UNIVERSITY of VIRGINIA

Privacy: Differential Privacy Part 2

Presented by

Group 7:
Dhriti Gampa(jwr9ew)
Jing-Ning Su (fzf9mg)
Shafat Shahnewaz(gsq2at)
Jackson Miskill (jcm4bsq)

Presentation Outline

- ***Introduction***
  - Real-world example
  - Problem Statement
- ***Part 1:*** Differential Privacy Mechanisms
  - Laplace Mechanism (Dhriti)
  - Exponential Mechanism (Dhriti)
  - Gaussian Mechanism (Shafat)
  - Approximate Differential Privacy (Shafat)
  - Sparse Vector Technique (Jackson)
  - Rényi Differential Privacy (Jackson)
- ***Discussion 1 – 10 minutes***
- ***Part 2***: Applying/Implementing Differential Privacy
  - Mechanism Design under Differential Privacy (Jing-Ning)
  - Coding up differential privacy feat. Google Colab
- ***Discussion 2 – 10 minutes***
- ***Part 3***: Critical Analysis (strengths, weaknesses, limitations, future work)
- ***Discussion 3 – 10 minutes***

Recap of Differential Privacy

- Explored risks of Database Reconstruction Attacks
  - Case studies: Community Age & Gender Inference, 911 Call Data Privacy Issues
- Noise Addition & DP Properties
  - Noise Addition to prevent data leakage
  - Key properties of Differential Privacy
    - **ε:** Defines the level of privacy, **smaller ε → stronger privacy, more noise**.
    - **Adjacent datasets**: Ensures similar output, with or without an individual's data.
    - **Privacy-accuracy tradeoff**: Balancing privacy and accuracy requires tuning ε.

Challenge of Applying DP in Real-world

- **How can we ensure privacy while extracting valuable insights?**
- Open challenge in real-world applications

Bidders want privacy in their pricing strategies

Patients worry about data traceability

Personalized results without full data exposure

Key challenge: Applying DP to protect privacy while ensuring useful outcomes

Problem Statement

- **Choosing the Right DP Mechanism**
  - Beyond Laplace Mechanism: Exponential Mechanism, Gaussian Mechanism…
- **DP in Mechanism Design**
  - Apply DP in systems and prove its truthfulness
- **Implementation & Practical Analysis**
  - Real-world DP implementation and its effect on accuracy

Differential Privacy Mechanisms

Laplace Mechanism

Dhriti

Overview

Purpose

- Add noise to numeric query results

How it works

- Compute sensitivity of the query
- Add Laplace noise proportional to sensitivity and privacy parameter ϵ

Formula

- $\mathrm{M}(X) = f(X) + Laplace(\frac{\Delta}{\epsilon})$

Properties

- Privacy Guarantee: Ensures ϵ-differential privacy
- Error Bound: Noise scales with sensitivity O(Δ/ϵ)
- More noise but stronger privacy

Sensitivity

- How much a function’s output changes when one individual's data changes
- L1 Sensitivity: maximum total change across all dimensions
  - $\Delta = max_{neighbors\ X,X'} ||f(X) - f(X')||_1$

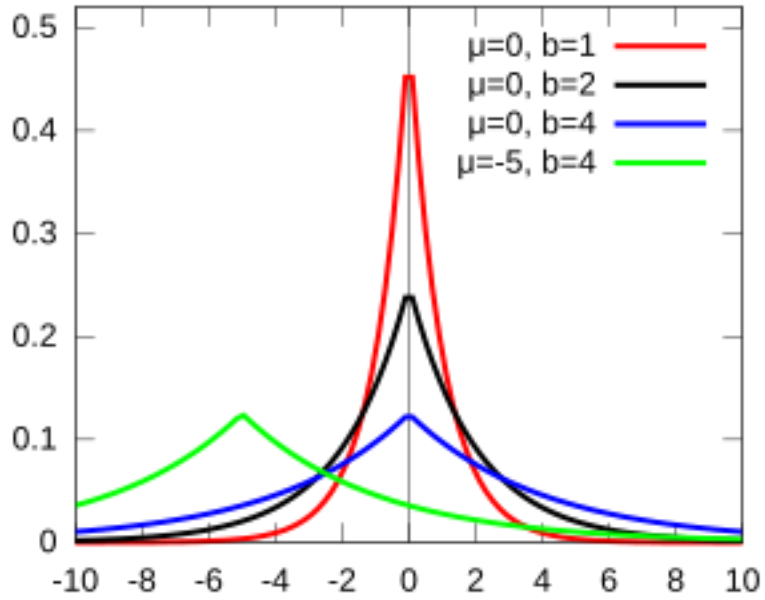
Example: Average age of patients in a hospital.

