Inference-proof Data Filtering for a Probabilistic Setting*

J. Biskup[1], P.A. Bonatti[2], and L. Sauro[2]

[1] Fakultät für Informatik, Technische Universität Dortmund
[2] Dip. Ing. Elet. e Tecnologie dell'Informazione, Università di Napoli Federico II

Abstract. In querying semantic data, access control must take into account the information that is implicitly entailed by the accessible part of triple stores and ontologies. While there exist inference control frameworks for this purpose, they still have a limitation: the confidentiality criterion does not take into account the probabilistic knowledge of the attacker. Therefore, the existing controlled query evaluation methods may return answers that actually reveal that a secret is true with very high probability. Given that such probabilistic knowledge is becoming more and more widely available as a result of analytics of various sorts, it is important to develop a refined confidentiality framework where probabilistic knowledge is taken into due account. Accordingly, in this paper, we extend and generalize an abstract data filtering framework for confidentiality-preserving, policy-based data publishing. The confidentiality requirement is strengthened so that the probability that a secret is true is bounded by a small constant ϵ. We formally define such a probabilistic setting, then we study two greedy data publishing methods based on refusals and lies, respectively. The refusal-based method is proved to be secure and maximally cooperative among a class of "reasonable" methods. We prove also that the natural generalization of the lying method is not secure. Furthermore, we extend the complexity hardness results from the deterministic framework to the probabilistic one.

Keywords: A priori knowledge, Confidentiality criterion, Confidentiality-preserving data publishing, Cooperativeness, Inference control, Lying, Privacy, Probabilistic methods, Refusal

1 Introduction

The need for inference-proof data publishing has been discovered well before the semantic web was born. Still, semantic (meta)data, that are expressed with standardized knowledge representation languages such as RDFS and OWL, are by design well-suited to the automated derivation of implied data. The applicable inference engines are becoming more and more powerful and scalable. This makes the need of inference-proof access control techniques particularly crucial in this area.

The literature is rich of *data filtering* methods for *confidentiality-preserving data publishing* (cf. [15]) that achieve secrecy properties of different kind and strength. In

* This work has been partially supported by the European Union's Horizon 2020 research and innovation programme under grant agreement N. 731601.

recent work [4], we have proposed an *abstract framework* that covers a wide range of deterministic filtering methods while not relying on any particular data model or query representation. Within this framework, we studied *possibilistic secrecy* regarding an observer-specific confidentiality *policy* which consists of (potential) *secrets* in form of yes–no queries.

For no secret query should a user be able to derive a "yes" answer, no matter how powerful is the rational reasoning method exploited by the user and no matter how many resources are employed. In this framework, confidentiality relies on the assumption that one counterexample suffices to prevent the user from believing that the secret is true.

However, reality may be different. The wide range of analytics tools for semantic data provide rich statistical information about the real world, that may make the available counterexamples unlikely. Moreover, if the majority of the interpretations of a knowledge base agree that a secret is true, the user may be inclined to believe it. To put this in general terms, a variety of probabilistic information may lead the user to believe that a secret is actually true *with high probability*. Then the aforementioned confidentiality framework should be refined by strengthening its confidentiality criterion so as to accommodate a priori probabilistic information about the domain of discourse.

Accordingly, in this paper, we generalize and extend the setting of [4] in the following ways. First, we consider probabilistic a priori knowledge in form of a *probability distribution* for the set of all possible data sources. Second, we *measure the information* regarding a secret query's validity learnt by observing filtered data and reasoning about it. This measurement is taken as the *conditional probability* of the secret query's validity under the observation. Third, we refine the secrecy criterion by requiring the observer to *believe in the truth of secrets with very low probability*, bounded by a suitable threshold ϵ.

Example 1. Let us consider a simple artificial situation where the set of data sources contains the 8 possible interpretations of 3 RDF triples p, q and r, represented by the words pqr, $pq\bar{r}$, $p\bar{q}r$, $p\bar{q}\bar{r}$, $\bar{p}qr$, $\bar{p}q\bar{r}$, $\bar{p}\bar{q}r$, $\bar{p}\bar{q}\bar{r}$. Intuitively speaking, the filtering should keep a joint validity of p and q confidential, formally expressed by the query $S = \{pqr, pq\bar{r}\}$, (the characteristic function of) which returns "yes" ($true$) for the two interpretations with representation $pq*$, and "no" ($false$) otherwise. Accordingly, the confidentiality policy is just the singleton $\{S\}$.

A straightforward filtering f hides the secret when applied to an interpretation by setting the first variable always to $\bar{p}$. Since in general we will allow uncertainty in a generated view, a filtering will return a set of data sources, and thus this filtering is formally defined by generating $f(xyz) := \{\bar{p}yz\}$. However, seeing the *verbatim* view $\{\bar{p}yz\}$, an intelligent and knowledgeable observer can easily determine the *inferred* view $\{pyz, \bar{p}yz\}$, each of whose elements might be the actual data source underlying the filtering. Nevertheless, since $\{pqr, pq\bar{r}\} \not\subseteq \{pyz, \bar{p}yz\}$, an observer will always believe in the possibility that the secret is *not* valid in the actual data source. For our probabilistic setting we are even more ambitious: we want the observer to believe in a probability of the non-validity of the secret not less than $1 - \epsilon$, for some security parameter $\epsilon \in [0, 1]$.

So let us further assume that the anticipated observer is knowing a priori the following probability distribution μ: each of the two protected interpretations occurs only

relatively rarely, with probability $\frac{1}{16}$; each of the two interpretations containing both $\bar{p}$ and $\bar{q}$ are more likely, having probability $\frac{3}{16}$; and each of the remaining four interpretations containing either p or q are equally distributed, with probability $\frac{2}{16}$. Then we have $\mu(S) = \mu(\{pqr, pq\bar{r}\}) = \frac{1}{16} + \frac{1}{16} = \frac{1}{8}$. By similar elementary calculations, for each case we can determine the a posteriori probability of the secret being valid under the condition of the inferred view based on the observation of the verbatim view. For the straightforward filtering f defined above and the assumed probability distribution μ the calculations show that these probabilities are either $\frac{1}{3}$ or 0, and thus the filtering is seen to be secure for each security parameter ϵ in the open interval $(\frac{1}{3}, 1]$.

This example is summarized in Table 1, and variants of it are serving throughout the paper.

Table 1. Summarized example, using the following notations: "$\rightarrow$" indicate that the interpretation d is mapped on the singleton containing the neighbored interpretation in the same line; "$\downarrow$" indicate that the interpretation d is mapped on the singleton containing itelf; [.] denotes the inferred view, formed as equivalence class consisting of the two interpretations in the respective line; $\mu(S \mid [.])$ denotes the derived conditional probability of the secret S under the inferred view [.]

