# Anonymity in probabilistic and nondeterministic systems [1]

Catuscia Palamidessi

*INRIA Futurs and LIX, École Polytechnique, Rue de Saclay, 91128 Palaiseau Cedex, FRANCE*

**Abstract**

Anonymity means that the identity of the user performing a certain action is maintained secret. The protocols for ensuring anonymity often use random mechanisms which can be described probabilistically. The user, on the other hand, may be selected either nondeterministically or probabilistically. We investigate various notions of anonymity, at different levels of strength, for both the cases of probabilistic and nondeterministic users. Probabilistic process algebra have been of great help in the development of our setting.

*Key words:* Anonymity, probability, nondeterminism.

Anonymity is the property of keeping secret the identity of the user performing a certain action. The need for anonymity may raise in a wide range of situations, like postings on electronic forums, voting, delation, donations, and many others.

The protocols for ensuring anonymity often use random mechanisms which can be described probabilistically. This is the case, for example, of the Dining Cryptographers [3], Crowds [7], and Onion Routing [11]. In contrast, we usually don't know anything about the users, so their behavior, and in particular, the choice of the user who performs the action with respect to which we want to ensure anonymity, should better be regarded as nondeterministic. (The same would hold for adversaries, although in this paper we do not consider them.) The whole system constituted by the protocol and the users presents therefore both probabilistic and nondeterministic aspects.

Various formal definitions and frameworks for analyzing anonymity have been developed in literature. They can be classified into approaches based on process-calculi [9,8], epistemic logic [10,5], and "function views" [6]. From the point of view of the concepts of probability and nondeterminism, however, all these approaches are either *purely nondeterministic* (also known as *possibilistic*) or *purely probabilistic*.

[1] This work has been partially supported by the Project Rossignol of the ACI Sécurité Informatique (Ministère de la recherche et nouvelles technologies).

*This paper is electronically published in*
*Electronic Notes in Theoretical Computer Science*
*URL:* `www.elsevier.nl/locate/entcs`

The purely nondeterministic approach in [9,8] is based on the so-called "principle of confusion": 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).

The purely probabilistic proposals can be classified under two different points of view: those which focus on the probability of the users, and those which focus on the effect that the observables have on the probability of the users. The distinction is subtle but fundamental. In the fist case, anonymity holds when (an observer knows that) all users have the same probability of having performed the action (cfr. *strong probabilistic anonymity* in [5]). In the second case, it holds when the for any user $i$ and any observable $o$ the conditional probability that $i$ has performed the action, given the observable, is the same as the (a priory) probability that the user has performed the action (cfr. the informal notion used in [3], and the *conditional probabilistic anonymity* in [5]).

The probabilistic approach also brings naturally to differentiate the notion of anonymity with respect to different levels of strength. Reiter and Robin [7] have proposed the following hierarchy:

*Beyond suspicion* The actual user (i.e. the user that performed the action) is not more likely (to have performed the action) than every other user.

*Probable innocence* The actual user has probability less than $1/2$.

*Possible innocence* There is a non trivial probability that another used could have performed the action.

These notions were only given informally in [7], and it is unclear to us whether the authors had in mind the first or the second of the "points of view" described above. On one hand, if we interpret the informal definitions literally, they correspond to the first point of view. This is the interpretation given by Halpern and O'Neill in [5]: they characterize probable innocence and possible innocence with the notion of (probabilistic) $\alpha$-anonymity, and beyond suspicion with their notion of strong probabilistic anonymity. On the other hand, the result of probable innocence proved in [7] for Crowds does not seem to fit with this interpretation, while it could fit with a suitable weakening of the anonymity notion illustrated above under the second perspective (i.e. what Halpern and O'Neill call conditional probabilistic anonymity).