- $f(X) = \frac{1}{n}\sum X_i$, where $X_i \in [0, 100]$
- Sensitivity $\Delta = \frac{100}{n}$

Noise Generation

- Scale Parameter: $b = \Delta/\epsilon$
- Smaller ϵ → larger b → More noise for privacy
- Larger Δ → Larger b → Compensates for sensitive functions

- PDF: $\frac{1}{2b} e^{-|x|/b}$
- Mean = 0
- Variance = $2b^2 = \frac{2\Delta^2}{\epsilon^2}$



Real-World Example

Scenario: Counting COVID cases in a city

- Query: How many people tested positive?
- Raw count: $f(X) = 5000$

Steps:

1. Sensitivity Δ=1 (one individual can change the count by at most 1)
2. Add Laplace noise Y ~ Laplace(1/ϵ)

Output: $M(X) = 5000 + Y$

Exponential Mechanism

Dhriti

Overview

Purpose

- Select the best object privately

Formula

- $\Pr[M_E(X) = h] \propto e^{\frac{\epsilon s(X,h)}{2\Delta}}$

Sensitivity

- L1 Sensitivity: $\Delta s = max_{h \in H} max_{neighbors\ X,X'} |s(X,h) - s(X',h)|$

Properties

- Accuracy: Selected object's score is close to optimal
- Flexibility: Works for non-numeric outputs
- Efficiency: Best for small/structured sets H

How it Works

1. Calculate Scores
   - For every object $h \in H$ calculate its score $s(X, h)$
2. Compute Sensitivity
3. Assign Probabilities
   - $\Pr[M_E(X) = h] \propto \exp(\frac{\epsilon(s(X,h))}{2\Delta})$
4. Sample an object h from this probability distribution

Real-World Example

Scenario: Digital Goods Auction

- A seller has unlimited copies of a digital item (e.g., movie, game)
- Buyers have valuations v_i for the item
- Seller wants to set a price p to maximize revenue

Steps:

1. Dataset: Buyers' valuations $X = [v_1, v_2, \dots, n]$
2. Set of Objects: Possible prices H$= [10, 50, \dots, 100]$
3. Score Function: Revenue at price p

$$s(X, p) = p \cdot |\{i : v_i \geq p\}|$$

Real-World Example

Price (p)	Buyers at ≥p	Revenue (s(X,p))	Probability Weight
$10	100	$1,000	$\exp(\frac{0.3 \cdot 1000}{2 \cdot 1})$
$50	20	$1,000	$\exp(\frac{0.3 \cdot 1000}{2 \cdot 1})$
$100	5	$500	$\exp(\frac{0.3 \cdot 500}{2 \cdot 1})$

Output: $Given\ privacy\ parameter\ \epsilon$ and $\Delta = 1$

- $higher\ revenue\ prices\ get\ exponentially\ higher\ selection\ probability$

UNIVERSITY of VIRGINIA

Approximate Differential Privacy

Shafat

Approximate (ϵ,δ)-DP Overview

Purpose of Approximate (ϵ, δ) DP:

- Pure DP can require ***too much noise***, reducing data utility
- Approximate DP ***relaxes*** pure DP slightly → More Flexibility
- Allows for ***higher utility*** with ***minimal risk*** of privacy breach

Definition:

