Differential Privacy for Belief Functions

Qiyu Li, Chunlai Zhou, B. Qin, Z. Xu

[1]Renmin University of China
[2]Mohamed bin Zayed University of Artificial Intelligence

BFTA2023, JAIST

Outline

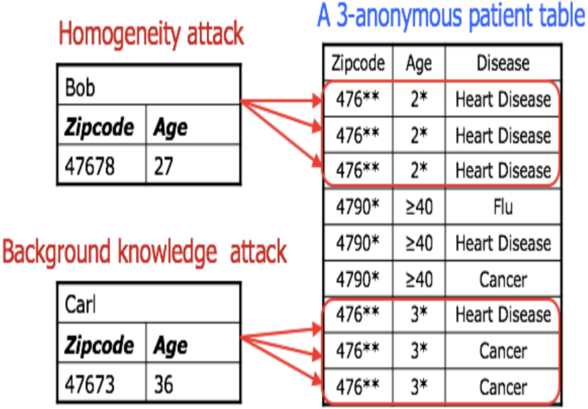
Outline

Motivation: Differential Privacy

- Differential Privacy (Cythia Dwork): "*Privacy comes from randomization*"

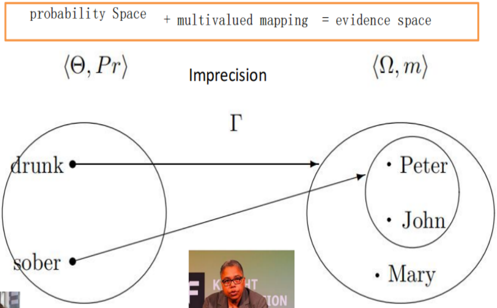
Motivation: k-anonymity

k-anonymity (Latanya Sweeney): the information for each person contained in the release cannot be distinguished from at least $k-1$ individuals whose information also appear in the release.



Motivation: what is the benefits of imprecision in privacy-preserving?

- Our work: Privacy comes from randomization + imprecision



Differential Privacy + k-Anonymoty = ????

Belief functions

Definition

Let Ω be a frame of discernment and $\mathcal{A} = 2^{\Omega}$ be the Boolean algebra of events. A mass assignment (or mass function) is a mapping $m : \mathcal{A} \to [0, 1]$ satisfying $\sum_{A \in \mathcal{A}} m(A) = 1$. A mass function m is called *normal* if $m(\emptyset) = 0$.

A belief function is a function $bel : \mathcal{A} \to [0, 1]$ satisfying the following conditions:

1. $bel(\emptyset) = 0$;
2. $bel(\Omega) = 1$; and
3. $bel(\bigcup_{i=1}^{n} A_i) \geq \sum_{\emptyset \neq I \subseteq \{1, \cdots, n\}} (-1)^{|I|+1} bel(\cap_{i \in I} A_i)$ where $A_i \in \mathcal{A}$ for all $i \in \{1, \cdots, n\}$.

Belief functions

Definition

Let Ω be a frame of discernment and $\mathcal{A} = 2^{\Omega}$ be the Boolean algebra of events. A mass assignment (or mass function) is a mapping $m : \mathcal{A} \to [0,1]$ satisfying $\sum_{A \in \mathcal{A}} m(A) = 1$. A mass function m is called *normal* if $m(\emptyset) = 0$.

A belief function is a function $bel : \mathcal{A} \to [0,1]$ satisfying the following conditions:

1. $bel(\emptyset) = 0$;
2. $bel(\Omega) = 1$; and
3. $bel(\bigcup_{i=1}^{n} A_i) \geq \sum_{\emptyset \neq I \subseteq \{1,\cdots,n\}} (-1)^{|I|+1} bel(\cap_{i \in I} A_i)$ where $A_i \in \mathcal{A}$ for all $i \in \{1, \cdots, n\}$.

Equivalent Characterizations

Theorem

A mapping $f : \mathcal{A} \to [0,1]$ is a belief function if and only if its Möbius transform is a mass assignment.

In other words,

- if $m : \mathcal{A} \to [0,1]$ is a mass assignment, then it determines a belief function $bel : \mathcal{A} \to [0,1]$ as follows:

$$bel(A) = \sum_{B \subseteq A} m(B) \text{ for all } A \in \mathcal{A}.$$

- Moreover, given a belief function bel, we can obtain its corresponding mass function m as follows:

$$m(A) = \sum_{B \subseteq A} (-1)^{|A \setminus B|} bel(B) \text{ for all } A \in \mathcal{A}.$$

Interpretations

Plausibility function pl can be defined as

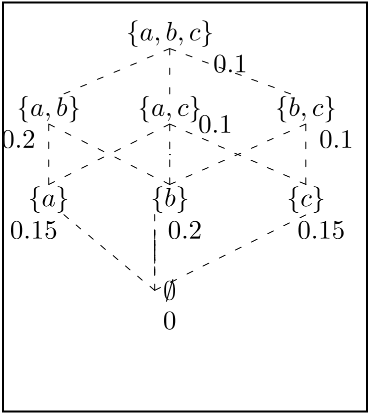
$$pl(A) := 1 - bel(\Omega \setminus A)$$

- $bel(A)$ measures the degree to which the evidence supports A,
- $pl(A)$ is the upper bound on the degree of support that could be assigned to A if more specific information became available
- $m(A)$ measures the belief that an agent commits exactly to A, not the total belief $bel(A)$ that an agent commits to A.

