J. Cryptol. (2015) 28: 623–640
DOI: 10.1007/s00145-013-9171-8

Journal of CRYPTOLOGY

# Provable Unlinkability Against Traffic Analysis with Low Message Overhead*

Ron Berman†
Haas School of Business, UC Berkeley, Berkeley, CA 94720-1900, USA
ron_berman@haas.berkeley.edu

Amos Fiat
Department of Computer Science, Tel Aviv University, Tel Aviv 69978, Israel
fiat@tau.ac.il

Marcin Gomułkiewicz, Marek Klonowski‡, and Mirosław Kutyłowski‡
Institute of Mathematics and Computer Science, Wrocław University of Technology, ul. Wybrzeże Wyspiańskiego 27, 50-370 Wrocław, Poland
marcin.gomulkiewicz@gmail.com; Marek.Klonowski@pwr.wroc.pl; Miroslaw.Kutylowski@pwr.wroc.pl

Tomer Levinboim†§
Viterbi School of Engineering, University of Southern California, Los Angeles, CA 90089, USA
levinboim.tomer@gmail.com

Amnon Ta-Shma§
Department of Computer Science, Tel Aviv University, Tel Aviv 69978, Israel
amnon@tau.ac.il

Communicated by Rafail Ostrovsky

Received 28 July 2010
Online publication 14 December 2013

**Abstract.** Rackoff and Simon proved that a variant of Chaum's protocol for anonymous communication, later developed as the Onion Routing Protocol, is unlinkable against a passive adversary that controls all communication links and most of the nodes in a communication system. A major drawback of their analysis is that the protocol is secure only if *(almost) all* nodes participate at *all* times. That is, even if only $n \ll N$ nodes wish to send messages, *all* $N$ nodes have to participate in the protocol at all times. This suggests necessity of sending dummy messages and a high message overhead.

Our first contribution is showing that this is unnecessary. We relax the adversary model and assume that the adversary only controls a certain fraction of the communication links in the communication network. We think this is a realistic adversary model.

* This paper is based on conference papers [5,15].

† This work was done while R. Berman and T. Levinboim were M.Sc. students at Tel Aviv University.

‡ M. Klonowski and M. Kutyłowski were partially supported by the EU within the 6th Framework Programme under contract 001907 (DELIS).

§ A. Ta-Shma and T. Levinboim were supported by Israel Science Foundation grant 217/05 and by USA Israel BSF grant 2004390.

For this adversary model we show that a low message overhead variant of Chaum's protocol is provably secure.

Furthermore, all previous security proofs assumed the a priori distribution on the messages is uniform. We feel this assumption is unrealistic. The analysis we give holds for any a priori information on the communication distribution. We achieve that by combining Markov chain techniques together with information theory tools in a simple and elegant way.

**Key words.** Mix protocol, Traffic analysis, Mixing time, Markov Chain, Unlinkability.

## 1. Introduction

We focus on communication protocols that allow *anonymous* communication even if the network is partially under an adversarial control. The anonymous communication problem is very basic, and models well the privacy issues that occur when exchanging messages over a public network such as Internet. It also serves as an underlying platform in several cryptographic protocols, most notably in some e-auctions and electronic voting protocols. Yet, up to date, there is no general satisfying solution to the problem.

The phrase "anonymous" can take several interpretations. First, we would like to hide the *content* of a sent message $m$ (this is sometimes called "confidentiality"). Second, we might want to have *senders and receivers* anonymity. And finally, we would like to have *unlinkability* meaning that even if an adversary knows the set $\{a_1, \ldots, a_n\}$ of senders and the set $\{b_1, \ldots, b_n\}$ of receivers, he cannot link the senders to the receivers.

The attack model also has several variants. In the *passive* model the adversary is curious but honest, i.e., it listens on the communication links under its control, but no node deviates from the protocol. We call such an adversary an *eavesdropper*. An *active* adversary might change, initiate or delete messages. Both the passive and the active adversaries can be *non-adaptive*, meaning that they determine the communication links under their control before the protocol begins, or *adaptive*, meaning that they may acquire communication links during the execution of the protocol and based on the communication so far.

Finally, there is the *cost* issue. Two common cost functions are *time delay* which is the time it takes a message to reach its destination, and *message overhead* which is the number of messages transmitted in the protocol per send request. More precisely, suppose at some time we have $n$ senders, and the protocol takes $t$ steps and $M$ messages to deliver the $n$ messages to their destination, then the time delay is $t$ and the message overhead is $M/n$. For simplicity we assume a *synchronous* communication model.

Current solutions can be divided into three groups:

Solutions assuming a trusted party. A simplified solution of this type is: "To send a message $m$ to $b$, send $(m, b)$ encrypted to the trusted party and ask it to send $m$ to $b$". For a survey of such solutions look at Danezis and Diaz [14, Sect. 2].

Heuristic solutions. Many papers offer a protocol, and sometimes even propose an attack model, but do not provide a security proof. The most notable example of this approach is Chaum's seminal paper from 1979 [6]. This short paper (only two page long) is full with bright ideas, and is a basis for many follow-ups, including this paper. We refer the reader to [14, Sect. 3] for a survey on the huge body of work building upon Chaum's seminal work. Chaum's paper suggests a rigorous attack model but

**Table 1.** Rigorous protocols with large time delay. The size of the network is $N$, and there are $n \ll N$ active nodes. $\alpha$ is any fixed fraction arbitrarily close to 1

| Protocol | Resources under adversary control | Time delay | Remarks |
|---|---|---|---|
| AMPC [20] | Fraction $\alpha$ of nodes, fraction $\alpha$ of links | $O(N)$ | Limited adversary<br>Large delay |
| Busses [4] | Fraction $\alpha$ of nodes, all links | $O(N)$ | Large delay |

gives no proof. This is also the typical situation for much of the work surveyed in [14, Sect. 3].

We believe informal work has many disadvantages and often leads to ad hoc solutions and wrong claims. For example, the protocol Chaum suggested in [6] uses RSA as an encryption method. In 1989, Pfitzmann and Pfitzmann [22] showed how to use the multiplicative homomorphism property of RSA to break security of the protocol for the attack model Chaum claimed. One can, of course, modify the protocol and make it immune to the attack suggested in [22]. Yet, other attacks exist, and we refer the interested reader to Danezis and Diaz [14, Sect. 3.1] for a nice survey. The bottom line, in our opinion, is that it is not enough to suggest a protocol with *heuristic* security, and instead one should look for a protocol with *provable* security.

Rigorous work. In 1988, Chaum suggested the DC-Nets protocol [8]. The protocol is information theoretic secure, and is a special case of secure computation. It guarantees both sender and receiver anonymity and unlinkability, and it is secure against passive adversaries, as well as some stronger forms of adversaries. The protocol uses shared secret keys and requires a secure and reliable broadcast mechanism. Furthermore, all nodes have to participate at each stage of the protocol, even if only few of them actually wish to send a message.

The buses protocol [4] has a rigorous proof, but has a large time delay and a high message overhead. Again, all nodes have to participate at each stage of the protocol, leading to a high message overhead when the number of active nodes $n$ is much smaller than the network size $N$.

