Probable Innocence and Independent Knowledge

Sardaouna Hamadou, Catuscia Palamidessi, Vladimiro Sassone, Ehab Elsalamouny

Probable innocence in the presence of independent knowledge

Sardaouna Hamadou[1], Catuscia Palamidessi[2], Vladimiro Sassone[1], and Ehab ElSalamouny[1]

[1] ECS, University of Southampton
[2] INRIA and LIX, Ecole Polytechnique

Abstract. We analyse the Crowds anonymity protocol under the novel assumption that the attacker has independent knowledge on behavioural patterns of individual users. Under such conditions we study, reformulate and extend Reiter and Rubin's notion of probable innocence, and provide a new formalisation for it based on the concept of protocol vulnerability. Accordingly, we establish new formal relationships between protocol parameters and attackers' knowledge expressing necessary and sufficient conditions to ensure probable innocence.

1 Introduction

Anonymity protocols often use random mechanisms. It is therefore natural to think of anonymity in probabilistic terms. Various notions of such probabilistic anonymity have been proposed and a recent trend of work in formalizing these notions is directed at exploring the application of information-theoretic concepts (e.g. [18, 4–6, 1, 15]). In our opinion, however, except a recent paper by Franz, Meyer, and Pashalidis [11] which addresses the advantage that an adversary could take of hints from the context in which the protocol operates, such approaches fail to account for the fact that in the real world, the adversary often have some extra information about the correlation between anonymous users and observables. Consider for example the following simple anonymous voting process. In a parliament composed by Labourists and Conservatives, one member voted against a proposal banning minimum wages. Without any additional knowledge it is reasonable to assume that the person is more likely to be in the most liberal political group. If however we know in addition that one Conservative voted against, then it is more reasonable to suspect the liberally-inclined Conservatives. Similarly, suppose that in a classroom of n students the teacher asks to tick one of two boxes on a piece of paper to indicate whether or not they are satisfied by her teaching. If n is small and the teacher noticed that the pupils use pens of different colours, then she can use these colours to partition the class so as to make the vote of some students more easily identifiable. Extra knowledge of this kind, independent of the logic of the protocol used, can affect dramatically its security. The extra knowledge can either arise from an independent source, as in the first example, or simply from the context in which the anonymity protocol is run, as in the second example.

A relevant point in case is Reiter and Rubin's Crowds protocol [16], which allows Internet users to perform anonymous web transactions. The idea is to send the message through a chain of users participating in the protocol. Each user in the '*crowd*' must

establish a path between her and a set of servers by selecting randomly some users to act as routers. The random selection process is performed in such a way that when a user in the path relays a message, she does not know whether or not the sender is the initiator of the message, or simply a forwarder like herself. Each user only has access to messages routed through her, and some participants may be corrupted, i.e., they may work together in order to uncover the identity of the initiator. It is well known that Crowds cannot ensure strong anonymity [16, 3] in presence of corrupted participants, but when the number of corrupted users is sufficiently small, it provides a weaker notion of anonymity known as *probable innocence*. Informally, a sender is probably innocent if to an attacker she is no more likely to be the originator than not to be.

Although Crowds has been widely analysed in the literature (e.g. [3, 15]), the fact that independent information may be available to the attacker has been so far ignored. We maintain that this is ultimately incompatible with achieving a comprehensive and reliable analysis of the protocol, as attackers' extra knowledge is inherent to Crowds. In particular, as any request routed through an attacker reveals the identity of the target server, a team of attackers will soon build up a host of observations suitable to classify the behaviour of honest participants.

This paper is to the best of our knowledge the first to investigate the impact of the attacker's independent knowledge on the anonymity in the Crowds protocol.

Related work. Quantitative approach to the foundations of information-hiding has became a very active and mature research field. Several various formal definitions and frameworks have been proposed for reasoning about *secure information flow analysis* (e.g. [19, 7–9]), *side-channel analysis* (e.g [14]) and *anonymity*. Our work follows a recent trend in the analysis of anonymity protocols directed to the application of information-theoretic notions (e.g. [17, 18, 4–6, 1, 15, 10, 2]).

The most related work to ours is the one of Reiter and Ruben [16], the one of Halpen and O'Neill [12], and the recent paper of Chatzikokolakis, and Palamidessi [3]. In [16] the authors propose a formal definition of probable innocence that considers the probability of observable events induced by actions of an anonymous user participating in the protocol. They require that the probability of an anonymous user producing any observable to be less than one half. In [12] the authors formalize probable innocence in terms of the adversary's confidence that a particular anonymous event happened, after performing an observation. Their definition requires that the probability of an anonymous events should be at most one half, under any observation. In [3] the authors argue that the definition of [16] makes sense only for systems satisfying certain properties while the definition of [12] depends on the probabilities of anonymous events which are not part of the protocol. They propose a definition of probable innocence that tries to combine the two previous ones by considering both the probability of producing some observable and the adversary's confidence after the observation.

Another recent work closely related to ours is the one of Smith's [19] which proposes a new metric for quantitative information flow based on the concept of *vulnerability* as an alternative to previous metrics based on Shannon entropy and mutual information. Informally, the idea is that the adversary *knows the a priori distributions* of the hidden (anonymous) events and always 'bets' on the *most likely culprit*. The *a priori* vulnerability then is the probability that the adversary guesses the true culprit based

only on the *a priori* distribution. The *a posteriori* vulnerability is the *average* probability that the adversary guesses the true culprit based on the *a posteriori* probability distribution on the agents after the observation.

The main difference between these approaches and ours is that they do not take into account the very likely additional knowledge of the adversary about the correlation between the anonymous events and some observables independent from the behaviour of the protocol. In this paper we first generalize the concepts of probable innocence and vulnerability. Instead than just comparing the probability of being innocent with the probability of being guilty, we consider the degree of the probability of being innocent. Informally a protocol is α-probable innocent if for any anonymous user the probability of being innocent is less than or equal to α. Similarly a protocol is α-vulnerable if the a posteriori vulnerability of the anonymous users is less than or equal to α. We prove that these two notions are related. In particular (α-)probable innocence implies (α-)vulnerability and in the specific case when the a priori distribution of the anonymous events is uniform, they are equivalent. We furthermore extend these definitions in order to cope with the extra independent knowledge of the adversary by computing the a posteriori probability and the a posteriori vulnerability w.r.t to both the protocol observables and the independent observables. We show that the presence of extra knowledge makes probable innocence (resp. vulnerability) more difficult to be achieved.