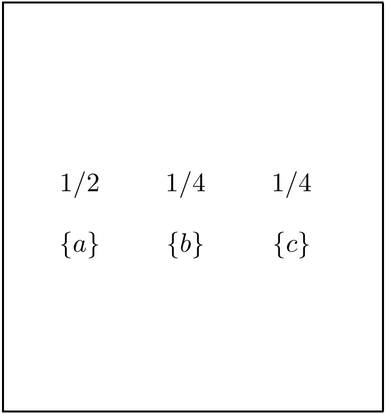
Belief function vs probability function

Let $\Omega := \{a, b, c\}$ be a frame.

mass function



probability function



Example

- A murder has been committed. There are three suspects $\Omega = \{\text{John, Mary, Peter}\}$.
- A witness saw the murderer going away in the darkness and he can only assert that it was a man. However, we know that the witness is drunk 20% of the time time.

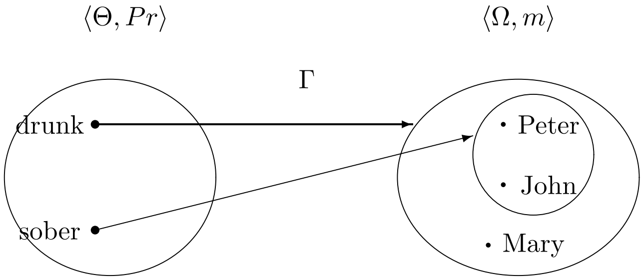
This piece of evidence can be represented by the following mass function:

$$m(\{John, Peter\}) = 0.8, \; m(\{\Omega\}) = 0.2.$$

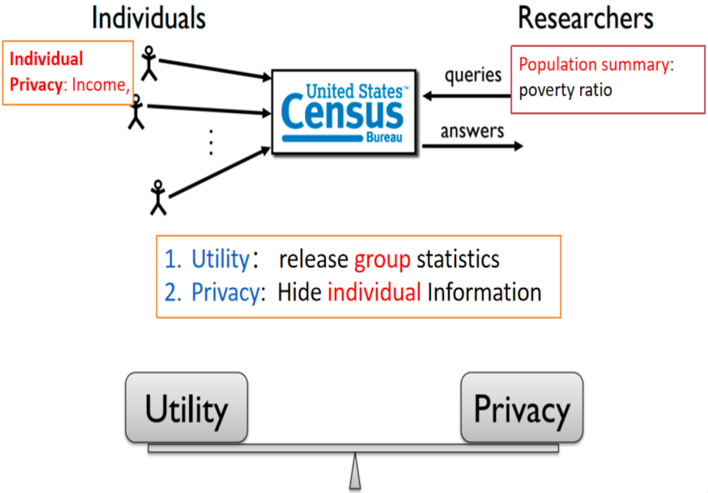
Represented as a Dempster model

The mass function m arises from the following Dempster model:

- a probability space $\langle \Theta, Pr \rangle$ where $\Theta = \{\text{drunk}, \text{sober}\}$ and $Pr(\text{drunk}) = 0.2$;
- a multivalued mapping $\Gamma : \Theta \to \Omega$ is illustrated as follows:

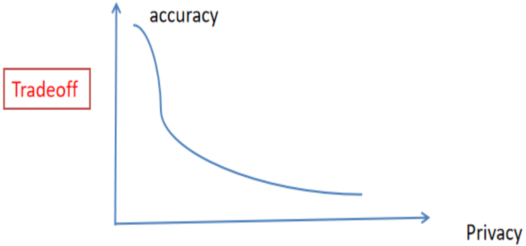


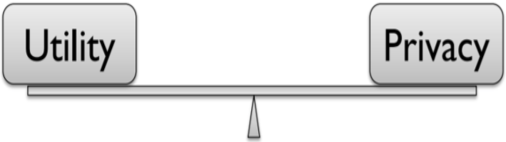
Two Goals in Data Analysis: Individual Privacy and Population Utility



Impossible Task: Perfect Privacy + Utility

- **Fundamental Law of Information Reconstruction** (Dwork and Roth 2014): "Overly accurate estimate of too many statistics is blatantly non-private".





Reconstruction attack (Dinur and Nissm 03)

$$\begin{bmatrix} f_1(X) \\ \vdots \\ f_k(X) \end{bmatrix} = \begin{bmatrix} \varphi_1(z_1) & \dots & \varphi_1(z_n) \\ \vdots & \ddots & \vdots \\ \varphi_k(z_1) & \dots & \varphi_k(z_n) \end{bmatrix} \begin{bmatrix} s_1 \\ \vdots \\ s_n \end{bmatrix}$$

Statistics public Privacy

Input: a set of query vectors $F_1, \ldots, F_k \in \{0,1\}^n$ and a set of answers $a_1, \ldots, a_k \in \mathbb{R}$
Output: a vector of secrets $\tilde{s} \in \{0,1\}^n$

Return $\tilde{s} \in \{0,1\}^n$ that *minimizes* the quantity $\max_{i \in [k]} |F_i \cdot \tilde{s} - a_i|$

Theorem 2.4 ([DN03]). *If all queries have error at most αn, then the reconstruction error (the number of entries on which $\tilde{s}$ and s disagree) is at most $4\alpha n$.*

Given answers to all these queries that are accurate to within 1%, we can recover the secret vector s that is correct for at least 95%.

