# Measuring Anonymity with Relative Entropy


Yuxin Deng[1], Jun Pang[2], Peng Wu[3]

[1] The University of New South Wales
School of Computer Science and Engineering, 2052 Sydney, Australia
yuxind@cse.unsw.edu.au

[2] Carl von Ossietzky Universität Oldenburg
Department für Informatik, 26111 Oldenburg, Germany
jun.pang@informatik.uni-oldenburg.de

[3] INRIA Futurs and LIX, École Polytechnique
Rue de Saclay, 91128 Palaiseau, France
wu@lix.polytechnique.fr



**Abstract.** Anonymity is the property of maintaining secret the identity of users performing a certain action. Anonymity protocols often use random mechanisms which can be described probabilistically. In this paper, we propose a probabilistic process calculus to describe protocols for ensuring anonymity, and we use the notion of relative entropy from information theory to measure the degree of anonymity these protocols can guarantee. Furthermore, we prove that the operators in the probabilistic process calculus are non-expansive, with respect to this measuring method. We illustrate our approach by using the example of the Dining Cryptographers Problem.


## 1 Introduction

With the growth and commercialisation of the Internet, users become more and more concerned about their anonymity and privacy in the digital world. Anonymity is the property of keeping secret the identity of the user who has performed a certain action. The need for anonymity may arise in a wide range of situations, from votings and donations to postings on electronic forums. Anonymity protocols often use random mechanisms. Typical examples are the Dining Cryptographers [4], Crowds [21], Onion Routing [26], SG-MIX [15], and many others.

Quantifying the degree of anonymity a protocol can guarantee is a line of active research. Various notions, like anonymity set and information theoretic metric, have been investigated in the literature [4, 21, 2, 24, 8, 5]. (See detailed discussions in Section 5.) In particular, [24, 8] used the notion of *entropy* from information theory as a measure for anonymity. It takes into account the probability distribution of the users performing certain actions, where the probabilities are assigned by an attacker after observing the system. However, it does not take into account the attacker's knowledge about the users before running a protocol.

In this paper we propose to use *relative entropy* as a general extension of the aforementioned approaches. Our method quantifies the amount of probabilistic information revealed by the protocol, i.e. how much information an attacker can obtain after observing the outcomes of the protocol, together with the information he has before the protocol running. For a protocol that contains both non-deterministic and probabilistic behaviours, we extend this measuring method to deal with two sets of probability distributions by using Hausdorff distance (see Definition 4).

Nowadays, the need for applying formal methods to security protocols has been widely recognised. To our knowledge, there have been several attempts to develop a formal framework for specifying and reasoning about anonymity properties. Schneider and Sidiropoulos [22] studied anonymity in CSP [13], but they only considered non-deterministic behaviour. Bhargava and Palamidessi [3] proposed a notion of probabilistic anonymity with careful distinction between non-deterministic and probabilistic behaviours and used a probabilistic $\pi$-calculus as a specification language. This work was extended by Deng, Palamidessi and Pang in [5], where a weak notion of probabilistic anonymity was defined to capture the amount of probabilistic information that may be revealed by a protocol. Other researchers define their notions of anonymity in terms of epistemic logic [12, 11] and "function views" [14].

In this paper, we follow the approach based on process calculi. Specifically, we propose a probabilistic extension of CCS [18], in the style of [1], to describe protocols for ensuring anonymity. It allows us to specify both non-deterministic and probabilistic behaviours. The operational semantics of a process is defined in terms of a probabilistic automaton [23]. Our formal characterisation of anonymity is then based on permutations over the traces of a probabilistic automaton. Inspired by [6, 7], we prove that except for the parallel composition all operators in our probabilistic CCS are non-expansive, with respect to the measuring method using relative entropy, which allows us to estimate the degree of anonymity of a complex system from its components, rather than analyse the system as a whole. We illustrate our ideas by using the example of the Dining Cryptographers Problem (DCP), in which a number of cryptographers cooperate to ensure that the occurrence of a certain action is visible, while the user who has performed it remains anonymous.

We summarise our main contributions of this work as follows:

- We propose to use relative entropy for measuring the degree of anonymity a protocol can guarantee. It is an extension of the results in [24, 8].
- We define a probabilistic CCS for specifying protocols, and prove the non-expansiveness of some operators.
- We show how to use our framework to reason about the degree of anonymity of protocols by the example of the Dining Cryptographers Problem.

*Plan of the paper* In next section we recall some basic notations which are used throughout the paper. In Section 3, we use relative entropy to measure anonymity, and we present the non-expansiveness proof for the operators in a probabilistic CCS. In Section 4, we apply our framework to the Dining Cryptographers Problem. In Section 5, we compare our approach with some related work. Finally, we conclude the paper by discussing our future work in Section 6.

*Acknowledgments* We are very grateful to Kostas Chatzikokolakis, Catuscia Palamidessi, and Prakash Panangaden. From discussions with them we learnt more about relative entropy, which helped us to improve the paper a lot.

## 2 Preliminaries

In this section, we present some basic definitions from probability theory, the notion of probabilistic automata, and a probabilistic CCS.

### 2.1 Probability measure

Let $\Omega$ be a set. A *$\sigma$-field* over $\Omega$ is a collection $\mathcal{F}$ of subsets of $\Omega$ containing $\emptyset$ and closed under complements and countable unions. A *probability measure* on a $\sigma$-field $\mathcal{F}$ is a function $\eta : \mathcal{F} \to [0,1]$ such that $\eta(\Omega) = 1$ and, for each family $\{Z_i\}_{i\in\mathbb{N}}$ of pairwise disjoint elements of $\mathcal{F}$, $\eta(\bigcup_{i\in\mathbb{N}} Z_i) = \sum_{i\in\mathbb{N}} \eta(Z_i)$. A *discrete probability measure* over $\Omega$ is a probability measure whose $\sigma$-field is the powerset of $\Omega$. A *discrete probability distribution* is a function $\eta : \Omega \to [0,1]$ such that $\sum_{s\in\Omega} \eta(s) = 1$. The *support* of $\eta$ is defined to be the set $supp(\eta) = \{s \in \Omega \mid \eta(s) \neq 0\}$. We denote by $\mathcal{D}(\Omega)$ the set of probability distributions over $\Omega$.

