Buying Private Data at Auction: The Sensitive Surveyor's Problem

AARON ROTH
University of Pennsylvania

In this letter, we survey some recent work on what we call the *sensitive surveyor's problem.* A curious data analyst wishes to survey a population to obtain an accurate estimate of a simple population statistic: for example, the fraction of the population testing positive for syphilis. However, because this is a statistic over sensitive data, individuals experience a cost for participating in the survey as a function of their loss in privacy. Agents must be compensated for this cost, and moreover, are strategic agents and will mis-report their cost if doing so is beneficial for them. The goal of the surveyor is to manage the inevitable tradeoff between the cost of the survey, and the accuracy of its results.

Categories and Subject Descriptors: J.4 [**Social and Behavioral Sciences**]: Economics

General Terms: Privacy, Mechanism Design

1. INTRODUCTION

Consider the following stylized problem of the sensitive surveyor Alice[1]. She is tasked with conducting a survey of a set of n individuals N, to determine what proportion of the individuals $i \in N$ satisfy some property $P(i)$. Her ultimate goal is to discover the true value of this statistic, $s = \frac{1}{n} \cdot |\{i \in N : P(i)\}|$, but if that is not possible, she will be satisfied with some estimate $\hat{s}$ such that the error, $|\hat{s} - s|$, is minimized. We will adopt a notion of accuracy based on large deviation bounds, and say that a surveying mechanism is α-accurate if $\Pr[|\hat{s} - s| \geq \alpha] \leq \frac{1}{3}$. The inevitable catch is that individuals will not participate in the survey for free. $P(i)$ may be an embarrassing or otherwise sensitive predicate (e.g. it may represent the presence of a disease, a juvenile taste in movies, a political stance, or any other property that might cause its owner to come to grief if it were to be revealed). Therefore, individuals will experience some *cost* as a function of their loss in privacy when they interact with Alice, and will insist on being compensated for this loss. To make matters worse, these individuals are rational (i.e. selfish) agents, and are apt to mis-report their costs to Alice if doing so will result in a financial gain. This places the sensitive surveyor's problem squarely in the domain of mechanism design, and requires Alice to develop a scheme for trading off statistical accuracy with cost, all while managing the incentives of the individuals in N.

Before we go on to describe the recent work in this area, we remark that this stylized problem is not only relevant to a surveyor, but to any organization that

[1] It is the data that is sensitive, not necessarily the surveyor herself. As we will see, Alice may or may not be sensitive.

Email address: aaroth@cis.upenn.edu. Supported in part by the National Science Foundation, award number CCF-1101389.

makes use of collections of potentially sensitive data to provide some service, or to otherwise extract value from it. This includes, for example, the use of search logs to provide search query completion and the use of browsing history to improve search engine ranking, the use of social network data to select display ads and to recommend new links, and the myriad other data-driven services now available on the web. In all of these cases, value is being derived from the statistical properties of collections of sensitive data in exchange for some payment[2].

Collecting data in exchange for some fixed price will inevitably lead to a biased estimate of population statistics, because such a scheme will result in collecting data only from those individuals who value their privacy less than the price being offered. To obtain an accurate estimate of the statistic, it is therefore natural to consider buying private data using an auction, which was recently considered by Ghosh and Roth [2011]. There are two obvious obstacles which one must confront when conducting an auction for private data, and one additional obstacle which is less obvious but more insidious. The first obstacle is that one must have a quantitative formalization of "privacy" which can be used to measure agents' costs under various operations on their data. We discuss this in the next section: the short answer is that we are fortunate that the internet-scale use of private data has dovetailed with the development of the notion of *differential privacy*, which is an excellent quantitative formalization for how one may trade off privacy for utility. The second obstacle is that the objective that we wish to trade off with cost is *statistical accuracy*, which is distinct from the objectives commonly studied in mechanism design.

The final, more insidious obstacle, is that we expect that an individual's cost for privacy loss should be highly correlated with his private data itself! Suppose, for example, that Bob reports a low value for privacy, but after being observed to visit an oncologist, Bob revises his value for privacy to be much higher. Although we only know Bob's value for privacy, and have not explicitly been shown his medical records, this is disclosive because Bob's cancer status is likely correlated with his value for privacy. More to the point, suppose that in the first step of a survey of cancer prevalence, we ask each individual to report their value for privacy, with the intention of then running an auction to choose which individuals to buy data from. If agents report truthfully, we may find that the reported values naturally form two clusters: low value agents, and high value agents. In this case, we may have learned something about the population statistic even before collecting any data or making any payments – and therefore, the agents will have already experienced a cost. As a result, the agents may not be incentivized to report their true values, and this could again serve to introduce a bias in the survey results. This phenomenon makes direct revelation mechanisms problematic in auctions for private data, and it is what most distinguishes this problem from classical mechanism design.