A priori $\mu(d)$	**Interpretation** d	**Filtering** $f(d)$	**Filtering** $f(d)$	**Interpretation** d	**A priori** $\mu(d)$	**Inferred view** $\mu([.])$	**Secret fraction** $\mu(S \cap [.])$	**A posteriori** $\mu(S \mid [.])$
$\frac{1}{16}$	pqr	$\rightarrow$	$\downarrow$	$\bar{p}qr$	$\frac{2}{16}$	$\frac{3}{16}$	$\frac{1}{16}$	$\frac{1}{3}$
$\frac{1}{16}$	$pq\bar{r}$	$\rightarrow$	$\downarrow$	$\bar{p}q\bar{r}$	$\frac{2}{16}$	$\frac{3}{16}$	$\frac{1}{16}$	$\frac{1}{3}$
$\frac{2}{16}$	$p\bar{q}r$	$\rightarrow$	$\downarrow$	$\bar{p}\bar{q}r$	$\frac{3}{16}$	$\frac{5}{16}$	0	0
$\frac{2}{16}$	$p\bar{q}\bar{r}$	$\rightarrow$	$\downarrow$	$\bar{p}\bar{q}\bar{r}$	$\frac{3}{16}$	$\frac{5}{16}$	0	0

Popular probabilistic models and approaches have already been introduced. *Differential privacy* [11,12,13,14] is currently one of the most important of them. Under suitable assumptions it is very effective and relatively easy to apply. Unfortunately, when (probabilistic) background knowledge such as record correlations is available, differential privacy may fail to preserve confidentiality [20]. For this reason, investigating methods that preserve confidentiality in a probabilistic sense, in the presence of probabilistic background knowledge, is still an important topic. Moreover, differential privacy is based on the perturbation of data or query answers, while we intend to consider (also) methods that do not report incorrect answers.

Accordingly, in this paper we provide as a further contribution a maximally cooperative, secure query-answering method based on a greedy refusal-based algorithm. We prove also that – differently from the deterministic case – its lying-based analogue is *not* secure, instead. Finally, we extend some of the computational hardness results from the deterministic framework to the probabilistic one.

This abstract framework should be regarded as a preliminary, general feasibility study that constitutes the first step towards inference-proof, confidential access control methods for semantic data. Thus the implementation of concrete mechanism still lies beyond the scope of this paper.

The paper is organized as follows. After formally defining the probabilistic setting in Section 2, we analyze the safety of the greedy filtering methods based on *refusal* and *lying* in Section 3 and Section 4, respectively. In Section 5, we investigate the degree of cooperativeness of the greedy refusal method (a form of optimality). Section 6 reports the complexity results. Finally, in Section 7, we briefly summarize and evaluate our achievements, discuss related work, and list some challenging open problems.

2 A Probabilistic Setting

As in the deterministic case presented in [4], the probabilistic framework is based on an arbitrary set $\mathscr{D}$ of *data sources*, where each data source $d \in \mathscr{D}$ is treated as an abstract entity. To enable reasoning about the probability of sets of data sources, we need to introduce a *σ-algebra* $\Sigma_{\mathscr{D}} \subseteq \wp(\mathscr{D})$, where $\wp(\mathscr{D})$ denotes the powerset of $\mathscr{D}$. By definition, $\Sigma_{\mathscr{D}}$ includes both $\mathscr{D}$ and $\emptyset$ and is closed under the set operations of complement, countable union and countable intersection. Like for perfect encryption,we assume that the attacking observer has probabilistic *a priori knowledge* about the owner's data source. Such a knowledge is formalized by a *probability measure* $\mu : \Sigma_{\mathscr{D}} \longrightarrow [0, 1]$, that is any function which satisfies the following conditions: *(i)* $\mu(\mathscr{D}) = 1$ and *(ii)* if $A_1, A_2, \ldots \in \Sigma_{\mathscr{D}}$ are pairwise disjoint, then $\mu(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mu(A_i)$.[3]

We refer to [1] for a wider introduction on Probability Theory, here we just report a few properties which will be used later on:

- (complementation) $\mu(\overline{A}) = 1 - \mu(A)$, where $\overline{A}$ is the complement of A;
- (monotonicity) if $A \subseteq A'$, then $\mu(A) \leq \mu(A')$;
- ($\cap$-continuity) if $A_1, A_2, \ldots$ is a descending chain, i.e. $A_{i+1} \subseteq A_i$ for all i, then $\mu(\bigcap_{i=1}^{\infty} A_i) = \lim_i \mu(A_i)$.

Some subsets Q of $\mathscr{D}$ can be regarded as a Boolean *query*, indicating whether a data source d *satisfies* Q, i.e., $d \in Q$, or not, i.e., $d \notin Q$. To comply with the probabilistic framework, each query must be measurable. Let $\mathscr{B} \subseteq \Sigma_{\mathscr{D}}$ be the set of queries considered. Clearly, since queries have to be expressed syntactically, $\mathscr{B}$ is countable.

To keep some aspects of his data source secret to the observer, the owner can specify a *confidentiality policy* consisting of a finite set $\mathscr{S} \subseteq \mathscr{B}$ of queries, in this context called (*potential*) *secrets*. Informally speaking, if the owner's data source satisfies a secret then the observer should not be able to learn this fact; conversely, the observer is allowed to know that a secret is not satisfied.

To preserve secrets, the owner applies a *filtering* $f : \mathscr{D} \longrightarrow \wp(\mathscr{D}) \setminus \{\emptyset\}$ on his actual data source d and publishes $f(d)$ as a (*verbatim*) secure *view*, making the observer uncertain about which data source in $f(d)$ is the actual one, and possibly even misleading the observer by not including the actual data source d in $f(d)$. Assuming that the observer is a rational agent that knows the filtering f, the observer can compute the *inferred filtering* $[\cdot]_f : \mathscr{D} \longrightarrow \wp\mathscr{D} \setminus \{\emptyset\}$, defined by

$$[d]_f := \{ d' \mid f(d) = f(d') \}. \tag{1}$$

[3] Function μ is the analogue of the *data generation function* P of [20].

So, given the published verbatim view $f(d)$, the observer can construct the *inferred view* $[\,d\,]_f$, which may still be "uncertain" (i.e., there are multiple data sources) but it definitely contains the actual data source d. To enforce the confidentiality policy, the filtering should prevent the observer from inferring that a secret is satisfied. For instance, in a deterministic setting, the filtering should enforce $[\,d\,]_f \not\subseteq S$ for all secrets $S \in \mathscr{S}$.

Finally, to complete the probabilistic setting, we require that $[\,d\,]_f \in \Sigma_{\mathscr{D}}$, i.e., each set $[\,d\,]_f$ of data sources *indistinguishable* under f is supposed to be measurable. Then we strengthen the security criterion by requiring that the *probability* that a secret is satisfied by d – given the observable view $f(d)$ – should not exceed a (small) threshold ϵ.