### 2.2 Probabilistic automata

We give a brief review of the formalism *probabilistic automata* [23].

**Definition 1.** *A* probabilistic automaton *is a tuple $M = (S, s_0, E, H, \to)$ where*

- *$S$ is a set of* states*,*
- *$s_0$ is the* start state*,*
- *$E$ is a set of* external *actions,*
- *$H$ is a set of* internal (hidden) *actions,*
- *$\to \subseteq S \times (E \cup H) \times \mathcal{D}(S)$ is a* transition relation.

We often write $s \xrightarrow{a} \eta$ for $(s, a, \eta) \in \to$. Informally, a probabilistic automaton is like an ordinary automaton except that a labelled transition leads to a probability distribution over a set of states instead of a single state. We will use probabilistic automata to give operational semantics for the probabilistic CCS that will be introduced in next section.

A *path* $\pi$ of $M$ is a (finite or infinite) sequence of the form $s_0 a_1 \eta_1 s_1 a_2 \eta_2 s_2 ...$ such that

1. each $s_i$ (resp. $a_i, \eta_i$) denotes a state (resp. action, distribution over states);
2. $s_0$ is the initial state;
3. if $\pi$ is finite, then it ends with a state;
4. $s_i \xrightarrow{a_{i+1}} \eta_{i+1}$ and $s_{i+1} \in supp(\eta_{i+1})$, for each non-final $i$.

The set of all paths of $M$ is denoted $Path(M)$, while the set of finite paths is denoted $Path^*(M)$. The last state of a finite path $\pi$ is written $last(\pi)$. A path $\pi$ is *maximal* if either it is infinite or it is a finite path without any outgoing transitions from $last(\pi)$.

A *scheduler* $\sigma$ of $M$ is a partial function of type $Path^*(M) \to (E \cup H) \times \mathcal{D}(S)$ such that (i) for each path $\pi$ that is not maximal $\sigma(\pi)$ is defined, (ii) $\sigma(\pi) = (a, \eta)$ implies $last(\pi) \xrightarrow{a} \eta$. A scheduler $\sigma$ of $M$ induces a discrete probability measure on the $\sigma$-field generated by cones of paths as follows. If $\pi$ is a finite path, then the *cone* generated by $\pi$ is the set of paths $C_\pi = \{\pi' \in Path(M) \mid \pi \preceq \pi'\}$, where $\preceq$ denotes the prefix ordering on sequences. The measure $\epsilon$ of a cone $C_\pi$ is defined by

$$\epsilon(C_\pi) = \begin{cases} 1 & \text{if } \pi = s_0 \\ \epsilon(C_{\pi'}) \cdot \eta(s') & \text{if } \pi = \pi' a \eta s' \text{ and } \sigma(\pi') = (a, \eta) \\ 0 & \text{otherwise.} \end{cases}$$

The measure $\epsilon$ is called a *probabilistic execution* of $M$.

The *trace* of a path $\pi$ of an automaton $M$, written $tr(\pi)$, is the sequence obtained by restricting $\pi$ to the set of external actions of $M$. A trace is *maximal* if it is so obtained from a maximal path. The cone of a finite trace $\gamma$ is defined by $C_\gamma = \{\gamma' \in E^\omega \mid \gamma \preceq \gamma'\}$. Given a probabilistic execution $\epsilon$, the *trace distribution* of $\epsilon$, $td(\epsilon)$, is the measure on the $\sigma$-field generated by cones of traces defined by

$$td(\epsilon)(C_\gamma) = \sum_{tr(\pi)=\gamma} \epsilon(C_\pi).$$

If there are only countably many maximal traces in a probabilistic automaton (which is the case in many applications including all examples in this paper), a trace distribution corresponds to a discrete probability distribution on the maximal traces of a fully probabilistic automaton resulted from resolving all non-determinism of the probabilistic automaton. We denote the set of trace distributions of probabilistic executions of a probabilistic automaton $M$ by $tds(M)$.

### 2.3 Probabilistic CCS

In this section we give a probabilistic extension of Milner's CCS [18] which is based on the calculus of [1] that allows for non-deterministic and probabilistic choice. We assume a countable set of variables, $Var = \{X, Y, ...\}$, and a countable set of atomic actions, $\mathcal{A} = \{a, b, ...\}$. Given a special action $\tau$ not in $\mathcal{A}$, we let $u, v, ...$ range over the set of *actions*, $Act = \mathcal{A} \cup \overline{\mathcal{A}} \cup \{\tau\}$. The class of expressions $\mathcal{E}$ is defined by the following syntax:

$$E ::= \mathbf{0} \;\Big|\; \sum_{i \in I} u_{p_i}.E_i \;\Big|\; E1 \boxplus E_2 \;\Big|\; E_1 \mid E_2 \;\Big|\; E \setminus A \;\Big|\; f[E] \;\Big|\; X \;\Big|\; \mu_X E$$

where $A \subseteq \mathcal{A}$, $f : Act \to Act$ is a renaming function, $I$ is a nonempty countable indexing set and $\{p_i\}_{i \in I}$ a family of probabilities such that $\sum_{i \in I} p_i = 1$. For finite indexing set $I = \{i_1, ..., i_n\}$ we also write $u_{p_{i_1}}.E_{p_{i_1}} + ... + u_{p_{i_n}}.E_{p_{i_n}}$ instead of $\sum_{i \in I} u_{p_i}.E_i$. The construction $E_1 \boxplus E_2$ stands for *non-deterministic choice*, which is denoted by $+$ in CCS. We use $|$ to denote the usual *parallel composition*. The *restriction* and *renaming* operators are as in CCS: $E \setminus A$ behaves like $E$ as long as $E$ does not perform an action $a \in A$; $f[E]$ behaves like $E$ where each action $a \in Act$ is replaced by $f(a)$. We let variables range over process expressions. The notation $\mu_X$ stands for a recursion which binds the variable $X$. We use $fv(E)$ for the set of free variables (i.e., not bound by any $\mu_X$) in $E$. As usual we identify expressions which differ only by a change of bound variables.