An algorithm M is (ϵ, δ)-DP if for all neighboring datasets X, X':

$$\Pr[M(X) \in T] \leq e^{\varepsilon} \Pr[M(X') \in T] + \delta.$$

Introduces relaxation parameter

- ϵ :Measures privacy loss
- δ: Probability of privacy failure
- Typically, $\delta \ll 1/|X|$ (much smaller than reciprocal of dataset size)

Properties of Approximate (ϵ,δ)-DP

- **Post-processing immunity:** Privacy is preserved after Processing

Let $M : \mathcal{X}^n \to \mathcal{Y}$ be (ε, δ)-differentially private, and let $F : \mathcal{Y} \to \mathcal{Z}$ be an arbitrary randomized mapping. Then $F \circ M$ is (ε, δ)-differentially private.

Example: If a histogram is privatized using DP, sorting or normalizing the histogram does not leak any additional information.

- **Group Privacy:** Protects small groups, not just individuals

Let $M : \mathcal{X}^n \to \mathcal{Y}$ be an (ε, δ)-differentially private algorithm. Suppose X and X' are two datasets which differ in exactly k positions. Then for all $T \subseteq \mathcal{Y}$, we have

$$\Pr[M(X) \in T] \leq \exp(k\varepsilon) \Pr[M(X') \in T] + ke^{(k-1)\varepsilon}\delta.$$

Example: A family sharing the same online shopping account could have their joint activity protected using group privacy bounds.

Properties of Approximate (ϵ,δ)-DP

- **Composition:** Privacy guarantees *degrade gracefully* when combining multiple queries. Performance decreases smoothly as privacy increases.

Suppose $M = (M_1, \ldots, M_k)$ is a sequence of algorithms, where M_i is $(\varepsilon_i, \delta_i)$-differentially private, and the M_i's are potentially chosen sequentially and adaptively.[2] *Then M is $(\sum_{i=1}^{k} \varepsilon_i, \sum_{i=1}^{k} \delta_i)$-differentially private.*

  - Basic: $\sum \varepsilon_i, \sum \delta_i$
  - Advanced: tighter bounds using $\sqrt{k}$ factors

Example: If a user's data is analyzed daily for 30 days, even with private analysis, the cumulative privacy loss grows and must be accounted for using composition theorems.

Real-World Example

- **US Census Bureau:** 2020 Census used (ϵ,δ)-DP for releasing statistics
- **Google's RAPPOR:** For telemetry and usage statistics
- **Meta's data analysis:** Processing user behavior with privacy guarantees
- **Healthcare data sharing:** Enabling research while protecting patient records
- **Federated learning:** Privacy-preserving machine learning across devices