Definition 1. *For a probabilistic setting with a priori knowledge $\mu : \Sigma_{\mathscr{D}} \longrightarrow [0,1]$, confidentiality policy $\mathscr{S}$, and threshold $\epsilon \in [0,1]$, a filtering $f : \mathscr{D} \longrightarrow \wp(\mathscr{D}) \setminus \{\emptyset\}$ with measurable inferred views is* $(\mu, \mathscr{S}, \epsilon)$-secure *iff for all data sources $d \in \mathscr{D}$, for all secrets $S \in \mathscr{S}$, it holds that*

$$\mu(S \cap [\,d\,]_f) \leq \epsilon \cdot \mu([\,d\,]_f)\,. \tag{2}$$

Note that inequality (2) means that, whenever defined (i.e., $\mu([\,d\,]_f) \neq 0$), the conditional probability of S under the inferred view $[d\,]_f$

$$\mu(\,S \,|\, [\,d\,]_f) := \frac{\mu(S \cap [\,d\,]_f)}{\mu([\,d\,]_f)}$$

is bounded by ϵ.

The abstract framework presented so far can be embodied in the context of Linked Data as follows: data sources consist of RDF-graphs, Boolean queries are represented by ASK forms of SPARQL[4] and secrets are specific Boolean queries, i.e. confidential pieces of information that can be retrieved through (a sequence of) ASK forms. Moreover, the a priori knowledge of the observer is represented by a probability distributions over possible RDF-graphs. More precisely, the set $\mathscr{D}$ is the set of all possible RDF-graphs over a given signature of interest and $\Sigma_{\mathscr{D}}$ is the whole powerset $\wp(\mathscr{D})$. Then, μ is determined by a discrete probability distribution δ over RDF-graphs such that for all $A \subseteq \mathscr{D}$,

$$\mu(A) = \sum_{d \in A} \delta(d)\,. \tag{3}$$

Finally, a Boolean query $Q_{\mathtt{a}}$ is the denotation of a ASK form `a`, that is the set of RDF-graphs where the form `a` returns `true`.

In the remainder of this paper, we will study the secure filterings f that result from a *greedy construction* of a decreasing sequence of sets $L_i(d)$ of data sources, for each input data source $d \in \mathscr{D}$:

$$\mathscr{D} =: L_0(d) \supseteq L_1(d) \supseteq \cdots \supseteq L_i(d) \supseteq \cdots \supseteq \bigcap_i L_i(d) := f(d)\,. \tag{4}$$

[4] Roughly speaking, an ASK form is a query that returns `true` if the RDF-graph satisfies a specified graph pattern, `false` otherwise. Graph patterns allow to verify, for example, whether a confidential set of RDF-triples or a specified node denoting some sensible individual occurs in the graph.

All constructions will be based on an exhaustive enumeration $\mathscr{B}_{en} = \langle Q_1, \ldots, Q_i, \ldots \rangle$ of all queries in $\mathscr{B}$. Stepwise, each query Q_i is submitted to a kind of stateful *censor* that keeps track of previous answers to determine whether Q_i should be answered correctly or *distorted*. This decision determines $L_i(d)$ from $L_{i-1}(d)$.

The sequence $\langle L_i(d) \rangle_i$ has also a *dynamic interpretation*. The observer chooses the queries $Q_1, Q_2, \ldots, Q_i, \ldots$ and submits them iteratively to the confidential data source. The query answering system returns for each Q_i the corresponding direct answer $A_i(d)$, leaving the computation of the accumulated information represented by $L_i(d)$ to the observer. In this dynamic setting, the confidentiality criterion is that at each step i, the accumulated information $L_i(d)$ of the direct answers to $Q_1, Q_2, \ldots, Q_i$ should not tell too much about the secrets.

3 Greedy Refusal

Under the *refusal* approach to data filtering, harmful queries are somehow explicitly *notified* to be hidden. In a dynamic query-answering environment, this may result in returning the special answer "mum" to a query whose correct answer would violate the confidentiality policy.

In the data filtering framework there is yet another interpretation of the refusal approach: some specific information about the actual data source is hidden by providing a view which contains not only the actual data source but also other data sources. If this generation of uncertainty is properly done, then the observer cannot know which element of the view is the actual one, therefore he is not able to infer that specific information.

A crucial point of the refusal approach is to block so-called *meta-inferences*. For instance, if a filtering f refused only the queries that entail a secret, then refused answers would *always* correspond to a secret satisfied by the actual data source d. The basic strategy for blocking this kind of attacks is to make the critical part of the censor decision independent from the actual data source. This strategy has already been proposed in the seminal work [27] on the refusal approach, further elaborated and refined for various models, e.g., [2,3,7,5], adopted for the abstract framework [4], and will also be exploited in the rest of this section for the probabilistic setting under consideration.

Using the notations of Definition 1 and instantiating (4), for an enumeration $\mathscr{B}_{en} = \langle Q_1, \ldots, Q_i, \ldots \rangle$ of all queries in $\mathscr{B}$, the (greedy) refusal filtering f_{re} results from the iterative application of the following distortion criterion to each query Q_i:

IF [censor criterion evaluating Q_i's harmfulness in a d-independent way]
for some secret $S \in \mathscr{S}$,
$\mu(\, S \cap\, L_{i-1}(d) \cap Q_i) > \epsilon \cdot \mu(\, L_{i-1}(d) \cap Q_i)$ or
$\mu(\, S \cap\, L_{i-1}(d) \cap \overline{Q_i}) > \epsilon \cdot \mu(\, L_{i-1}(d) \cap \overline{Q_i})$
THEN [distortion by refusal]
$L_i(d) := L_{i-1}(d)$
ELSE [honest answer w.r.t. d]
$A_i(d) :=$ IF $d \in Q_i$ THEN Q_i ELSE $\overline{Q_i}$
$L_i(d) := L_{i-1}(d) \cap A_i(d)$

Example 2. Resuming Example 1 with $\mu(S) = \mu(\{pqr, pq\bar{r}\}) = \frac{1}{16} + \frac{1}{16} = \frac{2}{16}$, we consider the refusal filtering for $d := pqr$. We choose $\epsilon := \frac{6}{16}$ as security parameter and take the powerset of the set of all interpretations over the propositional variables p, q and r as the set of all queries.

We start the enumeration of the 2^8 queries by "p?", intuitively asking whether p is valid, formalized as $Q_1 := \{pqr, pq\bar{r}, p\bar{q}r, p\bar{q}\bar{r}\}$ with $\mu(Q_1) = \frac{6}{16}$. Then $\overline{Q_1} := \{\bar{p}qr, \bar{p}q\bar{r}, \bar{p}\bar{q}r, \bar{p}\bar{q}\bar{r}\}$ with $\mu(\overline{Q_1}) = \frac{10}{16}$. Accordingly, $\mu(S \cap Q_1) = \mu(\{pqr, pq\bar{r}\}) = \frac{2}{16}$ and $\mu(S \cap \overline{Q_1}) = \mu(\emptyset) = 0$. Evaluating the censoring condition we get $\mu(S \,|\, Q_1) = \frac{2}{10} < \frac{6}{16}$ and $\mu(S \,|\, \overline{Q_1}) = 0 < \frac{6}{16}$, and thus the honest answer $A_1(pqr) := Q_1$ is processed to determine $L_1(d) := Q_1$.