We use $P, Q, ...$ to range over $\mathcal{P}r$, the set of expressions without free variables, called *processes*. The operational semantics of a process $P$ is defined as a probabilistic automaton whose states are the processes reachable from $P$ and the transition relation is defined by the rules in Figure 1, where $P \xrightarrow{u} \eta$ describes a transition that, by performing an action $u$, leaves from $P$ and leads to a distribution $\eta$ over $\mathcal{P}r$.

The presence of both probabilistic and non-deterministic choice in the probabilistic CCS allows us to specify systems that have both probabilistic and non-deterministic behaviour. Given a process $P$, we denote by $pa(P)$ the probabilistic automaton that represents the operational semantics of $P$ via the rules in Figure 1. If there is no occurrence of non-deterministic choice in $P$, the automaton $pa(P)$ is fully probabilistic. In this case $tds(pa(P))$ is a singleton set of trace distribution.

1. $\sum_{i \in I} u_{p_i}.P_i \xrightarrow{u} \eta$ where $\eta(P) = \sum\{p_i \mid i \in I, P_i = P\}$
2. $P_1 \boxplus P_2 \xrightarrow{u} \eta$ if $P_1 \xrightarrow{u} \eta$ or $P_2 \xrightarrow{u} \eta$
3. $P_1 \mid P_2 \xrightarrow{u} \eta$ if one the following four conditions is satisfied:
   (a) $P_1 \xrightarrow{u} \eta_1$ and $\eta(P) = \begin{cases} \eta_1(P_1') & \text{if } P = P_1' \mid P_2 \\ 0 & \text{otherwise} \end{cases}$
   (b) $P_2 \xrightarrow{u} \eta_2$ and $\eta(P) = \begin{cases} \eta_2(P_2') & \text{if } P = P_1 \mid P_2' \\ 0 & \text{otherwise} \end{cases}$
   (c) $u = \tau$ and there exists $a \in \mathcal{A}$ with $P_1 \xrightarrow{a} \eta_1$ and $P_2 \xrightarrow{\bar{a}} \eta_2$ such that
   $$\eta(P) = \begin{cases} \eta_1(P_1') \cdot \eta_2(P_2') & \text{if } P = P_1' \mid P_2' \\ 0 & \text{otherwise} \end{cases}$$
   (d) the symmetric case of (c)
4. $P \setminus A \xrightarrow{u} \eta$ if $u \notin A \cup \overline{A}$, $P \xrightarrow{u} \eta_1$, and $\eta(P) = \begin{cases} \eta_1(P') & \text{if } P = P' \setminus A \\ 0 & \text{otherwise} \end{cases}$
5. $f[P] \xrightarrow{u} \eta$ if $P \xrightarrow{v} \eta_1$, $f(v) = u$ and $\eta(P) = \begin{cases} \eta_1(P') & \text{if } P = f[P'] \\ 0 & \text{otherwise} \end{cases}$
6. $\mu_X E \xrightarrow{u} \eta$ if $E\{\mu_X E/X\} \xrightarrow{u} \eta$

**Fig. 1.** Operational semantics for Probabilistic CCS

## 3 Measuring anonymity

### 3.1 Relative entropy

We make a convention $0 \log \infty = 0$.

**Definition 2 (Relative entropy [16]).** *Let $\theta, \theta'$ be two discrete probability distributions on a set $S$. The* relative entropy *of $\theta$ w.r.t. $\theta'$ is defined by*

$$D(\theta, \theta') = \sum_{s \in S} \theta(s) \cdot \log \frac{\theta(s)}{\theta'(s)}.$$

In the sequel, whenever we write $D(\theta, \theta')$, it is implicitly assumed that the domains of $\theta$ and $\theta'$ are the same, i.e., $dom(\theta) = dom(\theta')$.

In general, we have $D(\theta, \theta') \neq D(\theta', \theta)$, so relative entropy is not a true metric. But it satisfies many important mathematical properties, e.g. it is always nonnegative, and equals zero only if $\theta = \theta'$. It plays an important role in quantum information theory, as well as statistical mechanics.

We now present a few properties of relative entropy.

**Proposition 3.** *Relative entropy $D$ has the following properties:*

1. (Nonnegativity) *$D(\eta, \eta') \geq 0$, with $D(\eta, \eta') = 0$ if and only if $\eta = \eta'$;*
2. (Possibility of extension) *$D(\eta_1, \eta_2) = D(\eta_1', \eta_2')$ where $dom(\eta_1') = dom(\eta) \cup \{s\}$ and $\eta_1'(s) = 0$, similarly for $\eta_2'$ w.r.t. $\eta_2$;*
3. (Additivity) *$D(\eta_1 \times \eta_2, \eta_1' \times \eta_2') = D(\eta_1, \eta_1') + D(\eta_2, \eta_2')$, where $(\eta_1 \times \eta_2)(s_1, s_2)$ is defined as $\eta_1(s_1) \cdot \eta_2(s_2)$;*
4. (Joint convexity) *For $0 \leq r \leq 1$, we have*
$$D(r\eta_1 + (1-r)\eta_2, r\eta_1' + (1-r)\eta_2') \leq rD(\eta_1, \eta_1') + (1-r)D(\eta_2, \eta_2').$$
5. (Strong additivity) *Let $dom(\eta_1) = dom(\eta_1') = S \cup \{s\}$, $dom(\eta_2) = dom(\eta_2') = S \cup \{s_1, s_2\}$ with $\eta_1(s) = \eta_2(s_1) + \eta_2(s_2)$ and $\eta_1'(s) = \eta_2'(s_1) + \eta_2'(s_2)$. Then it holds that $D(\eta_1, \eta_1') \leq D(\eta_2, \eta_2')$.*

*Proof.* Similar properties for Tsallis relative entropy have been proved in [10]; their proofs can be adapted for relative entropy. $\square$

We extend $D$ to sets of distributions by using Hausdorff distance.

