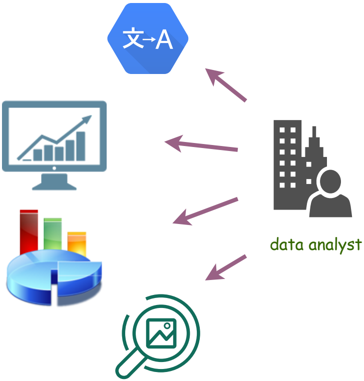
Contract Design for Purchasing Private Data Using a Biased Differentially Private Algorithm

Mohammad Mahdi Khalili, Xueru Zhang, Mingyan Liu

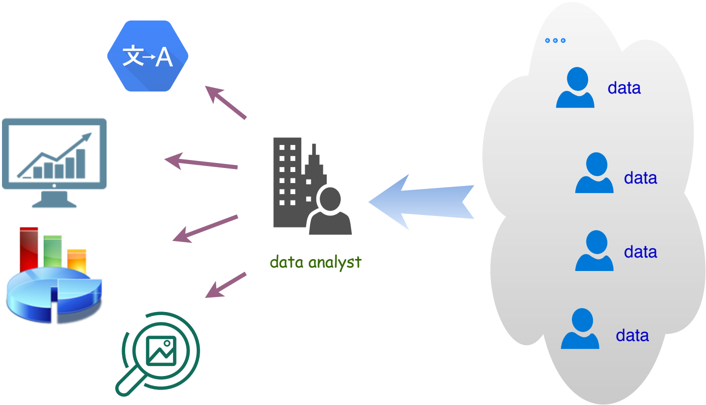
EECS Department, University of Michigan, Ann Arbor

Motivation

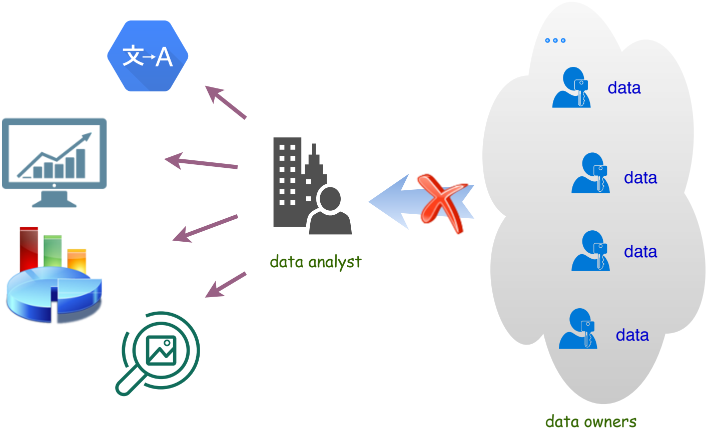
Motivation



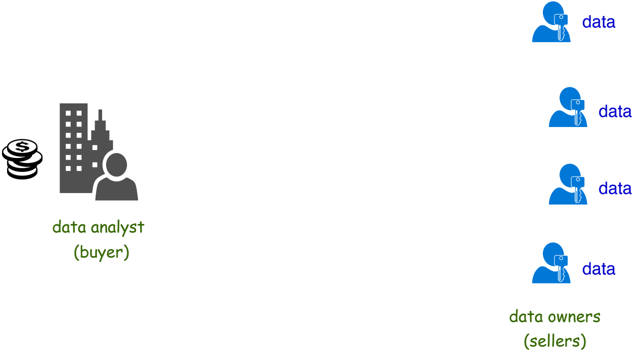
Motivation



Motivation



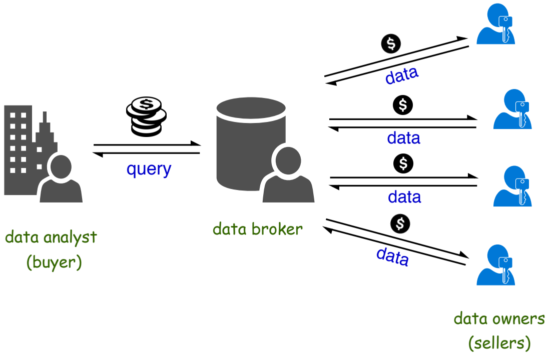
Motivation



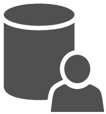
Motivation


data
data
data
data
data analyst
(buyer)
data broker
data owners
(sellers)