We continue the enumeration with the query "q?", intuitively asking whether q is valid, formalized as $Q_2 := \{pqr, pq\bar{r}, \bar{p}qr, \bar{p}q\bar{r}\}$. Evaluating the censoring condition, we then get $\mu(L_1(d) \cap Q_2) = \mu(\{pqr, pq\bar{r}\}) = \frac{2}{16}$ and $\mu(S \cap (L_1(d) \cap Q_2)) = \mu(\{pqr, pq\bar{r}\}) = \frac{2}{16}$, and thus $\mu(S \cap (L_1(d) \cap Q_2) \,|(L_1(d) \cap Q_2)\,) = 1$ leads to a refusal resulting in $L_2(d) := L_1(d) = \{pqr, pq\bar{r}, p\bar{q}r, p\bar{q}\bar{r}\}$.

Further continuing with "r?", i.e., $Q_3 := \{pqr, p\bar{q}r, \bar{p}qr, \bar{p}\bar{q}r\}$, we get $\mu(L_2(d) \cap Q_3) = \mu(\{pqr, p\bar{q}r\}) = \frac{3}{16}$ and $\mu(S \cap (L_2(d) \cap Q_3)) = \mu(\{pqr\}) = \frac{1}{16}$ and thus $\mu(S \cap (L_2(d) \cap Q_3) \,|(L_2(d) \cap Q_3)\,) = \frac{1}{3} < \frac{6}{16}$ as well as $\mu(L_2(d) \cap \overline{Q_3}) = \mu(\{pq\bar{r}, p\bar{q}\bar{r}\}) = \frac{3}{16}$ and $\mu(S \cap (L_2(d) \cap \overline{Q_3})) = \mu(\{pq\bar{r}\}) = \frac{1}{16}$ and thus $\mu(S \cap (L_2(d) \cap \overline{Q_3}) \,|(L_2(d) \cap \overline{Q_3})\,) = \frac{1}{3} < \frac{6}{16}$, leading to an honest reaction such that $L_3(d) := L_2(d) \cap A_3(d) = \{pqr, p\bar{q}r\}$.

Though somehow tedious to check, all further queries do not change that intermediate view and, thus, for the actual data source pqr we get $\{pqr, p\bar{q}r\}$ as the final verbatim view, leaving an observer uncertain whether or not q is valid. One may note that in Example 1 we got the verbatim view $\{pqr, \bar{p}qr\}$ leaving the status of p open; we would obtain this result by the refusal filtering if we exchanged the processing of the first two queries.

Clearly, in the context of RDF-graphs considered above, the presented approach requires that an ASK form returns a notification of the kind `query_refused` in case of refusal.

Example 3. Consider an RDF-graph containing sensible information about medical treatments in a small town of 10 000 citizens. For privacy reasons, we want to enforce anonymity by preventing the identification of any specific person with a confidence larger than 0.5. Then, Oreste Galli, who actually occurs in the data source, wants to verify that the refusal framework works properly by maliciously querying the system about himself. In particular, he considers two queries, the former asks whether there exists an individual in the graph whose name is Oreste, the latter asks for the presence of an individual who has the same address Machiavelli Street. The previous queries may look like:

$$Q_1 = \texttt{ASK}(\texttt{?x} \quad \texttt{ex:given\_name} \quad \text{“}\texttt{Oreste}\text{”})$$

$$Q_2 = \texttt{ASK}(\texttt{?x} \quad \texttt{ex:street\_address} \quad \text{“}\texttt{Machiavelli}\text{”})$$

The a priori knowledge is that (i) the data source contains the 1% of the whole population, (ii) 10 people are called Oreste, and (iii) 15 people live in Machiavelli Street.

Then, by the applying the Bayes' theorem, the conditional probability to infer the secret S that Oreste Galli occurs in the data source by querying Q_1 is given by

$$\mu(S \,|\, Q_1) = \frac{\mu(Q_1 \,|\, S) \cdot \mu(S)}{\mu(Q_1)} \,.$$

Note that the fact that Oreste Galli occurs in the graph implies that there exists a person called Oreste; in other words, since $S \subset Q_1$, we have that $\mu(Q_1|S) = 1$. Furthermore, $\mu(S)$ is the prior probability that Oreste Galli occurs in the graph, which is equal to the percentage of people occurring in the graph, i.e. 0.01. Finally, $\mu(Q_1)$ is the probability that, given a generic graph containing the 1% of the whole population, there exists at least one person in the graph called Oreste. This is given by the formula:

$$\mu(Q_1) = \frac{N - K}{N} \,,$$

where $N = \binom{10\,000}{100}$ is the number of all possible graphs containing the 1% of the whole population, and K is the number of possible graphs where no Oreste occurs. Since only 10 people over 10 000 are called Oreste, $K = \binom{9\,990}{100}$. Then, we straightforwardly have that $\mu(S \,|\, Q_1) = 0.104$.

Note that, since $S \subseteq Q_1$, it follows that $\mu(S \,|\, \overline{Q_1})$ is equal to zero. Consequently, since $\mu(S \,|\, Q_1) < 0.5$, the censor correctly answers to Q_1. Then, the user queries Q_2. By using again the Bayes' theorem, the censor estimates the conditional probability of S given Q_2, provided that Q_1 has been already answered and hence the prior probability is 0.104 instead of 0.01. By similar calculation, the resulting probability is 0.74, subsequently the censor refuses Q_2.

Theorem 1. *Let $\mathscr{B}_{en}$ be an enumeration of $\mathscr{B}$. For each μ, $\mathscr{S}$ and $\epsilon \in [0, 1]$, f_{re} is filtering function with measurable inferred views; moreover, f_{re} is $(\mu, \mathscr{S}, \epsilon)$-*secure*, provided the following precondition holds:*

$$\mu(S) = \frac{\mu(S \cap L_0(d))}{\mu(L_0(d))} \leq \epsilon \,, \textit{ for all } S \in \mathscr{S} \,.^5$$

Proof. We first prove the following sub-statements.

Fact 1: $f_{re}(d) \subseteq [d]_{f_{re}}$. Assume that $d' \in f_{re}(d)$, we will show by induction on the iterative construction of f_{re} that $f_{re}(d') = f_{re}(d)$; consequently, $d' \in [d]_{f_{re}}$. Since $L_0(d') = \mathscr{D} = L_0(d)$, the base case $i = 0$ trivially holds. Let $i > 0$ and assume by induction hypothesis that $L_{i-1}(d) = L_{i-1}(d')$. Clearly, then for all $S \in \mathscr{S}$, both $\mu(S \mid L_{i-1}(d) \cap Q_i) = \mu(S \mid L_{i-1}(d') \cap Q_i)$ and $\mu(S \mid L_{i-1}(d) \cap \overline{Q_i}) = \mu(S \mid L_{i-1}(d') \cap \overline{Q_i})$; this means that the refusal behavior of the censor to the query Q_i will be the same for both d and d'. Moreover, since $d' \in f_{re}(d)$, we also have that $d' \in A_i(d)$ and, hence, $A_i(d) = A_i(d')$. Consequently, $L_i(d) = L_i(d')$.

Fact 2: $d \in f_{re}(d)$. Again, by induction on the iterative construction of f_{re}, the base case, $d \in L_0(d) = \mathscr{D}$, holds. Furthermore, assume that $d \in L_{i-1}(d)$, where $i > 0$.