Rackoff and Simon [25] suggested several solutions, some building upon Chaum's work [6], giving a variant of Chaum's protocol a rigorous security proof. Again, all nodes have to participate at each stage of the protocol, so the protocol has a high message overhead when $n \ll N$. The analysis proves the protocol has a polylogarithmic time delay, and this can be improved to $O(\log^2 N)$ using the techniques of Czumaj et al. [10] and Czumaj and Kutyłowski [9].[1] We remark that our protocol achieves $O(\log n)$ time delay.

In all the above protocols all nodes have to participate at each stage of the protocol, leading to a high message overhead when $n \ll N$. This is a major drawback. Imagine for example a network with one million users, in which, on average, only 1000 users wish to send messages at a given time. A protocol that forces all the one million users to send messages at all times is clearly impractical. Our goal in this paper is to show a rigorous protocol having low-message overhead.

[1] It seems that $O(\log^2 N)$ is the correct time delay for the protocol suggested in [25]. This is because the protocol uses $O(\log n)$ stages, each requiring mixing messages within certain subgroups of nodes.

**Table 2.** Rigorous protocols with small time delay but high message overhead. Notation is as in Table 1

| Protocol | Resources under adversarial control | Time delay | Message overhead | Remarks |
|---|---|---|---|---|
| Mix-nets + sorting [25] | All links, no nodes | polylog($N$) | $N$polylog($N$) | Passive adversary<br>High message load |
| DC-nets [8] | All nodes, no links | $O(1)$ | $O(N^2)$ | Passive adversary<br>High message load |
| Our results | A constant fraction of links | polylog($n$) | 1 | Passive adversary<br>Low message load |

Some of the protocols above have both large time delay and high message overhead (those are summarized in Table 1) and some have small time delay and high message overhead (those are summarized in Table 2). The high message overhead and high delay are often pointed out as a great weakness and as a non-realistic assumption when considering email and Internet networks, see, e.g., Danezis and Diaz [14].

Another protocol with rigorous analysis is Crowds [24]. In Crowds a node takes a probabilistic decision whether to send the message to its final destination, or to forward it to another intermediate node. The security Crowds provides is very mild (it is proportional to the path length).

Our goal in this paper is to rigorously analyze a communication protocol based on Chaum's idea. Our aim is to show that it is provably unlinkable against passive, non-adaptive adversaries and has a low message overhead as well as a small time delay even when the number of active nodes $n$ is much smaller than the network size $N$.

### 1.1. *Mix-Based Systems*

Chaum's approach to communication anonymity [7] uses two fundamental building blocks.

*A Mix* A mix accepts batches of encrypted messages, each with its desired target address. It decrypts the messages and then forwards each message to its destination according to some *predefined* order, e.g., the lexicographic order.

*Onion Routing* Each sender randomly chooses a list of mixes through which the message is to be routed. The message is then encrypted multiple times, each additional encryption layer contains the needed information for a specific mix on the path. Chaum's protocol also allows returning messages back to the sender without revealing the sender's identity even to the receiver. This is achieved by having the message include two parts, the regular forward message "onion", and another separate "onion" containing the needed routing information and encryption keys that allow backward return of answers. Chaum's protocol (with some modifications) is the basis to the protocol given in Sect. 1.4.

### 1.2. *Traffic Analysis and Adversary Model*

Chaum's protocol hides the content of the message and its destination using encryption. Chaum's protocol can therefore be seen as a reduction from the unlinkability problem

to the *traffic analysis* problem. In the traffic analysis problem, $n$ packets are routed in the network. The $n$ packets are indistinguishable to the adversary and the only way the adversary can gain information on the communication is by analyzing the traffic rather than the messages content. Chaum does not give a formal proof of this reduction either. Nevertheless, in 2005, Camenisch and Lysyanskaya [11] defined and designed a provably secure onion routing scheme. Using their work one can see that Chaum's protocol is indeed a provable reduction from unlinkability to traffic analysis. In this paper we concern ourselves only with the traffic analysis problem.

The traffic analysis problem was not analyzed at all in [6]. In fact, Chaum's protocol does not withstand malicious adversaries [23] and other attacks (e.g., mix floods and replay attacks). In 1993, Rackoff and Simon gave a variant of Chaum's protocol a rigorous analysis. This forced some changes to the attack model. Most importantly, Rackoff and Simon mainly deal with passive adversaries.

Following Chaum, Rackoff and Simon assume that *all communication links* and some constant fraction of the nodes are under adversarial control. In this attack model mixing can happen only when a honest node receives two or more messages originally sent by honest senders in the same step. We call this situation *node mixing*. Rackoff and Simon set the number of mixes $M$ to equal the number of nodes $N$, for otherwise very few mixes take the burden of very many nodes. They also require that *all* nodes are active at *all* times, which leads to the huge message overhead mentioned before when $n \ll N$. We now explain why this choice is unavoidable in this adversarial model.

For simplicity, we assume that at each time there are $P$ nodes wishing to send a message, $H$ of them are honest and the rest are dishonest. Furthermore, we let the protocol know $N$, $P$ and $H$ in advance. As we do not have control over the adversary, we should be able to deal with the scenario where all the dishonest nodes are active at all times, and where the number $P - H$ of dishonest active nodes is $\Omega(N)$. Also, as we said before, because of load considerations the protocol chooses $M$ to equal $P$. If $M \gg H^2$ then, by the birthday paradox, we very rarely expect to see two honest messages reaching the same mix, and mixing will not take place. We therefore need the number $H$, of *active* honest players, to be at least $\Omega(\sqrt{M}) = \Omega(\sqrt{N})$. In fact, for the protocol to work well we should have $H = \Omega(N)$. Rackoff and Simon simply take $P = N$ and make all nodes active at all times.

We show that the problem disappears, if we slightly change the attack model (but still we keep it realistic). Specifically, we replace the assumption that the eavesdropper controls *all the communication links* with the assumption that the eavesdropper controls an arbitrarily large but *fixed fraction* of the communication links. We show that in this case the key parameter is only the number $n$ of active honest nodes, regardless of the number of active dishonest players.

The main difference between the two models is that if the eavesdropper controls *all communication links*, then it is necessary to send messages from most of the nodes (sending dummy ones if necessary), whereas if we assume the eavesdropper only controls *many communication links*, a small number of messages sent does not prevent unlinkability within the sets of senders and receivers, and the system may have low message overload.

Our analysis does not use node mixing. Instead we introduce a paradigm that we call *layer mixing*. Layer mixing occurs when two honest nodes communicate with two

other honest nodes using secure communication links, i.e. links not controlled by the adversary. Layer mixing can happen even when node mixing does not. To demonstrate this, assume there are only two active honest nodes. As Chaum pointed out [6], if at some point the two messages reach the same honest node, then thereafter the adversary cannot link the senders to the receivers. This is true even if the eavesdropper listens to all communication links. However, the expected number of rounds needed for this to happen is linear in the number of mixes in the network. Now we consider the same situation but for layer mixing: we assume that the adversary listens only to a constant fraction of the links. We shall see that we need on average $O(1)$ rounds to achieve unlinkability of these messages. If at some round $i$ the messages are at nodes $a$ and $b$, and in the next round they are at nodes $c$ and $d$ such that the eavesdropper does not listen to any of the four communication links $(a, c)$,$(a, d)$,$(b, c)$ and $(b, d)$, then thereafter the eavesdropper cannot link the senders to the receivers. Indeed, the adversary cannot distinguish if the messages were sent on the edges $(a, c)$,$(b, d)$ or on the edges $(a, d)$, $(b, c)$. The probability that the adversary does not listen to these four communication links at a given moment is constant. Therefore the expected number of rounds to achieve unlinkability is $O(1)$.[2]

We remark that since layer mixing is not done in the nodes, we may model a dishonest node by labeling all edges entering or leaving it insecure. Assume the fraction of insecure communication links is $b_{links}$, and the fraction of insecure nodes is $b_{nodes}$. The fraction of communication links that are labeled insecure because they either enter or leave an insecure node is at most $2b_{nodes}$. Thus, the total fraction of communication links in the network that are labeled insecure is at most $b = b_{links} + 2b_{nodes}$. Thus, from now on we will only consider the fraction of insecure *communication links* in the network.

