On the relation between Differential Privacy and Quantitative Information Flow*

Mário S. Alvim, Miguel E. Andrés,
Konstantinos Chatzikokolakis, and Catuscia Palamidessi

INRIA and LIX, Ecole Polytechnique, France.

Abstract. Differential privacy is a notion that has emerged in the community of statistical databases, as a response to the problem of protecting the privacy of the database's participants when performing statistical queries. The idea is that a randomized query satisfies differential privacy if the likelihood of obtaining a certain answer for a database x is not too different from the likelihood of obtaining the same answer on adjacent databases, i.e. databases which differ from x for only one individual.
Information flow is an area of Security concerned with the problem of controlling the leakage of confidential information in programs and protocols. Nowadays, one of the most established approaches to quantify and to reason about leakage is based on the Rényi min entropy version of information theory.
In this paper, we analyze critically the notion of differential privacy in light of the conceptual framework provided by the Rényi min information theory. We show that there is a close relation between differential privacy and leakage, due to the graph symmetries induced by the adjacency relation. Furthermore, we consider the utility of the randomized answer, which measures its expected degree of accuracy. We focus on certain kinds of utility functions called "binary", which have a close correspondence with the Rényi min mutual information. Again, it turns out that there can be a tight correspondence between differential privacy and utility, depending on the symmetries induced by the adjacency relation and by the query. Depending on these symmetries we can also build an optimal-utility randomization mechanism while preserving the required level of differential privacy. Our main contribution is a study of the kind of structures that can be induced by the adjacency relation and the query, and how to use them to derive bounds on the leakage and achieve the optimal utility.

1 Introduction

Databases are commonly used for obtaining statistical information about their participants. Simple examples of statistical queries are, for instance, the predominant disease of a certain population, or the average salary. The fact that

* This work has been partially supported by the project ANR-09-BLAN-0169-01 PANDA and by the INRIA DRI Equipe Associée PRINTEMPS. The work of Miguel E. Andrés has been supported by the LIX-Qualcomm postdoc fellowship 2010.

the answer is publicly available, however, constitutes a threat for the privacy of the individuals.

In order to illustrate the problem, consider a set of individuals *Ind* whose attribute of interest[1] has values in *Val*. A particular database is formed by a subset of *Ind*, where a certain value in *Val* is associated to each participant. A query is a function $f : \mathcal{X} \to \mathcal{Y}$, where $\mathcal{X}$ is the set of all possible databases, and $\mathcal{Y}$ is the domain of the answers.

For example, let *Val* be the set of possible salaries and let f represent the query "what is the average salary of the participants in the database". In principle we would like to consider the *global information* relative to a database x as *public*, and the *individual information* about a participant i as *private*. Namely, we would like to be able to obtain $f(x)$ without being able to infer the salary of i. However, this is not always possible. In particular, if the number of participants in x is known (say n), then the removal of i from the database would allow to infer i's salary by querying again the new database x', and by applying the formula $n\,f(x) - (n-1)\,f(x')$. Using an analogous reasoning we can argue that not only the removal, but also the addition of an individual is a threat for his privacy.

Another kind of private information we may want to protect is whether an individual i is participating or not in a database. In this case, if we know for instance that i earns, say $5K$ Euros/month, and all the other individuals in *Ind* earn less than $4K$ Euros/month, then knowing that $f(x) > 5K$ Euros/month will reveal immediately that i is in the database x.

A common solution to the above problems is to introduce some output perturbation mechanism based on *randomization*: instead of the exact answer $f(x)$ we report a "noisy" answer. Namely, we use some randomized function $\mathcal{K}$ which produces values in some domain[2] $\mathcal{Z}$ according to some probability distribution that depends on the input $x \in \mathcal{X}$. Of course for certain distributions it may still be possible to guess the value of an individual with a high probability of success. The notion of *differential privacy*, due to Dwork [10, 13, 11, 12], is a proposal to control the risk of violating privacy for both kinds of threats described above (value and participation). The idea is to say that $\mathcal{K}$ satisfies ϵ-differential privacy (for some $\epsilon > 0$) if the ratio between the probabilities that two adjacent databases give the same answer is bound by e^{ϵ}, where by "adjacent" we mean that the databases differ for only one individual (either for the value of an individual or for the presence/absence of an individual). Often we will abbreviate "ϵ-differential privacy" as ϵ-d.p.

Obviously, the smaller is ϵ, the greater is the privacy protection. In particular, when ϵ is close to 0 the output of $\mathcal{K}$ is nearly independent from the input (all distributions are almost equal). Unfortunately, such $\mathcal{K}$ is practically useless. The *utility*, i.e. the capability to retrieve accurate answers from the reported

[1] In general we could be interested in several attributes simultaneously, and in this case *Val* would be a set of tuples.

[2] The new domain $\mathcal{Z}$ may coincide with $\mathcal{Y}$, but not necessarily. It depends on how the randomization mechanism is defined.

ones, is the other important characteristic of $\mathcal{K}$, and it is clear that there is a trade-off between utility and privacy. On the other hand, these two notions are not the complete opposite of each other, because utility concerns the relation between the reported answer and the real answer, while privacy is concerns the relation between the reported answer and the information in the database. This asymmetry makes more interesting the problem of finding a good compromise between the two.