[5] This precondition is required since no secret which is violated *ex ante* can be protected by any filtering.

If the censor refuses Q_i, then $L_i(d) = L_{i-1}(d)$ and the induction hypothesis directly implies that $d \in L_i(d)$. Otherwise, $L_i(d) = L_{i-1}(d) \cap A_i(d)$. Since by induction hypothesis $d \in L_{i-1}(d)$ and by construction $d \in A_i(d)$, also in this case $d \in L_i(d)$.

Fact 3: $[d]_{f_{re}} \subseteq f_{re}(d)$. Assume $d' \in [\,d\,]_{f_{re}}$, i.e., $f_{re}(d) = f_{re}(d')$. By Fact 2, $d' \in f_{re}(d')$ so we have $d' \in f_{re}(d)$.

That f_{re} is a filtering function, i.e. $f_{re}(d) \neq \emptyset$ for all $d \in \mathscr{D}$, directly follows from Fact 2. Moreover, by Fact 1 and 3, $f_{re}(d) = [d]_{f_{re}}$. Note that, since $\mathscr{B} \subseteq \Sigma_{\mathscr{D}}$ and $f_{re}(d)$ consists of a countable intersection of elements of $\mathscr{B}$ or their complements, $f_{re}(d) \in \Sigma_{\mathscr{D}}$ and hence $[d]_{f_{re}} \in \Sigma_{\mathscr{D}}$ too. Finally, the statement's precondition and the definition of f_{re} immediately imply the following invariant:

$$\mu(S \cap L_i(d)) \leq \epsilon \cdot \mu(L_i(d))\,, \text{ for all } i \text{ and for all } S \in \mathscr{S}\,.$$

Hence, taking the limits, it holds that:

$$\lim_i \mu(S \cap L_i(d)) \leq \epsilon \cdot \lim_i \mu(L_i(d))\,.$$

On the other hand, by the $\cap$-continuity of μ, $\mu(f_{re}(d)) = \mu(\bigcap_i L_i(d)) = \lim_i \mu(L_i(d))$ and $\mu(S \cap f_{re}(d)) = \mu(S \cap \bigcap_i L_i(d)) = \mu(\bigcap_i (S \cap L_i(d))) = \lim_i \mu(S \cap L_i(d))$. Putting all together, we have that

$$\mu(S \cap f_{re}(d)) \leq \epsilon \cdot \mu(f_{re}(d))\,.$$

The theorem immediately follows by reminding that $f_{re}(d) = [d]_{f_{re}}$. □

Remark 1. The above theorem and its proof show that the refusal approach is *inference-proof* in a very strong sense. Namely, as $f_{re}(d)$ always coincides with $[d]_{f_{re}}$, inferring does not provide any information beyond what is immediately visible from the verbatim view. Moreover, the two views contains only limited information about the secrets, as specified by the security parameter ϵ.

4 Greedy Lying

Under the *lying* approach to data filtering a harmful piece of data is implicitly *distorted*, without any notification, by saying that it is false. In a dynamic query-answering environment, such a distortion consists in returning the complement of the correct answer to a query Q whenever the correct answer violates the confidentiality policy. In the framework of data filterings, lying aims at providing a view that contains a single data source (if necessary, different from the actual one) where all the secrets are false.

A crucial point of the lying approach is to make a current answer *consistent* with previously accumulated knowledge and to prepare for being able to provide definite (though possibly untrue) answers to any further queries. The basic strategy for achieving these goals is to always protect the *disjunction* of all secrets (in the abstract framework, their *union*). This strategy has already been proposed in the seminal work [9] on the lying approach, further elaborated and refined for various models, e.g., [2,3,7,8], adopted for the abstract framework [4], and will also be tentatively exploited next in the probabilistic setting.

Interestingly, the *natural* probabilistic generalization of the deterministic greedy lying approach defined in [4] is *not* secure, as shown in the following.

Using the notations of Definition 1 and instantiating (4), for an enumeration $\mathscr{B}_{en} = \langle Q_1, \ldots, Q_i, \ldots \rangle$ of queries in $\mathscr{B}$, the natural probabilistic generalization of the (greedy) lying filtering f_{ly} is obtained by iteratively applying the following distortion mechanism to all queries Q_i ($i = 1, 2, \ldots$):

$A_i(d) :=$ IF $d \in Q_i$ THEN Q_i ELSE $\overline{Q_i}$
IF [censoring by checking harmfulness of the correct (d-dependent) answer]
$\mu(\bigcup \mathscr{S} \cap L_{i-1}(d) \cap A_i(d)) > \epsilon \cdot \mu(L_{i-1}(d) \cap A_i(d))$
THEN [distortion by lying]
$L_i(d) := L_{i-1}(d) \cap \overline{A_i(d)}$
ELSE [honest answer w.r.t. d]
$L_i(d) := L_{i-1}(d) \cap A_i(d)$

The precondition for applying the greedy lying method, by analogy with the deterministic case, should be:

$$\mu(\bigcup \mathscr{S}) \leq \epsilon.$$

Proposition 1. *There exist μ, $\mathscr{S}$ and ϵ for which the above precondition is satisfied but f_{ly} is not $(\mu, \mathscr{S}, \epsilon)$*-secure.

Proof. Let us consider the following setting. The set of data sources $\mathscr{D} = \{d_1, d_2, d_3\}$. The a priori knowledge for d_i is described by the following measures: $\mu(d_1) = 0.7$, $\mu(d_2) = 0.2$ and $\mu(d_3) = 0.1$. The confidentiality policy $\mathscr{S}$ contains only the secret $S = \{d_2\}$, i.e., $\mathscr{S} = \{\{d_2\}\}$. The value of the threshold is given by $\epsilon = 0.5$. The set of queries $\mathscr{B} = \wp(\mathscr{D})$. We employ the following enumeration for the queries $\mathscr{B}_{en} = \langle Q_1, \ldots, Q_8 \rangle$, where the Q_i are defined as follows: $Q_1 = \{d_2\} = S$, $Q_2 = \{d_2, d_3\}$, $Q_3 = \emptyset$, $Q_4 = \{d_1\}$, $Q_5 = \{d_3\}$, $Q_6 = \{d_1, d_2\}$, $Q_7 = \{d_1, d_3\}$, $Q_8 = \{d_1, d_2, d_3\}$.

