Individualized Privacy Accounting via Subsampling

with Applications in Combinatorial Optimization

Badih Ghazi
Google Research
Mountain View, CA

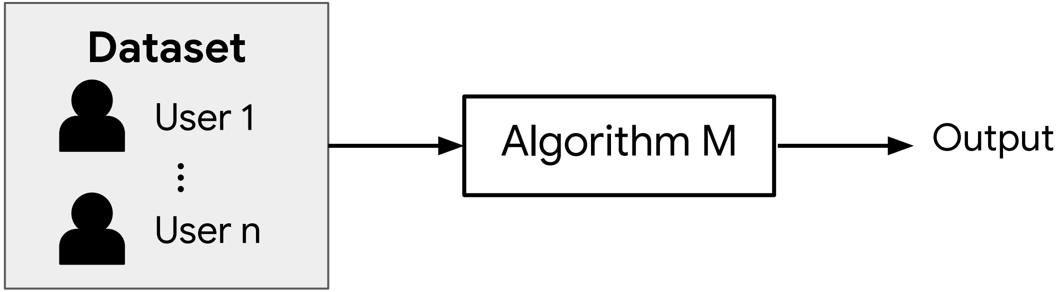
Pritish Kamath
Google Research
Mountain View, CA

Ravi Kumar
Google Research
Mountain View, CA

Pasin Manurangsi
Google Research
Bangkok, Thailand

Adam Sealfon
Google Research
New York, NY

Differential Privacy



(ε, δ)-Differential Privacy (DP)

[Dwork et al.'06]

For every datasets X, X' differing on a single record and every set S of outputs,

$\Pr[M(X) \in S] \le e^{\varepsilon} \cdot \Pr[M(X') \in S] + \delta$

Pure-DP

ε-DP ≡ (ε, 0)-DP

Approx-DP

$\delta > 0$

Previous Results: Combinatorial Optimization

Approx-DP Algorithms

Problem	**Approximation Ratio**	**Additive Error**	**Reference**
Set Cover	$O\left(\log n + \frac{\log m \log(1/\delta)}{\epsilon}\right)$	-	*[Gupta et al., SODA'10]*
Submodular maximization with cardinality constraint	$\left(1 - \frac{1}{e}\right)$	$O\left(\frac{k \log m \log(1/\delta)}{\epsilon}\right)$	*[Gupta et al., SODA'10]*
Metric k-means/k-median	$O(1)$	$O\left(\frac{k \log(mn) \log(1/\delta)}{\epsilon}\right)$	*[Jones et al., AAAI'21]*

Our Results: Combinatorial Optimization

"A **generic** recipe to make previous approx-DP algorithms pure-DP"