2. DIFFERENTIAL PRIVACY: QUANTIFYING PRIVACY'S UTILITY

"Differential Privacy" takes the position that privacy is a property of a process, and not a property of a piece of information. It asserts that people should care about

[2] The payment need not be explicit and dollar denominated. It may instead be, for example, the use of a "free" service.

the privacy of their data to the extent that the use of that data causes additional harm to befall them, compared to if their data was not used at all. The following definition of differential privacy is syntactically different from the one originally given in [Dwork et al. 2006; Dwork 2006], but is easily seen to be equivalent, and is particularly well suited to its use in mechanism design[3]. Let us suppose that each individual has a piece of private data drawn from some abstract domain X, and that a private database is a collection of n such records $D \in X^n$. Two databases $D, D' \in X^n$ are said to be *neighboring* if they differ in at most one coordinate (i.e. if they differ only in the data of a single individual). An algorithm is then a randomized mapping $M : X^n \to T$ to some abstract range T of outcomes.

Definition 2.1. An algorithm $M : X^n \to T$ is ϵ-differentially private if for all pairs of neighboring databases D and D', and for all utility functions $u : T \to \mathbb{R}$:

$$\exp(-\epsilon)\mathbb{E}[u(M(D'))] \leq \mathbb{E}[u(M(D))] \leq \exp(\epsilon)\mathbb{E}[u(M(D'))]$$

ϵ is typically taken to be some value < 1, and so $\exp(-\epsilon), \exp(\epsilon)$ should be thought of as factors of $(1 - \epsilon)$ and $(1 + \epsilon)$ respectively. In words, what a promise of differential privacy guarantees is that simultaneously for all agents, no matter what utility function they may have over the outcome of the mechanism, participation in a computation M cannot negatively (or positively) change that expected utility by more than a $(1 + \epsilon)$ factor.

This naturally motivates a way to measure the cost to an agent of allowing their private data to be used by some mechanism M: If D is the database that includes agent i's data, and D' is the identical database with agent i's data removed, then the cost to agent i of participating is $c_i = \mathbb{E}[u(M(D))] - \mathbb{E}[u(M(D'))]$. If M is differentially private, then this cost is guaranteed to be bounded by $\epsilon\mathbb{E}[u(M(D'))]$. This motivates a particularly simple linear form of privacy cost in terms of the differential privacy guarantee ϵ with which his data is protected: $c_i(\epsilon) = v_i\epsilon$, where v_i is a privately known value parameter[4].

3. DIRECT REVELATION MECHANISMS

Armed with a means of quantifying an agent i's loss for allowing his data to be used by an ϵ-differentially-private algorithm ($c_i(\epsilon) = \epsilon \cdot v_i$), we are almost ready to describe results for the sensitive surveyor's problem. It remains to define what exactly the data domain X is. We will consider two models. In both models, we will associate with each individual a bit $b_i \in \{0, 1\}$ which represents whether they satisfy the sensitive predicate $P(i)$, as well as a value for privacy $v_i \in \mathbb{R}^+$.

[3] McSherry and Talwar [2007] were the first to observe that differential privacy implied this alternative formulation and used it in the context of mechanism design. I first saw this formulation presented as the *definition* of differential privacy in a talk given by Kobbi Nissim on his work in privacy and mechanism design [Nissim et al. 2012; Nissim, Orlandi, and Smorodinsky 2012].

[4] Although this form of utility is simple and is what is used in Ghosh and Roth [2011], it can be problematic. Indeed, the models in [Ligett and Roth 2012; Fleischer and Lyu 2012; Nissim, Orlandi, and Smorodinsky 2012] use slightly modified models of privacy cost, and those in [Xiao 2011; Chen et al. 2011] use a significantly different measure. The "right" measure for privacy utility is still up for debate: see [Nissim, Orlandi, and Smorodinsky 2012; Ligett and Roth 2012] for further discussion on this issue.

(1) In the *insensitive value model*, we calculate the ϵ parameter of the private mechanism by letting its domain be $X^n = \{0,1\}^n$: i.e. we measure privacy cost only with respect to how the mechanism treats the sensitive bit b_i, and ignore how it treats the reported values for privacy, v_i.