Motivation


query
data
data
data
data
data analyst
(buyer)
data broker
data owners
(sellers)

Motivation



- Private algorithm to generate query $\mathcal{Q}$
- Optimal contract to minimize buyer's payment:
  - The data owners are compensated properly to share their data

An example

Datacoup: A new startup founded in 2012 and plays a role as a data broker

An example

Datacoup: A new startup founded in 2012 and plays a role as a data broker

- Datacoup offers a fixed monthly payment for having access to users' social media activities, credit card transactions, etc.

An example

Datacoup: A new startup founded in 2012 and plays a role as a data broker

- Datacoup offers a fixed monthly payment for having access to users' social media activities, credit card transactions, etc.
- Provides various computations for data analysts

An example

Datacoup: A new startup founded in 2012 and plays a role as a data broker

- Datacoup offers a fixed monthly payment for having access to users' social media activities, credit card transactions, etc.
- Provides various computations for data analysts
- Datacoup removes identifiable markers

Related work

Related work

- Gosh and Roth:[1] Fixed price auction mechanism

[1] Arpita Ghosh and Aaron Roth. 2011. Selling Privacy at Auction. In Proceedings of the 12th ACM Conference on Electronic Commerce (EC 11).

Related work

- Gosh and Roth:[1] Fixed price auction mechanism
  - Consider linear queries

[1] Arpita Ghosh and Aaron Roth. 2011. Selling Privacy at Auction. In Proceedings of the 12th ACM Conference on Electronic Commerce (EC 11).

Related work

- Gosh and Roth:[1] Fixed price auction mechanism
  - Consider linear queries
  - Ensure the same level of privacy for each individual who sells the data

[1] Arpita Ghosh and Aaron Roth. 2011. Selling Privacy at Auction. In Proceedings of the 12th ACM Conference on Electronic Commerce (EC 11).

Related work

- Gosh and Roth:[1] Fixed price auction mechanism
  - Consider linear queries
  - Ensure the same level of privacy for each individual who sells the data
- Xu *et. al*:[2] Contract design problem for purchasing privacy

[1] Arpita Ghosh and Aaron Roth. 2011. Selling Privacy at Auction. In Proceedings of the 12th ACM Conference on Electronic Commerce (EC 11).

[2] L. Xu, C. Jiang, Y. Chen, Y. Ren, and K. J. R. Liu. 2015. Privacy or Utility in Data Collection? A Contract Theoretic Approach. IEEE Journal of Selected Topics in Signal Processing 2015.

Related work

- Gosh and Roth:[1] Fixed price auction mechanism
  - Consider linear queries
  - Ensure the same level of privacy for each individual who sells the data
- Xu *et. al*:[2] Contract design problem for purchasing privacy
  - It is better to purchase from those who care the least about their privacy

[1] Arpita Ghosh and Aaron Roth. 2011. Selling Privacy at Auction. In Proceedings of the 12th ACM Conference on Electronic Commerce (EC 11).

[2] L. Xu, C. Jiang, Y. Chen, Y. Ren, and K. J. R. Liu. 2015. Privacy or Utility in Data Collection? A Contract Theoretic Approach. IEEE Journal of Selected Topics in Signal Processing 2015.

Related work

- Gosh and Roth:[1] Fixed price auction mechanism
  - Consider linear queries
  - Ensure the same level of privacy for each individual who sells the data
- Xu *et. al*:[2] Contract design problem for purchasing privacy
  - It is better to purchase from those who care the least about their privacy
  - They do not provide any algorithm to ensure privacy

[1] Arpita Ghosh and Aaron Roth. 2011. Selling Privacy at Auction. In Proceedings of the 12th ACM Conference on Electronic Commerce (EC 11).

[2] L. Xu, C. Jiang, Y. Chen, Y. Ren, and K. J. R. Liu. 2015. Privacy or Utility in Data Collection? A Contract Theoretic Approach. IEEE Journal of Selected Topics in Signal Processing 2015.

Our Contributions

Our Contributions

- Proposing an algorithm which can provide **personalized** privacy to the sellers: the user with higher privacy valuation can suffer the lower privacy loss.