**Definition 4.** *Given two sets of discrete probability distributions $\Theta = \{\theta_i\}_{i \in I}$ and $\Theta' = \{\rho_j\}_{j \in J}$, the relative entropy of $\Theta$ w.r.t. $\Theta'$ is defined by*

$$D(\Theta, \Theta') = \sup_{i \in I} \inf_{j \in J} D(\theta_i, \rho_j)$$

*where* $\inf \emptyset = \infty$ *and* $\sup \emptyset = 0$.

### 3.2 Anonymity systems

The concept of anonymity is relative to a certain set of anonymous actions, which we denote by $A$. Note that the actions in $A$ normally depend on the identity of users, and thus are not visible to the observer. However, for the purpose of defining and verifying anonymity we model the elements of $A$ as visible outcomes of the system. We write $F_A$ for the set of all renaming functions that are permutations on $A$ and identity elsewhere.

The idea of measuring anonymity is to consider a fully probabilistic automaton (resp. a probabilistic automaton) $M$ as a trace distribution (resp. a set of trace distributions) $tds(M)$, and then apply the distance defined in Definition 2 (resp. Definition 4). The interpretation of a probabilistic automaton as a set of trace distributions is given in Section 2.2. Usually we find it convenient to describe a system as a process in the probabilistic CCS. To measure the distance between two processes, we just view a process $P$ as its corresponding automaton $pa(P)$ and simply write $D(P, Q)$ for the distance between the set of trace distributions represented by $tds(pa(P))$ and that represented by $tds(pa(Q))$.

**Definition 5 ($\alpha$-anonymity[4]).** *Given $\alpha \in [0, 1]$, a process $P$ is* $\alpha$-anonymous *on a set of actions $A$ if*

$$\forall f \in F_A : D(P,\ f[P]) \leq \alpha$$

*In the particular case $\alpha = 0$, we say $P$ is* strongly anonymous *or $P$ provides* strong anonymity.

In [22] Schneider and Sidiropoulos consider a process as a set of traces, thus $P$ is strongly anonymous if $f[P]$, the process after the permutation of anonymous actions, represents the same set of traces as that of $P$. The non-determinism plays a crucial role in their formalism. A system is anonymous if the set of the possible outcomes is saturated with respect to the intended anonymous users, i.e. if one such user can cause a certain observable trace in one possible computation, then there must be alternative computations in which each other anonymous user can give rise to the same observable trace (modulo the identity of the anonymous users). In our case, $P$ is strongly anonymous if $P$ and $f[P]$ represent the same set of trace distributions. Thus, we extend their definition to the probabilistic setting in a natural way. We define $D_A(P)$ as $max\{D(P, f[P]) \mid f \in F_A\}$. Thus, $P$ is $\alpha$-anonymous if and only if $D_A(P) \leq \alpha$.

**Proposition 6 (Non-expansiveness).** *All the operators of the probabilistic CCS except for parallel composition are non-expansive.*

1. $D_A(\sum_{i\in I} u_{p_i}.P_i) \leq \sum_{i\in I} p_i D_A(P_i)$ *if* $u \notin A$;
2. $D_A(P_1 \boxplus P_2) \leq max\{D_A(P_1), D_A(P_2)\}$;

[4] The notion of $\alpha$-anonymity already appeared in [5] to describe weak probabilistic anonymity, but the measuring method used here is different and no explicit notion of schedulers is considered.

3. $D_A(P \setminus B) \leq D_A(P)$ *if* $A \cap B = \emptyset$*;*
4. $D_A(f[P]) \leq D_A(P)$ *if* $f(a) = a$ *for all* $a \in A$*;*
5. $D_A(\mu_X E) = D_A(E\{\mu_X E/X\})$.

*Proof.* We sketch the proof for each clause.

1. Given any $f \in F_A$, we show that for each $\eta \in tds(pa(\sum_{i\in I} u_{p_i}.P_i))$ there exists some $\eta' \in tds(pa(\sum_{i\in I} u_{p_i}.f[P_i]))$ such that $D(\eta, \eta') \leq \sum_{i\in I} p_i D_A(P_i)$. Note that $\eta$ is determined by a scheduler $\sigma$. Restricting $\sigma$ to $pa(P_i)$, for each $i \in I$, we have a scheduler $\sigma_i$ that resolves all non-deterministic choices in $P_i$, resulting in a trace distribution $\eta_i \in tds(pa(P_i))$. It is easy to see that

$$\eta(C_u) = 1 \text{ and } \eta(C_{u\gamma}) = \textstyle\sum_{i\in I} p_i \cdot \eta_i(C_\gamma)$$

for any trace $\gamma$. Observe that, as a graph, $pa(\sum_{i\in I} u_{p_i}.f[P_i])$ is isomorphic to $pa(\sum_{i\in I} u_{p_i}.P_i)$. Hence there is a scheduler $\sigma'$ of $\sum_{i\in I} u_{p_i}.f[P_i]$ that resolves all non-deterministic choices in the same way as $\sigma$ does for $\sum_{i\in I} u_{p_i}.P_i$. It follows that each scheduler $\sigma_i$ also has a counterpart $\sigma'_i$ that is a scheduler of $f[P_i]$, for each $i \in I$. Each $\sigma'_i$ determines a trace distribution $\eta'_i \in tds(pa(f[P_i]))$ satisfying

$$\eta'(C_u) = 1 \text{ and } \eta'(C_{u\gamma}) = \textstyle\sum_{i\in I} p_i \cdot \eta'_i(C_\gamma).$$

for some $\eta' \in tds(pa(\sum_{i\in I} u_{p_i}.f[P_i]))$. Therefore, it holds that

$$D(\eta, \eta') = D(\sum_{i\in I} p_i\eta_i, \sum_{i\in I} p_i\eta'_i) \leq \sum_{i\in I} p_i D(\eta_i, \eta'_i) \leq \sum_{i\in I} p_i D_A(P_i).$$