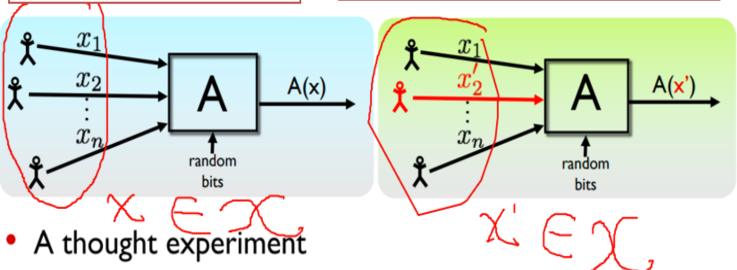
Preview of Differential Privacy

enough privacy → not too accurate estimates or not too many statistics

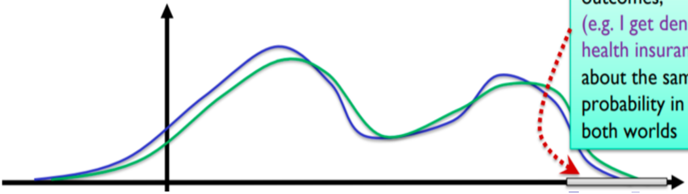
Intuition behind Differential Privacy

datasets x and x′ are neighbouring

Randomized algorithm. $A: \mathcal{X} \to \Delta(\mathcal{Y})$



- A thought experiment
  - Change one person's data (or add or remove them)
  - Will the **probabilities of outcomes** change?



For any set of outcomes, (e.g. I get denied health insurance) about the same probability in both worlds

Definition of Differential Privacy

- A randomized algorithm $A: \mathbb{X} \rightarrow \Delta(Y)$ is called ϵ-differentially private if, for all output y,

$$max_{x,x' \, neighbouring} \left| ln \frac{Pr[A(x)=y]}{Pr[A(x')=y]} \right| \leq \varepsilon$$

privacy loss on output y

Privacy budget

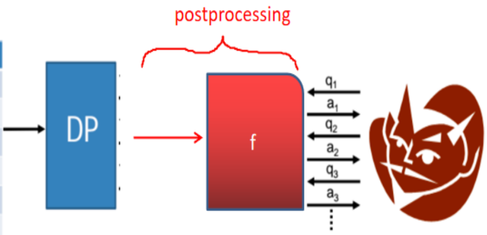
When ε is smaller, the privacy-preserving is better

Differential Privacy: Postprocessing

- If randomized algorithm A: $\mathbb{X} \rightarrow \Delta(Y)$ is ε-differentially private and f: $Y \rightarrow \Delta(Z)$ is a random mapping from Y to Z, then $f \circ A$: $\mathbb{X} \rightarrow \Delta(Z)$ is ε-Differentially private.

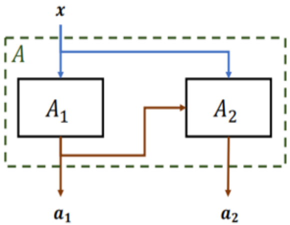
"A data analyst, without additional knowledge about the private database, cannot compute a function of the output of a private algorithm M and make it less differentially private"

Sex	Blood	⋯	HIV
F	B	⋯	Y
M	A	⋯	N
M	O	⋯	N
M	O	⋯	Y
F	A	⋯	N
M	B	⋯	Y



Differential Privacy: Composition

- If A_i:$\mathbb{X} \to \Delta(Y)$ is ε_i-differentially private, then A^n: $\mathbb{X} \to \Delta(Y)^n$ defined by A^n(x)= $(A_1(x), \cdots, A_n(x))$ is $\Sigma\varepsilon_i$-differentially private.

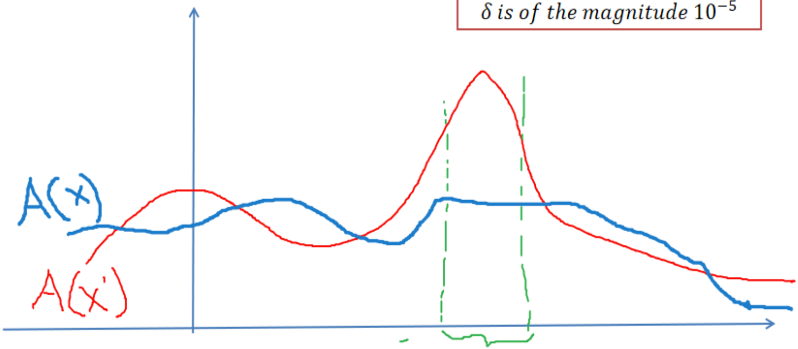


(Good news) It is important to construct more sophisticated privacy mechanisms.
(Bad news) When we compose more, the strength of the privacy guarantee degrade.