Our Contributions

- Proposing an algorithm which can provide **personalized** privacy to the sellers: the user with higher privacy valuation can suffer the lower privacy loss.
- Proposed contracts using our private algorithm improves the payment-accuracy tradeoff.

Our Contributions

- Proposing an algorithm which can provide **personalized** privacy to the sellers: the user with higher privacy valuation can suffer the lower privacy loss.
- Proposed contracts using our private algorithm improves the payment-accuracy tradeoff.
- We extend the algorithm to multi-dimensional data and non-linear queries.

Model

- Database $D = (d_1, d_2, \cdots, d_n)$ with $d_i \in [0, 1]$. d_i belongs to individual/seller i.
- Query $\mathcal{Q} : [0, 1]^n \to \mathbb{R}$: $\mathcal{Q}(D) = \sum_{i=1}^{n} d_i$.

Model

- Database $D = (d_1, d_2, \cdots, d_n)$ with $d_i \in [0,1]$.
  d_i belongs to individual/seller i.
- Query $\mathcal{Q} : [0,1]^n \to \mathbb{R}$: $\mathcal{Q}(D) = \sum_{i=1}^{n} d_i$.

How to quantify privacy? Differential Privacy

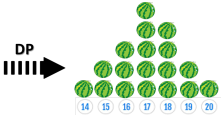
Model

- Database $D = (d_1, d_2, \cdots, d_n)$ with $d_i \in [0, 1]$.
  d_i belongs to individual/seller i.
- Query $\mathcal{Q} : [0, 1]^n \to \mathbb{R}$: $\mathcal{Q}(D) = \sum_{i=1}^{n} d_i$.

How to quantify privacy? Differential Privacy

- Obtain almost the same conclusion regardless of participation

Student ID	Last Name	Initial	Age	Program
ST348-245	Walton	L.	21	Drafting
ST348-246	Wilson	R.	19	Science
ST348-247	Thompson	G.	18	Business
ST348-248	James	L.	23	Nursing
ST348-249	Peterson	M.	37	Science
ST348-250	Graham	J.	20	Arts
ST348-251	Smith	F.	26	Business
ST348-252	Nash	S.	22	Arts



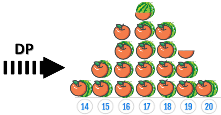
Model

- Database $D = (d_1, d_2, \cdots, d_n)$ with $d_i \in [0,1]$.
  d_i belongs to individual/seller i.
- Query $\mathcal{Q} : [0,1]^n \to \mathbb{R}$: $\mathcal{Q}(D) = \sum_{i=1}^{n} d_i$.

How to quantify privacy? Differential Privacy

- Obtain almost the same conclusion regardless of participation

Student ID	Last Name	Initial	Age	Program
ST348-245	Walton	L.	21	Drafting
ST348-246	Wilson	R.	19	Science
ST348-247	Thompson	G.	18	Business
ST348-248	James	L.	23	Nursing
~~ST348-249~~	~~Peterson~~	~~M.~~	~~37~~	~~Science~~
ST348-250	Graham	J.	20	Arts
ST348-251	Smith	F.	26	Business
ST348-252	Nash	S.	22	Arts

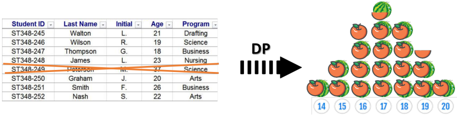


Model

- Database $D = (d_1, d_2, \cdots, d_n)$ with $d_i \in [0, 1]$.
  d_i belongs to individual/seller i.
- Query $\mathcal{Q} : [0, 1]^n \to \mathbb{R}$: $\mathcal{Q}(D) = \sum_{i=1}^{n} d_i$.