Let us then examine the execution of the filtering f_{ly} as defined above. We first show the computation for the data source d_2. We start by $L_0(d_2) = \mathscr{D} = \{d_1, d_2, d_3\}$. The answers to the queries are computed as follows, where – in order to easy the presentation – we will use the notation $\mu_i(d) = \mu(\bigcup \mathscr{S} \,|\, L_{i-1}(d) \cap A_i(d)) = \mu(\{d_2\} \,|\, L_{i-1}(d) \cap A_i(d))$:

- $Q_1 = \{d_2\}$: Since $d_2 \in Q_1$, it follows that $A_1(d_2) = Q_1 = \{d_2\}$ and, thus, $L_0(d_2) \cap A_1(d_2) = \{d_2\}$. We obtain that $\mu_1(d_2) = \mu(\bigcup \mathscr{S} \,|\, L_0(d_2) \cap A_1(d_2)) = \mu(\{d_2\} \,|\{d_2\}) = 1 \geq \epsilon = 0.5$. In this case, the filter lets

$$L_1(d_2) = L_0(d_2) \cap \overline{A_1(d_2)} = \{d_1, d_2, d_3\} \cap \{d_1, d_3\} = \{d_1, d_3\}\,.$$

- $Q_2 = \{d_2, d_3\}$: Since $d_2 \in Q_2$, it follows that $A_2(d_2) = Q_2 = \{d_2, d_3\}$ and, thus, $L_1(d_2) \cap A_2(d_2) = \{d_3\}$. We obtain that $\mu_2(d_2) = \mu(\{d_2\} \,|\{d_3\}) = 0$ and, thus,

$$L_2(d_2) = L_1(d_2) \cap A_2(d_2) = \{d_1, d_3\} \cap \{d_2, d_3\} = \{d_3\}\,.$$

– $Q_3 = \emptyset$: Since $d_2 \notin Q_3$, it follows that $A_3(d_2) = \overline{Q_3} = \{d_1, d_2, d_3\}$ and, thus, $L_2(d_2) \cap A_3(d_2) = \{d_3\}$. As before, $\mu_3(d_2) = \mu(\{d_2\} \,|\{d_3\}) = 0$ and, thus,

$$L_3(d_2) = L_2(d_2) \cap A_3(d_2) = \{d_3\} \cap \{d_1, d_2, d_3\} = \{d_3\}\,.$$

– $Q_4 = \{d_1\}$: Since $d_2 \notin Q_4$, it follows that $A_4(d_2) = \overline{Q_4} = \{d_2, d_3\}$ and, thus, $L_3(d_2) \cap A_4(d_2) = \{d_3\}$. As before, $\mu_4(d_2) = \mu(\{d_2\} \,|\{d_3\}) = 0$ and, thus,

$$L_4(d_2) = L_3(d_2) \cap A_4(d_2) = \{d_3\} \cap \{d_2, d_3\} = \{d_3\}\,.$$

– $Q_5 = \{d_3\}$: Since $d_2 \notin Q_5$, it follows that $A_5(d_2) = \overline{Q_5} = \{d_1, d_2\}$ and, thus, $L_4(d_2) \cap A_5(d_2) = \emptyset$. In this case the filter computes

$$L_5(d_2) = \emptyset\,.$$

The complete construction of f_{ly} is summarized in Table 2.

Table 2. Construction of f_{ly}

	d_1	d_2	d_3
$L_0()$	$\{d_1, d_2, d_3\}$	$\{d_1, d_2, d_3\}$	$\{d_1, d_2, d_3\}$
$L_1()$	$\{d_1, d_3\}$	$\{d_1, d_3\}$	$\{d_1, d_3\}$
$L_2()$	$\{d_1\}$	$\{d_3\}$	$\{d_3\}$
$L_3()$	$\{d_1\}$	$\{d_3\}$	$\{d_3\}$
$L_4()$	$\{d_1\}$	$\{d_3\}$	$\{d_3\}$
$L_5()$	$\{d_1\}$	$\emptyset$	$\emptyset$
$L_6()$	$\{d_1\}$	$\emptyset$	$\emptyset$
$L_7()$	$\{d_1\}$	$\emptyset$	$\emptyset$
$L_8()$	$\{d_1\}$	$\emptyset$	$\emptyset$

Given the above table, we can write that $[d_1]_{f_{ly}} = \{d_1\}$ and $[d_2]_{f_{ly}} = [d_3]_{f_{ly}} = \{d_2, d_3\}$. This means that $\mu(\bigcup \mathscr{S} \mid [d_2]_{f_{ly}}) = \mu(\{d_2\} \mid \{d_2, d_3\}) = \frac{2}{3} > \epsilon$, thus the proposition follows. □

Proposition 1 shows that the natural probabilistic generalization of the greedy lying approach is not secure. The existence of secure probabilistic variants of greedy lying is left as an open problem.

5 Cooperativeness

In this section we show that, under some natural assumptions, the greedy method based on refusals is *maximally cooperative*, that is, it hides a minimal amount of information. Maximal cooperativeness is formalized in [4] as follows:

Definition 2 ([4]). *A filtering f is* more cooperative *than a filtering g iff for all $d \in \mathscr{D}$, $[d]_f \subseteq [d]_g$. If f_1 is more cooperative than f_2 then we write $f_1 \succeq f_2$. If $f_1 \succeq f_2$ and $f_2 \not\succeq f_1$, then we write $f_1 \succ f_2$.*

Informally speaking, if f is more cooperative than g then f systematically refuses to answer less queries than g (or the same queries as g), because the partition induced by f is finer than (or equal to) g's. Maximally cooperative filterings are called *optimal*.

Definition 3 ([4]). *A secure filtering f is* optimal *iff there exists no secure filtering f' such that $f' \succ f$.*

Currently, we do not know whether the greedy refusal filtering f_{re} is optimal in this strong sense. It is difficult to compare f_{re} with arbitrary filterings g because g might exploit partitions (i.e. secure views) that cannot be denoted with the query language (while f_{re}, by construction, can only exploit the expressive power of $\mathscr{B}$). What we are going to prove is that f_{re} is maximally cooperative among the filterings whose views can be defined using $\mathscr{B}$. This class of filtering is formally defined as follows:

Definition 4. *A filtering g is* query-based *iff for all $d \in \mathscr{D}$, there exists a query $Q \in \mathscr{B}$ such that $[d]_g = Q$.*

To prove the optimality of f_{re} with respect to query-based filterings, we adopt some mild restrictions. We assume that data sources are countable:

$$\mathscr{D} = \{d_1, d_2, \ldots, d_i, \ldots\}.$$

For example, this is true of knowledge bases, as well as database tables whose values range over countable domains such as strings, integers and rational numbers. A property of this discrete framework is that all (possibly infinite) collections of pairwise disjoint subsets of $\mathscr{D}$ are countable.[6] The cooperativeness theorem is now formalized as follows:

Theorem 2. *If $\mathscr{D}$ is countable, then there exists no $(\mu, \mathscr{S}, \epsilon)$-secure query-based filtering g such that $g \succ f_{re}$.*

Proof. In the following, for the sake of readability, we abbreviate f_{re} with f. Suppose the theorem does not hold (we will derive a contradiction). Then there exists a $(\mu, \mathscr{S}, \epsilon)$-secure query-based filtering $g \succ f$. This means that for all $d \in \mathscr{D}$, $[d]_g \subseteq [d]_f$ and for some $d_0 \in \mathscr{D}$, $[d_0]_g \subset [d_0]_f$.