Finally, it should be acknowledged that our observations about the importance of additional knowledge of the adversary are not entirely new. Indeed, as already noticed above, Franz, Meyer, and Pashalidis [11] considered the fact that an adversary could take advantage of hints from the context in which a protocol operates. However, though that their approach is closely related to ours in spirit, it is not general in the sense that it assumes a deterministic correlation between the anonymous events and the observable hints and a uniform distribution on the anonymous events. Moreover, their metric is associated to Shannon entropy which is recently proven by Smith [19] of being less accurate than vulnerability-based metric.

Structure of the paper. The paper is organised as follows: in §2 we fix some basic notations and recall the fundamental ideas of the Crowds protocol and its properties, including the notion of probable innocence. In §3 we reformulate and extend probable innocence using the idea of protocol vulnerability; §4 and §5 deliver our core technical contribution by respectively extending probable innocence and vulnerability to the case of attacker's independent knowledge.

2 Preliminaries

This section describes our conceptual framework and briefly revises the Crowds protocol and the notion of probable innocence. We use capital letters A, B to denote discrete random variables and the corresponding small letters a, b and calligraphic letters $\mathcal{A}$, $\mathcal{B}$ for their values and set of values respectively. We denote by $p(a)$, $p(b)$ the probabilities of a and b respectively and by $p(a \wedge b)$ their *joint probability*. The *conditional probability* of a given b is defined as

$$p(a \,|\, b) = \frac{p(a \wedge b)}{p(b)}$$

The Bayes theorem relates the conditional probabilities $p(a \mid b)$ and $p(b \mid a)$ as follows

$$p(a \mid b) = \frac{p(b \mid a)\, p(a)}{p(b)}$$

2.1 The framework

In this paper we consider a framework similar to the probabilistic approaches to anonymity and information flow used in (for instance) [13], [5], [15], and [19]. We restrict ourselves to *total* protocols and programs with one *high level* (or *anonymous*) input A, a random variable over a finite set $\mathcal{A}$, and one *low level* output (observable) O, a random variable over a finite set $\mathcal{O}$. We represent a protocol/program by the matrix of the conditional probabilities $p(o_j \mid a_i)$, where $p(o_j \mid a_i)$ is the probability that the low output is o_j given that the high input is a_i. We assume that the high input is generated according to an *a priori* publicly-known probabilistic distribution. An adversary or eavesdropper can see the output of a protocol, but not the input, and he is interested in deriving the value of the input from the observed output *in one single try*.

In this paper we will also assume that the attacker has access to the value of a random variable S distributed over $\mathcal{S}$ that summarizes his additional knowledge (information) about A independent from the behavior of the protocol, as explained in the introduction. The matrix of the conditional probabilities $p(s_k \mid a_i)$ expresses the correlation between the anonymous events and the additional knowledge of the adversary.

When $|\mathcal{S}| = 1$ the adversary's additional information about A is a *trivial one* and cannot help his effort in determining the value of A. For example knowing the length of a password in a fixed-length password system is a trivial information since all passwords have the same length. Trivial information allows us to model the absence of additional information. The standard framework can therefore be seen as an instance of our framework.

2.2 The Crowds Protocol and the definition of probable innocence

The protocol. Crowds is a protocol proposed by Reiter and Rubin in [16] to allow Internet users performing anonymous web transactions by protecting their identity as originators of messages. The central idea to ensure anonymity is that the originator forwards the message to another, randomly-selected user, which in turn forwards the message to another user, and so on until the message reaches its destination (the end server). This routing process ensures that, even when a user is detected sending a message, there is a substantial probability that she is simply forwarding it on behalf of somebody else.

More specifically, a crowd is a *fixed* number of users participating in the protocol. Some members (users) in the crowd may be corrupted (the *attackers*), and they can collaborate in order to discover the originator's identity. The purpose of the protocol is to protect the identity of the message originator from the attackers. When an *originator* –also known as *initiator*– wants to communicate with a server, she creates a random *path* between herself and the server through the crowd by the following process.

- *Initial step:* the initiator selects uniformally at random a member of the crowd (possibly herself) and forwards the request to her. We refer to the latter user as the *forwarder*.

– *Forwarding steps:* a forwarder, upon receiving a request, flips a *biased* coin. With probability $1 - p_f$ she delivers the request to the end server. With probability p_f she selects uniformally at random a new forwarder (possibly herself) and forwards the request to her. The new forwarder repeats the same forwarding process.

The response from the server to the originator follows the same path in the opposite direction. Each user (including corrupted users) is assumed to have only access to messages routed through her, so that she only knows the identities of her immediate predecessor and successor in a path, and the end server.

Informal definition of Probable Innocence. In [16] Reiter and Rubin have proposed a hierarchy of anonymity notions in the context of Crowds. These range from '*absolute privacy*,' where the attacker cannot perceive the presence of communication, to '*provably exposed*,' where the attacker can prove the sender and receiver relationship. Clearly, as most protocols used in practice, Crowds cannot ensure absolute privacy in presence of attackers or corrupted users, but can only provide weaker notions of anonymity. In particular, in [16] the authors propose an anonymity notion called *probable innocence* and prove that, under some conditions on the protocol parameters, Crowds ensures the probable innocence property to the originator. Informally, they define it as follows:

> *A sender is probably innocent if, from the attacker's point of view, the sender appears no more likely to be the originator than to not be the originator.* (1)

In other words, the attacker may have reason to suspect the sender of being more likely than any other potential sender to be the originator, but it still appears at least as likely that she is not.

The formal property proved by Reiter and Rubin. Let m be the number of users participating in the protocol and let c and n be the number of the corrupted and honest users, respectively, with $m = n + c$. Since anonymity makes only sense for honest users, we define the set of anonymous events as $\mathcal{A} = \{a_1, a_2, \ldots, a_n\}$, where a_i indicates that user i is the initiator of the message. We define the set of observable events as $O = \{o_1, o_2, \ldots, o_n\}$, where o_i indicates that user i forwarded a message to a corrupted user. We also say that user i *is detected* by the attacker.