How to quantify privacy? Differential Privacy

- Obtain almost the same conclusion regardless of participation



Student ID	Last Name	Initial	Age	Program
ST348-245	Walton	L.	21	Drafting
ST348-246	Wilson	R.	19	Science
ST348-247	Thompson	G.	18	Business
ST348-248	James	L.	23	Nursing
~~ST348-249~~	~~Peterson~~	~~M.~~	~~37~~	~~Science~~
ST348-250	Graham	J.	20	Arts
ST348-251	Smith	F.	26	Business
ST348-252	Nash	S.	22	Arts

- A randomized algorithm $\mathcal{A}(\cdot)$ is ϵ_i-differentially private w.r.t. **individual** i if for any two datasets $D^{(i)}$, $\hat{D}^{(i)}$ differing in i's data and for any sets of possible outputs $S \subseteq \text{range}(\mathcal{A})$:

$$\frac{\Pr(\mathcal{A}(D^{(i)}) \in S)}{\Pr(\mathcal{A}(\hat{D}^{(i)}) \in S)} \leq \exp(\epsilon_i), \ \ \epsilon_i \in [0, +\infty)$$

Model

K-accurate algorithm

$\mathcal{A}(D)$ is K-accurate for query $\mathcal{Q}(D)$ if its mean squared error (MSE) is at most K for all $D \in [0,1]^n$:

$$\mathbb{E}[||\mathcal{A}(D) - \mathcal{Q}(D)||^2] \leq K, \ \forall D \in [0,1]^n$$

Relationship between K and $\epsilon = \sum_{i=1}^{n} \epsilon_i$

Relationship between K and $\epsilon = \sum_{i=1}^{n} \epsilon_i$

Theorem

A lower bound on total privacy loss $\epsilon = \sum_{i=1}^{n} \epsilon_i$: Consider an algorithm $\mathcal{A}(D)$ that is K-accurate for $\mathcal{Q}(D)$:

Relationship between K and $\epsilon = \sum_{i=1}^{n} \epsilon_i$

Theorem

A lower bound on total privacy loss $\epsilon = \sum_{i=1}^{n} \epsilon_i$: Consider an algorithm $\mathcal{A}(D)$ that is K-accurate for $\mathcal{Q}(D)$:

- *if $K < (\frac{n}{2})^2$, then $\epsilon = \sum_{i=1}^{n} \epsilon_i \geq \ln \frac{(n-\sqrt{K})^2}{K} = \epsilon^{lb}$.*

Relationship between K and $\epsilon = \sum_{i=1}^{n} \epsilon_i$

Theorem

A lower bound on total privacy loss $\epsilon = \sum_{i=1}^{n} \epsilon_i$: Consider an algorithm $\mathcal{A}(D)$ that is K-accurate for $\mathcal{Q}(D)$:

- *if $K < (\frac{n}{2})^2$, then $\epsilon = \sum_{i=1}^{n} \epsilon_i \geq \ln \frac{(n-\sqrt{K})^2}{K} = \epsilon^{lb}$.*
- *if $K < (\frac{m}{2})^2$, let $\mathcal{S} = \{i | \epsilon_i > 0, i = 1, 2, \cdots n\}$, then $|\mathcal{S}| \geq n - m + 1$.*

Two differentially private algorithms

Two differentially private algorithms

An unbiased algorithm (Laplace mechanism):

$$\mathcal{A}_u(D) = \sum_{i=1}^{n} d_i + N(b)$$

$$\boxed{f_N(x; b) = \frac{1}{2b} \exp\{-\frac{|x|}{b}\}}$$

Two differentially private algorithms

An unbiased algorithm (Laplace mechanism):

$$\mathcal{A}_u(D) = \sum_{i=1}^{n} d_i + N(b)$$

$$\boxed{f_N(x; b) = \frac{1}{2b} \exp\{-\frac{|x|}{b}\}}$$

A biased algorithm:

$$\mathcal{A}_{new}(D) = \sum_{i=1}^{n} a_i d_i + \sum_{i=1}^{n} \frac{1 - a_i}{2} + N(b), \quad a_i \in [0, 1], \forall i$$

Two differentially private algorithms

An unbiased algorithm (Laplace mechanism):

$$\mathcal{A}_u(D) = \sum_{i=1}^{n} d_i + N(b)$$

$$\boxed{f_N(x;b) = \frac{1}{2b}\exp\{-\frac{|x|}{b}\}}$$

A biased algorithm:

$$\mathcal{A}_{new}(D) = \sum_{i=1}^{n} a_i d_i + \sum_{i=1}^{n} \frac{1-a_i}{2} + N(b), \quad a_i \in [0,1], \forall i$$

	privacy ϵ_i	accuracy K	bias $\lvert\mathbb{E}[\mathcal{A}(D) - \mathcal{Q}(D)]\rvert$
$\mathcal{A}_u(D)$	$1/b$	$2b^2$	0
$\mathcal{A}_{new}(D)$	a_i/b	$(\sum_{i=1}^{n} \frac{1-a_i}{2})^2 + 2b^2$	$\lvert\sum_{i=1}^{n}(a_i-1)d_i + \frac{1-a_i}{2}\rvert \leq \sum_{i=1}^{n} \frac{1-a_i}{2}$

