ARTICLE IN PRESS

Discrete Applied Mathematics ▮ (▮▮▮▮) ▮▮▮–▮▮▮

Contents lists available at ScienceDirect

Discrete Applied Mathematics

journal homepage: www.elsevier.com/locate/dam

Differentially private response mechanisms on categorical data

Naoise Holohan [a], Douglas J. Leith [a], Oliver Mason [b,c,*]

[a] School of Computer Science and Statistics, Trinity College Dublin, Ireland

[b] Department of Mathematics and Statistics/Hamilton Institute, Maynooth University-National University of Ireland Maynooth, Co. Kildare, Ireland

[c] Lero, The Irish Software Research Centre, Ireland

ARTICLE INFO

Article history:
Received 28 April 2015
Received in revised form 11 March 2016
Accepted 16 April 2016
Available online xxxx

Keywords:
Data privacy
Differential privacy
Optimal mechanisms

ABSTRACT

We study mechanisms for differential privacy on finite datasets. By deriving *sufficient sets* for differential privacy we obtain necessary and sufficient conditions for differential privacy, a tight lower bound on the maximal expected error of a discrete mechanism and a characterisation of the optimal mechanism which minimises the maximal expected error within the class of mechanisms considered.

1. Introduction

Data privacy has been of interest to researchers for decades [4], but high-profile privacy breaches in recent years, such as those involving AOL [1] and Netflix [18], have renewed focus on the topic. The movement towards smart metering systems for electricity, water and other utilities and the greater use of data mining in so-called smart cities and transport have given rise to further concerns over personal data privacy.

The most traditional framework for the study of data privacy is that of tabular data. A simple model of this type considers the data to be arranged as individual records within a table, where each record contains entries from some underlying dataset, which may be continuous or discrete depending on the type of data being studied. Simple anonymisation techniques, such as removing names and social security numbers (so-called unique identifiers) from the data, have been shown to be inadequate [19]. More sophisticated frameworks such as k-anonymity [19] and ℓ-diversity [15] are also vulnerable to privacy attacks via the use of appropriate side-information or data from external sources [15,14].

Within the last decade, differential privacy [5] has emerged as a popular framework for research in the field of data privacy based on its capability to provide a quantifiable basis for privacy preserving data publishing and mining. This is a probabilistic approach to data privacy in which a suitably *randomised* version of the correct response to a query is released. The core idea is founded on the simple premise that the response to a user query should not be too tightly coupled with any one entry in the table. One widely-adopted implementation of differential privacy for real-valued databases is to add an appropriate amount of noise sampled from a Laplace distribution to each cell of the database [6].

Much research on differential privacy to date has been completed on real-valued databases [6], although a considerable body of literature also exists on discrete data [16,3]; in particular some recent work has focused on graph data relevant to applications in areas such as social networks [13,2].

* Corresponding author. Tel.: +353 01 7083672; fax: +353 501 7083913.
E-mail address: oliver.mason@nuim.ie (O. Mason).

http://dx.doi.org/10.1016/j.dam.2016.04.010
0166-218X/

Please cite this article in press as: N. Holohan, et al., Differentially private response mechanisms on categorical data, Discrete Applied Mathematics (2016), http://dx.doi.org/10.1016/j.dam.2016.04.010

Differentially private mechanisms can be divided into two distinct classes: sanitisation based mechanisms; and output perturbation based mechanisms. Our concern here is with the former class, which first constructs a sanitised version of the database and then answers queries on this. It has been shown in [11] that if the sanitised database satisfies the requirements of differential privacy, then any query can be answered on it in a differentially private manner.

In writing this paper, we have two aims: the first is to present a set of new results on the mathematical foundations of differential privacy for discrete data; the second is to bring the problems in this field to the attention of researchers in discrete applied mathematics.

We examine differentially private mechanisms for discrete data within the general probabilistic framework described in our previous paper [11]. As we deal with finite datasets here, many of the measure-theoretic details required for the more general setting can be suppressed. However, to properly set context, we include the more general definitions here in Section 2.