At this point, we would like to remark an intriguing analogy between the area of differential privacy and that of *quantitative information flow* (QIF), both in the motivations and in the basic conceptual framework. Information flow is concerned with the leakage of secret information through computer systems, and the attribute "quantitative" refers to the fact that we are interested in measuring the amount of leakage, not just its occurrence. One of the most established approaches to QIF is based on information theory: the idea is that a system is seen as a channel in the information-theoretic sense, where the secret is the input and the observables are the output. The entropy of the input represents its vulnerability, i.e. how easy it is for an attacher to guess the secret. We distinguish between the *a priori* entropy (before the observable) and the *a posteriori* entropy (given the observable). The difference between the two gives the *mutual information* and represents, intuitively, the increase in vulnerability due to the observables produced by the system, so it is naturally considered as a measure of the leakage. The notion of entropy is related to the kind of attack we want to model, and in this paper we focus on the Rényi min entropy [18], which represents the so-called *one-try attacks*. In recent years there has been a lot of research aimed at establishing the foundations of this framework [19, 7, 16, 3, 5]. It is worth pointing out that the a posteriori Rényi min entropy corresponds to the concept of Bayes risk, which has also been proposed as a measure of the effectiveness of attacks [8, 6, 17].

The analogy hinted above between differential privacy and QIF is based on the following observations: at the motivational level, the concern about privacy is akin the concern about information leakage. At the conceptual level, the randomized function $\mathcal{K}$ can be seen as an information-theoretic channel, and the limit case of $\epsilon = 0$, for which the privacy protection is total, corresponds to a 0-capacity channel[3] (the rows of the channel matrix are all identical), which does not allow any leakage. Another promising similarity is that the notion of utility (in the binary case) corresponds closely to the Bayes risk.

In this paper we investigate the notion of differential privacy, and its implications, in light of the min-entropy information theoretic framework developed for QIF. In particular, we wish to explore the following natural questions:

1. Does ϵ-d.p. induce a bound on the information leakage of $\mathcal{K}$?
2. Does ϵ-d.p. induce a bound on the information leakage *relative to an individual*?
3. Does ϵ-d.p. induce a bound on the utility?

[3] The channel capacity is the maximum mutual information over all possible input distributions.

4. Given f and ϵ, can we construct a $\mathcal{K}$ which satisfies ϵ-d.p. and maximum utility?

We will see that the answers to (1) and (2) are positive, and we provide bounds that are tight, in the sense that for every ϵ there is a $\mathcal{K}$ whose leakage reaches the bound. For (3) we are able to give a tight bound in some cases which depend on the structure of the query, and for the same cases, we are able to construct an oblivious[4] $\mathcal{K}$ with maximum utility, as requested by (4).

Part of the above results have already appeared in [1], and are based on techniques which exploit the graph structure that the adjacency relation induces on the domain of all databases $\mathcal{X}$, and on the domain of the correct answers $\mathcal{Y}$. The main contribution of this paper is an extension of those techniques, and a coherent graph-theoretic framework for reasoning about the symmetries of those domains. More specifically:

- We explore the graph-theoretic foundations of the adjacency relation, and point out various types of symmetries which allow us to establish a strict link between differential privacy and information leakage.
- We give a tight bound for the question (2) above, strictly smaller than the one in [1].
- We extend the structures for which we give a positive answer to the questions (3) and (4) above. In [1] the only case considered was the class of graphs with single-orbit automorphisms. Here we show that the results hold also for regular-distance graphs and a variant of vertex-transtive graphs.

In this paper we focus on the case in which $\mathcal{X}$, $\mathcal{Y}$ and $\mathcal{Z}$ are finite, leaving the more general case for future work.

2 Preliminaries

2.1 Database domain and Differential privacy

Let *Ind* be a finite set of individuals that may participate in a database and *Val* a finite set of possible values for the attribute of interest of these individuals. In order to capture in a uniform way the presence/absence of an individual in the database, as well as its value, we enrich the set of possible values with an element a representing the absence of the individual. Thus the set of all possible databases is the set $\mathcal{X} = V^{Ind}$, where $V = Val \cup \{a\}$. We will use u and v to denote the cardinalities of *Ind* and V, $|Ind|$ and $|V|$, respectively. Hence we have that $|\mathcal{X}| = v^u$. A database x can be represented as a u-tuple $v_0 v_1 \dots v_{u-1}$ where each $v_i \in V$ is the value of the corresponding individual. Two databases x, x' are *adjacent* (or *neighbors*), written $x \sim x'$, if they differ for the value of exactly one individual. For instance, for $u = 3$, $v_0 v_1 v_2$ and $v_0 w_1 v_2$, with $w_1 \neq v_1$, are adjacent. The structure $(\mathcal{X}, \sim)$ forms an undirected graph.

[4] A randomized function $\mathcal{K}$ is oblivious if its probability distribution depends only on the answer to the query, and not on the database.

Intuitively, differential privacy is based on the idea that a randomized query function provides sufficient protection if the ratio between the probabilities of two adjacent databases to give a certain answer is bound by e^ϵ, for some given $\epsilon > 0$. Formally:

Definition 1 ([12]). *A randomized function $\mathcal{K}$ from $\mathcal{X}$ to $\mathcal{Z}$ satisfies ϵ-differential privacy if for all pairs $x, x' \in \mathcal{X}$, with $x \sim x'$, and all $S \subseteq \mathcal{Z}$, we have that:*

$$Pr[\mathcal{K}(x) \in S] \leq e^\epsilon \times Pr[\mathcal{K}(x') \in S]$$

The above definition takes into account the possibility that $\mathcal{Z}$ is a continuous domain. In our case, since $\mathcal{Z}$ is finite, the probability distribution is discrete, and we can rewrite the property of ϵ-d.p. more simply as (using the notation of conditional probabilities, and considering both quotients):

$$\frac{1}{e^\epsilon} \leq \frac{Pr[Z = z|X = x]}{Pr[Z = z|X = x']} \leq e^\epsilon \qquad \text{for all } x, x' \in \mathcal{X} \text{ with } x \sim x', \text{ and all } z \in \mathcal{Z}$$

where X and Z represent the random variables associated to $\mathcal{X}$ and $\mathcal{Z}$, respectively.