The first inequality above is justified by the joint convexity property of relative entropy given in Proposition 3.
2. Given any $f \in F_A$, we let $\Theta = tds(pa(P_1 \boxplus P_2))$ and $\Theta' = tds(pa(f[P_1] \boxplus f[P_2]))$. Each $\eta \in \Theta$ is determined by a scheduler $\sigma$. We consider the interesting case in which $\sigma$ chooses an outgoing transition from $P_i$, for $i = 1, 2$. It follows from the isomorphism between $pa(P_1 \boxplus P_2)$ and $pa(f[P_1] \boxplus f[P_2])$ that $\sigma$ has a counterpart $\sigma'$ which chooses an outgoing transition from $f[P_i]$, and which determines a trace distribution $\eta' \in \Theta'$ satisfying

$$D(\eta, \eta') \leq D_A(P_i) \leq max\{D_A(P_1), D_A(P_2)\}.$$

3. Note that $pa(P \setminus B)$ is the same as $pa(P)$ except that all transitions labelled with actions in $B$ are blocked. If $\eta_1 \in tds(pa(P \setminus B))$, there is some $\eta_2 \in tds(pa(P))$ such that all probabilities assigned by $\eta_2$ to maximal traces of the form $\gamma a \gamma'$ with $a \in B$ are now assigned by $\eta_1$ to $\gamma$. Similar relation holds between the peer trace distribution $\eta'_1 \in tds(pa(f[P] \setminus B))$ of $\eta_1$ and the peer trace distribution $\eta'_2 \in tds(pa(f[P]))$ of $\eta_2$, for any $f \in F_A$. By the strong additivity property given in Proposition 3, we derive that

$$D(\eta_1, \eta'_1) \leq D(\eta_2, \eta'_2) \leq D_A(P).$$

4. If $f$ is an injective renaming function, i.e., $a \neq b$ implies $f(a) \neq f(b)$, then it is immediate that $D_A(f[P]) = D_A(P)$. Otherwise, two different actions may be renamed into the same one. As a result, two different maximal traces in $P$ may become the same in $f[P]$. We can then appeal to the strong additivity property of relative entropy to infer that $D_A(f[P]) \leq D_A(P)$.
5. The result follows from the fact that $\mu_X E$ and $E\{\mu_X E/X\}$ have the same transition graph. □

The above proposition shows a nice property of our approach using relative entropy to measure anonymity. The non-expansiveness of the operators in the probabilistic CCS allows us to estimate the degree of anonymity of a complex system from its components, rather than analyse the system as a whole.

*Remark 7.* Unfortunately, the parallel composition operator is expansive. For example, let $A = \{b, c\}$, $P = a_{\frac{1}{3}}.b + a_{\frac{2}{3}}.c$ and $Q = \bar{a}_{\frac{1}{3}}.b + \bar{a}_{\frac{2}{3}}.c$. We have

$$D_A(P) = D_A(Q) = \frac{1}{3}\log\frac{\frac{1}{3}}{\frac{2}{3}} + \frac{2}{3}\log\frac{\frac{2}{3}}{\frac{1}{3}} = \frac{1}{3}.$$

However, $(P \mid Q) \setminus a = \tau_{\frac{1}{9}}.(b \mid b) + \tau_{\frac{2}{9}}.(b \mid c) + \tau_{\frac{2}{9}}.(c \mid b) + \tau_{\frac{4}{9}}.(c \mid c)$ and

$$D_A((P \mid Q) \setminus a) = \frac{1}{9}\log\frac{\frac{1}{9}}{\frac{4}{9}} + \frac{2}{9}\log\frac{\frac{2}{9}}{\frac{2}{9}} + \frac{2}{9}\log\frac{\frac{2}{9}}{\frac{2}{9}} + \frac{4}{9}\log\frac{\frac{4}{9}}{\frac{1}{9}} = \frac{2}{3}.$$

It follows from Proposition 6 that $D_A(P \mid Q) \geq D_A((P \mid Q) \setminus a) = \frac{2}{3}$.

### 3.3 Small examples

We present some toy examples to show the basic ideas of our approach.

*Example 8.* Consider a communication system that provides anonymous email with 2 potential senders, a mix network and a recipient. The attacker wants to find out which sender sent an email to the recipient. By means of traffic analysis, the attacker obtains a communication system described by the process $P$.

$$P = \tau_p.\mathit{sender}(0).\mathit{email}.\mathit{receive}.\mathbf{0} + \tau_{1-p}.\mathit{sender}(1).\mathit{email}.\mathit{receive}.\mathbf{0}$$

The senders require anonymity, i.e., anonymity is required for the set $A = \{\mathit{sender}(0), \mathit{sender}(1)\}$. In this case, $F_A$ is a singleton set $\{f\}$ with $f[P]$ taking the form:

$$f[P] = \tau_{1-p}.\mathit{sender}(0).\mathit{email}.\mathit{receive}.\mathbf{0} + \tau_p.\mathit{sender}(1).\mathit{email}.\mathit{receive}.\mathbf{0}$$

It is easy to see that $D_A(P) = p\log\frac{p}{1-p} + (1-p)\log\frac{1-p}{p}$. If $p = \frac{1}{2}$, the attacker cannot distinguish the two senders, and indeed the system provides strong anonymity. If $p \to 0$ or $p \to 1$, we have $D_A(P) = +\infty$, which means that the system does not ensure any anonymity of the senders.

*Example 9.* Now suppose the actual system in Example 8 has a built-in non-determinism and behaves in a way described by the process Q.

$$Q = (\tau_{\frac{1}{3}}.sender(0).email.receive.\mathbf{0} + \tau_{\frac{2}{3}}sender(1).email.receive.\mathbf{0}) \boxplus$$
$$(\tau_{\frac{2}{3}}.sender(0).email.receive.\mathbf{0} + \tau_{\frac{1}{3}}.sender(1).email.receive.\mathbf{0})$$

We observe that $f[Q] = Q$ for $f \in F_A$, thus $D_A(Q) = 0$ and the system provides strong anonymity.

