18734: Foundations of Privacy

Differentially Private Recommendation Systems (cont'd)

Giulia Fanti

Slides by Anupam Datta, Jeremiah Blocki

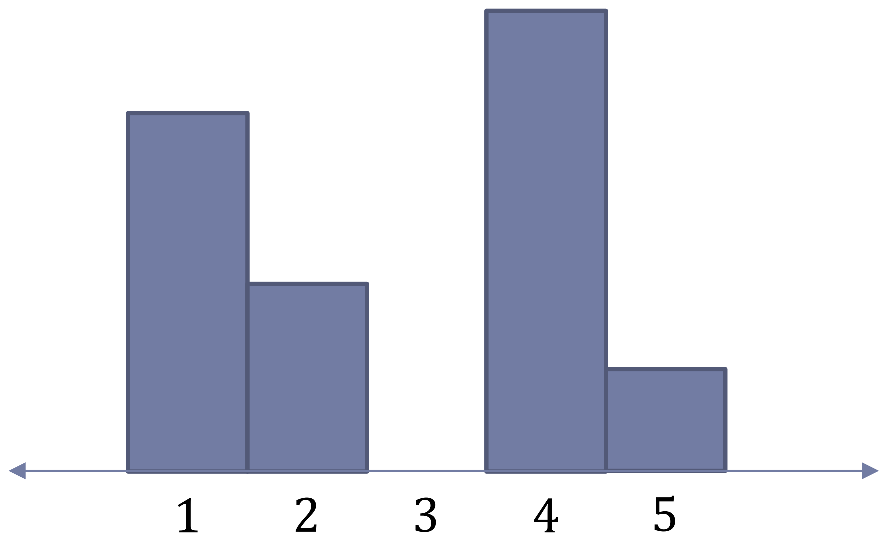
Fall 2019

Administrative

- HW2 out tonight
  - Differential privacy and deanonymization
- Project proposals
  - If you got marked down for your project, you can share new project idea with staff for feedback

Definition from last time…

- What is the support of a probability distribution?
- A: Set of values with nonzero probability mass



$\{1,2,4,5\}$

Canvas Quiz

- 10 minutes

Last time:

Differentially Private Recommender Systems:
Building Privacy into the Netflix Prize Contenders

Frank McSherry and Ilya Mironov

KDD 2019

Netflix Predictions – High Level

- Q(i,j) – "How would user i rate movie j?"

- Predicted rating may typically depend on
  - Global average rating over all movies and all users
  - Average movie rating of user i
  - Average rating of movie j
  - Ratings user i gave to *similar* movies
  - Ratings *similar users* gave to movie j

- Sensitivity may be small for many of these queries

What do we need to make predictions?

For a large class of prediction algorithms it suffices to have:

- Gavg – average rating for all movies by all users
- Mavg – average rating for each movie by all users
- Average Movie Rating for each user
- Movie-Movie Covariance Matrix (COV)

Differentially Private Recommender Systems (High Level)

To respect approximate differential privacy publish

- Gavg + NOISE
- Mavg + NOISE
- COV + NOISE

- GS(Gavg), GS(Mavg) are very small so they can be published with little noise (e.g., Laplacian)
- GS(COV) requires more care (our focus)

- Don't publish average ratings for users (used in per-user prediction phase using k-NN or other algorithms)

Source: *Differentially Private Recommender Systems(McSherry* and Mironov)

Movie-Movie Covariance Matrix

$$Cov = \sum_u (\tilde{r}_u)(\tilde{r}_u)^T$$

$$\tilde{r}_u = r_u - \bar{r}$$

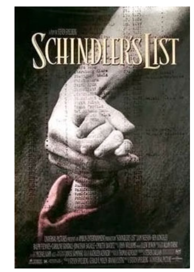
r_u: User u's rating for each movie

$\bar{r}$: Average rating for each movie

Movie-Movie Covariance Matrix

$$Cov = \sum_u (\tilde{r_u})(\tilde{r_u})^T$$

$$\bar{r} = \begin{pmatrix} 3.2 \\ 2 \\ 3 \end{pmatrix}$$



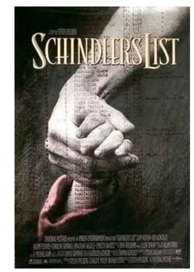
$$r_{u1} = \begin{pmatrix} 4.2 \\ 2 \\ 3 \end{pmatrix} \qquad r_{u2} = \begin{pmatrix} 1.5 \\ 4.5 \\ 2 \end{pmatrix}$$