Under full information: *a single seller* $D = d$

Under full information: *a single seller $D = d$*

- Individual's privacy attitude/valuation: v

Under full information: *a single seller $D = d$*

- Individual's privacy attitude/valuation: v
- Individual's cost function: $c(v, \epsilon)$

Under full information: *a single seller $D = d$*

- Individual's privacy attitude/valuation: v
- Individual's cost function: $c(v, \epsilon)$
  - increasing in v and ϵ

Under full information: *a single seller $D = d$*

- Individual's privacy attitude/valuation: v
- Individual's cost function: $c(v, \epsilon)$
  - increasing in v and ϵ
  - $c(v, 0) = 0$, $\forall v$

Under full information: *a single seller $D = d$*

- Individual's privacy attitude/valuation: v
- Individual's cost function: $c(v, \epsilon)$
  - increasing in v and ϵ
  - $c(v, 0) = 0, \forall v$
- query $\mathcal{Q}(D) = d$

Under full information: *a single seller $D = d$*

- Individual's privacy attitude/valuation: v
- Individual's cost function: $c(v, \epsilon)$
  - increasing in v and ϵ
  - $c(v, 0) = 0, \forall v$
- query $\mathcal{Q}(D) = d$
- Full information: v is known to broker and buyer

Under full information: *a single seller $D = d$*

- Individual's privacy attitude/valuation: v
- Individual's cost function: $c(v, \epsilon)$
  - increasing in v and ϵ
  - $c(v, 0) = 0, \forall v$
- query $\mathcal{Q}(D) = d$
- Full information: v is known to broker and buyer
- If the seller receives a payment (p) more than his privacy cost, he will share his own data
  - $p \geq c(v, \epsilon)$: Individual Rationality (IR)

Under full information: *a single seller* $D = d$

Contract design problem: Finding Contract (p, ϵ, K)

Under full information: *a single seller $D = d$*

Contract design problem: Finding Contract (p, ϵ, K)

- Buyer announces desired accuracy (K)

Under full information: *a single seller $D = d$*

Contract design problem: Finding Contract (p, ϵ, K)

- Buyer announces desired accuracy (K)
- The broker finds the right algorithm to generate $K-$accurate outcome with the minimum payment (p)

Under full information: *a single seller $D = d$*

Contract design problem: Finding Contract (p, ϵ, K)

- Buyer announces desired accuracy (K)
- The broker finds the right algorithm to generate $K-$accurate outcome with the minimum payment (p)
- The broker offers contract (p, ϵ) to the seller

Under full information: *a single seller $D = d$*

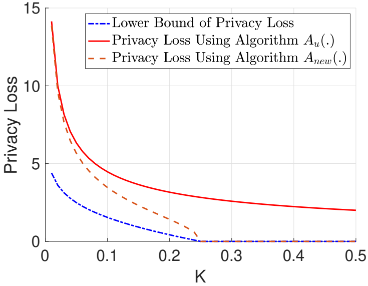
Contract design problem: Finding Contract (p, ϵ, K)

- Buyer announces desired accuracy (K)
- The broker finds the right algorithm to generate $K-$accurate outcome with the minimum payment (p)
- The broker offers contract (p, ϵ) to the seller

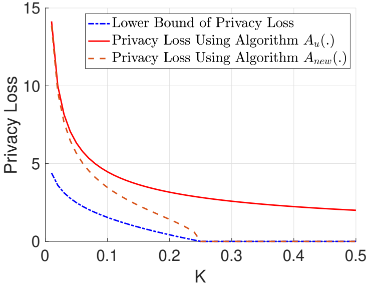
$\mathcal{A}_{new}(D) = a \cdot d + \frac{1-a}{2} + N(b)$

$$\min_{\{a \in [0,1], b>0, p\}} \quad p$$
$$s.t.(IR) \; p \geq c(v, \epsilon)$$
$$(AC) \; (\frac{1-a}{2})^2 + 2b^2 = K$$
$$\epsilon = \frac{a}{b}$$