2.2 Information theory and application to information flow

In the following, X, Y denote two discrete random variables with carriers $\mathcal{X} = \{x_0, \ldots, x_{n-1}\}$, $\mathcal{Y} = \{y_0, \ldots, y_{m-1}\}$, and probability distributions $p_X(\cdot)$, $p_Y(\cdot)$, respectively. An information-theoretic channel is constituted by an input X, an output Y, and the matrix of conditional probabilities $p_{Y|X}(\cdot \mid \cdot)$, where $p_{Y|X}(y \mid x)$ represent the probability that Y is y given that X is x. We shall omit the subscripts on the probabilities when they are clear from the context.

Rényi min-entropy In [18], Rényi introduced an one-parameter family of entropy measures, intended as a generalization of Shannon entropy. The Rényi entropy of order α ($\alpha > 0$, $\alpha \neq 1$) of a random variable X is defined as $H_\alpha(X) = \frac{1}{1-\alpha} \log_2 \sum_{x \in \mathcal{X}} p(x)^\alpha$. We are particularly interested in the limit of H_α as α approaches ∞. This is called *min-entropy*. It can be proven that $H_\infty(X) \stackrel{\text{def}}{=} \lim_{\alpha \to \infty} H_\alpha(X) = -\log_2 \max_{x \in \mathcal{X}} p(x)$.

Rényi defined also the α-generalization of other information-theoretic notions, like the Kullback-Leibler divergence. However, he did not define the α-generalization of the conditional entropy, and there is no general agreement on what it should be. For the case $\alpha = \infty$, we adopt here the definition of conditional entropy proposed by Smith in [19]:

$$H_\infty(X \mid Y) = -\log_2 \sum_{y \in \mathcal{Y}} p(y) \max_{x \in \mathcal{X}} p(x \mid y) \tag{1}$$

Analogously to the Shannon case, we can define the Rényi-mutual information I_∞ as $H_\infty(X) - H_\infty(X \mid Y)$, and the capacity C_∞ as $\max_{p_X(\cdot)} I_\infty(X; Y)$. It has been proven in [7] that C_∞ is obtained at the uniform distribution, and that it is equal to the sum of the maxima of each column in the channel matrix, i.e., $C_\infty = \sum_{y \in \mathcal{Y}} \max_{x \in \mathcal{X}} p(y \mid x)$.

Interpretation in terms of attacks: Rényi min-entropy can be related to a model of adversary who is allowed to ask exactly one question, which must be of the form "is $X = x$?" (one-try attacks). More precisely, $H_\infty(X)$ represents the (logarithm of the inverse of the) probability of success for this kind of attacks and with the best strategy, which consists, of course, in choosing the x with the maximum probability.

As for $H_\infty(X \mid Y)$, it represents the inverse of the (expected value of the) probability that the same kind of adversary succeeds in guessing the value of X *a posteriori*, i.e. after observing the result of Y. The complement of this probability is also known as *Bayes risk*. Since in general X and Y are correlated, observing Y increases the probability of success. Indeed we can prove formally that $H_\infty(X \mid Y) \leq H_\infty(X)$, with equality if and only if X and Y are independent. $I_\infty(X;Y)$ corresponds to the *ratio* between the probabilities of success a priori and a posteriori, which is a natural notion of leakage. Note that $I_\infty(X;Y) \geq 0$, which seems desirable for a good notion of leakage.

3 Graph symmetries

In this section we explore some classes of graphs that allow us to derive a strict correspondence between ϵ-d.p. and the a posteriori entropy of the input.

Let us first recall some basic notions. Given a graph $G = (\mathcal{V}, \sim)$, the *distance* $d(v, w)$ between two vertices $v, w \in \mathcal{V}$ is the number of edges in a shortest path connecting them. The *diameter* of G is the maximum distance between any two vertices in $\mathcal{V}$. The degree of a vertex is the number of edges incident to it. G is called *regular* if every vertex has the same degree. A regular graph with vertices of degree k is called a k-regular graph. An authomorphism of G is a permutation σ of the vertex set $\mathcal{X}$, such that for any pair of vertices x, x', if $x \sim x'$, then $\sigma(x) \sim \sigma(x')$. If σ is an authomorphism, and v a vertex, the orbit of v under σ is the set $\{v, \sigma(v), \ldots, \sigma^{k-1}(v)\}$ where k is the smallest positive integer such that $\sigma^k(v) = v$. Clearly, the orbits of the vertices under σ define a partition of $\mathcal{V}$.

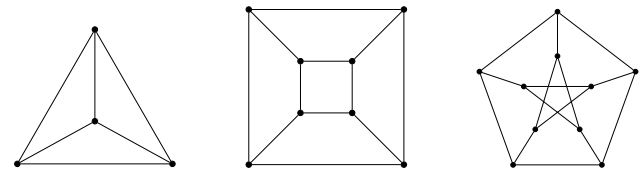
The following two definition introduce the classes of graphs that we are interested in. The first class is well known in literature.

Definition 2. *Given a graph $G = (\mathcal{V}, \sim)$, we say that G is distance-regular if there exist integers $b_i, c_i, i = 0, ..., d$ such that for any two vertices v, w in $\mathcal{V}$ with distance $i = d(v, w)$, there are exactly c_i neighbors of w in $G_{i-1}(x)$ and b_i neighbors of v in $G_{i+1}(x)$, where $G_i(x)$ is the set of vertices y of G with $d(x, y) = i$.*

Some examples of distance-regular graphs are illustrated in Figure 1.

The next class is a variant of the VT (vertex-transitive) class:

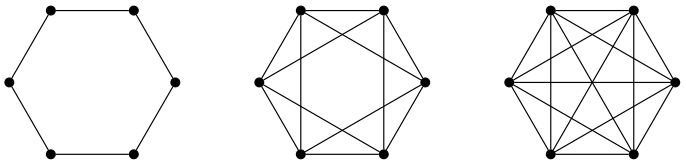
Definition 3. *A graph $G = (\mathcal{V}, \sim)$ is VT^+ (vertex-transitive +) if there are n automorphisms σ_0, σ_1, $\ldots \sigma_{n-1}$, where $n = |\mathcal{V}|$, such that, for every vertex $v \in \mathcal{V}$, we have that $\{\sigma_i(v) \mid 0 \leq i \leq n-1\} = \mathcal{V}$.*



(a) Tetrahedral graph (b) Cubical graph (c) Petersen graph

Fig. 1. Some distance-regular graphs with degree 3.