As it is usually the case in the analysis of Crowds, we assume that attackers will always deliver a request to forward immediately to the end server, since forwarding it any further cannot help them learn anything more about the identity of the originator.

In [16] Reiter and Rubin formalise their notion of probable innocence via the conditional probability $p(I \mid H)$ that the initiator is detected given that any user is detected at all. Here I denotes the event that it is precisely the initiator to forward the message to the attacker on the path, and H that there is an attacker in the path. Precisely, probable innocence holds if $p(I \mid H) \leq \frac{1}{2}$.

In our setting the probability that user j is detected given that user i is the initiator, can be written simply as $p(o_j \mid a_i)$. As we are only interested in the case in which a user is detected, for simplicity we do not write such condition explicitly. Therefore, the property proved in [16] (i.e. $p(I \mid H) \leq \frac{1}{2}$) translates in our setting as:

$$\forall i.\ p(o_i \mid a_i) \leq 1/2 \tag{2}$$

Reiter and Rubin proved in [16] that, in Crowds, the following holds:

$$p(o_j \mid a_i) \;=\; \begin{cases} 1 - \frac{n-1}{m} p_f & i = j \\ \frac{1}{m} p_f & i \neq j \end{cases}$$

Therefore, probable innocence (2) holds if and only if

$$n \geq \frac{c-1}{p_f - \frac{1}{2}} p_f$$

3 Probable innocence revisited and extended

In our opinion there is a mismatch between the idea of probable innocence expressed informally in (1) and the property actually proved by Reiter and Rubin, cf. (2). The former, indeed, seems to correspond to the following:

$$\forall i.\ p(a_i \mid o_i) \leq 1/2 \tag{3}$$

It is worth noting that this is also the interpretation given by Halpern and O'Neill [13].

The properties (2) and (3) however coincide under the assumption that the *a priori* distribution is uniform, i.e. that each honest user has equal probability of being the initiator. This is a standard asumption in Crowds.

Proposition 1. *If the* a priori *distribution is uniform, then* $\forall i, j.\ p(a_i \mid o_j) = p(o_j \mid a_i)$.

Proof. If the *a priori* distribution is uniform, then for every i we have $p(a_i) = 1/n$ where n is the number of honest users. The probability of user j being detected is also uniform, and hence equal to $1/n$. In fact, every initiator forwards the message to each other user with the same probability, and each forwarder does the same, hence each user has the same probability of being detected when she is the initiator, and the same probability of being detected when she is not the initiator. Therefore we have: $p(o_j \mid a_j) = p(o_k \mid a_k)$ and $p(o_j \mid a_i) = p(o_k \mid a_i)$ for every j, k and $i \neq j, k$, and hence:

$$\begin{aligned} p(o_j) &= p(o_j \wedge a_j) + \textstyle\sum_{i \neq j} p(o_j \wedge a_i) \\ &= p(o_j \mid a_j) p(a_j) + \textstyle\sum_{i \neq j} p(o_j \mid a_i) p(a_i) \\ &= p(o_k \mid a_k) p(a_k) + \textstyle\sum_{i \neq k} p(o_k \mid a_i) p(a_i) \quad \text{by symmetry} \\ &= p(o_k) \end{aligned}$$

Finally, by using the Bayes theorem, we have:

$$p(a_i \mid o_j) \;=\; \frac{p(o_j \mid a_i)\ p(a_i)}{p(o_j)} \;=\; \frac{p(o_j \mid a_i) \cdot 1/n}{1/n} \;=\; p(o_j \mid a_i)$$

□

Corollary 1. *If the* a priori *distribution is uniform, then (2) and (3) are equivalent.*

The following proposition points out that in presence of uniform *a priori* distribution, the matrix associated to the protocol, i.e. the array of the conditional probabilities $p(o_j|a_i)$, has equal elements everywhere except on the diagonal:

Proposition 2. *If the* a priori *distribution is uniform, then there exists a p such that*

$$p(o_j \mid a_i) = \begin{cases} p & i = j \\ \frac{1-p}{n-1} & i \neq j \end{cases}$$

Proof. As already noted in the proof of Proposition 1, for symmetry reasons we have $p(o_j \mid a_j) = p(o_k \mid a_k)$ and $p(o_j \mid a_i) = p(o_k \mid a_i)$ for every j, k and $i \neq j, k$. □

It is generally the case, in Crowds, that p is (much) greater than $(1 - p)/(n - 1)$, which means that the user which is detected is also the most likely culprit. This allows us to reformulate the property of probable innocence in terms of the (*a posteriori*) *vulnerability* [19] of a protocol, which coincides with the converse of the Bayes risk [6].

Let us briefly recall the definition of vulnerability. The idea is that the adversary *knows the a priori distributions* and always 'bets' on the *most likely culprit*. The *a priori* vulnerability then is the probability that the adversary guesses the true culprit based only on the *a priori* distribution $p(a)$. The *a posteriori* vulnerability is the *average* probability that the adversary guesses the true culprit based on the *a posteriori* probability distribution on the agents after the observation, i.e., $p(a \mid o)$. Formally:

Definition 1 ([19]).

- *The* a priori *vulnerability is* $V(A) = \max_i p(a_i)$
- *The* a posteriori *vulnerability is* $V(A \mid O) = \sum_j p(o_j) \max_i(p(a_i \mid, o_j))$

Using the Bayes theorem, we can reformulate $V(A \mid O)$ as follows:

$$V(A \mid O) = \sum_j \max_i(p(o_j|a_i)\, p(a_i)) \tag{4}$$

It is easy to see that probable innocence implies that the *a posteriori* vulnerability is smaller than $1/2$. The converse also holds, if the *a priori* distribution is uniform.

Proposition 3.