Under full information: *a single seller $D = d$*

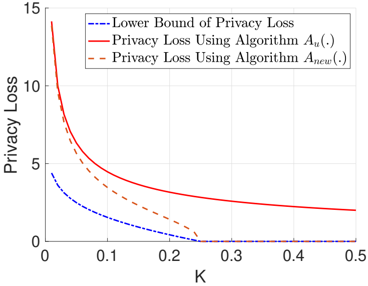


Under full information: *a single seller $D = d$*



- $\mathcal{A}_{new}(D)$ outperforms $\mathcal{A}_u(D)$ in terms of privacy(payment)-accuracy tradeoff.

Under full information: *a single seller $D = d$*



- $\mathcal{A}_{new}(D)$ outperforms $\mathcal{A}_u(D)$ in terms of privacy(payment)-accuracy tradeoff.
- when accuracy requirement is low ($K > 1/4$): best strategy is to report pure noise.

Under full information: n sellers $D = (d_1, d_2, \cdots, d_n)$

$\mathcal{A}_{new}(D) = \sum_{i=1}^{n} a_i \cdot d_i + \frac{1-a_i}{2} + N(b)$

$$\min_{\{a_i \in [0,1], b>0, p_i, i=1,\cdots,n\}} \sum_{i=1}^{n} p_i$$

$$s.t. (IR)\ p_i \geq c(v_i, \frac{a_i}{b}), i = 1, 2, \cdots n$$

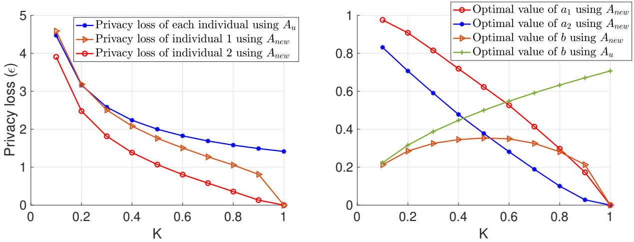
$$(AC)\ (\sum_{i=1}^{n} \frac{1-a_i}{2})^2 + 2b^2 = K$$

Theorem

The optimal solution under linear cost: *If $c(v_i, \epsilon_i) = v_i \cdot \epsilon_i\ \forall i$, then there is a closed form solution to contract design problem.*

A numerical example:

- Two sellers: $v_1 = 5$, $v_2 = 10$, $c(v_i, \epsilon_i) = v_i(e^{\epsilon_i} - 1)$



  – personalized privacy: $\mathcal{A}_{new}(D)$ assigns different ϵ_i to different individuals.
  – $\mathcal{A}_{new}(D)$ adds less noise than $\mathcal{A}_u(D)$: better privacy-accuracy tradeoff.

Motivation

Model

Private Algorithms

Contract under Full Information

Contract under Information Asymmetry

Conclusion

Under information asymmetry: *unknown privacy valuation*

- Two sellers $D = (d_1, d_2)$
- Privacy valuation is binary distributed:

$$v_i = \begin{cases} v_H, & w.p.\ \pi \\ v_L, & w.p.\ 1-\pi \end{cases}, \quad v_H > v_L$$

Design a **menu** of contracts $\{(p_H, \epsilon_H, K), (p_L, \epsilon_L, K)\}$:

- Incentive compatibility (IC):

$$p_H - c(v_H, \epsilon_H) \geq p_L - c(v_H, \epsilon_L) \ ,$$
$$p_L - c(v_L, \epsilon_L) \geq p_H - c(v_L, \epsilon_H) \ .$$

- Two options:
  - Offering **both** sellers the menu of contracts.
  - Offering only **one** seller the menu of contracts.