### 1.3. *Prior Information*

All the protocols mentioned so far deal only with unlinkability when the a priori probability distribution is *uniform*. In reality, however, the a priori distribution is very far from uniform. For example, people tend to communicate more often with people speaking their language. The prior information is often very significant and a protocol that is secure with a priori uniform distribution, might be insecure in general. Our approach is different in that it guarantees unlinkability for whatever a priori distribution. We believe that any reasonable definition for unlinkability should deal with prior knowledge.

### 1.4. *The Onion Protocol*

The protocol that we consider in this paper is almost identical to the routing protocol from Rackoff and Simon [25] based on onion-like encryption. Onion-routing can be considered as an extension of MIX -protocol from Chaum [7].

[2] Note that we rely on the assumption that the adversary is *non-adaptive*. An *adaptive* adversary may, e.g., track a packet $p$ sent at time 0 by $v_0$, by eavesdropping at time $t$ all communication links going out of $v_t$, the node where packet $p$ is at time $t$. If $v_t$ does not receive any other packet at time $t$ (an event that is likely in a system where the number of active nodes $n$ is much smaller than the network size $N$) then listening on the edges leaving $v_t$ determines $v_{t+1}$, and eventually the whole communication path.

We consider a fully connected network, in which every node can send a message directly to any other node of the network. Moreover, every protocol participant is aware of all nodes of the network.[3]

The protocol works as follows: if node $A$ wants to send a message $m$ to node $B$, then $A$ picks $T-1$ intermediate nodes $v_1, \ldots, v_{T-1}$ independently at random from the set of all nodes. Let $E_v$ denote encryption with the public key of node $v$. Node $A$ computes

$$a_0 := E_{v_1}\big(v_2, E_{v_2}\big(\ldots E_{v_{T-2}}\big(v_{T-1}, E_{v_{T-1}}\big(B, E_B(m)\big)\big)\big)\ldots\big),$$

that is $a_0$ is computed recursively:

$$a_i := E_{v_{i+1}}(v_{i+2}, a_{i+1}) \quad \text{for } 0 \le i < T-2 \quad \text{and} \quad a_{T-2} := E_{v_{T-1}}\big(B, E_B(m)\big).$$

The message $a_0$ is sent to the node $v_1$. This node can decrypt the message and retrieve the name of the next server on the path—i.e. $v_2$. Generally, node $v_i$ "peels off" one encryption layer, gets the name of the next node on the path and the ciphertext to be sent there. After $T-1$ steps, message $E_B(m)$ is delivered to the destination node $B$.

In fact, to provide high level of security, an implementation must take into account some details like for example applying an appropriate encryption method (see [22]). The actual implementation that we use is that of Camenisch and Lysyanskaya [11] who designed a provably secure onion routing scheme. Using their work we have a provable reduction from unlinkability to traffic analysis, and we can concern ourselves only with the traffic analysis problem.

We remark that one can also extend the protocol to handle return messages, e.g., by using the reversed paths for the return messages as is done in [6]. Also, the protocol can be adapted to a somewhat less synchronous setting, where all nodes have clocks, all the clocks are within $\Delta$ accuracy from a common time and there is some time bound $\Delta_{bound}$ on transmission latency.

### 1.5. *Summary of Results*

Most previous work considered passive adversaries that control all communication links and most communication nodes. We saw in Sect. 1.2 that protocols for such adversaries are forced to have high message overhead. We weaken the adversary model and consider passive adversaries that do not control some fraction of the communication links. We show the following properties of the onion routing protocol against such an adversary:

Low overhead: we do not require that most nodes send messages at all times. We show that traffic analysis does not provide substantial additional information regardless of the number of non-active players in the system.

Small delay: we get the upper bound $O(\log n)$ on the message delay instead of the former polylogarithmic bounds.

A priori distributions: unlike previous work, our analysis does not assume the receivers of the messages are chosen uniformly at random, and it applies for arbitrary a priori distributions.

[3] This is an important assumption; see [16] for security problems for the case when the assumption is not fulfilled.

Informally, a protocol is $\alpha$-unlinkable in an attack model (which in our case includes passive, non-adaptive adversaries that eavesdrop at most some fixed constant fraction of communication links in the system) if for any eavesdropper respecting that attack model, for any fixed public set of senders and receivers, the amount of information on the actual permutation linking senders to receivers is smaller than $\alpha$. In Sect. 3 we explain how we measure the information gain and we formally define *$\alpha$-unlinkability* against an attack model. We prove:

**Theorem 1.1.** *Assume the protocol of Sect. 1.4 runs in a fully connected network with $N$ nodes, and some constant fraction of the communication links cannot be monitored by the adversary. Let $\alpha(n)$ be an arbitrary function. Fix some $n \leq N$. Suppose the number of nodes on a path from a sender to a receiver is $T = \Omega(\log \frac{n}{\alpha(n)})$. Then for every $n$ honest vertices wishing to send a message, the protocol is $\alpha(n)$-unlinkable.*

The actual theorem we prove is stronger and deals, e.g., also with prior information.

The paper is organized as follows. After the preliminaries in Sect. 2, we give Rackoff and Simon's definition of unlinkability in Sect. 3 and prove an equivalent variant using mutual information. In Sect. 4 we prove that our protocol is unlinkable in the no-prior information case; in Sect. 5 we consider the prior information case. We conclude in Sect. 6 with some open problems.

## 2. Technical Preliminaries

### 2.1. *Information Theory*

A distribution $D$ over a finite set $\Lambda$ is a function $D{:}\Lambda \to [0, 1]$ such that $\sum_{x\in\Lambda} D(x) = 1$. For $S \subseteq \Lambda$, we denote $D(S) \stackrel{\text{def}}{=} \sum_{s\in S} D(s)$. If $A$ is a random variable that takes values from $\Lambda$, then for $x \in \Lambda$ by $A(x)$ we denote the probability that the random variable $A$ takes value $x$. We measure distance between random variables (and their distributions) defined over $\Lambda$ with the $\ell_1$ norm:

$$\|D_1 - D_2\|_1 \stackrel{\text{def}}{=} \sum_{x\in\Lambda} \big|D_1(x) - D_2(x)\big|.$$

The $\ell_1$ distance is twice the *variational distance*, namely

$$\|D_1 - D_2\|_1 = 2 \max_{S\subseteq\Lambda} \big(D_1(S) - D_2(S)\big).$$

Let $A$ and $B$ be random variables. By $A \otimes B$ we denote their product distribution, and by $(A, B)$ we denote their joint distribution. That is,

$$\Pr\big(A \otimes B = (a, b)\big) \stackrel{\text{def}}{=} \Pr(A = a) \cdot \Pr(B = b),$$

$$\Pr\big((A, B) = (a, b)\big) \stackrel{\text{def}}{=} \Pr(A = a \wedge B = b).$$

Let $(A|B = b)$ denote the random variable $A$ conditioned by the event that $B = b$.