- *If either (2) or (3) holds, then* $V(A \mid O) \leq 1/2$.
- *If* $V(A \mid O) \leq 1/2$ *and the* a priori *distribution is uniform, then (2) and (3) hold.*

We now generalize the concept of probable innocence. Instead than just comparing the probability of being innocent with the probability of being guilty, we consider the degree of the probability of being innocent. Similarly for the vulnerability.

Definition 2. *Given a real number* $\alpha \in [0, 1]$*, we say that a protocol satisfies*

- *α-probable innocence if and only if* $\forall i.\ p(a_i \mid o_i) \leq \alpha$
- *α-vulnerability if and only if* $V(A \mid O) \leq \alpha$.

Clearly α-probable innocence coincides with the standard probable innocence for $\alpha = 1/2$. It is also to be remarked that the minimum possible value of α is $1/n$, i.e., it is not possible for a protocol to satisfy α-probable innocence or α-vulnerability if α is smaller than this value.

4 Probable innocence in presence of extra information

We now consider the notion of probable innocence when we assume that the adversary has some extra information about the correlation between the culprit and the observable.

We express this extra information in terms of a random variable S, whose values $s_1 \dots s_\ell$ we assume to be observable, and the conditional probabilities $p(s_k \,|\, a_i)$. We assume that, the original observables O and the additional observables S are independent, for every originator.

Example 1. Consider an instance of the Crowds protocol in which there are two servers, and assume that the users are divided in two parts, A_1 and A_2. Assume that each user in A_1, when he is the initiator, has probability p_1 to address his message to the first server (as the final destination of the message). Conversely, assume that each user in A_2 has probability p_2 to address the second server. The address of the server appears in the message, and it is therefore observed by the adversary when he intercepts the message. It is clear that (because of the way Crowds works) the event that the message is intercepted is independent from the server to which the message is addressed.

If we indicate by s_1 the fact that the message is addressed to the first server, and by s_2 the fact that the message is addressed to the second server, the matrix of the conditional probabilities corresponding to this example is as follows:

$$p(s \,|\, a) \;=\; \begin{cases} p_1 & a \in A_1, s = s_1 \\ 1 - p_1 & a \in A_1, s = s_2 \\ 1 - p_2 & a \in A_2, s = s_1 \\ p_2 & a \in A_2, s = s_2 \end{cases}$$

We are interested in exploring how the extra information provided by S and the conditional probabilities of S given A affects the notion of probable innocence.

We take the point of view that the invariant property should be the one expressed by (3), generalized by Definition 2. We reformulate this definition to accommodate the presence of extra information in the observables.

Definition 3 (α-probable innocence in presence of extra information). *Given a real number $\alpha \in [0, 1]$, we say that a protocol satisfies α-probable innocence if and only if*

$$\forall i, k.\ p(a_i \,|\, o_i \wedge s_k) \leq \alpha$$

The following lemma expresses the relation between the conditional probabilities with respect to the new observables and the original ones

Lemma 1. $\forall i, j, k.\ p(a_i \,|\, o_j \wedge s_k) = p(a_i \,|\, o_j)\ \frac{p(s_k \,|\, a_i)}{p(s_k \,|\, o_j)}$

Proof. By Bayes theorem we have, for every i, j, k

$$p(a_i \,|\, o_j \wedge s_k) = \frac{p(o_j \wedge s_k \,|\, a_i)\ p(a_i)}{p(o_j \wedge s_k)}$$

Since we are assuming that, given any originator a_i, O and S are independent, we have $p(o_j \wedge s_k \,|\, a_i) = p(o_j \,|\, a_i)\; p(s_k \,|\, a_i)$, and therefore

$$p(a_i \,|\, o_j \wedge s_k) = \frac{p(o_j \,|\, a_i)\; p(s_k \,|\, a_i)\; p(a_i)}{p(o_j \wedge s_k)}$$

We can rewrite $p(o_j \wedge s_k)$ as $p(s_k \,|\, o_j)\; p(o_j)$. Hence:

$$p(a_i \,|\, o_j \wedge s_k) = \frac{p(o_j \,|\, a_i)\; p(s_k \,|\, a_i)\; p(a_i)}{p(s_k|o_j)\; p(o_j)}$$

Finally, using Bayes theorem again, we conclude. □

We can now prove the presence of extra information reduces the degree α of probable innocence by a factor $q = \min_{i,k}(p(s_k \,|\, o_i)/p(s_k \,|\, a_i))$:

Proposition 4.

- *In presence of extra information, a protocol satisfies α-probable innocence if*

$$\forall i.\; p(a_i \,|\, o_i) \leq q\, \alpha$$

- *If $\forall i, j.\; p(a_i \,|\, o_i) = p(a_j \,|\, o_j)$, then the above condition is also necessary, i.e. the protocol satisfies α-probable innocence only if*

$$\forall i.\; p(a_i \,|\, o_i) \leq q\, \alpha$$

Proof. Immediate from previous lemma, with $j = i$. □

In general the factor q in the above proposition is strictly greater than 0 and strictly smaller than 1. Note also that, in the case of Crowds, the protocol satisfies the required symmetry, i.e. the elements in the principal diagonal of the matrix of the conditional probabilities are all the same (cf. Prop. 2) and therefore the above factor q is strict.

Example 2. Consider an instance of the Crowds protocol where there are 6 members ($m = 6$). One of these members is an attacker ($c = 1$), and the others are honest ($n = 5$). Assume that $p_f = 3/4$ then we have

$$p(o_i \,|\, a_i) \;=\; 1 - \frac{n-1}{m} p_f \;=\; 1 - \frac{4}{6} \cdot \frac{3}{4} \;=\; \frac{1}{2}$$

and, for $i \neq j$,

$$p(o_j \,|\, a_i) \;=\; \frac{1}{m} p_f \;=\; \frac{1}{6} \cdot \frac{3}{4} \;=\; \frac{1}{8}$$

Now suppose that, as in Example 1, there are two servers and the honest members are divided into two groups A_1 and A_2, where $A_1 = \{1, 2\}$ (resp. $A_2 = \{3, 4, 5\}$) are the users which prefer the server 1 (resp. the server 2). Assume that the preference probabilities $p_1 = p_2 = 3/4$, i.e. that the conditional probabilities $p(s \,|\, a)$ are given by

$$p(s_k \,|\, a_i) = \begin{cases} \frac{3}{4} & a_i \in A_k \\ \frac{1}{4} & a_i \notin A_k \end{cases}$$