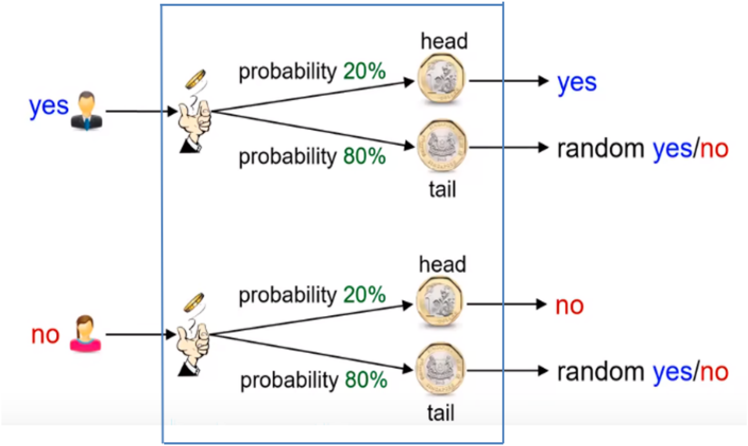
Approximate DP: (ϵ, δ)-DP

- A randomized algorithm A: $\mathbb{X} \rightarrow \Delta(Y)$ is (ε, δ)-DP if, for any neighbouring datasets x an x', and any observation $E \leq Y, Pr[A(x) \in E \,| \leq e^{\varepsilon} Pr[A(x') \in E] + \delta$.

δ *is of the magnitude* 10^{-5}



The First DP: Warner's Randomized Response



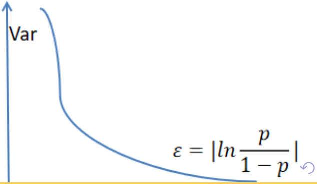
If we observe Yes, then the privacy loss $ln \frac{Pr[W(yes)=yes]}{\ldots} = ln \frac{0.6}{\ldots} = ln\, 1.5$

Warner's Mechanism: Distribution Estimation Problem

- **Step 1**: The percentage of the population has the sensitive property is π
- **Step 2**: Sample n individuals from the population and present them with the Warner's mechanism to protect their privacy.
- **Step 3**: Collect the noisy responses and apply the MLE.

$$Var[\hat{\pi}] = \frac{-(\pi - \frac{1}{2})^2 + \frac{1}{4}}{n} + \frac{\frac{1}{4(2p-1)^2} - \frac{1}{4}}{n}$$

where p is the property of answering truthfully



f-Diferential Privacy: Hypothesis-testing

f-differential privacy: **this talk**

- Interpreting privacy via hypothesis testing
- Privacy measure: type I and II errors *trade-off*
- Privacy *functional* parameter: $f : [0,1] \to [0,1]$
- How to achieve: adding *Gaussian* noise

(ϵ, δ)-differential privacy: **Dwork et al**

- Interpreting privacy via hypothesis testing
- Privacy measure: *worst-case* likelihood ratio
- Privacy parameters: $\epsilon \geqslant 0, 0 \leqslant \delta < 1$
- How to achieve: adding Laplace noise

f-Diferential Privacy: Hypothesis-testing (cont.)

$$H_0 : P \quad \textbf{vs} \quad H_1 : Q$$

For rejection rule $\phi \in [0,1]$, denote by $\alpha_\phi = \mathbb{E}_P[\phi]$ (type I error), and $\beta_\phi = 1 - \mathbb{E}_Q[\phi]$ (type II error)

Definition

For two probability distributions P and Q, define the trade-off function $T(P,Q) : [0,1] \to [0,1]$ as

$$T(P,Q)(\alpha) = \inf_{\phi} \{\beta_\phi : \alpha_\phi \leqslant \alpha\}$$

- The Neyman–Pearson lemma
- Function f is trade-off if and only if f is convex, continuous,

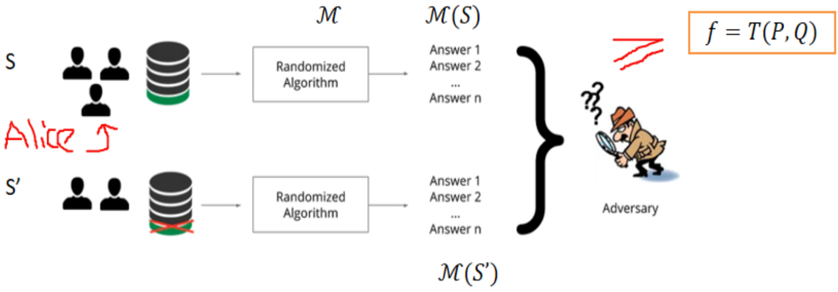
f-Diferential Privacy: Definition

Definition

A (randomized) algorithm $\mathcal{M}$ is said to be f-differentially private if

$$T(\mathcal{M}(S), \mathcal{M}(S')) \geqslant f$$

for all neighboring datasets S and S'