- Siri Personalization
- QuickType Keyboard
- **Speech Recognition Systems**

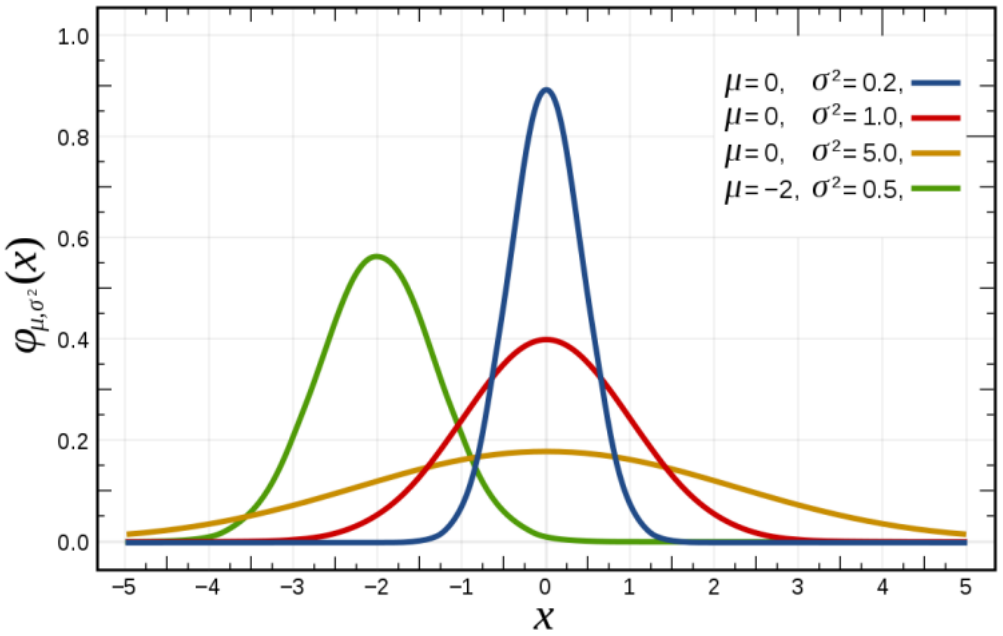
Gaussian Mechanism

Shafat

Gaussian DP overview

What is Gaussian Differential Privacy?

- Uses Gaussian noise to achieve privacy guarantees.
- Characterized by the addition of normally distributed noise calibrated to the sensitivity of a function



Visualizations of the density of the Gaussian distribution

Properties:

- Closely related to (ϵ,δ)-DP but with noise drawn from Gaussian distribution
- Characterized by a "privacy profile" rather than fixed (ϵ,δ) values
- Represents privacy guarantees in terms of trade-off function

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

How G-DP Works

The Gaussian mechanism:

For a function *f* with ℓ_2 sensitivity Δ_2, the G-DP adds normally distributed noise:

$$\mathcal{M}(X) = f(X) + \mathcal{N}(0, \sigma^2 I) \qquad \text{where, } \sigma^2 = \frac{2\ln(1.25/\delta) \cdot \Delta_2^2}{\varepsilon^2}$$

- Noise scale depends on query sensitivity and privacy parameters

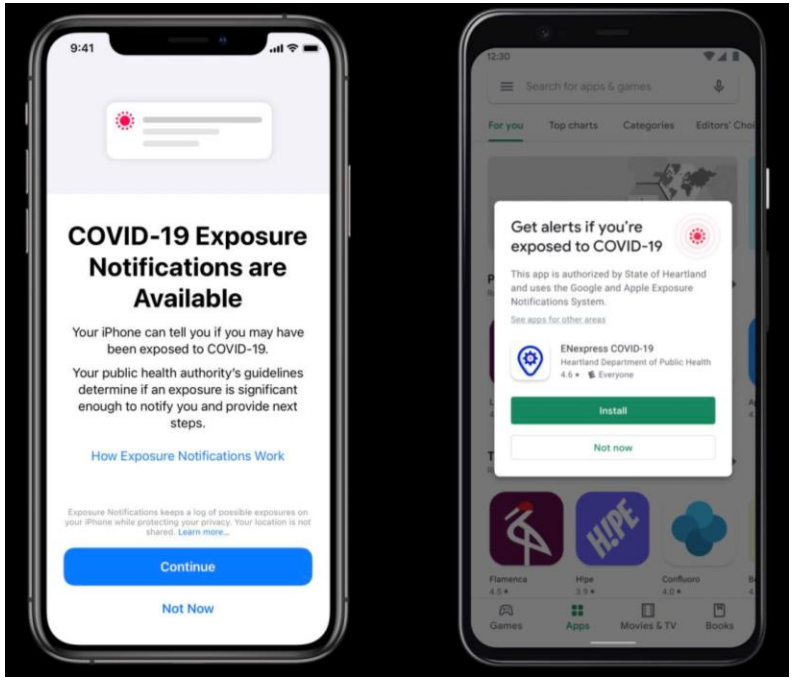
<u>$\ell 2$-sensitivity:</u>

$$\Delta_2(f) = \max_{X,X'} \|f(X) - f(X')\|_2$$

- Measures maximum change in output (Euclidean norm)
- Used to calibrate Gaussian noise
- **Example:** Mean of binary vectors $X \in \{0,1\}^{n \times d}$ has $\Delta_2 = \sqrt{d}/n$

Real-World Example of G-DP

COVID-19 Exposure Notifications: Apple and Google deployed a privacy-preserving protocol using Gaussian noise to anonymize contact-tracing Bluetooth beacon data shared between devices.