Because of the independence assumption, the conditional probabilities $p(o \wedge s \,|\, a)$ can be computed as the product $p(o \,|\, a)\; p(s \,|\, a)$ (see Fig. 1). From these we can compute the joint probabilities $p(o \wedge s)$ by using the formula

$$p(o_j \wedge s_k) \;=\; \sum_i p(o_j \wedge s_k \,|\, a_i)\; p(a_i)$$

Assuming that the *a priori* distribution is uniform ($p(a_i) = \frac{1}{5}$), we obtain the probabilities shown in Fig. 1. From these we can then calculate $p(s \,|\, o)$ using the definition

$$p(s_k \,|\, o_j) \;=\; \frac{p(s_k \wedge o_j)}{p(o_j)}$$

and the fact that if A is uniformly distributed then also O is uniformly distributed ($p(o_j) = \frac{1}{5}$). Finally, using Bayes theorem, we can calculate the probabilities $p(a \,|\, o \wedge s)$ from $p(o \wedge s \,|\, a)$, $p(o \wedge s)$, and $p(a)$.

Using the values of $p(s_k \,|\, o_i)$ and $p(s_k \,|\, a_i)$, the factor $q = \min_{i,k}(p(s_k \,|\, o_i)/p(s_k \,|\, a_i))$ in Proposition 4 is evaluated to 3/4. It is easy to see that Proposition 4 holds for this instance of Crowds, i.e. $\forall i, k.\; p(a_i \,|\, o_i \wedge s_k) \leq \alpha$ if and only if $\forall i.\; p(a_i \,|\, o_i) \leq q\,\alpha$. In fact $p(a_i \,|\, o_i) = 1/2$ and $\max_{i,k} p(a_i \,|\, o_i \wedge s_k) = 2/3$.

We note that in some cases the extra information may contradict the original observable. For instance it could be the case that user 1, when she is the originator, has a strong preference for the server 1. So if the attacker receives a message from user 1 addressed to the server 2, it may be better for him to assume that the originator is one (arbitrary) user from the group that favors the server 2, rather than user 1.

We argue, therefore, that the presence of extra information makes the property of probable innocence more difficult to satisfy, because the attacker can use the extra information to improve his guess about the culprit, and he may guess a user which is not necessarily the one who sent the message to him. Therefore it seems reasonable to consider the following definition:

Definition 4 (α-probable innocence in presence of extra information, safe version). *Given a real number $\alpha \in [0, 1]$, a protocol satisfies α-probable innocence if and only if*

$$\forall i, j, k.\; p(a_i \,|\, o_j \wedge s_k) \leq \alpha$$

However, it turns out that the relation with the original notion of probable innocence remains the same, and Proposition 4 still provides the appropriate bound:

Proposition 5.

- *In presence of extra information, a protocol satisfies the safe version of α-probable innocence if*
$$\forall i, j.\; p(a_i \,|\, o_j) \leq q\,\alpha$$
- *If $\forall i, j.\; p(a_i \,|\, o_i) \;=\; p(a_j \,|\, o_j)$, then the above condition is also necessary, i.e. the protocol satisfies the safe version of α-probable innocence only if*
$$\forall i.\; p(a_i \,|\, o_i) \leq q\,\alpha$$

$p(o \mid a)$	o_1	o_2	o_3	o_4	o_5
a_1	$\frac{1}{2}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$
a_2	$\frac{1}{8}$	$\frac{1}{2}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$
a_3	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{2}$	$\frac{1}{8}$	$\frac{1}{8}$
a_4	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{2}$	$\frac{1}{8}$
a_5	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{2}$

$p(s \mid a)$	s_1	s_2
a_1	$\frac{3}{4}$	$\frac{1}{4}$
a_2	$\frac{3}{4}$	$\frac{1}{4}$
a_3	$\frac{1}{4}$	$\frac{3}{4}$
a_4	$\frac{1}{4}$	$\frac{3}{4}$
a_5	$\frac{1}{4}$	$\frac{3}{4}$

$p(s \mid o)$	s_1	s_2
o_1	$\frac{9}{16}$	$\frac{7}{16}$
o_2	$\frac{9}{16}$	$\frac{7}{16}$
o_3	$\frac{3}{8}$	$\frac{5}{8}$
o_4	$\frac{3}{8}$	$\frac{5}{8}$
o_5	$\frac{3}{8}$	$\frac{5}{8}$

$p(o, s \mid a)$	o_1, s_1	o_2, s_1	o_3, s_1	o_4, s_1	o_5, s_1	o_1, s_2	o_2, s_2	o_3, s_2	o_4, s_2	o_5, s_2
a_1	$\frac{3}{8}$	$\frac{3}{32}$	$\frac{3}{32}$	$\frac{3}{32}$	$\frac{3}{32}$	$\frac{1}{8}$	$\frac{1}{32}$	$\frac{1}{32}$	$\frac{1}{32}$	$\frac{1}{32}$
a_2	$\frac{3}{32}$	$\frac{3}{8}$	$\frac{3}{32}$	$\frac{3}{32}$	$\frac{3}{32}$	$\frac{1}{32}$	$\frac{1}{8}$	$\frac{1}{32}$	$\frac{1}{32}$	$\frac{1}{32}$
a_3	$\frac{1}{32}$	$\frac{1}{32}$	$\frac{1}{8}$	$\frac{1}{32}$	$\frac{1}{32}$	$\frac{3}{32}$	$\frac{3}{32}$	$\frac{3}{8}$	$\frac{3}{32}$	$\frac{3}{32}$
a_4	$\frac{1}{32}$	$\frac{1}{32}$	$\frac{1}{32}$	$\frac{1}{8}$	$\frac{1}{32}$	$\frac{3}{32}$	$\frac{3}{32}$	$\frac{3}{32}$	$\frac{3}{8}$	$\frac{3}{32}$
a_5	$\frac{1}{32}$	$\frac{1}{32}$	$\frac{1}{32}$	$\frac{1}{32}$	$\frac{1}{8}$	$\frac{3}{32}$	$\frac{3}{32}$	$\frac{3}{32}$	$\frac{3}{32}$	$\frac{3}{8}$