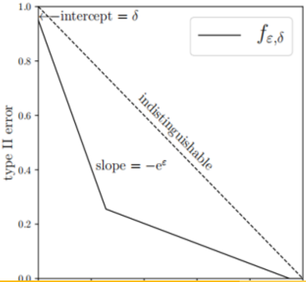
Connection to (ϵ, δ)-DP

(ϵ, δ)-DP requires that for any (measurable) event E

$$\mathbb{P}(\mathcal{M}(S) \in E) \leqslant e^{\epsilon}\, \mathbb{P}(\mathcal{M}(S') \in E) + \delta$$

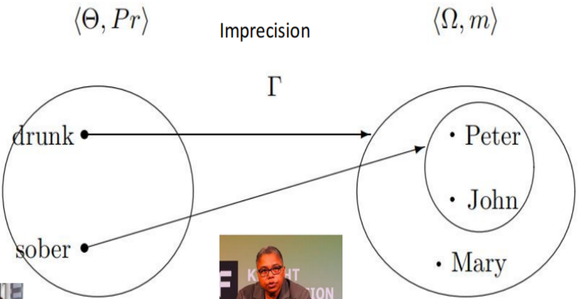
Proposition (Wasserman and Zhou '10)

Denote $f_{\epsilon,\delta}(\alpha) = \max\{0, 1-\delta-e^{\epsilon}\alpha, e^{-\epsilon}(1-\delta-\alpha)\}$. *An algorithm* $\mathcal{M}$ *is* (ϵ, δ)*-DP if and only if it is is* $f_{\epsilon,\delta}$*-DP*



It is natural to study DP for belief functions!

probability Space + multivalued mapping = evidence space



Differential Privacy + k-Anonymoty = ????

Evidential Privacy Mechanism

Let $X = \{x_1, \cdots, x_k\}$ be a private source of information and $Y = \{y_1, \cdots, y_l\}$ be an output alphabet.

- Q is called an *evidential* privacy mechanism if each row of the matrix Q is a mass function on Y
- In other words, each evidential privacy mechanism Q maps $X = x$ to $Y \in E$ with $Q(x)$ which can be represented by a mass $m_x^Q(E)$ (belief $bel_x^Q(E)$ or plausibility $pl_x^Q(E)$)

Total Belief Theorem (Smets1994, Zhou&Cuzzolin 2017)

For each $x \in X$, $bel_x^Q(E)$ can be identified with $\overrightarrow{(bel_x^Q)}^{X \times Y}$ and $bel^Q = \oplus_{x \in X} \overrightarrow{(bel_x^Q)}^{X \times Y}$. Let m^X be a prior mass function on X. So we obtain a total mass function $m^{X \times Y}$ over $X \times Y$ as $m^{X \times Y} = m^{X \uparrow (X \times Y)} \oplus m^Q$.

1. $bel^{X \times Y} \restriction_X = bel^X$, i.e., the marginal of the total belief function on X is the prior belief function on X;
2. The conditional total belief functions on individual input x (equivalently $\{x\} \times Y$) according to both Dempster's rule and geometric rule of conditioning are the same as the output belief function bel_x^Q associated with the input x in the privacy mechanism Q, i.e., for each $x \in X$, $bel_{\oplus}^Y(\cdot|\{x\}) = bel_x^Q$ and $bel_g^Y(\cdot|\{x\}) = bel_x^Q$.

An Uncertainty Framework for DP: Uncertainty Factor

- Let the adversary's prior uncertainty be represented with a set function U, which may be a mass, belief or plausibility function, and his posterior uncertainty after observing event $E \subseteq Y$ be represented by another set function U',
- the adversary's uncertainty change can be formulated as the following $F_{U',U}$ called uncertainty factor for the two inputs x and x'

$$F_{U',U}(x, x') := \frac{\frac{U'(x)}{U'(x')}}{\frac{U(x)}{U(x')}} \tag{2.1}$$

The denominator corresponds to the initial betting odds for x vs. x' before the observation. And the numerator is the betting odds afterwards.

DP in terms of evidential factors

DP has a very natural interpretation in terms of *evidential factor* when the updating in the factor here is on the geometric rule of conditioning. Indeed,

$$\begin{aligned}\frac{m_x^Q(E)}{m_{x'}^Q(E)} &= \frac{m_g^X(x|E)/m_g^X(x'|E)}{m^X(x)/m^X(x')} \\ &= F(m_g^X(\cdot|E), m^X)(x, x') \\ &\leq \epsilon\end{aligned} \tag{2.2}$$