In particular, the graphs for which there exists an authomorphism σ which induces only one orbit are VT^+: in fact it is sufficient to define $\sigma_i = \sigma^i$ for all i from 0 to $n-1$. Figure 2 illustrates some graphs with a single-orbit automorphism.



(a) Cycle: degree 2. (b) Degree 4. (c) Clique: degree 5.

Fig. 2. Some VT^+ graphs

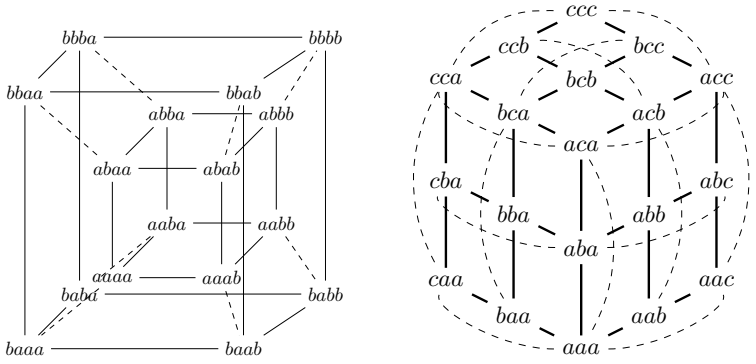
From graph theory we know that neither of the two classes subsumes the other. They have however a non-empty intersection, which contains in particular all the structures of the form $(V^{Ind}, \sim)$, i.e. the database domains.

Proposition 1. *The structure $(\mathcal{X}, \sim) = (V^{Ind}, \sim)$ is both a distance-regular graph and a VT^+ graph.*

Figure 3 illustrates some examples of structures $(V^{Ind}, \sim)$. Note that when $|Ind| = n$ and $|V| = 2$, $(V^{Ind}, \sim)$ is the n-dimentional hypercube.

The situation is summarized in Figure 4. We remark that in general the graphs $(V^{Ind}, \sim)$ do not have a single-orbit authomorphism. The only exceptions are the two simplest structures ($|V| = 2, |Ind| \leq 2$).

The two symmetry classes defined above, distance-regular and VT^+, will be used in the next section to transform a generic channel matrix into a matrix with a symmetric structure, while preserving the a posteriori min entropy and the ϵ-d.p.. This is the core of our technique to establish the relation between differential privacy and quantitive information flow, depending on the structure induced by the database adjacency relation.



(a) $|Ind| = 4, V = \{a, b\}$ (4-dimensional hypercube)

(b) $|Ind| = 3, V = \{a, b, c\}$ (for readability sake we show only part of the graph)

Fig. 3. Some $(V^{Ind}, \sim)$ graphs

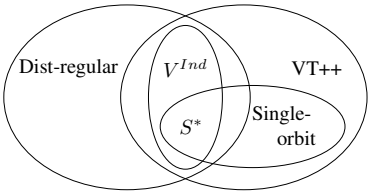


Fig. 4. Venn diagram for the classes of graphs considered in this section. Here, $S^* = \{V^{Ind} \mid |V| = 2, |Ind| \leq 2\}$

4 Deriving the relation between differential privacy and QIF on the basis of the graph structure

This section contains the main technical contribution of the paper: a general technique for determining the relation between ϵ-differential privacy and leakage, and between ϵ-differential privacy and utility, depending on the graph structure induced by $\sim$ and f. The idea is to use the symmetries of the graph structure to transform the channel matrix into an equivalent matrix with certain regularities, which allow to establish the link between ϵ-differential privacy and the a posteriori min entropy.

Let us illustrate briefly this transformation. Consider a channel whose matrix M has at least as many columns as rows. First, we transform M into a matrix M' in which each of the first n columns has a maximum in the diagonal, and the remaining columns are all 0's. Second, under the assumption that the input domain is distance-regular or VT^+, we transform M' into a matrix M'' whose diagonal elements are all the same, and coincide with the maximum element of M'', which we denote here by $\max^{M''}$. These steps are illustrated in Figure 5.

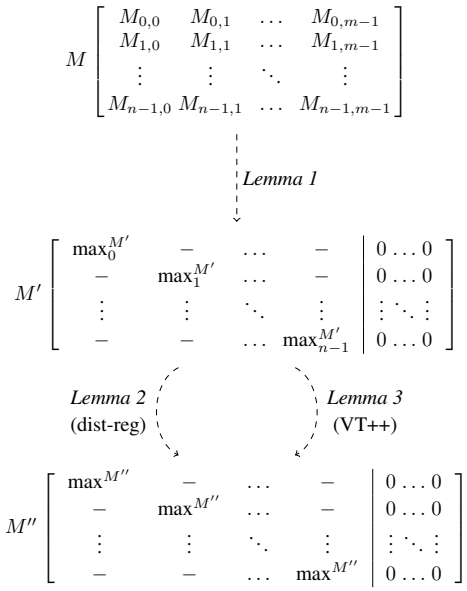


Fig. 5. Matrix transformations for distance-regular and VT^+ graphs

We are now going to present formally our the technique. Let us first fix some notation: In the rest of this section we consider channels with input A and output B, with carriers $\mathcal{A}$ and $\mathcal{B}$ respectively, and we assume that the probability distribution of A is uniform. Furthermore, we assume that $|\mathcal{A}| = n \leq |\mathcal{B}| = m$. We also assume an adjacency relation $\sim$ on $\mathcal{A}$, i.e. that $(\mathcal{A}, \sim)$ is an undirected graph structure. With a slight abuse of notation, we will also write $i \sim h$ when i and h are associated to adjacent elements of $\mathcal{A}$, and we will write $d(i, h)$ to denote the distance between the elements of $\mathcal{A}$ associated to i and h.