(2) In the *sensitive value model*, we calculate the ϵ parameter of the private mechanism by letting its domain be $X^n = (\{0,1\} \times \mathbb{R}^+)^n$: i.e. we measure privacy cost with respect to how it treats the pair (b_i, v_i) of bit/value pairs for each individual.

Intuitively, the insensitive value model treats individuals as ignoring the potential privacy loss due to correlations between their values for privacy and their private bits, whereas the sensitive value model treats individuals as assuming these correlations are worst-case, and that their values v_i are just as disclosive as their private bits b_i. The main results of Ghosh and Roth [2011] are that in the insensitive value model, it is possible to derive approximately optimal direct revelation mechanisms that achieve high accuracy and low cost. On the other hand, in the *sensitive value model*, no individually rational direct revelation mechanism can achieve any non-trivial accuracy.

Note that here we are considering a setting in which private data and costs are adversarially chosen. If we are willing to assume a known prior on agent costs (but still assume adversarially chosen private bits b_i), then it is possible to improve on the results of Ghosh and Roth [2011], and derive Bayesian optimal mechanisms for the sensitive survey problem. This is considered in Roth and Schoenebeck [2012], but we will not have space to discuss these results in the present note.

4. TAKE IT OR LEAVE IT MECHANISMS

Given the impossibility result proven in Ghosh and Roth [2011] for the sensitive value model, the immediate question is how we may circumvent it. In recent work, two methods have been proposed, which we briefly summarize here. Both approaches abandon direct revelation mechanisms in favor of mechanisms which offer individuals binary take-it-or-leave-it offers, but both also require subtle changes in how individuals are modeled as valuing privacy. We will not have space to explain these subtleties here, but readers are directed to the papers Fleischer and Lyu [2012; Ligett and Roth [2012] for more details.

4.1 Circumventing Impossibility with a Sensitive Surveyor

Suppose an individual is approached with a take it or leave it offer: "If you let us use your bit b_i in an ϵ-differentially private manner, I will give you $10." An individual might be reluctant to respond to such an offer, because the very act of responding might reveal whether his value v_i is such that $v_i \geq 10/\epsilon$ or not, and if values are correlated with private data, this might reveal something about his bit b_i beyond that which is revealed through the differentially private computation. To model such correlations, Fleischer and Lyu [2012] assume that each individual's value v_i is drawn independently from one of two known priors: $v_i \sim F_0$ if $b_i = 0$, and $v_i \sim F_1$ if $b_i = 1$. Alice, the surveyor, knows both priors, but does not know whether $b_i = 0$ or $b_i = 1$. Using this assumption, Fleischer and Lyu use an elegant idea which allows Alice to make a take-it-or-leave-it offer which an agent can truthfully decide to accept or reject without revealing *anything* about his private bit! The idea is this: Alice may choose some acceptance probability $q \in [0, 1]$. Although she does not know the bit of the agent she is surveying, she can pick two values p_0, p_1 such that $\Pr_{v \sim F_0}[v \leq p_0/\epsilon] = q$ and $\Pr_{v \sim F_1}[v \leq p_1/\epsilon] = q$. Alice can then offer the following take-it-or-leave-it offer to each agent: "If you accept the offer and your (verifiable) bit is 0, I will pay you p_0 dollars. If you accept the offer and your bit is 1 I will pay you p_1 dollars." The beauty of this solution is that no matter what private bit the agent has, he will accept the offer with probability q (where the probability is taken over the draw of his value from the corresponding prior) and reject the offer with probability $(1 - q)$. Therefore, *nothing* can be learned about his private bit from his participation decision, and so he has no incentive not to respond to the offer truthfully. Using this idea, Fleischer and Lyu [2012] develop approximately optimal truthful take-it-or-leave-it mechanisms that can be used whenever exact priors F_0 and F_1 are known. The solution is to make Alice's query to each agent more sensitive to their privacy concerns so that their participation decisions do not reveal their private data.