$p(o, s)$	o_1, s_1	o_2, s_1	o_3, s_1	o_4, s_1	o_5, s_1	o_1, s_2	o_2, s_2	o_3, s_2	o_4, s_2	o_5, s_2
	$\frac{9}{80}$	$\frac{9}{80}$	$\frac{6}{80}$	$\frac{6}{80}$	$\frac{6}{80}$	$\frac{7}{80}$	$\frac{7}{80}$	$\frac{10}{80}$	$\frac{10}{80}$	$\frac{10}{80}$

$p(a \mid o, s)$	o_1, s_1	o_2, s_1	o_3, s_1	o_4, s_1	o_5, s_1	o_1, s_2	o_2, s_2	o_3, s_2	o_4, s_2	o_5, s_2
a_1	$\frac{2}{3}$	$\frac{1}{6}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{2}{7}$	$\frac{1}{14}$	$\frac{1}{20}$	$\frac{1}{20}$	$\frac{1}{20}$
a_2	$\frac{1}{6}$	$\frac{2}{3}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{14}$	$\frac{2}{7}$	$\frac{1}{20}$	$\frac{1}{20}$	$\frac{1}{20}$
a_3	$\frac{1}{18}$	$\frac{1}{18}$	$\frac{1}{3}$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{3}{14}$	$\frac{3}{14}$	$\frac{3}{5}$	$\frac{3}{20}$	$\frac{3}{20}$
a_4	$\frac{1}{18}$	$\frac{1}{18}$	$\frac{1}{12}$	$\frac{1}{3}$	$\frac{1}{12}$	$\frac{3}{14}$	$\frac{3}{14}$	$\frac{3}{20}$	$\frac{3}{5}$	$\frac{3}{20}$
a_5	$\frac{1}{18}$	$\frac{1}{18}$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{3}$	$\frac{3}{14}$	$\frac{3}{14}$	$\frac{3}{20}$	$\frac{3}{20}$	$\frac{3}{5}$

Fig. 1. The matrices of the conditional probabilities of Example 2. We use here the notation o, s to represent $o \wedge s$.

where $q = \min_{i,j,k}(p(s_k|o_j)/p(s_k|a_i))$.

Example 3. Consider again the instance of Crowds like in Example 2, but assume now that the preference probabilities are much higher than before, namely

$$p(s_k \mid a_i) = \begin{cases} \frac{9}{10} & a_i \in A_k \\ \frac{1}{10} & a_i \notin A_k \end{cases}$$

We can compute the probabilities $p(o \wedge s \,|\, a)$, $p(o \wedge s)$, $p(s \,|\, o)$ and $p(a \,|\, o \wedge s)$ like before. The results are shown in Fig. 2.

We note that in certain cases the extra knowledge dominates over the original observables. For instance, if the adversary receives a message from user 3 addressed to server 1, it is better for him to bet that a sender of group 1 is the originator, rather than user 3. In fact the *a posteriori* probability of the latter is $p(a_3 \,|\, o_3 \wedge s_1) = 1/6$ while the *a posteriori* probability of (say) user 1 is $p(a_1 \,|\, o_3 \wedge s_1) = 3/8$.

Using the values of $p(s_k \,|\, o_i)$ and $p(s_k \,|\, a_i)$, the factor $q = \min_{i,k}(p(s_k \,|\, o_i)/p(s_k \,|\, a_i))$ in Proposition 4 is evaluated to $2/3$, and we can see that Proposition 5 holds for this instance of Crowds, i.e. $\forall i, k.\ p(a_i \,|\, o_i \wedge s_k) \leq \alpha$ if and only if $\forall i.\ p(a_i \,|\, o_i) \leq q\,\alpha$. In fact $p(a_i \,|\, o_i) = 1/2$ and $\max_{i,j,k} p(a_i \,|\, o_j \wedge s_k) = 3/4$.

5 Vulnerability in presence of extra information

In this section we explore how the definition of α-vulnerability is affected by the presence of extra information. Let us start with the definition of α-vulnerability in presence of the new observables. It is natural to extend the notion of α-vulnerability by considering the (*a posteriori*) vulnerability when the observables are constituted by the joint random variables O, S, which is given by

$$V(A \,|\, O, S) = \sum_{j,k} p(o_j \wedge s_k)\ \max_i p(a_i \,|\, o_j \wedge s_k)$$

Hence we extend α-vulnerability as follows:

Definition 5 (α-vulnerability in presence of extra information). *Given a real number $\alpha \in [0, 1]$, a protocol satisfies α-vulnerability if and only if $V(A \,|\, O, S) \leq \alpha$.*

For the next proposition, we consider the specific case in which the protocol satisfies the symmetry of Crowds.

Proposition 6. *Let $\ell = |\mathcal{S}|$ denote the cardinality of the extra observables. Assume that, for each i, $p(o_i \,|\, a_i) = p = \max_{i,j} p(o_j \,|\, a_i)$ and let $q = \max_{i,k} p(s_k \,|\, a_i)$. We have:*

1. *$V(A \,|\, O, S) \leq \alpha$ if $V(A \,|\, O) \leq \ell\, q\, \alpha$.*
2. *If the* a priori *distribution is uniform and $\frac{(1-p)}{n-1} q \leq p \frac{(1-q)}{\ell-1}$, then $V(A \,|\, O, S) \leq \alpha$ if and only if $V(A \,|\, O) \leq \alpha$.*

Proof. By definition we have:

$$V(A \,|\, O, S) = \sum_{j,k} p(o_j \wedge s_k)\ \max_i p(a_i \,|\, o_j \wedge s_k)$$

Using Bayes theorem we derive:

$$V(A \,|\, O, S) = \sum_{j,k} \max_i (p(o_j \wedge s_k \,|\, a_i)\ p(a_i))$$

$p(o \mid a)$	o_1	o_2	o_3	o_4	o_5
a_1	$\frac{1}{2}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$
a_2	$\frac{1}{8}$	$\frac{1}{2}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$
a_3	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{2}$	$\frac{1}{8}$	$\frac{1}{8}$
a_4	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{2}$	$\frac{1}{8}$
a_5	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{2}$