Let $D$ be a random variable with values in a finite set $\Lambda$. The entropy of the random variable $D$ is

$$H(D) \stackrel{\text{def}}{=} \sum_{x \in \Lambda} D(x) \cdot \log \frac{1}{D(x)}$$

(where we assume that $0 \cdot \log \frac{1}{0} = 0$).

Conditional entropy is defined as follows:

$$H\big(A|B\big) \stackrel{\text{def}}{=} \mathbb{E}_{b \in B} H\big(A|B = b\big).$$

Let us recall that the entropy function is continuous. Moreover, if $A$ and $A'$ are distributed over $\Lambda$ and $\|A - A'\|_1 \leq \alpha$ for some $\alpha < e^{-1}$, then

$$\big|H(A) - H\big(A'\big)\big| \leq \alpha\big(\log\big(|\Lambda|\big) + \log\big(\alpha^{-1}\big)\big) \tag{1}$$

(see, e.g., [21, Box 11.2]).

The mutual information of random variables $A$ and $B$ is

$$\mathrm{I}(A; B) \stackrel{\text{def}}{=} H(A) + H(B) - H(A, B). \tag{2}$$

Similarly,

$$\mathrm{I}\big(A; B|C\big) \stackrel{\text{def}}{=} H\big(A|C\big) + H\big(B|C\big) - H\big(A, B|C\big).$$

Recall that the mutual information function is always positive. One way to think about the mutual information $\mathrm{I}(A; B)$ is that it measures the amount of information contained in $A$ about $B$. It shows how much knowing the value of $A$ affects our knowledge of $B$. The chain rule states that

$$\mathrm{I}(A; B, C) = \mathrm{I}(A; B) + \mathrm{I}\big(A; C|B\big). \tag{3}$$

In particular, the mutual information is monotone: for every random variables $A, B$ and $C$, $\mathrm{I}(A; B, C) \geq \mathrm{I}(A; B)$.

Another important phenomenon expressed by mathematical properties of mutual information is that knowledge cannot increase without communication. This is captured in the data processing inequality (see [21, Sect. 11.2.4]). It says that for every deterministic or probabilistic function $f$,

$$\mathrm{I}\big(f(A, C); B|C\big) \leq \mathrm{I}\big(A; B|C\big).$$

We stress that $f$ should be a function of $A$ and $C$ alone.

The relative entropy of two random variables $A$ and $B$ distributed over the same domain $\Lambda$, and having the property that for every $x$ for which $\Pr(A = x) > 0$ we also have $\Pr(B = x) > 0$, is defined as follows:

$$D\big(A||B\big) \stackrel{\text{def}}{=} \sum_{x \in \Lambda} \Pr(A = x) \cdot \log \frac{\Pr(A = x)}{\Pr(B = x)},$$

where if the quantity $0 \log 0$ appears in the formula, it is interpreted as 0 (see [12,21]). Relative entropy is not symmetric, i.e., $D(A||B)$ is usually different than $D(B||A)$. Relative entropy is always non-negative and respects the following inequality:

$$D\big(A||B\big) \geq \frac{1}{2 \ln 2} \|A - B\|_1^2. \tag{4}$$

In particular, $D(A||B) = 0$ if and only if $A = B$. Another simple fact is that

$$\mathrm{I}(A; B) = D\big((A, B)||A \otimes B\big). \tag{5}$$

These and other basic facts of information theory appear, e.g., in [12] and [21, Chap. 11].

### 2.2. *Markov Chains*

We use standard notions from the theory of finite Markov chains, as appearing, e.g., in [19]. Let $M$ be a homogeneous Markov chain with a finite state space $\mathbf{S}$ and a unique stationary distribution $\mu$. Abusing notation the transition matrix of $M$ will be also denoted by $M$. Let $Y^0 = Y$ be the initial distribution of the chain and $Y^t = M^t Y^0$ the distribution at time $t$. A standard measure of convergence to the stationary distribution is the *mixing time* defined as

$$\tau_M(\epsilon) \stackrel{\text{def}}{=} \min\left\{T : \forall_{Y^0}, \forall_{t \geq T} \left\| Y^t - \mu \right\|_1 \leq \epsilon\right\}.$$

#### 2.2.1. *Coupling*

Coupling (and path coupling) is a very elegant technique to estimate from above the mixing time of many Markov chains. We define coupling and path coupling below. We refer the interested reader to Guruswami [18] for examples where coupling is used and for a comparison of the coupling proof technique with other proof techniques for proving rapid mixing of Markov chains.

Let $M$ be a homogeneous Markov chain with a finite state space $\mathbf{S}$. Let $D$ be a homogeneous Markov chain with state space $\mathbf{S} \times \mathbf{S}$. Define $(Y', \widetilde{Y}') = D(Y, \widetilde{Y})$, i.e., the transition matrix of $D$ applied on the distribution on states defined by $(Y, \widetilde{Y})$. We say $D$ marginally agrees with $M$ over $\Gamma \subseteq \mathbf{S} \times \mathbf{S}$, if for every $(s_1, s_2) \in \Gamma$:

$$\big(Y'|(Y, \widetilde{Y}) = (s_1, s_2)\big) = \big(MY|Y = s_1\big),$$
$$\big(\widetilde{Y}'|(Y, \widetilde{Y}) = (s_1, s_2)\big) = \big(M\widetilde{Y}|\widetilde{Y} = s_2\big).$$

We say $D$ is a *coupling* [1] for $M$, if $D$ marginally agrees with $M$ over $\mathbf{S} \times \mathbf{S}$. Note that $D$ may introduce arbitrary dependencies between $Y'$ and $\widetilde{Y}'$ as long as marginally $Y'$ and $\widetilde{Y}'$ develop as $M$.

The mixing time of $M$ may be bounded using the following lemma:

**Lemma 2.1** (The Coupling Lemma). *Suppose $D$ is a coupling for a Markov chain $M$, $(Y^t, \widetilde{Y}^t) = D^t(Y^0, \widetilde{Y}^0)$. If for every initial state $(y_0, \tilde{y}_0)$ for $(Y^0, \widetilde{Y}^0)$ and $t \geq T$,*

$$\Pr\big[Y^t \neq \widetilde{Y}^t | \big(Y^0, \widetilde{Y}^0\big) = (y_0, \tilde{y}_0)\big] \leq \epsilon,$$

*then $\tau_M(\epsilon) \leq T$.*

2.2.2. *Path Coupling*

The path coupling construction [3] reduces the task of finding a coupling that works on *all* pairs of states, to that of finding one that needs to work only on states in a subset $\Gamma$. Formally,

**Lemma 2.2** (The Path Coupling Lemma). *Let $\Gamma \subseteq \mathbf{S} \times \mathbf{S}$ be a symmetric relation whose transitive closure is $\mathbf{S} \times \mathbf{S}$. For $(Y, \widetilde{Y}) \in \mathbf{S} \times \mathbf{S}$, let $d(Y, \widetilde{Y})$ be the length of the shortest path from $Y$ to $\widetilde{Y}$ via $\Gamma$, and define*

$$K = \max_{s_1, s_2 \in \mathbf{S}} d(s_1, s_2).$$

*Let $M$ be a homogeneous Markov chain with state space $\mathbf{S}$, and $D$ a homogeneous Markov chain with state space $\mathbf{S} \times \mathbf{S}$, such that $D$ marginally agrees with $M$ over $\Gamma \subseteq \mathbf{S} \times \mathbf{S}$. Assume there exists a constant $\beta < 1$ such that for any $(y_0, \tilde{y}_0) \in \Gamma$:*

$$\mathbb{E}\big[d\big(D(Y, \widetilde{Y})\big) | (Y, \widetilde{Y}) = (y_0, \tilde{y}_0)\big] < \beta.$$