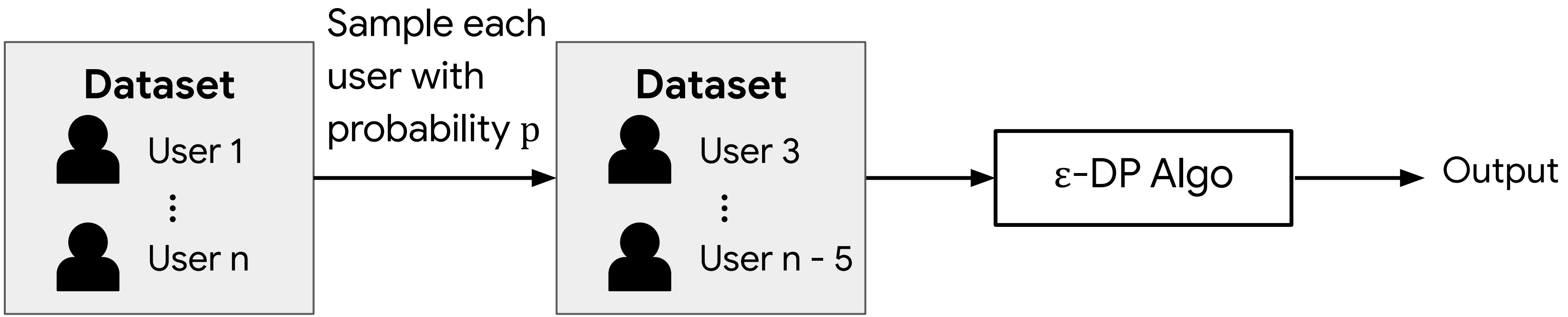
Pure-DP Algorithms

Problem	Approximation Ratio	Additive Error	Reference
Set Cover	$O\left(\log n + \frac{\log m}{\epsilon}\right)$	-	*[This work]*
Submodular maximization with cardinality constraint	$\left(1 - \frac{1}{e} - \eta\right)$	$O_\eta\left(\frac{k \log m}{\epsilon}\right)$	
Metric k-means/k-median	$O(1)$	$O\left(\frac{k \log(mn)}{\epsilon}\right)$	

* More results on submodular maximization with matroid constraint and shifting heavy hitters in the paper

Amplification by Subsampling

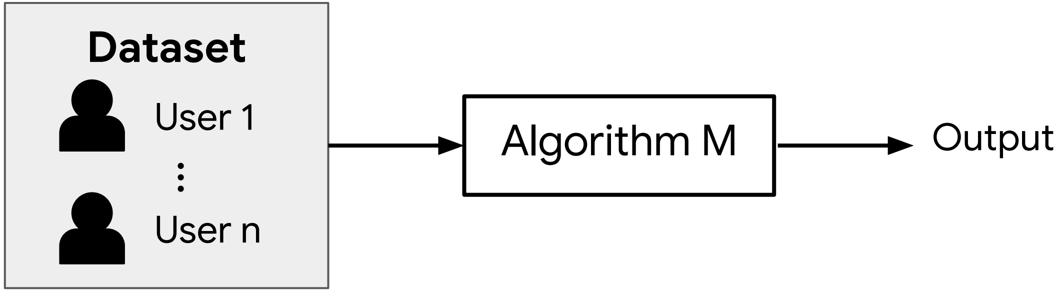
"Subsampling makes the algorithm more private."



Amplification-by-subsampling Theorem

For $\varepsilon \leq 1$, the above mechanism is $O(p \cdot \varepsilon)$-DP

Differential Privacy



(ε, δ)-Differential Privacy (DP)

[Dwork et al.'06]

For every datasets X, X' differing on a single record and every set S of outputs,

$\Pr[M(X) \in S] \leq e^{\varepsilon} \cdot \Pr[M(X') \in S] + \delta$

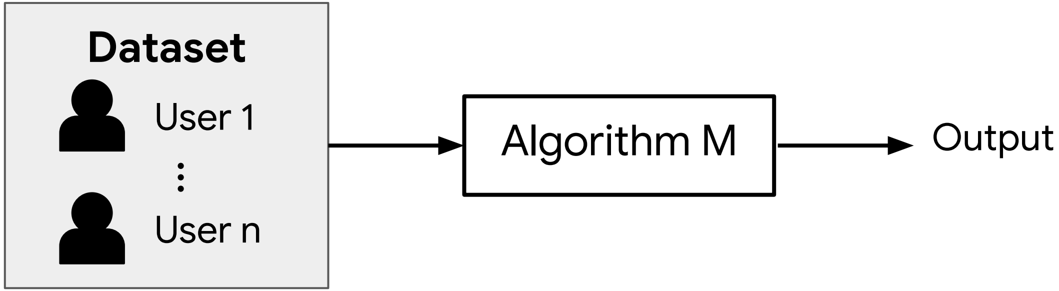
Pure-DP

ε-DP ≡ (ε, 0)-DP

Approx-DP

$\delta > 0$

One-Sided DP



"Two-sided DP"

(ε, δ)-Differential Privacy (DP)
[Dwork et al.'06]
For every datasets X, X' differing on a single record and every set S of outputs,
$\Pr[M(X) \in S] \leq e^{\varepsilon} \cdot \Pr[M(X') \in S] + \delta$