- **Google's TensorFlow Privacy:** Machine learning with privacy guarantees
- **Microsoft's SmartNoise:** Statistical analysis with differential privacy
- **Apple's device analytics:** Collecting user data while preserving privacy
- **Recommendation systems:** Privacy-preserving personalized recommendations
- **Genomic data analysis:** Protecting sensitive genetic information

Sparse Vector Technique

Jackson

Sparse Vector Technique

1. Not all queries are created equal.
2. What about yes/no questions (index)?
3. Sparse vector technique comes in handy because it looks for records that only satisfy that condition and then releases those data noisily and under a noisy threshold
4. We will see this in the ***demo***

1. The threshold is established (this is set by whoever will be running the queries)
2. Add noise to the threshold (we can assume Laplacian noise here)
3. Run for all queries
4. Halt when a query is above the threshold and return that query. Otherwise, return None.
5. Privacy cost is epsilon (or n*epsilon where we need to return n records above threshold)

Important

The sparse vector technique operates on a stream of sensitivity-1 queries over a dataset; it releases the *identity* of the first query in the stream which passes a test, and nothing else.

Rényi Differential Privacy (RDP)

Jackson

RDP and zCDP

- ***Goal***: Let's find tighter bounds for our privacy budget.
- ***Technique***: ***Divergences*** (probability), but more specifically the Rényi divergence. zCDP is based on Rényi
  - Utilizes a parameter alpha, which was taken from the Renyi divergence.
- ***Why:***
  - You want tighter bounds (accounting), you need to track privacy), or plan to convert to ADP later). More practical.

In 2017, Ilya Mironov proposed Rényi differential privacy (RDP) [12]. A randomized mechanism F satisfies $(\alpha, \bar{\epsilon})$-RDP if for all neighboring datasets x and x'