*Then,*

$$\tau_M(\epsilon) \leq \lceil \log\big(K\epsilon^{-1}\big) / \log\big(\beta^{-1}\big) \rceil.$$

2.3. *Graph Theory*

Let $G = (V, E)$ be a graph. We say $(v_1, v_2, v_3, v_4) \in V^4$ is a *crossover*, if $(v_1, v_3)$, $(v_1, v_4)$, $(v_2, v_3)$, $(v_2, v_4) \in E$. We will use the following lemma:

**Fact 2.3** [2, Corollary 2.1]. *Let $G = (V, E)$ be a graph and assume that $|E| \geq f \cdot \binom{|V|}{2}$. If we choose vertices $v_1, v_2, v_3, v_4$ uniformly at random, then* $\Pr[(v_1, v_2, v_3, v_4)$ *is a crossover*$] \geq f^4$.

## 3. Unlinkability

3.1. *Unlinkability Measures*

Let $A$ and $B$ be two possibly correlated random variables.

**Definition 3.1.** We say $A$ and $B$ are $\alpha$-*independent*, if $\|(A, B) - A \otimes B\|_1 \leq \alpha$.

Note that

$$\big\|(A, B) - A \otimes B\big\|_1 = \mathbb{E}_{a \in A}\big\|\big(B|A = a\big) - B\big\|_1 = \mathbb{E}_{b \in B}\big\|\big(A|B = b\big) - A\big\|_1.$$

So two random variables are $\alpha$-independent, if on average knowing one does not affect much the marginal distribution of the other.

**Definition 3.2.** We say $A$ and $B$ are $\alpha$-*unlinkable*, if $\mathrm{I}(A; B) \leq \alpha$.

The following lemma asserts that the two definitions are in a certain sense equivalent. This equivalence turns out to be very useful, since it enables to use information theoretic tools and stochastic tools interchangeably in the proofs.

**Lemma 3.1.** *Let $A$ and $B$ be two random variables defined over $\Lambda$.*

- *If $A$ and $B$ are $\alpha$-unlinkable, then $A$ and $B$ are $\sqrt{2\ln 2}\sqrt{\alpha}$-independent.*
- *If $A$ and $B$ are $\alpha$-independent for $\alpha \leq e^{-1}$, then $A$ and $B$ are $\delta$-unlinkable for $\delta = \alpha(\log|\Lambda| + \log\alpha^{-1})$.*

**Proof.** For the first assertion, by (4) and (5)

$$\big\|(A,B) - A\otimes B\big\|_1 \leq \sqrt{2\ln(2)\cdot D\big((A,B)||A\otimes B\big)} = \sqrt{2\ln(2)\cdot \mathrm{I}(A;B)}.$$

For the second assertion denote $(A',B') = A\otimes B$. We have $\|(A',B')-(A,B)\|_1 \leq \alpha$ so by (1) we get $|H(A',B') - H(A,B)| \leq \delta$. On the other hand, by (2) and $\mathrm{I}(A';B') = 0$, we have

$$\begin{aligned}&\big|H\big(A',B'\big) - H(A,B)\big| \\ &\quad = \big|\big(H\big(A'\big) + H\big(B'\big) - \mathrm{I}\big(A';B'\big)\big) - \big(H(A) + H(B) - \mathrm{I}(A;B)\big)\big| = \mathrm{I}(A;B). \quad \square\end{aligned}$$

### 3.2. *Unlinkable Communication*

Informally, a protocol is unlinkable, if for every set of $n$ senders, $n$ receivers and any passive eavesdropper that listens to at most $1-f$ fraction of links, the random variable that describes the actual permutation $\pi$ between senders and receivers has very little mutual information with the information the eavesdropper knows.

Formally, fix an eavesdropper that at each time step eavesdrops at most $1-f$ fraction of the communication links. The eavesdropper is *non-adaptive*, i.e., it has to choose which communication links are wiretapped before the protocol starts. We let $E$ denote the information the eavesdropper has gathered. Specifically, $E$ is a matrix with rows indexed by time steps $t$, columns indexed by communication links $e$, and values taken from $\{*, 0, 1, \ldots, n\}$, where value $i \in \{0,\ldots,n\}$ means positive knowledge that $i$ messages were sent on that link at time step $t$, and $*$ means lack of such knowledge.

We now run the protocol. We have $N$ nodes and $n$ honest senders. Let $Q^0 = (w_1^0,\ldots,w_n^0)$ be the list of senders. First the senders choose receivers according to the a priori distribution $\Pi^T$, say $Q^T$ is the list of $n$ receivers (a node occurs multiple times on the list, if it receives more than one message). Then, each sender $w_i^0$ chooses a random path $w_i^0, w_i^1, \ldots, w_i^T$ starting with him and ending at the receiver he chose (we assume that the senders do not know the adversary's choices). In this way, for every time $t = 0,\ldots,T$, the $n$ senders determine the following two lists of active nodes (possibly with repetitions on each list):

- $Q^t = (w_1^t,\ldots,w_n^t)$—the list of the $n$ active nodes at time $t$ ordered by the original senders, so that at time $t$ the $i$th message is at $w_i^t$.
- $P^t = (v_1^t,\ldots,v_n^t)$—the list of the $n$ active nodes at time $t$ ordered lexicographically.

Furthermore, there is a permutation $\pi^t$ such that $w_i^t = v_{\pi^t(i)}^t$ linking between a message's location at time $t$ and its original sender. Note that we may assume $Q^0$ is lexicographically ordered, so that $Q^0 = P^0$ and $\pi^0 = id$. Informally, $P^t$ is a stripped-down version of $Q^t$ that knows the active nodes at time $t$, but forgets the correspondence to the original senders, while the permutation $\pi^t$ holds exactly the knowledge needed to link between a sender and the location of his message at time $t$. As we explained before, we do not try to conceal the identities of the senders or receivers, nor the intermediate nodes. We thus assume the lists $P^0, \ldots, P^T$ are public. Let $\overline{P} = (P^1, \ldots, P^{T-1})$.

Thus, so far, we have made public the list of $n$ honest senders $P^0$, the list of receivers $P^T$, the a priori distribution $\Pi^T$ and all the information the adversary (controlling only $1 - f$ fraction of communication links) knows. During protocol execution also the set of intermediate active nodes $\overline{P}$ becomes public. All this data are public. Moreover, for technical reasons which will become clear later, we shall also assume that the adversary has complete knowledge of the communication occurring in all *odd* time steps.

We now define a joint distribution $(\Pi^t, C^t)$ as the distribution obtained by the following sampling process. We pick at random an execution of the protocol. This determines $\pi^0, \ldots, \pi^T$ and the information $E$ the eavesdropper has learned, where $E$ also includes all the public data (i.e., the set of $N$ nodes, $n$ senders, $n$ receivers, intermediate nodes, communication occurring in odd time steps, a priori distribution). For every $t$, we let $\sigma^t$ be a permutation chosen at random from the set of all permutations consistent with $E$ (and therefore also with the public data). We then output $(\pi^t, \sigma^t)$.