Our first results concern an adaptation for discrete data of the exponential mechanism introduced by McSherry and Talwar. In particular, we consider the problem of *sufficient sets* for differential privacy for this mechanism. This problem is motivated by the practical issue of testing whether or not a mechanism is differentially private and arises from the following simple considerations.

For a sanitisation to be differentially private, certain inequalities (described formally later) must hold on all subsets of the database space, which can necessitate checking a prohibitively large collection of sets in order to test for differential privacy. The question of sufficient sets asks whether it is sufficient for the differential privacy condition to hold on a collection of these subsets for it to hold on all subsets. We can therefore reduce the workload required to check that a mechanism satisfies differential privacy. In Section 3, we present results characterising sufficient sets for the discrete exponential mechanism. We then use these to give necessary and sufficient conditions for differential privacy for this mechanism.

A major concern of privacy research is the trade-off between privacy and accuracy. For the current setting, in the absence of a given metric on the dataset, we measure the error of a sanitisation using hamming distance; in Theorem 5 we derive a tight lower bound on the maximal expected error of a discrete exponential mechanism.

In Section 4 we consider a seemingly unrelated approach to database sanitisation: product sanitisations. We show that these are in fact equivalent to the discrete exponential mechanism constructed using the hamming distance and, building on results in [11], we characterise differential privacy and the error for these in Theorems 8 and 9 respectively. Finally in Theorem 10 we provide a characterisation of the optimal product sanitisation mechanism, which minimises the maximal expected error within the class of product sanitisations (and hence within the class of discrete exponential mechanisms). Concluding remarks are given in Section 5.

1.1. Related work

Before the advent of differential privacy, Fienberg examined the use of data swapping and cell suppression for privacy protection on categorical data [9]. Dwork then presented the notion of differential privacy in [5], and its limitations were discussed by Dankar in [7], including its applicability to categorical data.

Dwork's work was closely followed by McSherry and Talwar who proposed the exponential mechanism in [16]. An instantiation of this was used by Hardt and Talwar [10] in examining the geometry of differential privacy. Mohammed made use of the exponential mechanism for releasing count queries in [17,3], while a more recent contribution has looked at differential privacy on counts using a combination of the Laplace and exponential mechanisms [21].

2. Preliminaries

2.1. Database model

We consider a finite data set D with $(m+1)$ elements ($m \geq 1$). A database $\mathbf{d}$ with n rows drawn from this data set is represented by a vector $\mathbf{d} = (d_1, \ldots, d_n) \in D^n$. D is equipped with a σ-algebra, in this case the power set 2^D and D^n inherits the product σ-algebra, 2^{D^n}. We are therefore considering all subsets of D and D^n.

We will consider hamming distance on D^n. Recall that the hamming distance, $\mathrm{h} : D^n \times D^n \to \{0, 1, \ldots, n\}$, between two databases is the number of rows on which they differ:

$$\mathrm{h}(\mathbf{d}, \mathbf{d}') = |\{i : d_i \neq d'_i\}|. \tag{1}$$

Definition 1 (*Neighbours*). Two databases $\mathbf{d}, \mathbf{d}' \in D^n$ are said to be *neighbours*, written $\mathbf{d} \sim \mathbf{d}'$ if $\mathrm{h}(\mathbf{d}, \mathbf{d}') = 1$.

Informally, two databases are neighbours if they differ on exactly one row.

2.2. Query model

We make use of the generalised query model introduced in [11], adapted to the discrete setting. A query $Q : D^n \to E_Q$ outputs a response in E_Q, the structure of which is not specified (it may be numeric, categorical, functional, etc.). E_Q is, however, equipped with a σ-algebra $\mathcal{A}_Q$. We require that all queries be measurable, which is trivial in this setting since $Q^{-1}(A) \subseteq D^n$ for all $A \in \mathcal{A}_Q$.

Please cite this article in press as: N. Holohan, et al., Differentially private response mechanisms on categorical data, Discrete Applied Mathematics (2016), http://dx.doi.org/10.1016/j.dam.2016.04.010