Movie-Movie Covariance Matrix

$$Cov = \sum_{u} (\widetilde{r_u})(\widetilde{r_u})^{T}$$

$$\bar{r} = \begin{pmatrix} 3.2 \\ 2 \\ 3 \end{pmatrix}$$

$$\widetilde{r_{u1}} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \qquad \widetilde{r_{u2}} = \begin{pmatrix} -1.7 \\ 2.5 \\ -1 \end{pmatrix}$$



Example

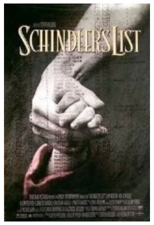
$$\widetilde{r_{u1}}(\widetilde{r_{u1}})^T = \left\langle \begin{matrix} \boxed{-1.7} \\ 2.5 \\ -1 \end{matrix} \right\rangle \left\langle \begin{matrix} -1.7 & \boxed{2.5} & -1 \end{matrix} \right\rangle$$

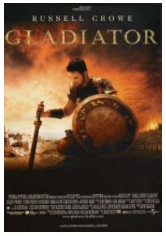
$$\mathbf{-4.25 = -1.7 \times 2.5}$$

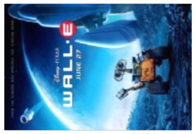
$$= \begin{bmatrix} 2.89 & \boxed{-4.25} & 1.7 \\ -4.25 & 6.25 & -2.5 \\ 1.7 & -2.5 & 1 \end{bmatrix}$$

Example

$$Cov = \widetilde{r_{u1}}\left(\widetilde{r_{u1}}\right)^T + \widetilde{r_{u2}}\left(\widetilde{r_{u2}}\right)^T$$







$$= \begin{bmatrix} 3.89 & -4.25 & 1.7 \\ -4.25 & 6.25 & -2.5 \\ 1.7 & -2.5 & 1 \end{bmatrix}$$

Goal

- Come up with differentially-private method of computing these covariance matrices
- How should we do this?

Covariance Matrix Sensitivity

$$\mathrm{Cov} = \sum_u \tilde{r}_u \tilde{r}_u^T$$

$$\|\mathrm{Cov}^{\mathrm{a}} - \mathrm{Co}v^b\| = \|\tilde{r}_u^a \tilde{r}_u^{a^T} - \tilde{r}_u^b \tilde{r}_u^{b^T}\|$$

$$\leq \left\|\tilde{r}_u^a - \tilde{r}_u^b\right\| \times (\|\tilde{r}_u^a\| + \left\|\tilde{r}_u^b\right\|)$$

- Prove this with a neighbor
- Could be large if a user's rating has large spread or if a user has rated many movies

Covariance Matrix Trick I

- Center and clamp all ratings around averages. If we use clamped ratings then we reduce the sensitivity of our function.

$$\widehat{r}_{ui} = \begin{cases} -B, & \text{if } r_{ui} - \overline{r}_u < -B, \\ r_{ui} - \overline{r}_u, & \text{if } -B \leq r_{ui} - \overline{r}_u < B, \\ B, & \text{if } B \leq r_{ui} - \overline{r}_u. \end{cases}$$

Example (B = 1)

User 1:

$$r_{u1} = \langle \boxed{4.2} \quad 2 \quad 3 \rangle$$

$$\overline{r_{u1}} = \frac{4.2 + 2 + 3}{3} \approx \boxed{3.07}$$

$$\widehat{r_{u1}} = \langle \boxed{1} \quad -1 \quad -.07 \rangle$$

$$\min\{B, 4.2 - 3.07\}$$

$$\max\{-B, 2 - 3.07\}$$

Covariance Matrix Trick II

- Carefully weight the contribution of each user to reduce the sensitivity of the function. Users who have rated more movies are assigned lower weight.

$$\text{Cov} = \sum_u w_u \widehat{r}_u \widehat{r}_u^T + \text{Noise}^{d \times d}$$