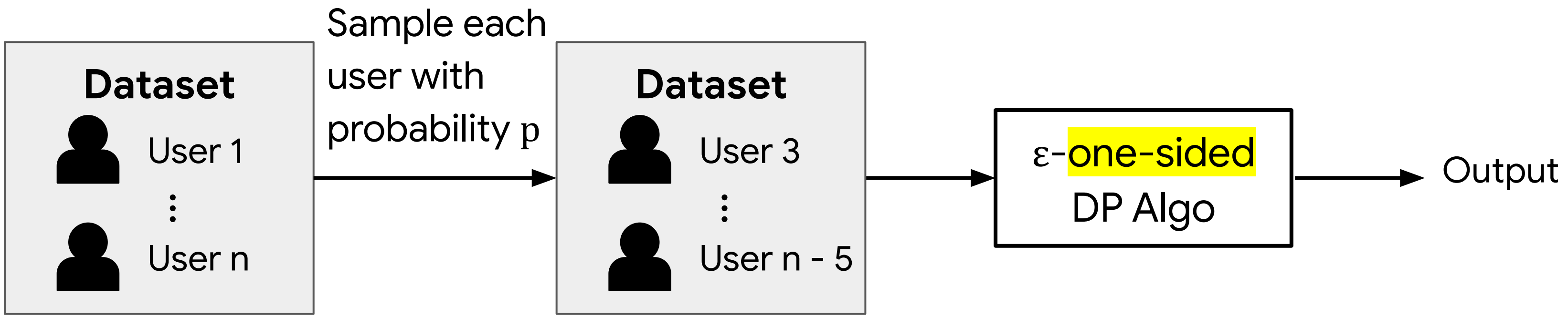
Pure-DP
ε-DP ≡ (ε, 0)-DP

Approx-DP
$\delta > 0$

(ε, δ)-one-sided DP
[Dwork et al.'06]
For every X, X' s.t. X results from adding a record to X' and every set S of outputs,
$\Pr[M(X) \in S] \leq e^{\varepsilon} \cdot \Pr[M(X') \in S] + \delta$

Our Amplification by Subsampling

"Subsampling makes one-sided-DP algorithm two-sided DP."



Amplification-by-subsampling Theorem

For $\varepsilon \leq 1$, the above mechanism is $O(p)$-DP

For combinatorial opt. problems: suffices to give one-sided-DP algorithm

Submodular Maximization & Greedy Algo

Submodular Maximization with Cardinality Constraint

- **Input:**
  - integer k,
  - dataset X,
  - for each $x \in X$, monotone submodular $f_x : U \rightarrow [0, 1]$
- **Output:** $S \subseteq U$ of size k that maximizes $F(S) := \sum_{x \in X} f_x(S)$

Submodular Maximization & Greedy Algo

Submodular Maximization with Cardinality Constraint

- **Input:**
  - integer k,
  - dataset X,
  - for each $x \in X$, monotone submodular $f_x : U \rightarrow [0, 1]$
- **Output:** $S \subseteq U$ of size k that maximizes $F(S) := \sum_{x \in X} f_x(S)$

Non-private Greedy Algorithm

$S \leftarrow \emptyset$

Repeat k times:

    Find u such that $F(S \cup \{u\})$ is maximized

Return S

Gives (1 - 1/e)-approximation

Repeated Exponential Mechanism

Submodular Maximization with Cardinality Constraint

- **Input:**
  - integer k,
  - dataset X,
  - for each $x \in X$, monotone submodular $f_x : U \to [0, 1]$
- **Output:** $S \subseteq U$ of size k that maximizes $F(S) := \sum_{x \in X} f_x(S)$

Private Greedy Algorithm

$S \leftarrow \emptyset$

Repeat k times:

Find u such that $F(S \cup \{u\})$ is maximized using ε_0-DP exponential mechanism

Return S

Basic composition ⇒ $k\varepsilon_0$-DP

Theorem *[Gupta et al.'10]*

Private Greedy is $(\varepsilon_0 \cdot \log(1/\delta), \delta)$-DP

Theorem *[This work]*

Private Greedy is ε_0-one-sided-DP

Conclusion

- Pure-DP algorithms for Combinatorial Optimization
  - Observation: Subsampling makes one-sided DP into two-sided DP
  - Suffices to give one-sided DP algorithms
    - Repeated Exponential Mechanism
    - Repeated AboveThreshold
- Open Problem: Can we make our technique work without monotonicity?

Thank you!