$$D_\alpha(F(x) \| F(x')) \leq \bar{\epsilon} \tag{15}$$

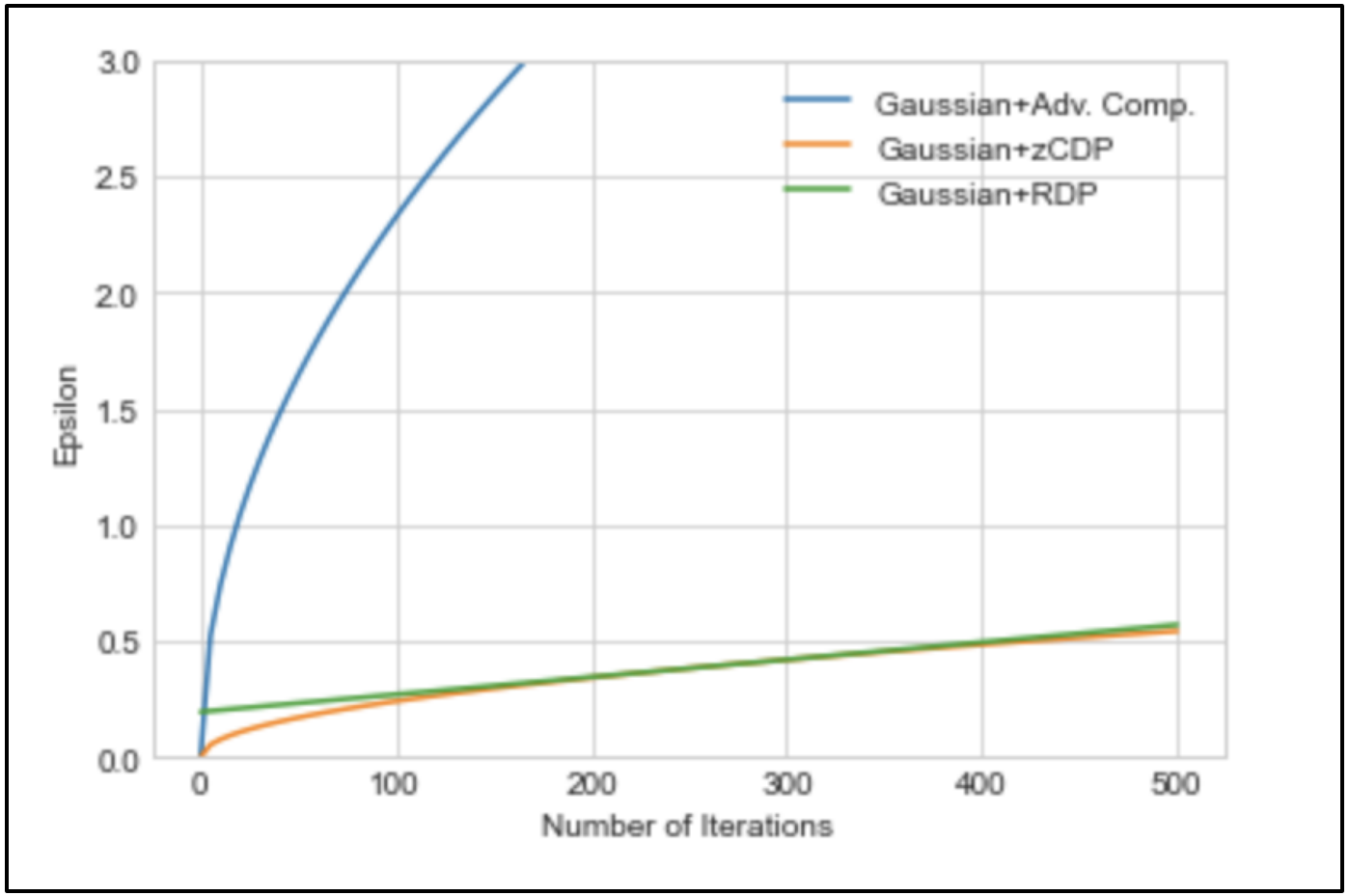
Definition 23 (Renyi Gaussian)

$$F(x) = f(x) + \mathcal{N}(\sigma^2) \text{ where } \sigma^2 = \frac{\Delta f^2 \alpha}{2\bar{\epsilon}}$$

Definition 24 (zCDP Gaussian)

$$F(x) = f(x) + \mathcal{N}(\sigma^2) \text{ where } \sigma^2 = \frac{\Delta f^2}{2\rho}$$

RDP/zCDP in action



Mechanisms Summary Table

Aspect	Laplace Mechanism	Exponential Mechanism	Gaussian Mechanism	ADP ($ \varepsilon, \delta $)-DP	Sparse Vector Technique	Renyi Differential Privacy
Output Type	Numeric ($\mathbb{R}^k$)	Object (e.g., price, classifier)	Numeric ($\mathbb{R}^k$)	Numeric	Thresholder values	Numeric
Noise	Additive Laplace noise	Probabilistic selection based on scores	Gaussian (Normal)	Gaussian (Normal)	Laplace	Gaussian (Normal)
Sensitivity	ℓ_1-sensitivity of f	Sensitivity of score function s	$ \ell_2 $	$ \Delta2(f) = |max\{X, X'\} |f(X) - f(X')|_2$	Depends	$L2 = \Delta f$
Key Use Case	Answering low-sensitivity numeric queries	Selecting high-quality objects privately	High-dimensional queries	Same as regular DP	Thresholder queries	Keeping track of privacy budget in real time
Error Bound	$O(\Delta / \epsilon)$	$O(\frac{\Delta \log \mathcal{H}}{\epsilon})$	$O(\frac{\Delta f \sqrt{log(1/\delta)}}{\epsilon})$	Same as Gaussian	$O(\frac{1}{\epsilon})$	$O(\frac{\delta}{\sqrt{\epsilon}})$
Example Application	Histograms, counting queries	PAC learning, revenue-maximizing pricing	Releasing height data	Computing statistics on user locational data	Find user with > 100 logins	Training a patient health model on patient data