We note that a channel matrix M satisfies ϵ-d.p. if for each column j and for each pair of rows i and h such that $i \sim h$ we have that:

$$\frac{1}{e^\epsilon} \leq \frac{M_{i,j}}{M_{h,j}} \leq e^\epsilon.$$

The a posteriori entropy of a channel with matrix M will be denoted by $H_\infty^M(A|B)$.

Next Lemma is relative to the first step of the transformation.

Lemma 1. *Consider a channel with matrix M. Assume that M satisfies ϵ-d.p.. Then it is possible to transform M into a matrix M' such that:*

- *Each of the first n columns has a maximum in the diagonal, i.e. $M'_{i,i} = \max_i^{M'} = \max_h M'_{h,i}$ for each i from 0 to $n-1$.*
- *The rest of the columns contain only 0's, i.e. $M'_{i,j} = 0$ for each i from 0 to $n-1$ and each j from n to $m-1$.*
- *M' satisfies ϵ-d.p.*

– $H_\infty^{M'}(A|B) = H_\infty^{M}(A|B)$.

Next lemma is relative to the second step of the transformation, for the case of distance-regular graphs.

Lemma 2. *Consider a channel with matrix M'. Assume that M' satisfies ϵ-d.p., and the first n columns have maxima in the diagonal, and the rest of the columns contain only 0's. Assume that $(\mathcal{A}, \sim)$ is distance-regular. Then it is possible to transform M' into a matrix M'' such that:*

– *The elements of the diagonal are all the same, and are equal to the maximum of the matrix, i.e. $M''_{i,i} = \max^{M''} = \max_{h,i} M''_{h,i}$ for each i from 0 to $n-1$.*
– *The rest of the columns contain only 0's.*
– *M'' satisfies ϵ-d.p.*
– *$H_\infty^{M''}(A|B) = H_\infty^{M'}(A|B)$.*

Next lemma is relative to the second step of the transformation, for the case of VT^+ graphs.

Lemma 3. *Consider a channel with matrix M' satisfying the assumptions of Lemma 2, except for the assumption about distance-regularity, which we replace by the assumption that $(\mathcal{A}, \sim)$ is VT^+. Then it is possible to transform M' into a matrix M'' with the same properties as in Lemma 2.*

Note that the fact that in M'' the diagonal elements are all equal to the maximum $\max^{M''}$ implies that $H_\infty^{M''}(A|B) = \max^{M''}$.

Once we have a matrix with the properties of M'', we can use again the graph structure of $\mathcal{A}$ to determine a bound on $H_\infty^{M''}(A|B)$.

First we note that the property of ϵ-d.p. induces a relation between the ratio of elements at any distance:

Remark 1. Let M be a matrix satisfying ϵ-d.p.. Then, for any column j, and any pair of rows i and h we have that:

$$\frac{1}{e^{\epsilon\, d(i,h)}} \leq \frac{M_{i,j}}{M_{h,j}} \leq e^{\epsilon\, d(i,h)}$$

In particular, if we know that the diagonal elements of M are equal to the maximum element $\max^M$, then for each element $M_{i,j}$ we have that:

$$M_{i,j} \geq \frac{\max^M}{e^{\epsilon\, d(i,j)}} \tag{2}$$

Let us fix a row, say row r. For each distance d from 0 to the diameter of the graph, let n_d be the number of elements $M_{r,j}$ that are at distance d from the corresponding diagonal element $M_{j,j}$, i.e. such that $d(r,j) = d$. (Clearly, n_d depends on the structure of the graph.) Since the elements of the row i represent a probability distribution, we obtain the following dis-equation:

$$\max^M \sum_d \frac{n_d}{e^{\epsilon\, d}} \leq 1$$

from which we derive immediately a bound on the min a-posteriori entropy.

Putting together all the steps of this section, we obtain our main result.

Theorem 1. *Consider a matrix M, and let r be a row of M. Assume that $(\mathcal{A}, \sim)$ is either distance-regular or VT$^+$, and that M satisfies ϵ-d.p. For each distance d from 0 to the diameter of $(\mathcal{A}, \sim)$, let n_d be the number of nodes j at distance d from r. Then we have that:*

$$H_\infty^M(A|B) \leq -\log_2 \frac{1}{\sum_d \frac{n_d}{e^{\epsilon\, d}}} \tag{3}$$

Note that this bound is tight, in the sense that we can build a matrix for which (3) holds with equality. It is sufficient to define each element $M_{i,j}$ according to (2) (with equality instead of dis-equality, of course).

In the next section, we will see how to use this theorem for establishing a bound on the leakage and on the utility.

5 Application to leakage

As already hinted in the introduction, we can regard $\mathcal{K}$ as a channel with input X and output Z. From Proposition 1 we know that $(\mathcal{X}, \sim)$ is both distance-regular and VT$^+$, we can therefore apply Theorem 1. Let us fix a particular database $x \in \mathcal{X}$. The number of databases at distance d from x is

$$n_d = \binom{u}{d} (v-1)^d \tag{4}$$

where $u = |Ind|$ and $v = V$. In fact, recall that x can be represented as a u-tuple with values in V. We need to select d individuals in the u-tuple and then change their values, and each of them can be changed in $v-1$ different ways.