## 4 The Dining Cryptographers

The general Dining Cryptographers Problem [4] is described as follows: A number of cryptographers sitting around a table are having dinner. The representative of their organisation (master) may or may not pay the bill of the dinner. If he does not, then he will select exactly one cryptographer and order him to pay the bill. The master will tell secretly each cryptographer whether he has to pay or not. The cryptographers would like to reveal whether the bill is paid by the master or by one of them, but without knowing who among them, if any, is paying. In this paper we consider a DCP with three cryptographers connected by a ring. It is not difficult to extend it to the general case.

A possible solution to this problem, as described in [4], is to associate a coin to every two neighbouring cryptographers. The result of each coin-tossing is only visible to the adjacent cryptographers. Each cryptographer examines the two adjacent coins: If he is not paying, he announces "agree" if the results are the same, and "disagree" otherwise. If he is paying, he says the opposite. If the number of "disagree" is even, then the master is paying. Otherwise, one of the cryptographers is paying.

### 4.1 Fully probabilistic users

We consider the case in which the master probabilistically select one cryptographer to pay. We formalise the DCP as a process in the probabilistic CCS, as illustrated in Figure 2 [5], where $\Pi$ is the parallel composition. We use $\oplus$ (resp. $\ominus$) to represent the sum (resp. the subtraction) modulo 3. Messages $p$ and $n$ are the instructions sent by the master, requiring each cryptographer to pay or not to pay, respectively. The set of anonymous actions is $A = \{pay(i) \mid i = 0, 1, 2\}$. The restriction operator $\backslash$ over the action sequences $\overrightarrow{c}$ and $\overrightarrow{m}$ enforces these actions into internal communications. The traces of $\mathcal{DCP}$ are in the form of $pay(i)xyz$ with $i \in \{0, 1, 2\}$ and $x, y, z \in \{a, d\}$ ($a$ for "agree" and $d$ for "disagree"). $F_A$ contains two elements, one renames $pay(i)$ according to the permutation $f_1 = \{0 \mapsto 1, 1 \mapsto 2, 2 \mapsto 0\}$ and the other $f_2 = \{0 \mapsto 2, 1 \mapsto 0, 2 \mapsto 1\}$.

We assume that all the coins are uniform. With a probabilistic master, $tds(pa(\mathcal{DCP}))$ contains only one trace distribution. Each maximal trace of $\mathcal{DCP}$ can only contain one of the following sequences: $ddd$, $aad$, $ada$, and $daa$.

[5] For the sake of brevity, we formalise the DCP in a value-passing version of the probabilistic CCS, which can be encoded into the probabilistic CCS in the standard way [18]; incorporating the "if-then-else" construct is also straightforward.

$$
\begin{aligned}
Master &= \textstyle\sum_{i=0}^{2} \tau_{p_i}.\ \overline{m}_i(p).\ \overline{m}_{i\oplus 1}(n).\ \overline{m}_{i\oplus 2}(n).\ \mathbf{0} \\
Coin_i &= \tau_{p_h}.\ Head_i +\ \tau_{p_t}.\ Tail_i \\
Head_i &= \overline{c}_{i,i}(head).\ \overline{c}_{i\ominus 1,i}(head).\ \mathbf{0} \\
Tail_i &= \overline{c}_{i,i}(tail).\ \overline{c}_{i\ominus 1,i}(tail).\ \mathbf{0} \\
Crypt_i &= m_i(x).\ c_{i,i}(y).\ c_{i,i\oplus 1}(z). \\
&\quad \text{if } x = p \text{ then } \overline{pay}_i. \\
&\qquad \text{if } y = z \text{ then } \overline{out}_i(disagree) \text{ else } \overline{out}_i(agree) \\
&\quad \text{else if } y = z \text{ then } \overline{out}_i(agree) \text{ else } \overline{out}_i(disagree) \\
\mathcal{DCP} &= (Master \mid (\Pi_{i=0}^{2} Crypt_i \mid \Pi_{i=0}^{2} Coin_i) \setminus \overrightarrow{c}) \setminus \overrightarrow{m}
\end{aligned}
$$

**Fig. 2.** Specification of the DCP in the probabilistic CCS

*Fair coins* With fair coins, if the master assigns the initial probabilities $p_0 = \frac{1}{3}$, $p_1 = \frac{1}{3}$ and $p_2 = \frac{1}{3}$, i.e., each cryptographer has an equal chance to pay, then it is easy to see that $f_1[\mathcal{DCP}] = \mathcal{DCP}$ and $f_2[\mathcal{DCP}] = \mathcal{DCP}$. Therefore, $D_A(\mathcal{DCP}) = 0$ and the DCP provides strong anonymity.

If the master assigns the initial probabilities $p_0 = \frac{1}{2}$, $p_1 = \frac{1}{3}$ and $p_2 = \frac{1}{6}$. The probabilities of traces with *ddd*, *aad*, *ada*, and *daa* are all $\frac{1}{4}$. By the definition of $D$, we can check that

$$D(\mathcal{DCP}, f_1[\mathcal{DCP}]) = 0.431 \quad \text{and} \quad D(\mathcal{DCP}, f_2[\mathcal{DCP}]) = 0.362$$

Hence, the degree of anonymity of such DCP is 0.431.

| | $crypt_0$ pays $(i = 0)$ | $crypt_1$ pays $(i = 1)$ | $crypt_2$ pays $(i = 2)$ |
|---|---|---|---|
| $p(pay(i)ddd)$ | 0.120 | 0.080 | 0.040 |
| $p(pay(i)aad)$ | 0.120 | 0.080 | 0.047 |
| $p(pay(i)ada)$ | 0.120 | 0.093 | 0.040 |
| $p(pay(i)daa)$ | 0.140 | 0.080 | 0.040 |

**Table 1.** A trace distribution (with biased coins: $p_h = \frac{2}{5}$).

*Biased coins* We assume the coins are biased, e.g. $p_h = \frac{2}{5}$. We also consider the case that $p_0 = \frac{1}{2}$, $p_1 = \frac{1}{3}$ and $p_2 = \frac{1}{6}$. Then the probabilities of traces can be calculated as in Table 1. We have

$$D(\mathcal{DCP}, f_1[\mathcal{DCP}]) = 0.209 \quad \text{and} \quad D(\mathcal{DCP}, f_2[\mathcal{DCP}]) = 0.878$$