Bayesian Data and Evidential Mechanism

If bel^X is Bayesian, $bel^X = pl^X$ and all of focal elements in $\mathcal{E}^X$ are singletons. All of $\epsilon^{pl}, \epsilon^{bel}$ and ϵ^m-DPs have a natural semantics as evidential factor with Dempster's rule of conditioning on observations on the output space Y. In other words, for any $x \in X$ and $E \subseteq Y$,

$$\frac{bel^X_{\oplus}(x|E)}{bel^X_{\oplus}(x'|E)} = \frac{\sum_{E' \subseteq E} m^X(x) m^Q_x(E')}{\sum_{E' \subseteq E} m^X(x') m^Q_{x'}(E')} = \frac{bel^X(x)}{bel^X(x')} \frac{bel^Q_x(E)}{bel^Q_{x'}(E)},$$

$$\frac{pl^X_{\oplus}(x|E)}{pl^X_{\oplus}(x'|E)} = \frac{\sum_{E' \cap E \neq \emptyset} m^X(x) m^Q_x(E')}{\sum_{E' \cap E \neq \emptyset} m^X(x') m^Q_{x'}(E')} = \frac{pl^X(x)}{pl^X(x')} \frac{pl^Q_x(E)}{pl^Q_{x'}(E)}, \text{ and}$$

$$\frac{m^X_{\oplus}(x|E)}{m^X_{\oplus}(x'|E)} = \frac{m^X(x) m^Q_x(E')}{m^X(x') m^Q_{x'}(E')} = \frac{m^X(x)}{m^X(x')} \frac{m^Q_x(E)}{m^Q_{x'}(E)}.$$

Shafer's random-coded messages semantics

Assume a list of codes $c_1, \cdots, c_n$ and the chance of c_i being chosen is p_i

- chooses a code at random , uses the code to encode a message, and then sends us the result
- decode the encoded message using each of the codes and find that this always produces a message of the form "the truth is in A" for some non-empty subset A

Let A_i denote the subset we get when we decode using c_i, and set $m(A) = \sum\{p_i : 1 \leq i \leq n, A_i = A\}$ for each $A \subseteq \Omega$. The number $m(A)$ is the sum of the chances for those codes that indicate A was the true message; it is, in a sense, the total chance that the true message was A.

SLDP : LDP w.r.t. Shafer' Semantics

For an evidential privacy mechanism Q, let

$r_S^Q = max_{x,x' \in X, E \subseteq Y} \frac{m_x^Q(E)}{m_{x'}^Q(E)}$ and $\epsilon_S^Q = ln(r_S^Q)$

Definition

For any $\epsilon > 0$, the mechanism Q is called *ϵ-locally differential private* according to Shafer (ϵ-SLDP for short) if $-\epsilon \leq \epsilon_S^Q \leq \epsilon$.

And ϵ_S^Q is called the *privacy loss* of Q according to Shafer and ϵ is a *privacy budget*.

Composition and Processing

If we have several building blocks for designing differentially private algorithms, it is important to understand how we can combine them to design more sophisticated algorithms.

Lemma

(Composition) Let Q_1 be an ϵ_1-SLDP evidential privacy mechanism from X to Y_1 and Q_2 be an ϵ_2-SLDP evidential privacy mechanisms from X to Y_2. Then their combination $Q_{1,2}$ defined by $Q_{1,2}(x) = (Q_1(x), Q_2(x))$ is $\epsilon_1 + \epsilon_2$-SLDP.

Lemma

(Post-processing) Let Q be an ϵ-SLDP mechanism from X to Y and f is a randomized algorithm from Y to another finite alphabet set Z. Then $f \circ Q$ is an ϵ-SLDP mechanism from X to Z.

Hypothesis-Testing Framwork for SLDP

Hypothesis-Testing Interpretation

From an attacker's perspective, the privacy requirement can be formalized as:

H_0: the underlying dataset is x vs. H_1: the underlying is x'.

$\mathcal{P}_x^Q = \{pr \in \Delta(Y) : pr \geq bel_x^Q\}$
$,\mathcal{P}_{x'}^Q = \{pr \in \Delta(Y) : pr \geq bel_{x'}^Q\}$.

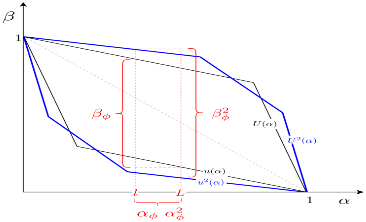
- type I error $\alpha_\phi = \sup_{pr \in \mathcal{P}_x^Q} \mathbb{E}_{pr}(\phi)$
- type II error $\beta_\phi = 1 - \inf_{pr \in \mathcal{P}_{x'}^Q} \mathbb{E}_{pr}(\phi)$
- A test ϕ is called a *level-α minimax test* if $\phi = argmin\{\beta_\phi : \alpha_\phi \leq \alpha\}$.

Trade-off between Type I and II Errors in SLDP

Theorem

The following two statements are equivalent:

1. *Q is ϵ-SLDP;*
2. *If type I error $\alpha_\phi \in [l, L]$, then type II error $\beta_\phi \in [u(L), U(l)]$ where $u(\alpha) := max\{e^{-\epsilon}(1-\alpha), 1-\alpha e^{\epsilon}\}$ and $U(\alpha) := min\{e^{\epsilon}(1-\alpha), 1-\alpha e^{-\epsilon}\}$.*



Discrete Distribution Estimation Problem

Assume that

- $X_1, \cdots, X_n$ are drawn i.i.d. according to π
- A privacy mechanism Q is then applied independently to each sample X_i to produce $Y^n = (Y_1; \cdots, Y_n)$.

Our goal is to estimate the distribution vector π from Y^n *within a certain privacy budget requirement*.