- Where e_{ui} is 1 if user u rated movie i

  and $w_u = \frac{1}{||e_u||_2}$

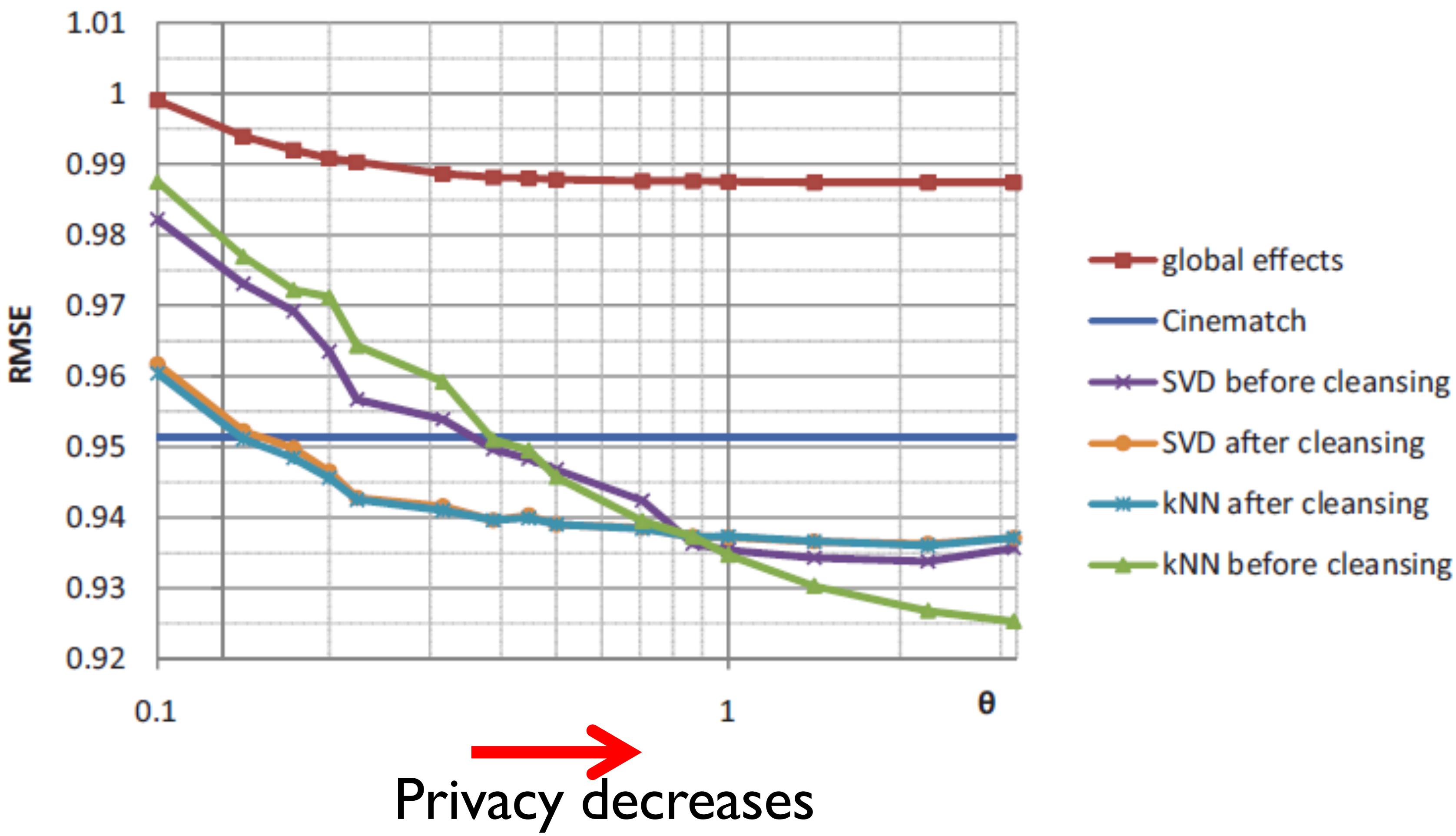
Publishing the Covariance Matrix

- Theorem 5 from paper: If ratings vectors r^a and r^b have at most one rating different, then for appropriate parameter settings, we have:

$$\|w_u^a \widehat{r}_u^a \widehat{r}_u^{aT} - w_u^b \widehat{r}_u^b \widehat{r}_u^{bT}\|_2 \quad \leq \quad (1 + 2\sqrt{2})B^2$$