Hence, the degree of anonymity of such DCP is 0.878, which is greater than 0.431. Therefore, the biased coins leak more information to the attacker than fair coins. If $p_h \to 1.0$, then $D(\mathcal{DCP}, f_1[\mathcal{DCP}]) = D(\mathcal{DCP}, f_2[\mathcal{DCP}]) = +\infty$. Hence, the degree of anonymity of such DCP is $+\infty$, in such case the DCP does not provide any anonymity.

| | $crypt_0$ pays $(i=0)$ | $crypt_1$ pays $(i=1)$ | $crypt_2$ pays $(i=2)$ |
|---|---|---|---|
| $p(pay(i)ddd)$ | 0.24 | 0.24 | 0.24 |
| $p(pay(i)aad)$ | 0.24 | 0.24 | 0.28 |
| $p(pay(i)ada)$ | 0.24 | 0.28 | 0.24 |
| $p(pay(i)daa)$ | 0.28 | 0.24 | 0.24 |

**Table 2.** Three trace distributions (with biased coins: $p_h = \frac{2}{5}$).

### 4.2 Non-deterministic users

We now consider the case in which the master non-deterministically choose a cryptographer to pay, i.e., the master is of the form

$$\begin{aligned} \mathit{Master} = {} & \overline{m}_0(p).\ \overline{m}_1(n).\ \overline{m}_2(n).\ \mathbf{0}\ \boxplus\ \overline{m}_0(n).\ \overline{m}_1(p).\ \overline{m}_2(n).\ \mathbf{0} \boxplus \\ & \overline{m}_0(n).\ \overline{m}_1(n).\ \overline{m}_2(p).\ \mathbf{0} \end{aligned}$$

*Fair coins* With fair coins, it is easy to see that $f_1[\mathcal{DCP}] = \mathcal{DCP}$ and $f_2[\mathcal{DCP}] = \mathcal{DCP}$, i.e., $tds(pa(f_1[\mathcal{DCP}]))$ and $tds(pa(f_2[\mathcal{DCP}]))$ represent the same set of trace distributions as $tds(pa(\mathcal{DCP}))$. Therefore, $D_A(\mathcal{DCP}) = 0$ and the DCP provides strong anonymity.

*Biased coins* We assume the coins are biased, e.g. $p_h = \frac{2}{5}$. Then $tds(pa(\mathcal{DCP}))$ contains the three trace distributions shown in the last three columns of Table 2. It can then be checked that

$$D(\mathcal{DCP}, f_1[\mathcal{DCP}]) = D(\mathcal{DCP}, f_2[\mathcal{DCP}]) = 0.009$$

Hence, the degree of anonymity of such DCP is 0.009

*Remark 10.* The master of a DCP models the *a priori* knowledge of the attacker. In the particular case that the master is purely non-deterministic, the attacker has no *a priori* knowledge of the users. The attacker simply assumes that there is a uniform probability distribution among the users, we then get an ideal situation of anonymity similar to that considered in [8].

## 5 Related Work

In his seminal paper, Chaum [4] used the size of an anonymity set to indicate the degree of anonymity provided by a DC network. An anonymity set is defined as the set of participants who could have sent a particular message as observed by the attacker. Berthold *et al.* [2] defined the degree of anonymity as $\ln(N)$, where $N$ is the number of users of the protocols. Both [4] and [2] only deal with non-deterministic cases, and do not consider the probabilistic information of the users the attacker can gain by observing the system.

Reiter and Rubin [21] defined the degree of anonymity as $1-p$, where $p$ is the probability assigned to a particular user by the attacker. Halpern and O'Neill have proposed in [12] several notions of probabilistic anonymity. Their basic notion is formulated as a requirement on the knowledge of the attacker about the probability of the user. They have given both strong and weak version of this notion, proposing a formal interpretation of the three levels of the hierarchy proposed in [21]. Deng, Palamidessi and Pang [5] proposed a weak notion of probabilistic anonymity as an extension of [3] to measure the leaked information, which can be used by an attacker to infer the likeliness that the action has been performed by a certain user. Thus, the degree of anonymity is formalised as an factor by which the probability the attacker attributes to a user as the performer of the anonymous action has increased, after observing the system. All these methods focus on the probability of the users. Thus, they do not give any information on how distinguishable the user is within the anonymity set.

Serjantov and Danezis [24] and Claudia *et al.* [8] independently proposed an information theoretic metric based on the idea of measuring probability distributions. They used entropy to define the quality of anonymity and to compare different anonymity systems. Compared to [8], [24] does not normalise the degree in order to get a value relative to the anonymity level of the ideal system for the same number of users. Both [24] and [8] take into account the probabilities of the users performing certain actions which are assigned by an attacker after observing the system. However, they do not take into account the *a priori* information that the attacker might have. The attacker simply assumes a uniform distribution among the users before observation. Our method uses relative entropy, and it quantifies the amount of probabilistic information revealed by the protocol, i.e. how much information an attacker can achieve after observing the outcomes of the protocol, together with the information he has before the protocol running. Furthermore, we extend the measuring method to two sets of probability distributions using Hausdorff distance for protocols containing both non-deterministic and probabilistic behaviours.

Moskowitz *et al.* [19] proposed to use a related notion of mutual information to measure the capacity of covert channels. They have applied it to the analysis of a wide range of Mix-networks [20]. Our work is still diffrent from them, in the sense that we use relative entropy instead of mutual information, and we focus on the non-expansiveness of the operators of the probabilistic CCS, which potentially allows for compositional analysis.

## 6 Conclusion and Future Work

In this paper, we have proposed to use relative entropy as a distance of two discrete probability distributions to measure anonymity for protocols which can be interpreted as a fully probabilistic automaton. This definition has been extended for two sets of probability distributions to also capture the non-deterministic aspect of these protocols. We have proved that based on this measuring method, most of the operators in the probabilistic CCS are non-expansive. We have demonstrated our approach by using the example of the Dining Cryptographers Problem.