**Definition 3.3** ($\alpha(n)$-unlinkability). We say that the protocol run for $T$ steps is $\alpha(n)$-unlinkable and $\beta(n)$-independent, if for every set of $n$ senders, $n$ receivers, prior distribution $\Pi^T$ and any passive eavesdropper that listens to at most $1 - f$ fraction of links, the random variables $\Pi^T$ and $C^T$ are $\alpha(n)$-unlinkable and $\beta(n)$-independent.

We find this definition pretty strong.

Two remarks are in place. First, we mention that previous definitions did not allow prior information. This omission is explicit in the work of Rackoff and Simon, and implicit in the vast body of work on "applied" protocols. It seems clear that the assumption the eavesdropper has no prior information is typically false, e.g., an eavesdropper might know that residents of China tend to correspond more with other Chinese. We believe that any reasonable definition for unlinkability should deal with prior knowledge.

Also, we defined unlinkability as the amount of information that leaks given that the set of senders and receivers is public. However, we do not try to conceal the set of senders and receivers themselves. It is well known that if the protocol is run several times and the a priori distribution is not uniform, then the public data of senders and receivers itself might easily reveal a sender (see, e.g., [13]). Unlinkability means that the eavesdropper does not gain (much) information beyond this.

## 4. Unlinkability Without Prior Information

In this section we consider the situation where each sender chooses message destination uniformly at random. In Sect. 5 we deal with the more general case where messages are

picked according to some known a priori distribution. All products in this section are products in the symmetric group $\mathbb{S}_n$.

**Theorem 4.1.** *Let $\epsilon > 0$ and assume a fraction $f$ of communication links are secure. Suppose the number of nodes on a path from a sender to a receiver is $T = 2\lceil \ln(2n\epsilon^{-1})/\ln\frac{1}{1-f^4}\rceil$. Then the protocol is $\epsilon$-independent and therefore $O(\epsilon(n\log n + \log\epsilon^{-1}))$-unlinkable.*

**Proof.** We define a path coupling process. Let $\mathbf{S}$ be the state space $\mathbb{S}_n \times \mathbb{S}_n$. Let $Y^t$ be distributed according to $(\Pi^{2t}, C^{2t})$ as defined in Sect. 3. Thus, each step of $Y = \{Y^t\}_{t\in\mathbb{N}}$ corresponds to two steps of the protocol. $Y^0 = (id, id)$ corresponds to the initial state where $\pi^0 = id$ and the eavesdropper has complete knowledge on it. To get $Y^{t+1} = (\pi^{2t+2}, \sigma^{2t+2})$ from $Y^t = (\pi^{2t}, \sigma^{2t})$ we pick two random permutations $\kappa, \pi \in \mathbb{S}_n$ corresponding to the odd and even time steps respectively, and let $\pi^{2t+2} = \pi\kappa\pi^{2t}$. We let $\kappa$ be known to the adversary (odd time step communications are public) and proceed to examine the communication links that were used in the even time step, to see which were wiretapped and which were not. We then pick a *random* permutation $\sigma$ that is consistent with the communication on the wiretapped communication links at time $2t+2$ and let $\sigma^{2t+2} = \sigma\kappa\sigma^{2t}$. Thus, we see that $Y$ develops according to a (homogeneous) Markov chain $M$ whose unique stationary distribution is $U_{\mathbb{S}_n} \times U_{\mathbb{S}_n}$.

Building up towards a path coupling argument for $M$, we define the set of *adjacent* states $\Gamma$ to contain all pairs $((\pi,\sigma),(\widetilde{\pi},\widetilde{\sigma})) \in \mathbf{S}\times\mathbf{S}$ such that there exists $i \neq j \in \{1,\ldots,n\}$ for which either $\pi = \widetilde{\pi}$ and $\sigma = \widetilde{\sigma}(i,j)$ or $\pi = \widetilde{\pi}(i,j)$ and $\sigma = \widetilde{\sigma}$.

Note that $\Gamma$ is symmetric and the transitive closure of $\Gamma$ is indeed $\mathbf{S}\times\mathbf{S}$. We let $d((\pi,\sigma),(\widetilde{\pi},\widetilde{\sigma}))$ be the length of the shortest path between $(\pi,\sigma)$ and $(\widetilde{\pi},\widetilde{\sigma})$ via $\Gamma$. Clearly $d((\pi,\sigma),(\widetilde{\pi},\widetilde{\sigma})) \leq 2(n-1)$.

We now define a path coupling $D$ that marginally agrees with $M$ over $\Gamma$. Given $((\pi^0,\sigma^0),(\widetilde{\pi}^0,\widetilde{\sigma}^0)) \in \Gamma$, we define

$$\left(Y^1, \widetilde{Y}^1\right) = D\left(\left(\pi^0,\sigma^0\right),\left(\widetilde{\pi}^0,\widetilde{\sigma}^0\right)\right)$$

as follows. The transition from $Y^0$ to $Y^1$ is performed according to the protocol. That is, when $Y^0$ is in a state $(\pi^0,\sigma^0)$, then $Y^1 = (\pi\kappa\pi^0, \sigma\kappa\sigma^0)$.

We now define $\widetilde{Y}^1$. We know that $((\pi^0,\sigma^0),(\widetilde{\pi}^0,\widetilde{\sigma}^0)) \in \Gamma$, and therefore there exist $i < j \in \{1,\ldots,n\}$ such that either $\pi = \widetilde{\pi}$ and $\sigma = \widetilde{\sigma}(i,j)$ or $\pi = \widetilde{\pi}(i,j)$ and $\sigma = \widetilde{\sigma}$. Consider the locations of the $i$th and $j$th messages at time steps 1 and 2, i.e.,

$$v_1 = P^1_{\kappa\pi^0(i)},$$

$$v_2 = P^1_{\kappa\pi^0(j)},$$

$$v_3 = P^2_{\pi\kappa\pi^0(i)},$$

$$v_4 = P^2_{\pi\kappa\pi^0(j)}.$$

We call $(v_1, v_2, v_3, v_4)$ a *crossover*, if all of the links $(v_1, v_3)$, $(v_2, v_4)$, $(v_1, v_4)$, $(v_2, v_3)$ are secure. We now have two cases (guaranteed by the assumption that $((\pi^0, \sigma^0), (\widetilde{\pi}^0, \widetilde{\sigma}^0)) \in \Gamma$):

Case 1: $\pi^0 = \widetilde{\pi}^0$, $\sigma^0 = \widetilde{\sigma}^0(i, j)$.
We let $\widetilde{Y}^1 = (\pi\kappa\widetilde{\pi}^0, \widetilde{\sigma}\kappa\widetilde{\sigma}^0)$, where $\widetilde{\sigma}$ is defined as follows:

- If $(v_1, v_2, v_3, v_4)$ is not a crossover, then we set $\widetilde{\sigma} = \sigma$,
- otherwise, we set $\widetilde{\sigma} = (v_3 v_4)\sigma$, i.e., $\widetilde{\sigma}$ first acts according to $\sigma$, and then swaps the locations of $v_3$ and $v_4$ in the second step of the protocol.