Since g is query-based, for some step k of the greedy construction the query Q_k satisfies $Q_k = [d_0]_g$. Note that in f's construction, Q_k is refused (otherwise $[d_0]_f \subseteq [d_0]_g$ would hold, which is a contradiction). This means that for some secret $S_0 \in \mathscr{S}$, either

$$\mu(S_0 \cap L_{k-1}(d_0) \cap Q_k) > \epsilon \cdot \mu(L_{k-1}(d_0) \cap Q_k), \text{ or} \tag{5}$$

$$\mu(S_0 \cap L_{k-1}(d_0) \cap \overline{Q_k}) > \epsilon \cdot \mu(L_{k-1}(d_0) \cap \overline{Q_k}). \tag{6}$$

Note that $L_{k-1}(d_0) \supseteq [d_0]_f \supset [d_0]_g = Q_k$, therefore (5) entails $\mu(S_0 \cap [d_0]_g) > \epsilon \cdot \mu([d_0]_g)$. This disequality cannot hold, because g is $(\mu, \mathscr{S}, \epsilon)$-secure by assumption, so (5) does not hold, and hence (6) must hold.

Finally, we prove that the security of g implies that (6) should *not* hold, instead, which proves the theorem.

[6] To see this, associate each $X \subseteq \mathscr{D}$ in the collection to the integer $\min\{i \mid d_i \in X\}$.

Let $\mathscr{C} = \{[d]_g \mid d \in L_{k-1}(d_0) \cap \overline{Q_k}\}$. Recall that $\mathscr{C}$ can be enumerated, because its elements are pairwise disjoint: so let $\mathscr{C} = \{X_1, X_2, \ldots, X_i, \ldots\}$, where each X_i is an equivalence class induced by g.

Since g is $(\mu, \mathscr{S}, \epsilon)$-secure, for all $X_i \in \mathscr{C}$ we have

$$\mu(S_0 \cap X_i) \leq \epsilon \cdot \mu(X_i). \tag{7}$$

It follows that

$$\begin{aligned}
\mu(S_0 \cap L_{k-1}(d_0) \cap \overline{Q}_k) &= \mu(S_0 \cap \bigcup \mathscr{C}) \\
&= \sum_{i=1}^{\infty} \mu(S_0 \cap X_i) \\
&\leq \sum_{i=1}^{\infty} \epsilon \cdot \mu(X_i) \qquad \text{by (7)} \\
&= \epsilon \cdot \mu(\bigcup \mathscr{C}) \\
&= \epsilon \cdot \mu(\, L_{k-1}(d_0) \cap \overline{Q_k}\,)\,.
\end{aligned}$$

This contradicts (6). □

6 Computational Complexity

The computational complexity of the non-probabilistic framework has been studied in [4] using finite $\mathscr{D}$ (the abstract analogue of propositional logic frameworks). For such frameworks, it can be assumed that all subsets of $\mathscr{D}$ are measurable and that $\mathscr{B} = \wp(\mathscr{D})$. Such finite non-probabilistic frameworks can be seen as a special case of the probabilistic framework: For all $X \subseteq \mathscr{D}$, let

$$\mu_0(X) = \frac{|X|}{|\mathscr{D}|} \quad \text{and} \quad \epsilon_0 = 1 - \frac{1}{|\mathscr{D}|}\,.$$

Then it can be proved that:

Proposition 2. *For all $\mathscr{S} \subseteq \mathscr{B}$, a filtering is $(\mu_0, \mathscr{S}, \epsilon_0)$-secure iff it is secure in the sense of [4].*

Thanks to this proposition, all the computational hardness results of [4] can immediately be extended to the probabilistic framework. In particular, for any given μ, $\mathscr{S}$, and ϵ:

- (Optimality checking) Deciding whether a given filtering f is an optimal $(\mu, \mathscr{S}, \epsilon)$-secure filtering is coNP-hard.
- (Query availability) Given a data source $d \in \mathscr{D}$ and a query of interest $Q \in \mathscr{B}$, deciding whether there exists a $(\mu, \mathscr{S}, \epsilon)$-secure filtering f that preserves Q on d (that is, $[d]_f \subseteq Q$ iff $d \in Q$) is NP-complete.

7 Conclusions and related work

Extending recent work [4], we presented an abstract *probabilistic* setting for data *filterings* to achieve *confidentiality* according to a *policy*, in the presence of *probabilistic background knowledge*. The *security property* is probabilistic, too: a rational and knowledgeable observer should not be able to believe in the truth of a secret with a probability larger than a given parameter ϵ.

We proved that a natural, probabilistic generalization of the *refusal-based* controlled query evaluation approach is both secure and maximally cooperative among a class of "reasonable" filterings, that define their secure views using the query language. Interestingly, the corresponding *lying* approach is not secure (differently from the deterministic framework). Finally, computational hardness results for some decision problems have been extended from the deterministic case to the probabilistic case.

The security criterion can be easily modified by associating a different threshold ϵ_i to each secret S_i, and adapting the precondition of Theorem 1 accordingly. In this way one can simulate relativistic privacy preservation requirements, where the inferred probability of each secret S remains close to the prior probability $\mu(S)$ of S, just by defining the thresholds as $\epsilon_i = \mu(S_i) + \tilde{\epsilon}_i$. Then the precondition of Theorem 1 is satisfied by the definitions. Proofs do not require any significant changes.

Our work continues a long line of research on imposing confidentiality constraints on computing system, including logic-oriented information systems used for query answering and data publishing. The research line on refusals and controlled interaction execution has been devoted to non-probabilistic methods, so far [27,3,7,6,5]. Several of these papers additionally and [9,8] explicitly deal with lies in non-probabilistic settings.

Similarly, the *non-interference* property for inference control in non-logical, operational settings has originally not been probabilistic [16], but some later and recent work also considers probabilities, e.g. [17,26,10,19,29], in particular to measure the (average, entropy-based) information flow from security high (secret) inputs to security low (open) outputs. Moreover, formalized in a more general framework of value transmission over a randomized channel, based on [10] the authors of [19] emphasize that the actual probabilities might differ from those believed by the attacker, and they then study both the (decrease of) belief-uncertainty – defined as min-entropy of belief vulnerability – and the impact of the posterior inaccuracy of the attacker's belief. These studies, however, leave constructive methods to minimize secrecy-violating effects to future work. Also inspired by [10] – and rediscovering fundamental insight about refusals [27] –, now within a framework of procedural programs whose execution semantics are seen as a transformation of probabilities of states over numeric variables, the authors of [23] propose to employ efficiently updatable probabilistic polyhedra to approximate the currently assumed probabilistic knowledge of the attacker.

Other popular methods, such as k-anonymity with l-diversity [25,28,22], are partly covered by the deterministic abstract framework [4]. In some sense, though, k-anonymity lies in between possibilistic and probabilistic methods, since its first confidentiality criterion is aimed at ensuring a *sufficient number* of alternative states of the world (governed by the k parameter) to prevent re-identification. The approach of weakened relational views [5] pursues a closely related goal within a more general setting.

Moreover, probabilistic refinements of l-diversity of values associated with a k-block of worlds attempt to let each of the associated values appear to be equally plausible.

Halpern and O'Neill studied probabilistic secrecy in the context of dynamic systems [18]. Their approach is based on a modal logic of beliefs whose models encode a streamlined account of system runs. Our technical analysis differs in several respects. Their focus is on the definition and logical characterization of security properties; they do not show how to develop secure systems, nor do they deal with cooperativeness. Moreover, they do not estimate the complexity of achieving security.