Model checking anonymity protocols in the logic of knowledge was considered in [27]. It would also be interesting to investigate the problem in a probabilistic setting. The probabilistic model checker PRISM [17] was used to find novel attacks on Crowds [25], and to analyse cascase networks [9]. We intend to integrate our method with PRISM to build an automatic tool to assist the calculation of the degree of anonymity as defined in the paper. We also plan to apply our approach to more complex and real protocols to justify its usefulness.

## References


1. C. Baier and M. Z. Kwiatkowaska. Domain equations for probabilistic processes. *Mathematical Structures in Computer Science*, 10(6):665–717, 2004.
2. O. Berthold, A. Pfiztmann, and R. Standtke. The disavantages of free mix routes and how to overcome them. In *Proc. Workshop on Design Issues in Anonymity and Unobservability*, volume 2009 of *LNCS*, pages 30–45. Springer, 2000.
3. M. Bhargava and C. Palamidessi. Probabilistic anonymity. In *Proc. 16th Conference on Concurrency Theory*, volume 3653 of *LNCS*, pages 171–185. Springer, 2005.
4. D. Chaum. The dining cryptographers problem: Unconditional sender and recipient untraceability. *Journal of Cryptology*, 1:65–75, 1988.
5. Y. Deng, C. Palamidessi, and J. Pang. Weak probabilistic anonymity. In *Proc. 3rd Workshop on Security Issues in Concurrency*, ENTCS, 2006. To appear.
6. J. Desharnais, R. Jagadeesan, V. Gupta, and P. Panangaden. The metric analogue of weak bisimulation for probabilistic processes. In *Proc. 17th IEEE Symposium on Logic in Computer Science*, pages 413–422. IEEE Computer Society, 2002.
7. J. Desharnais, R. Jagadeesan, V. Gupta, and P. Panangaden. Metrics for labelled Markov processes. *Theoretical Computer Science*, 318(3):323–354, 2004.
8. C. Díaz, S. Seys, J. Claessens, and B. Preneel. Towards measuring anonymity. In *Proc. 2nd Workshop on Privacy Enhancing Technologies*, volume 2482 of *LNCS*, pages 54–68. Springer, 2002.
9. R. Dingledine, V. Shmatikov, and P. F. Syverson. Synchronous batching: From cascades to free routes. In *Proc. 4th Workshop on Privacy Enhancing Technologies*, volume 3424 of *LNCS*, pages 186–206. Springer, 2004.
10. S. Furuichi, K. Yanagi, and K. Kuriyama. Fundamental properties of Tsallis relative entropy. *Journal of Mathematical Physics*, 45(12):4868–4877, 2004.
11. F. D. Garcia, I. Hasuo, W. Pieters, and P. van Rossum. Provable anonymity. In *Proc. 3rd ACM Workshop on Formal Methods in Security Engineering*, pages 63–72. ACM Press, 2005.
12. J. Y. Halpern and K. R. O'Neill. Anonymity and information hiding in multiagent systems. In *Proc. 16th IEEE Computer Security Foundations Workshop*, pages 75–88. IEEE Computer Society, 2003.
13. C. A. R. Hoare. *Communicating Sequential Processes*. Prentice Hall, 1985.
14. D. Hughes and V. Shmatikov. Information hiding, anonymity and privacy: a modular approach. *Journal of Computer Security*, 12(1):3–36, 2004.
15. D. Kesdogan, J. Egner, and R. Büschkes. Stop-and-go MIXes: Providing probabilistic anonymity in an open system. In *Proc. 2nd Workshop on Information Hiding*, volume 1525 of *LNCS*, pages 83–98. Springer, 1998.

16. S. Kullback and R. A. Leibler. On information and sufficiency. *Annals of Mathematical Statistics*, 22(1):79–86, 1951.
17. M. Kwiatkowska, G. Norman, and D. Parker. PRISM: Probabilistic symbolic model checker. In *Proc. 12th Conference on Computer Performance Evaluation, Modelling Techniques and Tools*, volume 2324 of *LNCS*, pages 200–204. Springer, 2002.
18. R. Milner. *Communication and Concurrency.* Prentice-Hall, 1989.
19. I. S. Moskowitz, R. E. Newman, D. P. Crepeau, and A. R. Miller. Covert channels and anonymizing networks. In *Proceedings of 2nd ACM Workshop on Privacy in the Electronic Society*, pages 79–88. ACM, 2003.
20. R. E. Newman, V. R. Nalla, and I. S. Moskowitz. Anonymity and covert channels in simple timed mix-firewalls. In *Proc. 4th Workshop on Privacy Enhancing Technologies*, volume 3424 of *LNCS*, pages 1–16. Springer, 2004.
21. M. K. Reiter and A. D. Rubin. Crowds: Anonymity for Web transactions. *ACM Transactions on Information and System Security*, 1(1):66–92, 1998.
22. S. Schneider and A. Sidiropoulos. CSP and anonymity. In *Proc. 4th European Symposium on Research in Computer Security*, volume 1146 of *LNCS*, pages 198–218. Springer, 1996.
23. R. Segala. *Modeling and Verification of Randomized Distributed Real-Time Systems*. PhD thesis, MIT, Deptartment of EECS, 1995.
24. A. Serjantov and G. Danezis. Towards an information theoretic metric for anonymity. In *Proc. 2nd Workshop on Privacy Enhancing Technologies*, volume 2482 of *LNCS*, pages 41–53. Springer, 2002.
25. V. Shmatikov. Probabilistic model checking of an anonymity system. *Journal of Computer Security*, 12(3/4):355–377, 2004.
26. P. F. Syverson, D. M. Goldschlag, and M. G. Reed. Anonymous connections and onion routing. In *Proc. 18th IEEE Symposium on Security and Privacy*, pages 44–54. IEEE Computer Society, 1997.
27. R. van der Meyden and K. Su. Symbolic model checking the knowledge of the dining cryptographers. In *Proc. 17th IEEE Computer Security Foundations Workshop*, pages 280–291. IEEE Computer Society, 2004.