Case 2: $\pi^0 = \widetilde{\pi}^0(i, j)$, $\sigma^0 = \widetilde{\sigma}^0$.
We let $\widetilde{Y}^1 = (\widetilde{\pi}\kappa\widetilde{\pi}^0, \sigma\kappa\widetilde{\sigma}^0)$, where $\widetilde{\pi}$ is defined as follows:

- If $(v_1, v_2, v_3, v_4)$ is not a crossover, then we choose $\widetilde{\pi} = \pi$,
- otherwise we choose $\widetilde{\pi} = (v_3 v_4)\pi$.

We now claim,

**Claim 1.** *$D$ marginally agrees with $M$ over $\Gamma$.*

**Proof.** We only need to show that $\widetilde{Y}^1$ is a faithful copy of $M$ since this is trivial for $Y^1$. Assume $(Y^0, \widetilde{Y}^0) \in \Gamma$. Let $\widetilde{Y}^1 = (\widetilde{\pi}\kappa\widetilde{\pi}^0, \widetilde{\sigma}\kappa\widetilde{\sigma}^0)$, and suppose $Y^1 = (\pi\kappa\pi^0, \sigma\kappa\sigma^0)$ for some $\kappa, \pi, \sigma$ selected according to $M$. In Case 1, $\widetilde{\pi} = \pi$ and it is easy to see $\widetilde{\sigma}$ is selected uniformly at random among all permutations consistent with $\pi$ and the wire-tapped links. Case 2 is similar. □

**Claim 2.** *For $(Y^0, \widetilde{Y}^0) \in \Gamma$, $\mathbb{E}[d(Y^1, \widetilde{Y}^1)] \leq 1 - f^4$.*

**Proof.** In both cases, if $(v_1, v_2, v_3, v_4)$ is not a crossover at time step 2 of the protocol, then $d(Y^1, \widetilde{Y}^1)$ remains 1. Otherwise step 2 yields $d(Y^1, \widetilde{Y}^1) = 0$. The adversary is fixed before the active nodes at steps 1 and 2 of the protocol are chosen and the odd steps ensure these nodes are chosen independently and uniformly at random. Therefore, by Fact 2.3,

$$\mathbb{E}\big[d\big(Y^1, \widetilde{Y}^1\big)\big] = \Pr\big[(v_1, v_2, v_3, v_4) \text{ is not a crossover}\big] \leq 1 - f^4. \qquad \square$$

Finally, Using Lemma 2.2 with $\beta = 1 - f^4$ and $K = 2n$ we obtain

$$\tau_M(\epsilon) \leq T/2.$$

This shows $\|(\Pi^T, C^T) - U_{\mathbb{S}_n} \times U_{\mathbb{S}_n}\|_1 \leq \epsilon$. Since we are in the no-prior information case, $\Pi^T \times C^T = U_{\mathbb{S}_n} \times U_{\mathbb{S}_n}$ and therefore the protocol is $\epsilon$-independent. The bound on the mutual information follows from Lemma 3.1. □

## 5. Unlinkability with Prior Information

We now deal with the general case where the a priori distribution $\Pi$ is not necessarily uniform. Technically, we show that our protocol is unlinkable by concentrating on the

*middle* layer. This is intuitively natural, because the eavesdropper knows the initial permutation $\Pi^0$ at the beginning, and has partial information about the final permutation $\Pi^T$ given by the prior, but the permutation at the middle layer $\Pi^{T/2}$ is masked by the random choices made throughout the protocol.

**Lemma 5.1.** *Let $\Pi^T$ be an arbitrary distribution and $T = 4\lceil \ln(2n\epsilon^{-1})/\ln \frac{1}{1-f^8} \rceil$. Then $C^{T/2}$ and $\Pi^{T/2}$ are $\epsilon$-independent and therefore $O(n \log n \cdot \epsilon)$-unlinkable.*

**Proof.** We say a node $v_i^t \in P^t$ is associated with a node $w_j^{T-t} \in P^{T-t}$, if the message that $v_i^t$ forwards eventually arrives at $w_j^{T-t}$. We also say the communication link $(w, v)$ is associated with the communication link $(v', w')$ if $w$ is associated with $w'$, and $v$ is associated with $v'$.

For the proof, we give the eavesdropper the extra knowledge about which node at level $t$ is associated with which node at level $T - t$, for every $0 \le t \le \frac{T}{2}$. Let $\widehat{E}$ be all the information known to the eavesdropper including the additional information we reveal to the eavesdropper. Let $\widehat{C}^t$ be as in Sect. 3, defined with respect to $\widehat{E}$. Let us look at the first $T/2$ steps in the protocol. From the eavesdropper's point of view, $n$ honest nodes started a no-prior information protocol (it is no-prior information because $\Pi^{T/2}$ is uniform) and each communication link $(v^t, v^{t+1})$ is secure, if both the link $(v^t, v^{t+1})$ and its associated link are secure. Clearly, when a link is secure, the eavesdropper (even with the additional information we give him) does not know if there was communication on that link or not.

Furthermore, let $a, b, c, d$ be nodes and $a', b', c', d'$ their associated nodes. $(a, b, c, d)$ is a crossover if and only if both $(a, b, c, d)$ and $(a', b', c', d')$ were crossovers before. Each event happens with independent probability at least $f^4$. Altogether, the probability of a crossover is at least $f^8$. Thus, from the eavesdropper point of view there are $n$ honest nodes that run the protocol for $T/2$ steps, and the probability of a crossover is at least $f^8$. We are now in back to the no-prior information case! We therefore can proceed as in the proof of Theorem 4.1 and conclude that $\widehat{C}^{T/2}$ and $\Pi^{T/2}$ are $O(\epsilon)$ independent and $\mathrm{I}(C^{T/2}; \Pi^{T/2}) \le O(n \log n \cdot \epsilon)$. In particular, $\mathrm{I}(C^{T/2}; \Pi^{T/2}) \le \mathrm{I}(\widehat{C}^{T/2}; \Pi^{T/2}) \le O(n \log n \cdot \epsilon)$. □

To complete the proof we show that since the eavesdropper gains very little information about the middle layer, it must be the case that the eavesdropper does not gain much information about the last layer. We claim:

**Lemma 5.2.** $\mathrm{I}(C^T; \Pi^T) \le \mathrm{I}(C^{T/2}; \Pi^{T/2})$.

**Proof.** Let $E_1$ denote the random variable that contains the communication seen by the eavesdropper throughout the first $T/2$ steps. Similarly, $E_2$ is the random variable that contains the communication seen by the eavesdropper throughout the last $T/2$ steps. We define a probabilistic function $f(\sigma, e_2)$ that given $\sigma \in \mathbb{S}_n$ and $e_2$ chooses a permutation $\pi$ according to the distribution $(\Pi^T | \Pi^{T/2} = \sigma \wedge E_2 = e_2)$.

Note that $f(\Pi^{T/2}, E_2) = \Pi^T$ because we may think of it as first picking $\pi^{T/2}, e_1, e_2$ according to the correlated distributions $(\Pi^{T/2}, E_1, E_2)$, and then picking $\pi^T$ accord-

ing to the distribution $(\Pi^T|\Pi^{T/2} = \pi^{T/2}, E_1 = e_1, E_2 = e_2) = (\Pi^T|\Pi^{T/2} = \pi^{T/2}, E_2 = e_2)$ which is what $f(\pi^{T/2}, e_2)$ does.