The methods based on lies (which have been extensively studied in the above non-probabilistic contexts) can be regarded as the mainstream method in probabilistic methods, in the form of random answer perturbations [15]. Concerning differential privacy [11,12,13,14], the main difference from our framework are two: (i) The differential privacy model does not embody prior knowledge such as record correlation, that may affect confidentiality [20]; in our model the a priori knowledge modeled by μ is used to ensure confidentiality even in the presence of record correlations and the like. (ii) The filterings for differential privacy are probabilistic while our filterings are not. Extending our framework to probabilistic filterings is an interesting topic for further research.

Recent more thorough comparisons and discussion of various probabilistic approaches to enforce confidentiality of sensitive information in provider-consumer interactions are provided by [21] and [24].

The challenges for future work include dealing with the inherent complexity of some decision problems of interest (see, e.g., [4,29]), working safely with approximate estimates of μ (see, e.g., [10,19,14]), and designing efficient implementations of the secure methods (see, e.g., [23]).

References

1. P. Billingsley. *Probability and Measure, 3rd edition*. John Wiley & Sons, New York, NY, USA, 1995.
2. J. Biskup and P. A. Bonatti. Lying versus refusal for known potential secrets. *Data Knowl. Eng.*, 38(2):199–222, 2001.
3. J. Biskup and P. A. Bonatti. Controlled query evaluation for enforcing confidentiality in complete information systems. *Int. J. Inf. Sec.*, 3(1):14–27, 2004.
4. J. Biskup, P. A. Bonatti, C. Galdi, and L. Sauro. Optimality and complexity of inference-proof data filtering and CQE. In M. Kutylowski and J. Vaidya, editors, *European Symposium on Research in Computer Security, ESORICS 2014, Part II*, volume 8713 of *Lecture Notes in Computer Science*, pages 165–181. Springer, 2014.
5. J. Biskup and M. Preuß. Information control by policy-based relational weakening templates. In I. G. Askoxylakis, S. Ioannidis, S. K. Katsikas, and C. A. Meadows, editors, *Computer Security - ESORICS 2016 - 21st European Symposium on Research in Computer Security, Part II*, volume 9879 of *Lecture Notes in Computer Science*, pages 361–381. Springer, 2016.
6. J. Biskup and C. Tadros. Preserving confidentiality while reacting on iterated queries and belief revisions. *Ann. Math. Artif. Intell.*, 73(1-2):75–123, 2015.
7. J. Biskup and T. Weibert. Keeping secrets in incomplete databases. *Int. J. Inf. Sec.*, 7(3):199–217, 2008.
8. J. Biskup and L. Wiese. A sound and complete model-generation procedure for consistent and confidentiality-preserving databases. *Theoretical Computer Science*, 412:4044–4072, 2011.

9. P. A. Bonatti, S. Kraus, and V. S. Subrahmanian. Foundations of secure deductive databases. *IEEE Trans. Knowl. Data Eng.*, 7(3):406–422, 1995.
10. M. R. Clarkson, A. C. Myers, and F. B. Schneider. Quantifying information flow with beliefs. *Journal of Computer Security*, 17(5):655–701, 2009.
11. I. Dinur and K. Nissim. Revealing information while preserving privacy. In F. Neven, C. Beeri, and T. Milo, editors, *22nd ACM SIGACT-SIGMOD-SIGART Symposium on Principles of Database Systems, PODS 2003*, pages 202–210. ACM, 2003.
12. C. Dwork. Differential privacy. In M. Bugliesi, B. Preneel, V. Sassone, and I. Wegener, editors, *Automata, Languages and Programming, 33rd International Colloquium, ICALP 2006, Part II*, volume 4052 of *Lecture Notes in Computer Science*, pages 1–12. Springer, 2006.
13. C. Dwork. Differential privacy: A survey of results. In M. Agrawal, D. Du, Z. Duan, and A. Li, editors, *Theory and Applications of Models of Computation, 5th International Conference, TAMC 2008*, volume 4978 of *Lecture Notes in Computer Science*, pages 1–19. Springer, 2008.
14. C. Dwork and G. N. Rothblum. Concentrated differential privacy. *CoRR*, abs/1603.01887, 2016.
15. B. C. M. Fung, K. Wang, A. W.-C. Fu, and P. S. Yu. *Introduction to Privacy-Preserving Data Publishing – Concepts and Techniques*. Chapman & Hall/CRC, Boca Raton, FL, 2010.
16. J. A. Goguen and J. Meseguer. Unwinding and inference control. In *IEEE Symposium on Security and Privacy*, pages 75–87, 1984.
17. J. W. Gray III. Toward a mathematical foundation for information. *Journal of Computer Security*, 1(3-4):255–294, 1992.
18. J. Y. Halpern and K. R. O'Neill. Secrecy in multiagent systems. *ACM Trans. Inf. Syst. Secur.*, 12(1):5.1–5.47, 2008.
19. S. Hamadou, C. Palamidessi, and V. Sassone. Quantifying leakage in the presence of unreliable sources of information. *J. Comput. Syst. Sci.*, 88:27–52, 2017.
20. D. Kifer and A. Machanavajjhala. No free lunch in data privacy. In T. K. Sellis, R. J. Miller, A. Kementsietsidis, and Y. Velegrakis, editors, *ACM SIGMOD International Conference on Management of Data, SIGMOD 2011*, pages 193–204. ACM, 2011.
21. J. Liu, L. Xiong, and J. Luo. Semantic security: Privacy definitions revisited. *Trans. Data Privacy*, 6(3):185–198, 2013.
22. A. Machanavajjhala, D. Kifer, J. Gehrke, and M. Venkitasubramaniam. *L*-diversity: privacy beyond *k*-anonymity. *TKDD*, 1(1), 2007.
23. P. Mardziel, S. Magill, M. Hicks, and M. Srivatsa. Dynamic enforcement of knowledge-based security policies using probabilistic abstract interpretation. *Journal of Computer Security*, 21(4):463–532, 2013.
24. C. Palamidessi. Quantitative approaches to the protection of private information: State of the art and some open challenges. In R. Focardi and A. C. Myers, editors, *Principles of Security and Trust - 4th International Conference, POST 2015*, volume 9036 of *Lecture Notes in Computer Science*, pages 3–7. Springer, 2015.
25. P. Samarati. Protecting respondents' identities in microdata release. *IEEE Trans. Knowl. Data Eng.*, 13(6):1010–1027, 2001.
26. T. Santen. Preservation of probabilistic information flow under refinement. *Inf. Comput.*, 206(2-4):213–249, 2008.
27. G. L. Sicherman, W. de Jonge, and R. P. van de Riet. Answering queries without revealing secrets. *ACM Trans. Database Syst.*, 8(1):41–59, 1983.
28. L. Sweeney. *k*-anonymity: A model for protecting privacy. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 10(5):557–570, 2002.
29. H. Yasuoka and T. Terauchi. Quantitative information flow as safety and liveness hyperproperties. *Theor. Comput. Sci.*, 538:167–182, 2014.