- Add independent Gaussian noise proportional to this sensitivity bound to each entry in covariance matrix

Experimental Results



Source: *Differentially Private Recommender Systems(McSherry* and Mironov)

Note About Results

- Granularity: One *rating* present in D_1 but not in D_2
  - Accuracy is much lower when one user is present in D_1 but not in D_2
  - Intuition: Given query Q(i, j) the database D-u[i] gives us no history about user i.
- Approximate Differential Privacy
  - Gaussian Noise added according to L_2 Sensitivity
  - Clamped Ratings (B = 1) to further reduce noise

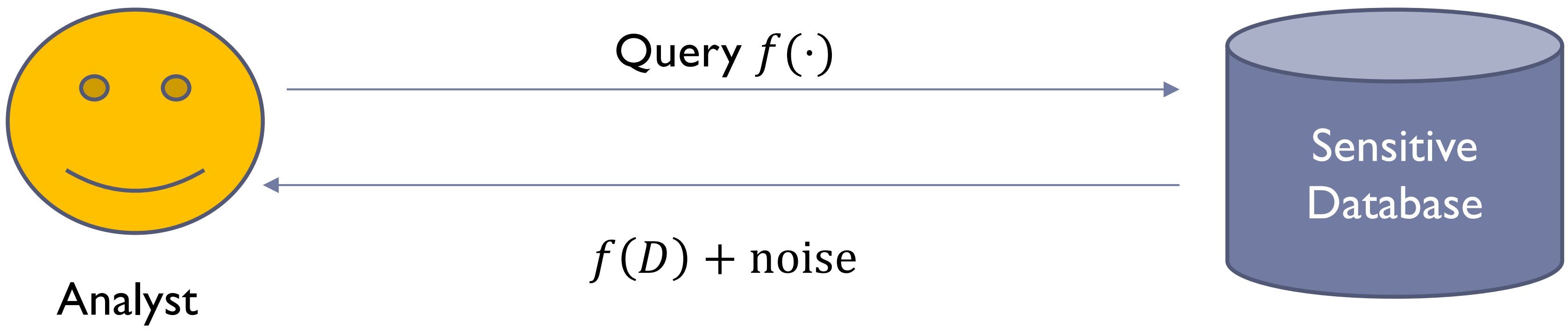
Summary

- Why did we talk about this paper?
  - Takes a complicated task (DP recommendation system)
  - Turns it into well-defined simpler task (DP covariance matrix)
- In general, you need to either
  - Bound the sensitivity of your desired function
  - Change the model to have bounded sensitivity
- What was their approach?
  - Use a bound on the sensitivity of covariance
  - Use the bound to design tools for limiting sensitivity