$p(s \mid a)$	s_1	s_2
a_1	$\frac{9}{10}$	$\frac{1}{10}$
a_2	$\frac{9}{10}$	$\frac{1}{10}$
a_3	$\frac{1}{10}$	$\frac{9}{10}$
a_4	$\frac{1}{10}$	$\frac{9}{10}$
a_5	$\frac{1}{10}$	$\frac{9}{10}$

$p(s \mid o)$	s_1	s_2
o_1	$\frac{6}{10}$	$\frac{4}{10}$
o_2	$\frac{6}{10}$	$\frac{4}{10}$
o_3	$\frac{3}{10}$	$\frac{7}{10}$
o_4	$\frac{3}{10}$	$\frac{7}{10}$
o_5	$\frac{3}{10}$	$\frac{7}{10}$

$p(o, s \mid a)$	o_1, s_1	o_2, s_1	o_3, s_1	o_4, s_1	o_5, s_1	o_1, s_2	o_2, s_2	o_3, s_2	o_4, s_2	o_5, s_2
a_1	$\frac{9}{20}$	$\frac{9}{80}$	$\frac{9}{80}$	$\frac{9}{80}$	$\frac{9}{80}$	$\frac{1}{20}$	$\frac{1}{80}$	$\frac{1}{80}$	$\frac{1}{80}$	$\frac{1}{80}$
a_2	$\frac{9}{80}$	$\frac{9}{20}$	$\frac{9}{80}$	$\frac{9}{80}$	$\frac{9}{80}$	$\frac{1}{80}$	$\frac{1}{20}$	$\frac{1}{80}$	$\frac{1}{80}$	$\frac{1}{80}$
a_3	$\frac{1}{80}$	$\frac{1}{80}$	$\frac{1}{20}$	$\frac{1}{80}$	$\frac{1}{80}$	$\frac{9}{80}$	$\frac{9}{80}$	$\frac{9}{20}$	$\frac{9}{80}$	$\frac{9}{80}$
a_4	$\frac{1}{80}$	$\frac{1}{80}$	$\frac{1}{80}$	$\frac{1}{20}$	$\frac{1}{80}$	$\frac{9}{80}$	$\frac{9}{80}$	$\frac{9}{80}$	$\frac{9}{20}$	$\frac{9}{80}$
a_5	$\frac{1}{80}$	$\frac{1}{80}$	$\frac{1}{80}$	$\frac{1}{80}$	$\frac{1}{20}$	$\frac{9}{80}$	$\frac{9}{80}$	$\frac{9}{80}$	$\frac{9}{80}$	$\frac{9}{20}$

$p(o, s)$	o_1, s_1	o_2, s_1	o_3, s_1	o_4, s_1	o_5, s_1	o_1, s_2	o_2, s_2	o_3, s_2	o_4, s_2	o_5, s_2
	$\frac{6}{50}$	$\frac{6}{50}$	$\frac{3}{50}$	$\frac{3}{50}$	$\frac{3}{50}$	$\frac{4}{50}$	$\frac{4}{50}$	$\frac{7}{50}$	$\frac{7}{50}$	$\frac{7}{50}$

$p(a \mid o, s)$	o_1, s_1	o_2, s_1	o_3, s_1	o_4, s_1	o_5, s_1	o_1, s_2	o_2, s_2	o_3, s_2	o_4, s_2	o_5, s_2
a_1	$\frac{3}{4}$	$\frac{3}{16}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$	$\frac{1}{32}$	$\frac{1}{56}$	$\frac{1}{56}$	$\frac{1}{56}$
a_2	$\frac{3}{16}$	$\frac{3}{4}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{32}$	$\frac{1}{8}$	$\frac{1}{56}$	$\frac{1}{56}$	$\frac{1}{56}$
a_3	$\frac{1}{48}$	$\frac{1}{48}$	$\frac{1}{6}$	$\frac{1}{24}$	$\frac{1}{24}$	$\frac{9}{32}$	$\frac{9}{32}$	$\frac{9}{14}$	$\frac{9}{56}$	$\frac{9}{56}$
a_4	$\frac{1}{48}$	$\frac{1}{48}$	$\frac{1}{24}$	$\frac{1}{6}$	$\frac{1}{24}$	$\frac{9}{32}$	$\frac{9}{32}$	$\frac{9}{56}$	$\frac{9}{14}$	$\frac{9}{56}$
a_5	$\frac{1}{48}$	$\frac{1}{48}$	$\frac{1}{24}$	$\frac{1}{24}$	$\frac{1}{6}$	$\frac{9}{32}$	$\frac{9}{32}$	$\frac{9}{56}$	$\frac{9}{56}$	$\frac{9}{14}$

Fig. 2. The matrices of the conditional probabilities of Example 3. We use here the notation o, s to represent $o \wedge s$.

Because of the independence of O and S for any given originator, we deduce:

$$V(A \mid O, S) = \sum_{j,k} \max_i (p(o_j \mid a_i)\ p(s_k \mid a_i)\ p(a_i)) \tag{5}$$

1. Since $q = \max_{i,k} p(s_k \mid a_i)$, from (5) we derive:

$$V(A \mid O, S) \leq \sum_{j,k} \max_i (p(o_j \mid a_i)\, q\, p(a_i)) = \ell\, q \sum_j \max_i (p(o_j \mid a_i)\ p(a_i)) = \ell\, q\, V(A \mid O)$$

2. Since the input distribution is uniform:

$$V(A \mid O,S) = \frac{1}{n} \sum_{j,k} \max_i (p(o_j \mid a_i)\ p(s_k \mid a_i))$$

If $\frac{(1-p)}{n-1} q \leq p \frac{(1-q)}{\ell-1}$ then $\max_i (p(o_j \mid a_i)\ p(s_k \mid a_i)) = p(o_j \mid a_j)\ p(s_k \mid a_j) = p\, p(s_k \mid a_j)$. Hence

$$V(A \mid O,S) \ = \ \frac{1}{n} \sum_{j,k} p\ p(s_k \mid a_j) \ = \ \frac{1}{n} \sum_j p \ \sum_k p(s_k \mid a_j) \ = \ \frac{1}{n} \sum_j p \ = \ V(A|O)$$

□

It is interesting to note that, in the part (2) of Proposition 6, the extra knowledge does not make the protocol more vulnerable. This is because the additional knowledge is sometimes in accordance with the best guess based on the original observable, and sometimes in conflict, but the original observable always dominates, and therefore the additional knowledge is either redundant or disregarded. In any case, it is not used to make the guess. In the general case (represented by the first part of the proposition), however, the additional knowledge may dominate the original observable, and induce the adversary to change his bet, thus increasing his chances. For this reason, the vulnerability increases in general of a factor ℓq.