Now, by the chain rule (3): $\mathrm{I}(\Pi^T; E_1, E_2) = \mathrm{I}(\Pi^T; E_2) + \mathrm{I}(\Pi^T; E_1|E_2)$. Also, $\mathrm{I}(\Pi^T; E_2) = 0$. This is so because one way to view the protocol is that the $n$ nodes first pick $\pi \in \Pi^T$, then *independently* pick random paths for the top $T-1$ levels (thus determining $E_2$) and then complete the first layer to implement $\pi$. Thus, $E_2$ is independent of $\Pi^T$.[4] Thus, using the data-processing inequality we get

$$\begin{aligned}\mathrm{I}\big(\Pi^T; E_1, E_2\big) &= \mathrm{I}\big(\Pi^T; E_1|E_2\big) = \mathrm{I}\big(f\big(\Pi^{T/2}, E_2\big); E_1|E_2\big) \\ &\le \mathrm{I}\big(\Pi^{T/2}; E_1|E_2\big) \le \mathrm{I}\big(\Pi^{T/2}; E_1, E_2\big).\end{aligned}$$

□

We are now ready to prove

**Theorem 5.3.** *Assume the protocol of Sect. 1.4 runs in a fully connected network with $N$ nodes, and some constant fraction of the communication links cannot be monitored by the adversary. Let $\alpha(n)$ be an arbitrary function. Then for every $n < N$, and every prior information on the communication, the protocol is $\alpha(n)$-unlinkable when $T = \Omega(\log \frac{n}{\alpha(n)})$, where $T$ stands for the number of nodes on the path from the sender to the receiver.*

**Proof.** Combining Lemmas 5.1 and 5.2 we see that the protocol is $(n \log n \cdot \epsilon)$-unlinkable after $T = c \log \frac{n}{\epsilon}$ steps, for some constant $c$. Taking $\epsilon = \frac{\alpha(n)}{n \log n}$ we see that the protocol is $\alpha(n)$-unlinkable after $O(\log \frac{n}{\epsilon(n)}) = O(\log \frac{n^2 \log n}{\alpha(n)}) = O(\log \frac{n}{\alpha(n)})$ steps. □

We believe the proof clearly demonstrates the advantages one gets when quantifying unlinkability using information theoretic tools.

## 6. Extensions and Open Problems

We now briefly discuss *active* adversaries. Chaum [6] suggests to check the behavior of possibly dishonest nodes, and Rackoff and Simon make that concrete by using secure computation and zero knowledge. It would be nice to have a variant of our protocol (even using secure computation and zero knowledge) that is secure against active adversaries and yet has low message overhead.

In our protocol (and many other protocols) we assume the underlying graph is complete. However, in reality, the actual underlying graph is sparse. Simulating the complete graph with the actual underlying sparse graph is not good, because for some graphs, the eavesdropper may gain control of most of the communication links in the complete graph by taking over a few communication links in the underlying graph. It is an interesting problem to find a provably secure protocol with low message overhead when the underlying graph has short mixing time. Gogolewski et al. [17] go in this direction using *node mixing*.

[4] Note that the above argument does not work for $E_2$ and $\Pi^{T-1}$, and indeed $E_2$ and $\Pi^{T-1}$ can be dependent.

## References


[1] D.J. Aldous, Random walks on finite groups and rapidly mixing Markov chains, in *Séminaire de Probabilités de Strasbourg*, vol. 17, (1983), pp. 243–297

[2] N. Alon, Testing subgraphs in large graphs, in *FOCS*, (2001), pp. 434–439

[3] R. Bubley, M. Dyer, Path coupling: a technique for proving rapid mixing in Markov chains, in *FOCS*, (1997), pp. 223–231

[4] A. Beimel, S. Dolev, Buses for anonymous message delivery. *J. Cryptol.* **16**(1), 25–39 (2003)

[5] R. Berman, A. Fiat, A. Ta-Shma, Provable unlinkability against traffic analysis, in *Financial Cryptography (FC)*. LNCS, vol. 3110, (2004), pp. 266–280

[6] D. Chaum, Untraceable electronic mail, return addresses, and digital pseudonyms. Thesis (M.S. in Computer Science), University of California, Berkeley (1979)

[7] D. Chaum, Untraceable electronic mail, return addresses, and digital pseudonyms. *Commun. ACM* **24**(2), 84–88 (1981)

[8] D. Chaum, The Dining Cryptographers Problem: unconditional sender and recipient untraceability. *J. Cryptol.* **1**(1), 65–75 (1988)

[9] A. Czumaj, M. Kutyłowski, Delayed path coupling and generating random permutations. *Random Struct. Algorithms* **17**(3–4), 238–259 (2000)

[10] A. Czumaj, P. Kanarek, M. Kutyłowski, K. Loryś, Delayed path coupling and generating random permutations via distributed stochastic processes, in *SODA*, (1999), pp. 271–280

[11] J. Camenisch, A. Lysyanskaya, A formal treatment of onion routing, in *CRYPTO*. LNCS, vol. 3621, (2005), pp. 169–187

[12] T.M. Cover, J.A. Thomas, *Elements of Information Theory* (Wiley, New York, 1991)

[13] G. Danezis, Statistical disclosure attacks: traffic confirmation in open environments, in *Security and Privacy (SEC)*, (2003), pp. 421–426

[14] G. Danezis, C. Diaz, A survey of anonymous communication channels. Microsoft Technical report MSR-TR-2008-35. Available at: http://research.microsoft.com/apps/pubs/default.aspx?id=70553

[15] M. Gomułkiewicz, M. Klonowski, M. Kutyłowski, Provable unlinkability against traffic analysis already after $O(\log(n))$ steps! in *International Workshop on Information Security*. LNCS, vol. 3225, (2004), pp. 354–366

[16] M. Gogolewski, M. Klonowski, M. Kutyłowski, Local view attack on anonymous communication, in *European Symposium on Research in Computer Security (ESORICS)*. LNCS, vol. 3679, (2005), pp. 475–488

[17] M. Gogolewski, M. Kutyłowski, T. Łuczak, Mobile mixing, in *Information Security and Cryptology (ICISC)*. LNCS, vol. 3506, (2004), pp. 380–393

[18] V. Guruswami, Rapidly mixing Markov chains: a comparison of techniques (2000)

[19] O. Häggström, *Finite Markov Chains and Algorithmic Applications*, vol. 52 (Cambridge University Press, Cambridge, 2002)

[20] D. Malkhi, E. Pavlov, Anonymity without 'cryptography' (extended abstract), in *Financial Cryptography (FC)*. LNCS, vol. 2339, (2001), pp. 117–135

[21] M. Nielsen, I. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, 2000)

[22] B. Pfitzmann, A. Pfitzmann, How to break the direct RSA-implementation of MIXes, in *Eurocrypt*. LNCS, vol. 434 (Springer, Berlin, 1989), pp. 373–381

[23] J. Raymond, Traffic analysis: protocols, attacks, design issues, and open problems, in *Designing Privacy Enhancing Technologies: Workshop on Design Issues in Anonymity and Unobservability*, ed. by H. Federrath. LNCS, vol. 2009, (2001), pp. 10–29

[24] M.K. Reiter, A.D. Rubin, Crowds: anonymity for Web transactions. *ACM Trans. Inf. Syst. Secur.* **1**(1), 66–92 (1998)

[25] C. Rackoff, D.R. Simon, Cryptographic defense against traffic analysis, in *STOC*, (1993), pp. 672–681