March 3, 2005, Revised Papers*, volume 3570 of *Lecture Notes in Computer Science*, pages 93–108. Springer Verlag, Berlin, 2005.
14. I. Damgård. Payment systems and credential mechanisms with provable security against abuse by individuals. In S. Goldwasser, editor, *Advances in Cryptology — CRYPTO '88: Proceedings*, number 403 in Lecture Notes in Computer Science, pages 328–335. Springer Verlag, Berlin, 1990.
15. G. Danezis and A. Serjantov. Statistical disclosure or intersection attacks on anonymity systems. In J. Fridrich, editor, *Information Hiding, 6th International Workshop, IH 2004, Toronto, Canada, May 23–25, 2004, Revised Selected Papers*, volume 3200 of *Lecture Notes in Computer Science*, pages 293–308. Springer Verlag, Berlin, May 2004.
16. C. Díaz and B. Preneel. Reasoning about the anonymity provided by pool mixes that generate dummy traffic. In J. Fridrich, editor, *Information Hiding, 6th International Workshop, IH 2004, Toronto, Canada, May 23–25, 2004, Revised Selected Papers*, volume 3200 of *Lecture Notes in Computer Science*, pages 309–325. Springer Verlag, Berlin, May 2004.
17. C. Díaz and B. Preneel. *Security, Privacy and Trust in Modern Data Management*, chapter Accountable Anonymous Communication. Springer Verlag, Berlin, 2006. in print.
18. C. Díaz, S. Seys, J. Claessens, and B. Preneel. Towards measuring anonymity. In R. Dingledine and P. F. Syverson, editors, *Privacy Enhancing Technologies, Second International Workshop, PET 2002, San Francisco, CA, USA, April 14–15, 2002, Revised Papers*, number 2482 in Lecture Notes in Computer Science, pages 54–68. Springer Verlag, Berlin, 2002.
19. R. Dingledine, N. Mathewson, and P. F. Syverson. Tor: The second-generation onion router. In *Proceedings of the 13th USENIX Security Symposium, August 9–13, 2004, San Diego, CA, USA*, pages 303–320. USENIX, 2004.
20. J. Douceur. The Sybil attack. In P. Druschel, F. Kaashoek, and A. Rowstron, editors, *Peer-to-Peer Systems, First International Workshop, IPTPS 2002, Cambridge, MA, USA, March 7–8, 2002, Revised Papers*, number 2429 in Lecture Notes in Computer Science, pages 251–260. Springer Verlag, Berlin, 2002.
21. P. Erdős and A. Rényi. On random graphs I. *Publicationes Mathematicae Debrecen*, 6:290–297, 1959.
22. R. L. Graham, D. E. Knuth, and O. Patashnik. *Concrete Mathematics: A Foundation for Computer Science*, chapter 6.1, pages 257–267. Addison-Wesley, 2nd edition, 1994.
23. S. Janson, T. Łuczak, and A. Ruciński. *Random Graphs.* Interscience Series in Discrete Mathematics and Optimization. John Wiley & Sons, Inc., 2000.

24. D. Kagan. The origin and purposes of ostracism. *Hesperia*, 30(4):393–401, October 1961.
25. D. Kesdogan, D. Agrawal, and S. Penz. Limits of anonymity in open environments. In F. Petitcolas, editor, *Information Hiding, 5th International Workshop, IH 2002, Noordwijkerhout, The Netherlands, October 7–9, 2002, Revised Papers*, volume 2578 of *Lecture Notes in Computer Science*, pages 53–69. Springer Verlag, Berlin, 2003.
26. D. Kesdogan, J. Egner, and R. Büschkes. Stop-and-go-mixes providing probabilistic anonymity in an open system. In D. Aucsmith, editor, *Information Hiding, Second International Workshop, Portland, Oregon, USA, April 14–17, 1998, Proceedings*, volume 1525 of *Lecture Notes in Computer Science*, pages 83–98. Springer Verlag, Berlin, 1998.