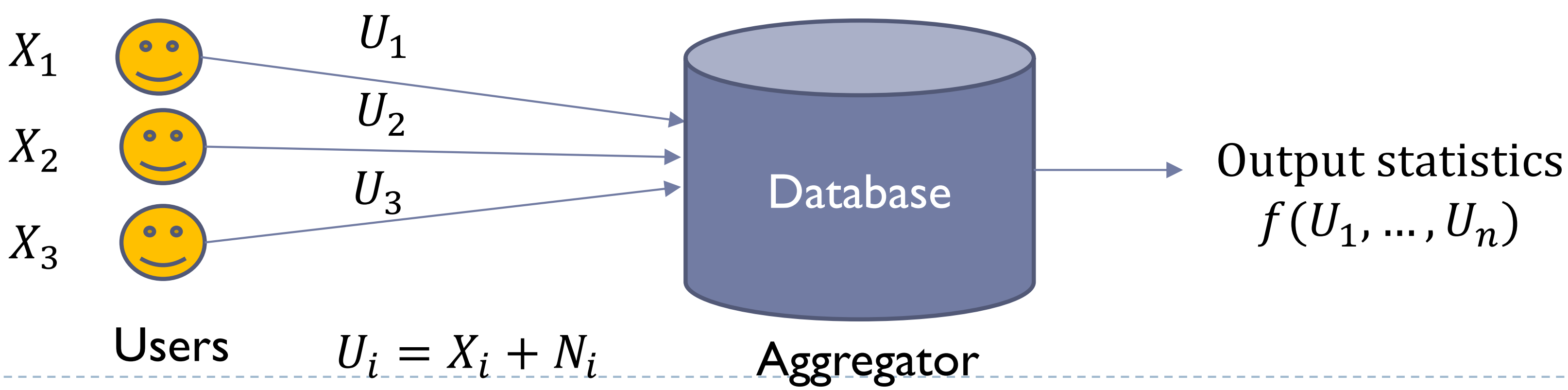
Next: Local Differential Privacy

Different models

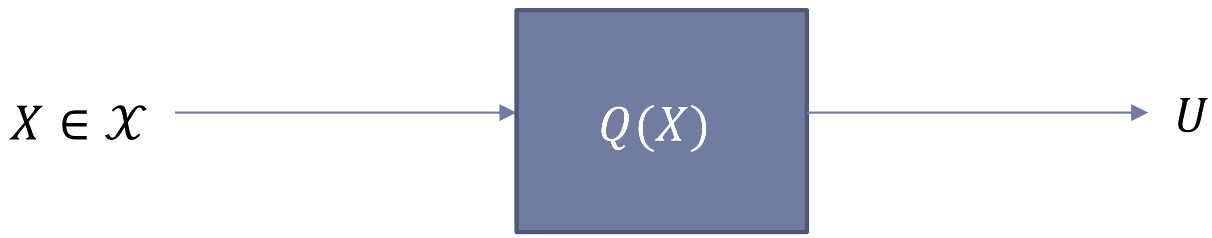
- Global (database) differential privacy



- Local differential privacy



Local Differential Privacy

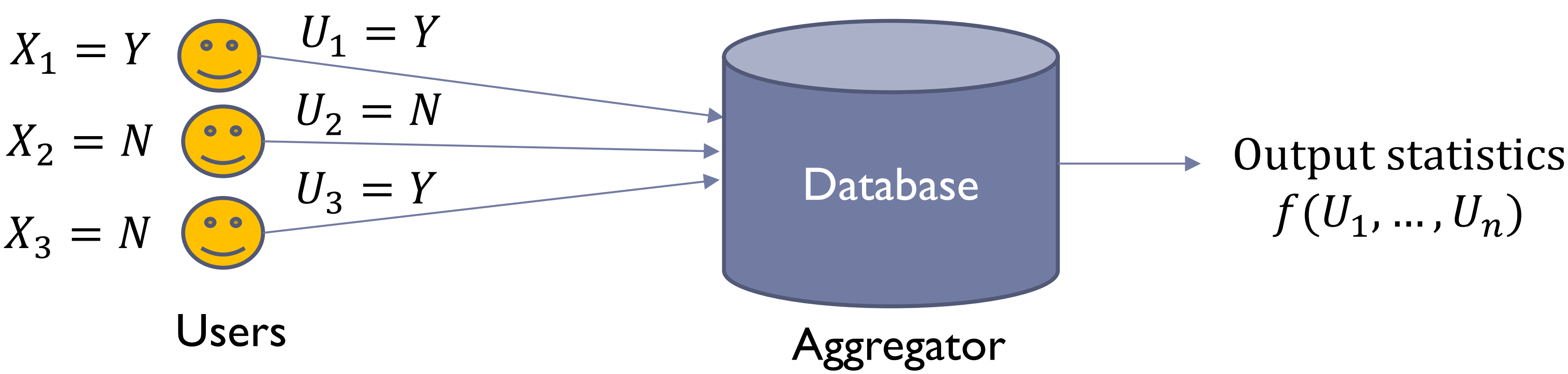


- We say mechanism Q is ϵ-locally differentially private if

$$\sup_{S, x, x' \in \mathcal{X}} \frac{Q(S|X = x)}{Q(S|X = x')} \le e^{\epsilon},$$

Example: Measuring Drug Use

Question: Have you consumed illegal drugs in the last week?



- Randomized response (Warner):
  - If heads, answer truth
  - If tails, random answer

$$\frac{P(U = Y \mid X = Y) = 0.75}{P(U = Y \mid X = N) = 0.25} = e^{\log 3}$$

Local Differential Privacy

- Widely used in practice
  - Google
  - Apple
- Mechanism is applied to privatize data itself
  - I.e., query function $f(x) = x$
- No notion of neighboring databases anymore
  - Compare P(output | input)
- Plausible deniability protects users from:
  - aggregator
  - hackers
  - surveillance