Trade-off Between Privacy and Utility for SLDP

Theorem

$Var(\hat{\pi}|N_1+N_2 \neq 0) = \frac{1}{(q-p)^2}[\pi p+(1-\pi)q][\pi q+(1-\pi)p]A = [-(\pi-\frac{1}{2})^2+\frac{1}{4}(\frac{p+q}{p-q})^2]A$ *where* $A = \sum_{0\leq N_3<n}\frac{1}{n-N_3}\binom{n}{N_3}(1-q_3)^{n-N_3}q_3^{N_3}$.

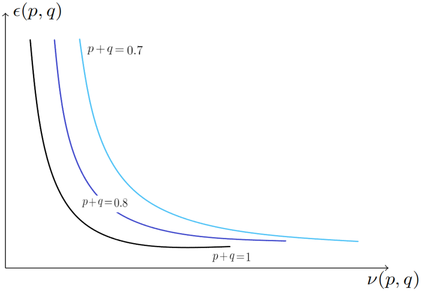


Figure: The trade-off in Shafer's semantics

LDP according to Walley

For an evidential privacy mechanism Q, let $r_Q^W = max_{pr_x \in \mathcal{P}_{bel_x^Q}, pr_{x'} \in \mathcal{P}_{bel_{x'}^Q}} \frac{pr_x(E)}{pr_{x'}(E)}$. And the logarithm $\epsilon_Q^W = ln(r_Q^W)$ quantifies the privacy loss of the privacy mechanism Q in Walley's semantics of imprecise probabilities.

Definition

For any $\epsilon > 0$, Q is called ϵ-locally differential private according to Walley (ϵ-*WLDP* for short) if, $-\epsilon \leq \epsilon_Q^W \leq \epsilon$.

And ϵ_W^Q is called the *privacy loss* of Q according to Walley and ϵ is a *privacy budget*.

Compsoition and Postprocessing

Lemma

(Composition) Let Q_1 be an ϵ_1-WLDP evidential privacy mechanism from X to Y_1 and Q_2 be an ϵ_2-WLDP evidential privacy mechanisms from X to Y_2. Then their combination $Q_{1,2}$ defined by $Q_{1,2}(x) = (Q_1(x), Q_2(x))$ is $\epsilon_1 + \epsilon_2$-WLDP.

Lemma

(Post-processing) Let Q be an ϵ-WLDP mechanism from X to Y and f is a data-independent randomized algorithm from Y to another finite alphabet set Z. Then $f \circ Q$ is an ϵ-WLDP mechanism from X to Z.

Hypothesis-Testing Interpretation

- For the rejection rule ϕ, the *pessimistic* type I and II are defined as $\alpha_\phi^{pe} = \sup_{pr \in \mathcal{P}_{bel_x^Q}} \mathbb{E}_{pr}(\phi)$ and $\beta_\phi^{pe} = \sup_{pr \in \mathcal{P}_{bel_{x'}^Q}} \mathbb{E}_{pr}(1-\phi)$, respectively.
- Also we define the *optimistic* type I and II errors as $\alpha_\phi^{op} := \inf_{pr \in \mathcal{P}_{bel_x^Q}} \mathbb{E}_{pr}(\phi)$ and $\beta_\phi^{op} := \inf_{pr \in \mathcal{P}_{bel_{x'}^Q}} \mathbb{E}_{pr}(1-\phi)$, respectively.

Definition

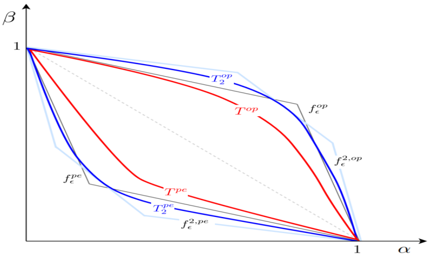
For the above pessimistic errors, the following function is called the pessimistic trade-off function:
$T^{pe}(Q(x), Q(x'))(\alpha) := inf\{\beta_\phi^{pe} : \alpha_\phi^{pe} \leq \alpha\}$. For the above optimistic errors, the following function is called the optimistic trade-off function:
$T^{op}(Q(x), Q(x'))(\alpha) := sup\{\beta_\phi^{op} : \alpha_\phi^{op} \leq \alpha\}$.

Trade-off Between Type I and II erros for WLDP

Theorem

The following two statements are equivalent:

1. *Q is ϵ-WLDP;*
2. *For any $\alpha \in [0,1]$, $T^{pe}(Q(x), Q(x'))(\alpha) \geq f_\epsilon^{pe}(\alpha)$ and $T^{op}(Q(x), Q(x'))(\alpha) \leq f_\epsilon^{op}(\alpha)$ where $f_\epsilon^{pe}(\alpha) = max\{1 - \alpha e^\epsilon, 0, e^{-\epsilon}(1-\alpha)\}$ and $f_\epsilon^{op}(\alpha) = min\{1 - \alpha e^{-\epsilon}, e^\epsilon(1-\alpha)\}$.*