Discussion-1 (10 min)

1. A company is querying their user usage data to figure out which of their products is the most popular so that they can release to the world to hopefully gain more traction. Which mechanism would be best to use to maintain the privacy of their users' preferences and why?
2. You are training a deep learning model and need to ensure that you have a measure for privacy at each phase of the training process. Which mechanism would be best and why?
3. A hospital wants to check if any patient's blood pressure exceeds a threshold (e.g., 140 mmHg) across 1,000 patients. Only a few patients are above the threshold. What mechanism would be best to use here and why?
4. A hospital wants to release the number of patients diagnosed with a rare disease while protecting individual privacy. The count is small (e.g., 10–100), and the hospital requires strict privacy guarantees. What mechanism would be best here and why?
5. If you finish the scenarios: what sorts of details would you consider about your project/system when you are considering what mechanism to use to maintain privacy?

Mechanism Design under Differential Privacy

Differential Privacy in Mechanism Design

- Design systems where participants act in self-interest while achieving desirable outcomes.
- Example: A auction where bidders are incentivized to bid honestly.

Core Focus and Traits of Mechanism Design		Why DP Matters
Incentive Compatibility	Truthful reporting should be the best strategy	Reduces the risk of strategic manipulation
Information Revelation	Players must disclose private data for optimal results	Allows data aggregation while protecting individual privacy
Efficiency	Maximizes welfare but requires truthful information, creating a privacy challenge	Balances data utility and confidentiality

- Importance of Differential Privacy
  - Helps resolve the tension between accuracy and privacy.
  - Ensures fairness, trust, and efficiency in mechanism design.

Differential Privacy as a Solution Concept

- Truthfulness in Mechanism Design
  - Introduction to Truthfulness:

$$u_i(t_i, M(t_i, t_{-i})) \geq u_i(t'_i, M(t'_i, t_{-i}))$$

  - Approximate truthfulness:

$$u_i(t_i, M(t_i, t_{-i})) \geq u_i(t'_i, M(t'_i, t_{-i})) - \epsilon$$

- **Problem: How does differential privacy help us achieve approximate truthfulness?**
- Quick Review of DP Formula

$$\frac{\Pr[M(D) \in S]}{\Pr[M(D') \in S]} \leq e^{\epsilon}$$

Proving DP Implies Approximate Truthfulness

1. Define Utility Difference:
   - Compare expected utility for truthful and misleading reports.

$$u_i(t_i, M(t_i, t_{-i})) \geq u_i(t_i', M(t_i', t_{-i})) - \epsilon$$

2. Apply Differential Privacy
   - ϵ-DP ensures output distribution changes by at most e^{ϵ} when one entry changes.

$$\mathbb{E}\left[u_i(t_i, M(t_i, t_{-i}))\right] \leq e^{\epsilon} \cdot \mathbb{E}\left[u_i(t_i, M(t_i', t_{-i}))\right]$$

3. Establish Approximate Truthfulness

$$e^{\epsilon} \leq 1 + 2\epsilon \text{ (valid for } \epsilon \leq 1)$$

$$\mathbb{E}\left[\left|u_i(t_i, M(t_i, t_{-i})) - u_i(t_i, M(t_i', t_{-i}))\right|\right] \leq 2\epsilon$$

4. The mechanism is 2ϵ-approximately truthful.

How Differential Privacy Helped in Auction Example

- Why Does the Exponential Mechanism Preserve Privacy?
  - Selects price **probabilistically**, not based on exact valuations
  - **One buyer's change** only slightly affects selection probability
  - **Low sensitivity** ensures only minimal noise is needed
- Does It Hurt Accuracy?
  - Output revenue is still **close to optimal**
  - Privacy cost is **small and controllable**
  - Balances **utility and privacy** effectively

Revenue formula:

$$\mathrm{Rev}(p, v) = p \cdot |\{i : v_i \geq p\}|$$

Optimal revenue:

$$\mathrm{OPT} = \max_p \mathrm{Rev}(p, v)$$

Privacy-preserving revenue formula:

$$\mathrm{Rev}(p, v) \geq (\mathrm{OPT} - \alpha n) - O\left(\frac{\epsilon}{\ln \alpha}\right)$$

Even with differential privacy, revenue remains near-optimal!