4.2 Circumventing Impossibility with an Insensitive Surveyor

What if agent costs are again determined adversarially, and there are no known priors? Ligett and Roth [2012] give an alternative solution for this case, again based on making take-it-or-leave-it offers. To circumvent the impossibility result of Ghosh and Roth [2011], Alice is here granted with one additional power: the ability to accost random members of the population on the street, and present them with a take-it-or-leave-it offer. Once individuals are presented with an offer, they are free to accept it or refuse it however they see fit. But they may not choose to have never even heard the offer, and if they reject the offer (perhaps just by walking away), their non-participation decision is observed by Alice. This can be seen as a weakening of the individual rationality condition: because costs may be correlated with private data, merely by rejecting an offer and walking away, Alice may learn something about the surveyed individual. If the individual did not accept the offer, he receives no payment, and yet still experiences some cost! This ends up giving a semi-truthfulness guarantee. Whenever Alice makes an offer of p dollars in exchange for ϵ-differential privacy, a rational agent will accept whenever $p \geq \epsilon v_i$. On the other hand, rational agents may or may not accept offers that are below their cost – because they will still experience some cost by walking away. But these deviations away from "truthfulness" are in only one direction, and only help Alice, whose aim it is to compute an accurate population statistic, and does not necessarily care about protecting privacy for its own sake. Here, Ligett and Roth [2012] are able to again obtain non-trivial accuracy (circumventing the impossibility result of Ghosh and Roth [2011]) even in the sensitive value model by making Alice insensitive to the privacy concerns of the agents she surveys, by making offers that they can refuse (but can't avoid).

5. OTHER RELATED WORK

This note has discussed only the sensitive surveyors problem, but it should be noted that there is a significant amount of work on other problems at the intersection of privacy and mechanism design. Differential privacy was first used as a tool in mechanism design (and proposed as a solution concept) by McSherry and Talwar [2007]. Gupta et al. [2010] also used differential privacy as a solution concept. These results produced only approximately truthful mechanisms: recently Nissim et al. [2012] showed how to combine differentially private mechanisms with "imposition mechanisms" to derive new mechanisms which are exactly truthful and do not require payments. In recent insightful work, Xiao [2011] studied the problem of mechanism design for agents who explicitly value privacy, and asked whether mechanisms of the sort studied in Nissim et al. [2012] could be made truthful even in the presence of such agents. Recently, Nissim, Orlandi, and Smorodinsky [2012] and Chen et al. [2011] gave positive answers to this question, each in a slightly different model.

REFERENCES

Chen, Y., Chong, S., Kash, I., Moran, T., and Vadhan, S. 2011. Truthful mechanisms for agents that value privacy. *Arxiv preprint arXiv:1111.5472*.

Dwork, C. 2006. Differential privacy. In *Proceedings of the 33rd International Colloquium on Automata, Languages and Programming.* LECTURE NOTES IN COMPUTER SCIENCE, vol. 4052. Springer, 1.

Dwork, C., McSherry, F., Nissim, K., and Smith, A. 2006. Calibrating noise to sensitivity in private data analysis. In *Proceedings of the Third Theory of Cryptography Conference TCC.* Lecture Notes in Computer Science, vol. 3876. Springer, 265.

Fleischer, L. and Lyu, Y.-H. 2012. Approximately optimal auctions for selling privacy when costs are correlated with datan. In *EC 2012: Proceedings of the 13th ACM conference on Electronic commerce.*

Ghosh, A. and Roth, A. 2011. Selling privacy at auction. In *EC 2011: Proceedings of the 12th ACM conference on Electronic commerce.* ACM, 199–208.

Gupta, A., Ligett, K., McSherry, F., Roth, A., and Talwar, K. 2010. Differentially Private Combinatorial Optimization. In *Proceedings of the ACM-SIAM Symposium on Discrete Algorithms.*

Ligett, K. and Roth, A. 2012. Take it or leave it: Running a survey when privacy comes at a cost. *Arxiv preprint arXiv:1202.4741*.

McSherry, F. and Talwar, K. 2007. Mechanism design via differential privacy. In *Proceedings of the 48th Annual Symposium on Foundations of Computer Science.*

Nissim, K., Orlandi, C., and Smorodinsky, R. 2012. Privacy-aware mechanism design. In *EC 2012: Proceedings of the 13th ACM conference on Electronic commerce.*

Nissim, K., Smorodinsky, R., and Tennenholtz, M. 2012. Approximately optimal mechanism design via differential privacy. In *ITCS 2012: Proceedings of the 3rd Innovations in Computers Science Conference.*

Roth, A. and Schoenebeck, G. 2012. Conducting truthful surveys, cheaply. In *EC 2012: Proceedings of the 13th ACM conference on Electronic commerce.*

Xiao, D. 2011. Is privacy compatible with truthfulness. Tech. rep., Cryptology ePrint Archive, Report 2011/005, 2011. http://eprint. iacr. org.