Using the n_d from (4) in Theorem 1 we obtain a binomial expansion in the denominator, namely:

$$H_\infty^M(X|Z) \leq -\log_2 \frac{1}{\sum_{d=0}^{u} \binom{u}{d} (v-1)^d \frac{e^{\epsilon(u-d)}}{e^{\epsilon\, u}}} = -u \log_2 \frac{e^\epsilon}{v-1+e^\epsilon}$$

which gives the following result:

Theorem 2. *If $\mathcal{K}$ satisfies ϵ-d.p., then for the uniform input distribution the information leakage is bound from above as follows:*

$$I_\infty(X;Z) \leq -u \log_2 \frac{v\, e^\epsilon}{v-1+e^\epsilon}$$

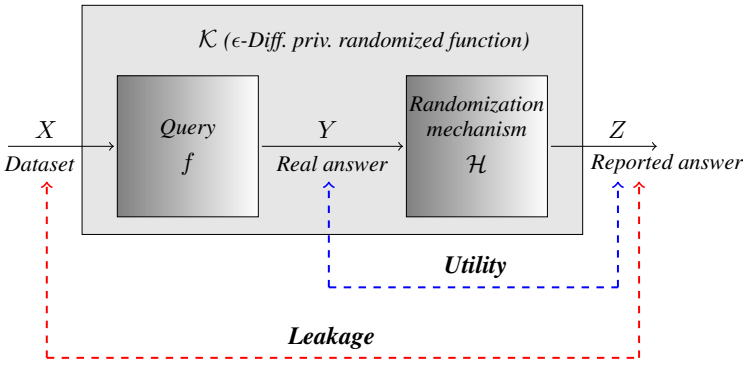


Fig. 6. Schema of an oblivious randomized function

We consider now the *leakage for a single individual*. Let us fix a database x, and a particular individual i in Ind. The possible ways in which we can change the value of i in x are $v-1$. All the new databases obtained in this way are adjacent to each other, i.e. the graph structure associated to the input is a clique of v nodes. Therefore we obtain $n_d = 1$ for $d = 0$, $n_d = v-1$ for $d = 1$, and $n_d = 0$ otherwise. By substituting this value of n_d in Theorem 1, we get

$$H_\infty^{ind}(Val|Z) \leq -\log_2 \frac{1}{1+\dfrac{v-1}{e^\epsilon}} = -\log_2 \frac{e^\epsilon}{v-1+e^\epsilon}$$

which leads to the following result:

Proposition 2. *Assume that $\mathcal{K}$ satisfies ϵ-d.p.. Then for the uniform distribution on V the information leakage for an individual is bound from above as follows:*

$$I_\infty^{ind}(Val; B) \leq \log_2 \frac{v\, e^\epsilon}{v-1+e^\epsilon}$$

Note that the bound on the leakage for an individual does not depend on the size of Ind, nor on the database x that we fix.

6 Application to utility

We turn now our attention to the issue of *utility*. We focus on the case in which $\mathcal{K}$ is *oblivious*, which means that it depends only on the (exact) answer to the query, i.e. on the value of $f(x)$, and not on x.

An oblivious function can be decomposed in the concatenation of two channels, one representing the function f, and the other representing the randomization mechanism $\mathcal{H}$ added as output perturbation. The situation is illustrated in Figure 6.

The standard way to define utility is by means of *guess* and *gain* functions. The functionality of the first is $guess : \mathcal{Z} \to \mathcal{Y}$, and it represents the user's

strategy to retrieve the correct answer form the reported one. The functionality of the latter is $gain : \mathcal{Y} \times \mathcal{Y} \to \mathbb{R}$. the value $gain(y, y')$ represents the reward for guessing the answer y when the correct answer is y'. The utility $\mathcal{U}$ can then be defined as the expected gain:

$$\mathcal{U}(Y, Z) = \sum_{y,z} p(y, z)\, gain(guess(z), y)$$

We focus here on the so-called *binary* gain function, which is defined as

$$gain(y, y') = \begin{cases} 1 & \text{if } y = y' \\ 0 & \text{otherwise} \end{cases}$$

This kind of function represents the case in which there is no reason to prefer an answer over the other, except if it is the *right* answer. More precisely, we get a gain if and only if we guess the right answer.

If the gain function is binary, and the *guess* function represents the user's best strategy, i.e. it is chosen to optimize utility, then there is a well-known correspondence between $\mathcal{U}$ and the Bayes risk / the a posteriori min entropy. Such correspondence is expressed by the following proposition:

Proposition 3. *Assume that gain is binary and guess is optimal. Then:*

$$\mathcal{U}(Y, Z) = \sum_{z} \max_{y}(p(z|y)\, p(y)) = 2^{-H_\infty(Y|Z)}$$

In order to analyze the implications of the ϵ-d.p. requirement on the utility, we need to consider the structure that the adjacency relation induces on $\mathcal{Y}$. Let us define $\sim$ on $\mathcal{Y}$ as follows: $y \sim y'$ if there are $x, x' \in \mathcal{X}$ such that $y = f(x)$, $y' = f(x')$, and $x \sim x$,. Note that $\mathcal{K}$ satisfies ϵ-d.p. if and only if $\mathcal{H}$ satisfies ϵ-d.p.

If $(\mathcal{Y}, \sim)$ is distance-regular or VT^+, then we can apply Theorem 1 to find a bound on the utility. In the following, we assume that the distribution of Y is uniform.