Coding up differential privacy- DP Demo

Jackson

Link to demo

Demo

Post-Implementation: A Checklist for Privacy Deployment

1. Data classification
2. Data preparation
3. Define Privacy Budget
4. Data Analysis Requirements
5. ***Implement DP***
6. ***Noise Generation***
7. Testing and Verification
8. Performance Evaluation
9. Documentation and compliance
10. Additional Security Measures
11. User Education
12. Continuous Monitoring and Maintenance

Discussion-2 (10 min)

- How does knowing that data is protected by differential privacy affect participants' strategies in mechanism design? For example, in an auction, could bidders change their bids because of the added noise? How does this noise impact the predictability of behavior in real-world decision systems, like auctions? What are the technical effects on efficiency and performance?
- Let's consider the case that you work for the Social Security Administration on the developer team responsible for maintaining data integrity and piping to other teams. What privacy mechanisms might you use on this developer team and why? How would you frame the privacy-utility tradeoff to a higher up who may not understand the more complex mathematical terms?
- In what real world use case would the sparse vector technique be advantageous in comparison to standard noise addition?

Strengths and Weaknesses

Strengths

Laplace/Gaussian:
- Simple implementation, strong theoretical guarantees, well suited for numerical queries

ADP:
- Allows for failure of pure differential privacy, satisfies other DP properties

Exponential Mechanism
- Versatility, Theoretical soundness, Preserves utility

Sparse Vector
- Reduces the privacy budget by providing a more granular query definition

RDP
- Provides a tool for more accurate privacy budget tracking and allocation

Weaknesses

Laplace/Gaussian:
- Noise depends on sensitivity, does not extend to non-numerical queries

ADP:
- A looser bounds on privacy error, the failure parameter is arbitrarily made up

Exponential Mechanism
- Computational Complexity, Sensitivity estimation, Privacy budget constraints

Sparse Vector
- Only applies to very specific thresholder questions

RDP
- Relies on Renyi distribution, the alpha parameter s typically iteratively decided iteratively decided

Future Work and Contributions to RAI

Future Work

1. Improvements to efficiency and accuracy
2. Continuously refining the privacy vs utility tradeoff
   - Exploring more divergences
   - Innovation on different privacy methods

Contributions to RAI

1. Privacy (the theme of this section of presentations)
2. Alternatives to pure DP (especially with categorical queries)))
3. Shows how research and mechanisms evolve over time!

Discussion-3 (10 min)

- How should software engineers/data scientists and product managers go about prioritizing privacy within an Agile sprint cycle? In what ways can we incentivize adoption of privacy practices?
- Think back to research, internship, or work experience (software engineering, data engineering, ML, etc). How would you use these 6 mechanisms to **implement privacy** into your workflows? Try to develop one specific example (without breaking confidentiality of the business/research) of where in the architecture/pipeline you would add in DP and how you would do it.
- Brainstorm methods for explaining DP implementation and **why it's important** to implement to product/program managers so that they can understand why they should incorporate it.
- Brainstorm scenarios in which it would **NOT be necessary** to implement *differential privacy*.

Resources

1. Cynthia Dwork and Aaron Roth. 2014. The Algorithmic Foundations of Differential Privacy. Found. Trends Theor. Comput. Sci. 9, 3–4 (Aug 2014), 211–407. https://doi.org/10.1561/0400000042
2. Near, J. P., & Abuah, C. (2021). Programming differential privacy. URL: https://programming-dp.com/cover.html .
3. Kamath, Gautam. Lectures 5-8 (notes). CS 860 - Algorithms for Private Data Analysis, 9/22/2022-10/4/2022. Url: http://www.gautamkamath.com/courses/CS860-fa2022.html