6 Conclusion

In this paper we focussed on the Crowds anonymity protocol and asked the question of how its existing analyses are affected by taking into account that attackers may have independent knowledge about users' behaviours. This amounts to providing the attackers with information about the correlation between a set of observables $s_1, \ldots, s_\ell$ and the event that user i is the originator of a message, as formalised by the conditional probability $p(s_k \mid a_i)$. We formalised the idea of probable innocence for such systems, both in standard terms and via the notion of protocol vulnerability, and identified a simple and neat measure of the impact of independent knowledge. Namely, it makes probable innocence (resp. vulnerability) more difficult to achieve by a factor q (resp. ℓq) which depends on the ratio between the probability of the observables conditional to the originator and conditional to the user detected (and, in the case of vulnerability, also from the cardinality of the random variable that represents the extra knowledge).

In conclusion, we remark that although the scenario in which attackers possess or can acquire extra knowledge is highly likely, it has so far been ignored. In the near future, we plan to work on the even more interesting scenario in which the attackers use their 'beliefs' about users behaviour to raise the vulnerability of anonymity protocols such as Crowds.

References

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

2. Christelle Braun, Konstantinos Chatzikokolakis, and Catuscia Palamidessi. Compositional methods for information-hiding. In Roberto M. Amadio, editor, *FoSSaCS*, volume 4962 of *Lecture Notes in Computer Science*, pages 443–457. Springer, 2008.
3. Konstantinos Chatzikokolakis and Catuscia Palamidessi. Probable innocence revisited. *Theor. Comput. Sci.*, 367(1-2):123–138, 2006.
4. Konstantinos Chatzikokolakis, Catuscia Palamidessi, and Prakash Panangaden. Probability of error in information-hiding protocols. In *CSF*, pages 341–354. IEEE Computer Society, 2007.
5. Konstantinos Chatzikokolakis, Catuscia Palamidessi, and Prakash Panangaden. Anonymity protocols as noisy channels. *Inf. Comput.*, 206(2-4):378–401, 2008.
6. Konstantinos Chatzikokolakis, Catuscia Palamidessi, and Prakash Panangaden. On the Bayes risk in information-hiding protocols. *Journal of Computer Security*, 16(5):531–571, 2008.
7. Han Chen and Pasquale Malacaria. Quantitative analysis of leakage for multi-threaded programs. In *PLAS '07: Proceedings of the 2007 workshop on Programming languages and analysis for security*, pages 31–40, New York, NY, USA, 2007. ACM.
8. David Clark, Sebastian Hunt, and Pasquale Malacaria. A static analysis for quantifying information flow in a simple imperative language. *Journal of Computer Security*, 15(3):321–371, 2007.
9. Michael R. Clarkson, Andrew C. Myers, and Fred B. Schneider. Belief in information flow. In *CSFW*, pages 31–45. IEEE Computer Society, 2005.
10. Yuxin Deng, Jun Pang, and Peng Wu 0002. Measuring anonymity with relative entropy. In Theodosis Dimitrakos, Fabio Martinelli, Peter Y. A. Ryan, and Steve A. Schneider, editors, *Formal Aspects in Security and Trust*, volume 4691 of *Lecture Notes in Computer Science*, pages 65–79. Springer, 2006.
11. Matthias Franz, Bernd Meyer, and Andreas Pashalidis. Attacking unlinkability: The importance of context. In Nikita Borisov and Philippe Golle, editors, *Privacy Enhancing Technologies*, volume 4776 of *Lecture Notes in Computer Science*, pages 1–16. Springer, 2007.
12. Joseph Y. Halpern and Kevin R. O'Neill. Anonymity and information hiding in multiagent systems. *Journal of Computer Security*, 13(3):483–512, 2005.
13. Joseph Y. Halpern and Kevin R. O'Neill. Anonymity and information hiding in multiagent systems. *Journal of Computer Security*, 13(3):483–512, 2005.
14. Boris Köpf and David A. Basin. An information-theoretic model for adaptive side-channel attacks. In Peng Ning, Sabrina De Capitani di Vimercati, and Paul F. Syverson, editors, *ACM Conference on Computer and Communications Security*, pages 286–296. ACM, 2007.
15. Pasquale Malacaria and Han Chen. Lagrange multipliers and maximum information leakage in different observational models. In Úlfar Erlingsson and Marco Pistoia, editors, *PLAS*, pages 135–146. ACM, 2008.
16. M. K. Reiter and A. D. Rubin. Crowds: Anonymity for web transactions. *ACM Transactions on Information and Systems Security*, 1(1):66–92, 1998.
17. Andrei Serjantov and George Danezis. Towards an information theoretic metric for anonymity. In Roger Dingledine and Paul F. Syverson, editors, *Privacy Enhancing Technologies*, volume 2482 of *Lecture Notes in Computer Science*, pages 41–53. Springer, 2002.
18. Vitaly Shmatikov and Ming-Hsiu Wang. Measuring relationship anonymity in mix networks. In Ari Juels and Marianne Winslett, editors, *WPES*, pages 59–62. ACM, 2006.
19. Geoffrey Smith. On the foundations of quantitative information flow. In Luca De Alfaro, editor, *Proceedings of the Twelfth International Conference on Foundations of Software Science and Computation Structures (FOSSACS 2009)*, volume 5504 of *Lecture Notes in Computer Science*, pages 288–302, York, UK, March 2009 2009. Springer.