Theorem 3. *Consider a randomized mechanism $\mathcal{H}$, and let y be an element of $\mathcal{Y}$. Assume that $(\mathcal{Y}, \sim)$ is either distance-regular or VT^+ and that $\mathcal{H}$ satisfies ϵ-d.p. For each distance d from 0 to the diameter of $(\mathcal{Y}, \sim)$, let n_d be the number of nodes y' at distance d from y. Then we have that:*

$$\mathcal{U}(Y, Z) \leq \frac{1}{\sum_{d} \frac{n_d}{e^{\epsilon\, d}}} \tag{5}$$

The above bound is tight, in the sense that (provided $(\mathcal{Y}, \sim)$ is distance-regular or VT^+) we can construct a mechanism $\mathcal{H}$ which satisfies (5) with equality. More precisely, define

$$c = \frac{1}{\sum_{d} \frac{n_d}{e^{\epsilon\, d}}}$$

Then define $\mathcal{H}$ (here identified with its channel matrix for simplicity) as follows:

$$\mathcal{H}_{i,j} = \frac{c}{e^{\epsilon\, d(i,j)}} \tag{6}$$

Theorem 4. *Assume $(\mathcal{Y}, \sim)$ is distance-regular or VT^+. Then the matrix $\mathcal{H}$ defined in (6) satisfies ϵ-d.p. and has maximal utility:*

$$\mathcal{U}(Y, Z) = \frac{1}{\sum_d \frac{n_d}{e^{\epsilon\, d}}}$$

Note that we can always define $\mathcal{H}$ as in (6): the matrix so defined will be a legal channel matrix, and it will satisfy ϵ-d.p.. However, if $(\mathcal{Y}, \sim)$ is neither distance-regular nor VT^+, then the utility of such $\mathcal{H}$ is not necessarily optimal.

We end this section with an example (borrowed from [1]) to illustrate our technique.

Example 1. Consider a database with electoral information where each row corresponds to a voter and contains the following three fields:

- *Id*: a unique (anonymized) identifier assigned to each voter;
- *City*: the name of the city where the user voted;
- *Candidate*: the name of the candidate the user voted for.

Consider the query *"What is the city with the greatest number of votes for a given candidate cand?"*. For such a query the binary utility function is the natural choice: only the right city gives some gain, and all wrong answers are equally bad. It is easy to see that every two answers are neighbors, i.e. the graph structure of the answers is a clique.

Let us consider the scenario where $City = \{A, B, C, D, E, F\}$ and assume for simplicity that there is a unique answer for the query, i.e., there are no two cities with exactly the same number of individuals voting for candidate *cand*. Table 1 shows two alternative mechanisms providing ϵ-differential privacy (with $\epsilon = \log 2$). The first one, M_1, is based on the truncated geometric mechanism method used in [14] for counting queries (here extended to the case where every pair of answers is neighbor). The second mechanism, M_2, is obtained by applying the definition (6). From Theorem 4 we know that for the uniform input distribution M_2 gives optimal utility.

For the uniform input distribution, it is easy to see that $\mathcal{U}(M_1) = 0.2242 < 0.2857 = \mathcal{U}(M_2)$. Even for non-uniform distributions, our mechanism still provides better utility. For instance, for $p(A) = p(F) = 1/10$ and $p(B) = p(C) = p(D) = P(E) = 1/5$, we have $\mathcal{U}(M_1) = 0.2412 < 0.2857 = \mathcal{U}(M_2)$. This is not too surprising: the geometric mechanism, as well as the Laplacian mechanism proposed by Dwork, perform very well when the domain of answers is provided with a metric and the utility function is not binary[5]. It also works well when

[5] In the metric case the gain function can take into account the proximity of the reported answer to the real one, the idea being that a close answer, even if wrong, is better than a distant one.

(a) M_1: truncated geometric mechanism

In/Out	A	B	C	D	E	F
A	0.535	0.060	0.052	0.046	0.040	0.267
B	0.465	0.069	0.060	0.053	0.046	0.307
C	0.405	0.060	0.069	0.060	0.053	0.353
D	0.353	0.053	0.060	0.069	0.060	0.405
E	0.307	0.046	0.053	0.060	0.069	0.465
F	0.267	0.040	0.046	0.052	0.060	0.535

(b) M_2: our mechanism

In/Out	A	B	C	D	E	F
A	2/7	1/7	1/7	1/7	1/7	1/7
B	1/7	2/7	1/7	1/7	1/7	1/7
C	1/7	1/7	2/7	1/7	1/7	1/7
D	1/7	1/7	1/7	2/7	1/7	1/7
E	1/7	1/7	1/7	1/7	2/7	1/7
F	1/7	1/7	1/7	1/7	1/7	2/7

Table 1. Mechanisms for the city with higher number of votes for candidate *cand*

$(\mathcal{Y}, \sim)$ has low connectivity, in particular in the cases of a ring and of a line. But in this example, we are not in these cases, because we are considering *binary gain functions* and *high connectivity*.

7 Related work

As far as we know, the first work to investigate the relation between differential privacy and information-theoretic leakage *for an individual* was [2]. In this work, a channel is relative to a given database x, and the channel inputs are all possible databases adjacent to x. Two bounds on leakage were presented, one for teh Rényi min entropy, and one for Shannon entropy. Our bound in Proposition 2 is an improvement with respect to the (Rényi min entropy) bound in [2].

Barthe and Köpf [4] were the first to investigates the (more challenging) connection between differential privacy and the Rényi min-entropy leakage *for the entire universe of possible databases*. They consider the "end-to-end differentially private mechanisms", which correspond to what we call $\mathcal{K}$ in our paper, and propose, like we do, to interpret them as information-theoretic channels. They provide a bound for the leakage, but point out that it is not tight in general, and show that there cannot be a domain-independent bound, by proving that for any number of individual u the optimal bound must be at least a certain expression $f(u, \epsilon)$. Finally, they show that the question of providing optimal upper bounds for the leakage of ϵ-differentially private randomized functions in terms of rational functions of ϵ is decidable, and leave the actual function as an open question. In our work we used rather different techniques and found (independently) the same function $f(u, \epsilon)$ (the bound in Theorem 1), but we actually proved that $f(u, \epsilon)$ is the optimal bound[6]. Another difference is that [4] captures the case in which the focus of differential privacy is on hiding *participation* of individuals in a database. In our work, we consider both the participation and the *values* of the participants.