27. M. Klonowski and M. Kutylowski. Provable anonymity for networks of mixes. In M. Barni, J. Herrera-Joancomartí, S. Katzenbeisser, and F. Pérez-González, editors, *Information Hiding, 7th International Workshop, IH 2005, Barcelona, Spain, June 6–8, 2005, Revised Selected Papers*, volume 3727 of *Lecture Notes in Computer Science*, pages 26–38. Springer Verlag, Berlin, 2005.
28. R. F. Ling. The expected number of components in random linear graphs. *The Annals of Probability*, 1(5):876–881, 1973.
29. A. Lysyanskaya, R. L. Rivest, A. Sahai, and S. Wolf. Pseudonym systems. In H. M. Heys and C. M. Adams, editors, *Selected Areas in Cryptography, 6th Annual International Workshop, SAC'99, Kingston, Ontario, Canada, August 9-10, 1999, Proceedings*, volume 1758 of *Lecture Notes in Computer Science*, pages 184–199. Springer Verlag, Berlin, 2000.
30. N. Mathewson and R. Dingledine. Practical traffic analysis: Extending and resisting statistical disclosure. In D. Martin and A. Serjantov, editors, *Privacy Enhancing Technologies, 4th International Workshop, PET 2004, Toronto, Canada, May 26–28, 2004, Revised Selected Papers*, volume 3424 of *Lecture Notes in Computer Science*, pages 17–34. Springer Verlag, Berlin, 2004.
31. A. Pashalidis and B. Meyer. Linking anonymous transactions: The consistent view attack. In *Proceedings of Privacy Enhancing Technologies, 6th International Workshop, PET 2006*, number 4258 in Lecture Notes in Computer Science (to be published). Springer Verlag, Berlin, 2006.
32. G. Persiano and I. Visconti. An efficient and usable multi-show non-transferable anonymous credential system. In A. Juels, editor, *Proceedings of the Eighth International Financial Cryptography Conference (FC '04)*, volume 3110 of *Lecture Notes in Computer Science*, pages 196–211. Springer Verlag, Berlin, 2004.
33. M. K. Reiter and A. D. Rubin. Crowds: anonymity for Web transactions. *ACM Transactions on Information and System Security*, 1(1):66–92, 1998.
34. M. Rennhard and B. Plattner. Introducing morphmix: peer-to-peer based anonymous internet usage with collusion detection. In S. Jajodia and P. Samarati, editors, *Proceedings of the 2002 ACM Workshop on Privacy in the Electronic Society, WPES 2002, Washington, DC, USA, November 21, 2002*, pages 91–102. ACM, 2002.
35. G. C. Rota. The number of partitions of a set. *American Mathematical Monthly*, 71:498–504, 1964.
36. A. Serjantov and G. Danezis. Towards an information theoretic metric for anonymity. In R. Dingledine and P. F. Syverson, editors, *Privacy Enhancing Technologies, Second International Workshop, PET 2002, San Francisco, CA, USA, April 14–15, 2002, Revised Papers*, volume 2482 of *Lecture Notes in Computer Science*, pages 41–53. Springer Verlag, Berlin, 2002.

37. S. Steinbrecher and S. Köpsell. Modelling unlinkability. In R. Dingledine, editor, *Privacy Enhancing Technologies, Third International Workshop, PET 2003, Dresden, Germany, March 26-28, 2003, Revised Papers*, volume 2760 of *Lecture Notes in Computer Science*, pages 32–47. Springer Verlag, Berlin, 2003.
38. E. R. Verheul. Self-blindable credential certificates from the Weil pairing. In C. Boyd, editor, *ASIACRYPT '01: Proceedings of the 7th International Conference on the Theory and Application of Cryptology and Information Security*, volume 2248 of *Lecture Notes in Computer Science*, pages 533–551. Springer Verlag, Berlin, 2001.