Trade-off between Privacy and Utility for WLDP

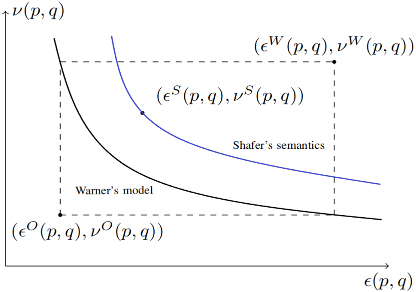
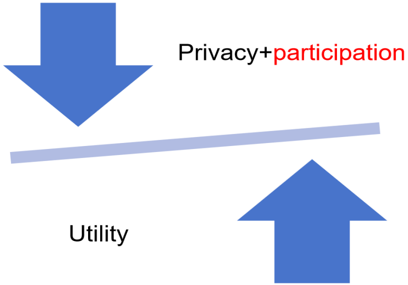


Figure: Comparison of trade-offs in the two semantics

Survey Experiments

We have conducted an online survey to show

- whether our mechanism increase participants' willingness to provide private information in survey?
- different levels of willingness to disclose privacy of different degrees.



A Simple and Direct Motivation

Our research was motivated by

- In surveys, people may prefer not to response or say "I don't know" to withhold sensitive information which *minimizes the questionable ethical consequences of lying*
- Dempster-Shafer theory improves the root concepts of probabilities "yes" and "no" that sum to one, by appending a third probability of "don't know"
- Generally, explore different aspects of uncertainties in privacy preserving

Experiment Results

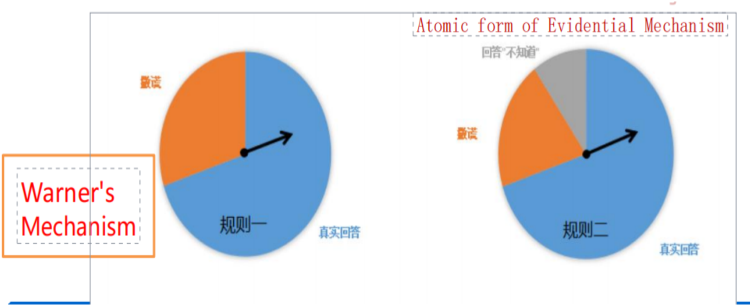


Table: Comparison of survey results

2*Questions	Warner's Mechanism		Our Mechanism		Undecided	
	Num	Per	Num	Per	Num	Per
Willingness to share	160	20.81	**387**	**50.33**	222	28.87
User experience	154	20.03	**399**	**51.89**	216	28.09
Data utility	226	29.39	310	40.31	233	30.30

Comparison at Different Sensitivity Levels

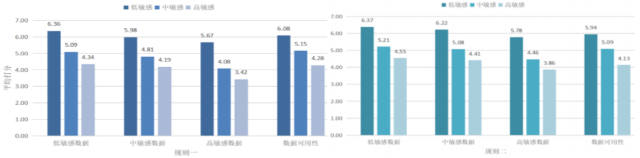


图 6-6 不同敏感程度用户披露意愿和数据可用性打分

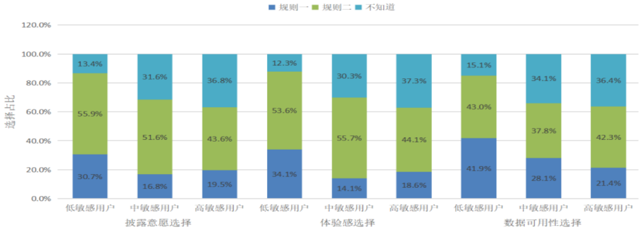


图 6-7 不同敏感程度用户偏好选择占比

References

- C Dwork, A Roth, The algorithmic foundations of differential privacy *Foundations and Trends® in Theoretical Computer Science* 9 (3–4), 211-407
- Denoeux, T. 2014. Likelihood-based belief function: Justification and some extensions to low-quality data. *Int. J. Approx. Reasoning*, 55(7): 1535–1547.
- Warner, S. 1965. Randomized Response: A Survey Technique for Eliminating Evasive Answer Bias. *Journal of the American Statistical Association*, 60(309): 63–69
- Qiyu Li, Chunlai Zhou, Biao Qin, Zhiqiang Xu: Local Differential Privacy for Belief Functions. *AAAI 2022*: 10025-10033.
- Philippe Smets: Belief functions: The disjunctive rule of combination and the generalized Bayesian theorem. *Int. J.*

- Huber, P. J.; and Strassen, V. 1973. Minimax tests and Neyman-Pearson tests for capacities. *The Annals of Statistics*, 1(2): 251–263
- Wasserman, L.; and Zhou, S. 2010. A statistical framework for differential privacy. *Journal of the American Statistical Association*, 105(489): 375–389
- Chunlai Zhou, Fabio Cuzzolin: The Total Belief Theorem. *UAI 2017*
- Shafer, G.; and Tversky, A. 1985. Languages and Designs for Probability Judgment. *Cogn. Sci.*, 9(3): 309–339.
- Walley, P. 1990. *Statistical Reasoning with Imprecise Prob- abilities*. Chapman and Hall. ISBN 3-54029586-0.

THANK YOU.