Clarkson and Schneider also considered differential privacy as a case study of their proposal for quantification of integrity [9]. There, the authors analyze

[6] When discussing our result with Barthe and Köpf, they said that they also conjectured that $f(u, \epsilon)$ is the optimal bound.

database privacy conditions from the literature (such as differential privacy, k-anonymity, and l-diversity) using their framework for utility quantification. In particular, they study the relationship between differential privacy and a notion of leakage (which is different from ours - in particular their definition is based on Shannon entropy) and they provide a tight bound on leakage.

Heusser and Malacaria [15] were among the first to explore the application of information-theoretic concepts to databases queries. They proposed to model database queries as programs, which allows for statical analysis of the information leaked by the query. However [15] did not attempt to relate information leakage to differential privacy.

In [14] the authors aim at obtaining optimal-utility randomization mechanisms while preserving differential privacy. The authors propose adding noise to the output of the query according to the geometric mechanism. Their framework is very interesting in the sense it provides a general definition of utility for a mechanism M that captures any possible side information and preference (defined as a loss function) the users of M may have. They prove that the geometric mechanism is optimal in the particular case of counting queries. Our results in Section 6 do not restrict to counting queries, but on the other hand we only consider the case of binary loss function.

8 Conclusion and future work

In this paper we have investigated the relation between ϵ-differential privacy and leakage, and between ϵ-differential privacy and utility. Our main contribution is the development of a general technique for determining these relations depending on the graph structure induced by the adjacency relation and by the query. We have considered two particular structures, the distance-regular graphs, and the VT^+ graphs, which allow to obtain tight bounds on the leakage and on the utility, and to construct the optimal randomization mechanism satisfying ϵ-differential privacy.

As future work, we plan to extend our result to other kinds of utility functions. In particular, we are interested in the case in which the the answer domain is provided with a metric, and we are interested in taking into account the degree of accuracy of the inferred answer.

References

1. Mário S. Alvim, Miguel E. Andrés, Konstantinos Chatzikokolakis, Pierpaolo Degano, and Catuscia Palamidessi. Differential privacy: on the trade-off between utility and information leakage. Technical report, 2011. `http://hal.inria.fr/inria-00580122/en/`.
2. Mário S. Alvim, Konstantinos Chatzikokolakis, Pierpaolo Degano, and Catuscia Palamidessi. Differential privacy versus quantitative information flow. Technical report, 2010.

3. Miguel E. Andrés, Catuscia Palamidessi, Peter van Rossum, and Geoffrey Smith. Computing the leakage of information-hiding systems. In *Proc. of TACAS*, volume 6015 of *LNCS*, pages 373–389. Springer, 2010.
4. Gilles Barthe and Boris Köpf. Information-theoretic bounds for differentially private mechanisms. In *Proc. of CSF*, 2011. To appear.
5. Michele Boreale, Francesca Pampaloni, and Michela Paolini. Asymptotic information leakage under one-try attacks. In *Proc. of FOSSACS*, volume 6604 of *LNCS*, pages 396–410. Springer, 2011.
6. Christelle Braun, Konstantinos Chatzikokolakis, and Catuscia Palamidessi. Compositional methods for information-hiding. In *Proc. of FOSSACS*, volume 4962 of *LNCS*, pages 443–457. Springer, 2008.
7. Christelle Braun, Konstantinos Chatzikokolakis, and Catuscia Palamidessi. Quantitative notions of leakage for one-try attacks. In *Proc. of MFPS*, volume 249 of *ENTCS*, pages 75–91. Elsevier, 2009.
8. Konstantinos Chatzikokolakis, Catuscia Palamidessi, and Prakash Panangaden. Probability of error in information-hiding protocols. In *Proc. of CSF*, pages 341–354. IEEE, 2007.
9. M. R. Clarkson and F. B. Schneider. Quantification of integrity, 2011. Tech. Rep.. `http://hdl.handle.net/1813/22012`.
10. Cynthia Dwork. Differential privacy. In *Automata, Languages and Programming, 33rd Int. Colloquium, ICALP 2006, Venice, Italy, July 10-14, 2006, Proc., Part II*, volume 4052 of *LNCS*, pages 1–12. Springer, 2006.
11. Cynthia Dwork. Differential privacy in new settings. In *Proc. of the Twenty-First Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2010, Austin, Texas, USA, January 17-19, 2010*, pages 174–183. SIAM, 2010.
12. Cynthia Dwork. A firm foundation for private data analysis. *Communications of the ACM*, 54(1):86–96, 2011.
13. Cynthia Dwork and Jing Lei. Differential privacy and robust statistics. In *Proc. of the 41st Annual ACM Symposium on Theory of Computing, STOC 2009, Bethesda, MD, USA, May 31 - June 2, 2009*, pages 371–380. ACM, 2009.
14. Arpita Ghosh, Tim Roughgarden, and Mukund Sundararajan. Universally utility-maximizing privacy mechanisms. In *Proc. of the 41st annual ACM symposium on Theory of computing*, STOC '09, pages 351–360. ACM, 2009.
15. Jonathan Heusser and Pasquale Malacaria. Applied quantitative information flow and statistical databases. In *Proc. of the Int. Workshop on Formal Aspects in Security and Trust*, volume 5983 of *LNCS*, pages 96–110. Springer, 2009.
16. Boris Köpf and Geoffrey Smith. Vulnerability bounds and leakage resilience of blinded cryptography under timing attacks. In *Proc. of CSF*, pages 44–56. IEEE, 2010.
17. Annabelle McIver, Larissa Meinicke, and Carroll Morgan. Compositional closure for bayes risk in probabilistic noninterference. In *Proc. of ICALP*, volume 6199 of *LNCS*, pages 223–235. Springer, 2010.
18. Alfréd Rényi. On Measures of Entropy and Information. In *Proc. of the 4th Berkeley Symposium on Mathematics, Statistics, and Probability*, pages 547–561, 1961.
19. Geoffrey Smith. On the foundations of quantitative information flow. In *Proc. of FOSSACS*, volume 5504 of *LNCS*, pages 288–302. Springer, 2009.