Option 1: offering both sellers the menu of contracts

Due to the uncertainty of v_1, v_2:

$$
\begin{aligned}
A_{new}(D) &= a_1 d_1 + a_2 d_2 + \frac{1-a_1}{2} + \frac{1-a_2}{2} + N(b), a_i \in \{a_H, a_L\} \\
&= \begin{cases}
a_H d_1 + a_H d_2 + \frac{1-a_H}{2} + \frac{1-a_H}{2} + N(b) \text{ w.p. } \pi^2 \\
a_H d_1 + a_L d_2 + \frac{1-a_H}{2} + \frac{1-a_L}{2} + N(b) \text{ w.p. } \pi(1-\pi) \\
a_L d_1 + a_H d_2 + \frac{1-a_L}{2} + \frac{1-a_H}{2} + N(b) \text{ w.p. } \pi(1-\pi) \\
a_L d_1 + a_L d_2 + \frac{1-a_L}{2} + \frac{1-a_L}{2} + N(b) \text{ w.p. } (1-\pi)^2
\end{cases}
\end{aligned}
$$

- $\epsilon_H = \frac{a_H}{b}$, $\epsilon_L = \frac{a_L}{b}$
- Expected accuracy

$$
\begin{aligned}
e(a_L, a_H, b) &= \pi^2 \cdot (2b^2 + (1-a_H)^2) + (1-\pi)^2 \cdot (2b^2 + (1-a_L)^2) \\
&+ 2\pi \cdot (1-\pi) \cdot (2b^2 + ((1-a_H)/2 + (1-a_L)/2)^2)
\end{aligned}
$$

Option 1: offering both sellers the menu of contracts

Design the menu of contracts:

$$
\begin{aligned}
\min_{\{p_i, a_i, b\}, i \in \{H,L\}} \quad & \mathbb{E}(p) = \pi^2 \cdot (2p_H) + (1-\pi)^2 \cdot (2p_L) + 2\pi(1-\pi) \cdot (p_H + p_L) \\
\text{s.t.} \qquad (IR) \quad & p_i \geq c(v_i, a_i/b), \;\; i \in \{H, L\} \\
(IC) \quad & p_i - c(v_i, a_i/b) \geq p_j - c(v_i, a_j/b), \;\; i, j \in \{H, L\} \\
(AC) \quad & e(a_L, a_H, b) \leq K, \;\; i \in \{H, L\} \\
& 0 \leq a_i \leq 1, p_i \geq 0, b > 0, \;\; i \in \{H, L\}
\end{aligned}
$$

Option 2: offering only one seller the menu of contracts

Due to the uncertainty of v_1, v_2:

$$A_{new}(D) = \begin{cases} a_H d_1 + \frac{1-a_H}{2} + \frac{1}{2} + N(b) & \text{w.p.} \quad \pi \\ a_L d_1 + \frac{1-a_L}{2} + \frac{1}{2} + N(b) & \text{w.p.} \quad 1-\pi \end{cases},$$

- $\epsilon_H = \frac{a_H}{b}$, $\epsilon_L = \frac{a_L}{b}$
- Expected accuracy

$$\begin{aligned} e(a_L, a_H, b) &= \pi \cdot (2b^2 + ((2-a_H)/2)^2) \\ &+ (1-\pi) \cdot (2b^2 + ((2-a_L)/2)^2) \ . \end{aligned}$$

Option 2: offering only one seller the menu of contracts

Design the menu of contracts:

$$
\begin{aligned}
\min_{\{a_i, p_i, b_i,\ i \in \{H,L\}\}} \quad & \mathbb{E}(p) = \pi \cdot p_H + (1-\pi) \cdot p_L \\
\text{s.t.} \quad (IR) \quad & p_i \geq c(v_i, a_i/b),\ \ i \in \{H, L\} \\
(IC) \quad & p_i - c(v_i, a_i/b) \geq p_j - c(v_i, a_j/b),\ \ i, j \in \{H, L\} \\
(AC) \quad & e(a_L, a_H, b) \leq K,\ \ i \in \{H, L\} \\
& 0 \leq a_i \leq 1, p_i \geq 0, b > 0,\ \ i \in \{H, L\}
\end{aligned}
$$

Simplifying the optimization

- (IR) Constraint is binding for users with high valuation.
- (IR) Constraint is redundant for users with low valuation.
- (IC) Constraint is binding for users with low valuation.