In our approach we assume that the users may be nondeterministic, i.e. that nothing may be known about the relative frequency by which each user perform the anonymous action. More precisely, the users can in principle be totally unpredictable and change intention every time, so that their behavior cannot be thought of as probabilistic [2] . The internal mechanisms of the systems, on the contrary, like

[2] In the areas of concurrency theory there has been a long-standing discussion on whether nondeterminism can be thought of as a situation in which the probabilities are unknown and can change very time the experiment is repeated. Nowadays the prevailing opinion, which we share, is that

coin tossing in the dining philosophers, or the random selection of a nearby node in Crowds, are supposed to exhibit a certain regularity and obey a probabilistic distribution. Correspondingly, we explore a notion of probabilistic anonymity that focuses on the internal mechanism of the system, i.e. their non-leakage of probabilistic information, and it is in a sense independent from the users in case they are nondeterministic. The counterpart of our definition in the case the users are probabilistic (with possibly unknown probabilities), can be shown to correspond to a generalized version of Halpern and O'Neill's conditional probabilistic anonymity, where "generalized" here means that the anonymity holds for any probability distribution to the users.

The formalism that we use for the description and the analysis of the anonymity protocols is a process algebra with both nondeterministic and probabilistic choice. This kind of process algebra constitute a rich framework that allows describing a variety of phenomena, from concurrent and distributed systems to security protocols which involve random primitives. Despite the fact that this subject is relatively recent and its theoretical foundations are still under development, it has been of great help in this project: once the protocols have been expressed in the familiar formalism of process algebra, the author has got a much better understanding of the concept of (probabilistic) anonymity, and the the development of the formal setting has been a quite natural process.

This abstract is based on the papers [1], [2] and [4].

## References

[1] Mohit Bhargava and Catuscia Palamidessi. Probabilistic anonymity. In Martín Abadi and Luca de Alfaro, editors, *Proceedings of CONCUR*, volume 3653 of *Lecture Notes in Computer Science*, pages 171–185. Springer-Verlag, 2005.

[2] Kostantinos Chatzikokolakis and Catuscia Palamidessi. Probable innocence revisited. In *Proceedings of the 3rd International Workshop on Formal Aspects in Security and Trust (FAST)*, Lecture Notes in Computer Science. Springer-Verlag, 2005.

[3] David Chaum. The dining cryptographers problem: Unconditional sender and recipient untraceability. *Journal of Cryptology*, 1:65–75, 1988.

[4] Yuxin Deng, Catuscia Palamidessi, and Jun Pang. Weak probabilistic anonymity. In *Proceedings of the 3rd International Workshop on Security Issues in Concurrency (SecCo)*, Electronic Notes in Theoretical Computer Science. Elsevier Science Publishers, 2005.

[5] Joseph Y. Halpern and Kevin R. O'Neill. Anonymity and information hiding in multiagent systems. In *Proc. of the 16th IEEE Computer Security Foundations Workshop*, pages 75–88, 2003.

---

nondetermistism is a different notion. In the formalisms used in concurrency the difference is made explicit by different laws.

[6] Dominic Hughes and Vitaly Shmatikov. Information hiding, anonymity and privacy: a modular approach. *Journal of Computer Security*, 12(1):3–36, 2004.

[7] Michael K. Reiter and Aviel D. Rubin. Crowds: anonymity for Web transactions. *ACM Transactions on Information and System Security*, 1(1):66–92, 1998.

[8] Peter Y. Ryan and Steve Schneider. *Modelling and Analysis of Security Protocols*. Addison-Wesley, 2001.

[9] Steve Schneider and Abraham Sidiropoulos. CSP and anonymity. In *Proc. of the European Symposium on Research in Computer Security (ESORICS)*, volume 1146 of *Lecture Notes in Computer Science*, pages 198–218. Springer-Verlag, 1996.

[10] Paul F. Syverson and Stuart G. Stubblebine. Group principals and the formalization of anonymity. In *World Congress on Formal Methods (1)*, pages 814–833, 1999.

[11] P.F. Syverson, D.M. Goldschlag, and M.G. Reed. Anonymous connections and onion routing. In *IEEE Symposium on Security and Privacy*, pages 44–54, Oakland, California, 1997.