Broker's decision

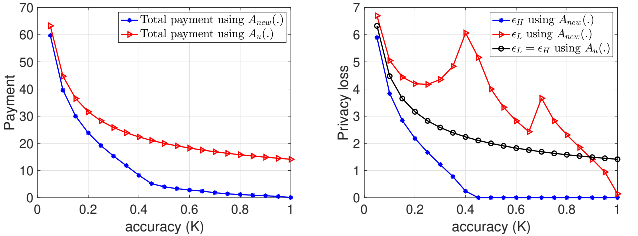
- $\mathcal{A}_u(D)$: a single contract $(a_H = a_L = 1)$ and $\begin{cases} b^* = \sqrt{K/2} \\ \epsilon^* = 1/b^* \\ p^* = c(v_H, 1/b^*) \end{cases}$
- $\mathcal{A}_{new}(D)$: a menu of contracts via **Option 1** or **Option 2**.

Broker's decision

- $\mathcal{A}_u(D)$: a single contract $(a_H = a_L = 1)$ and $\begin{cases} b^* = \sqrt{K/2} \\ \epsilon^* = 1/b^* \\ p^* = c(v_H, 1/b^*) \end{cases}$
- $\mathcal{A}_{new}(D)$: a menu of contracts via **Option 1** or **Option 2**.

A numerical example:

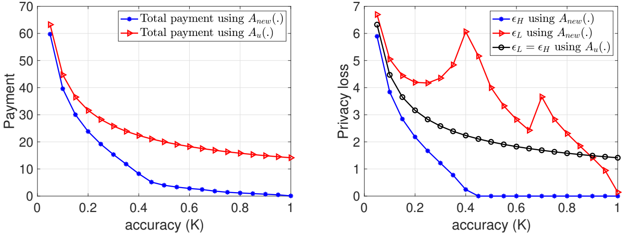
- $c(v_i, \epsilon_i) = v_i \cdot \epsilon_i,\ v_H = 5,\ v_L = 1,\ \pi = 0.5$



Broker's decision

A numerical example:

- $c(v_i, \epsilon_i) = v_i \cdot \epsilon_i, \;\; v_H = 5, \;\; v_L = 1, \;\; \pi = 0.5$



  - $\mathcal{A}_{new}(\cdot)$ lowers the payment significantly.
  - $\mathcal{A}_{new}(\cdot)$ can differentiate sellers of different types: $\epsilon_H < \epsilon_L$.

$$\begin{cases} K \leq 0.4 : \text{offer both sellers the menu of contracts} \\ 0.4 < K \leq 0.65 : \text{offer both sellers a single contract of low privacy type} \\ 0.65 < K : \text{offer a randomly seleted seller a single contract of low privacy type} \end{cases}$$

Conclusions

Conclusions

- A novel biased differentially private algorithm is proposed:
  - personalized privacy is preserved to data owners according to their privacy attitudes.
  - privacy/payment-accuracy tradeoff can be improved significantly.

Conclusions

- A novel biased differentially private algorithm is proposed:
  - personalized privacy is preserved to data owners according to their privacy attitudes.
  - privacy/payment-accuracy tradeoff can be improved significantly.
- An optimal contract is designed for a buyer aiming at purchasing private data to perform certain computations.
  - under full information
  - under information asymmetry

Conclusions

- A novel biased differentially private algorithm is proposed:
  - personalized privacy is preserved to data owners according to their privacy attitudes.
  - privacy/payment-accuracy tradeoff can be improved significantly.
- An optimal contract is designed for a buyer aiming at purchasing private data to perform certain computations.
  - under full information
  - under information asymmetry
- Generalization to non-linear queries and multi-dimensional data is available online.

Questions?

An example of nonlinear queries: polynomial queries

$$
\begin{aligned}
D &= (d_1, d_2)\\
Q(D) &= d_1^2 + d_1 \cdot d_2 + d_2^2\\
A_{new}(D) &= a_1 d_1^2 + a_{12} \cdot d_1 \cdot d_2 + a_2 d_2\\
&+ \frac{1-a_1}{2} + \frac{1-a_{12}}{2} + \frac{1-a_3}{2} + N(b)\\
\epsilon_1 &= \frac{a_1 + a_{12}}{b},\ \epsilon_2 = \frac{a_{12} + a_2}{b}\\
K &= (\frac{1-a_1}{2} + \frac{1-a_{12}}{2} + \frac{1-a_3}{2})^2 + 2b^2
\end{aligned}
$$