The Price of Privacy In Untrusted Recommendation Engines

Siddhartha Banerjee, Nidhi Hegde & Laurent Massoulié
UT Austin Technicolor

technicolor

Privacy – efficiency trade-offs

- ❑ Google & FaceBook track online browsing behaviour
- ❑ Apple & Android phones track geographical location

- ❑ Official reason for harvesting user data: better service results
  - ❑ Amazon's "You might also like"
  - ❑ Netflix's cinematch engine

- ❑ Privacy ≠ Anonymity: Netflix sued for disclosing anonymized "Prize" dataset

→ What trade-offs between recommendation accuracy and user privacy when service providers are untrusted?

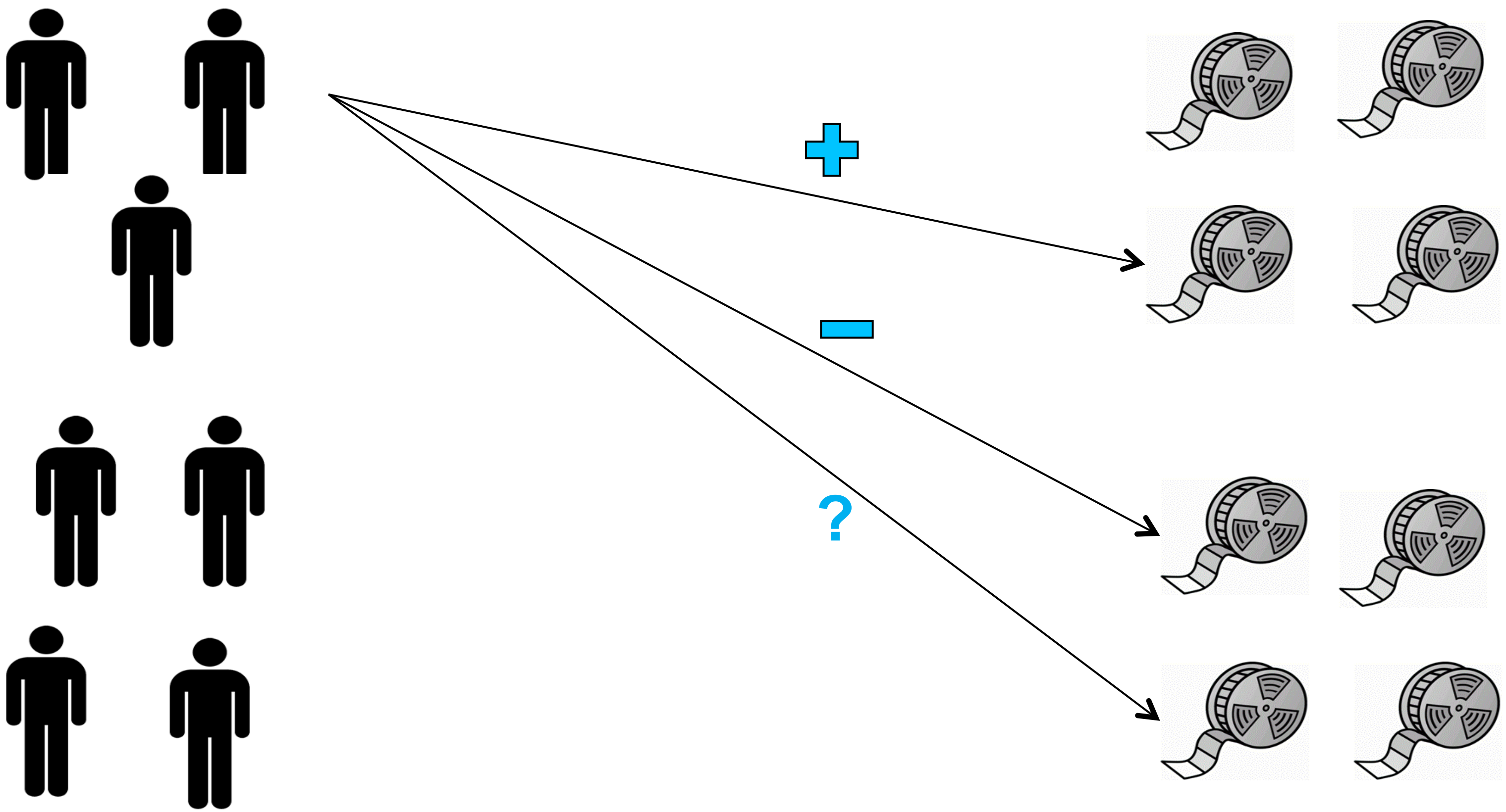
technicolor

Roadmap

- Recommendation as Learning
- "Local" Differential Privacy
- Query Complexity Bounds
  - Mutual Information and Fano's Inequality
  - Information-Rich Regime: Optimal Complexity via Spectral Clustering
  - Information-Scarce Regime: Complexity Gap and Optimality of "MaxSense"

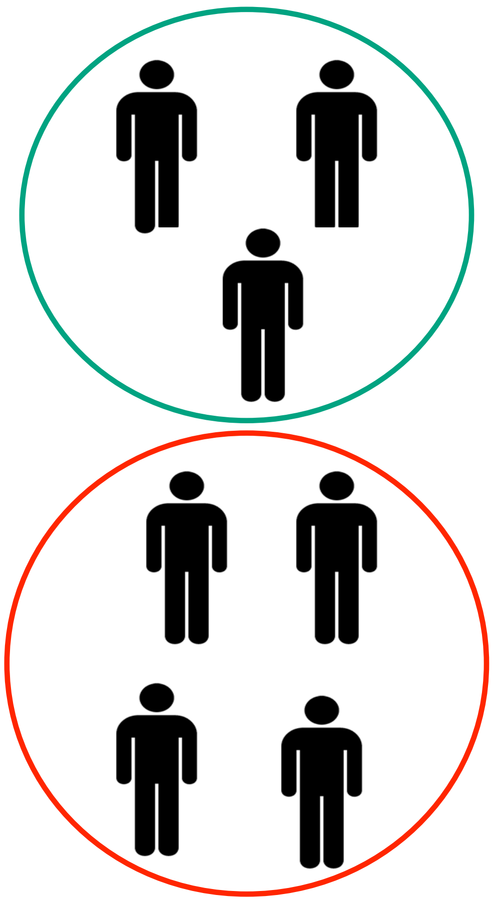
Recommendation

- ❑ Users watch and rate items (movies)
- ❑ Engine predicts unobserved ratings & recommends items with highest predicted ratings

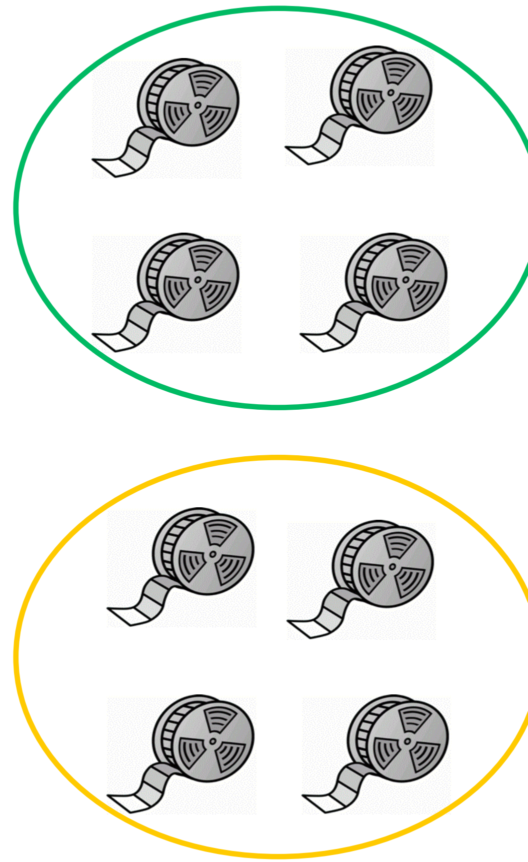


technicolor

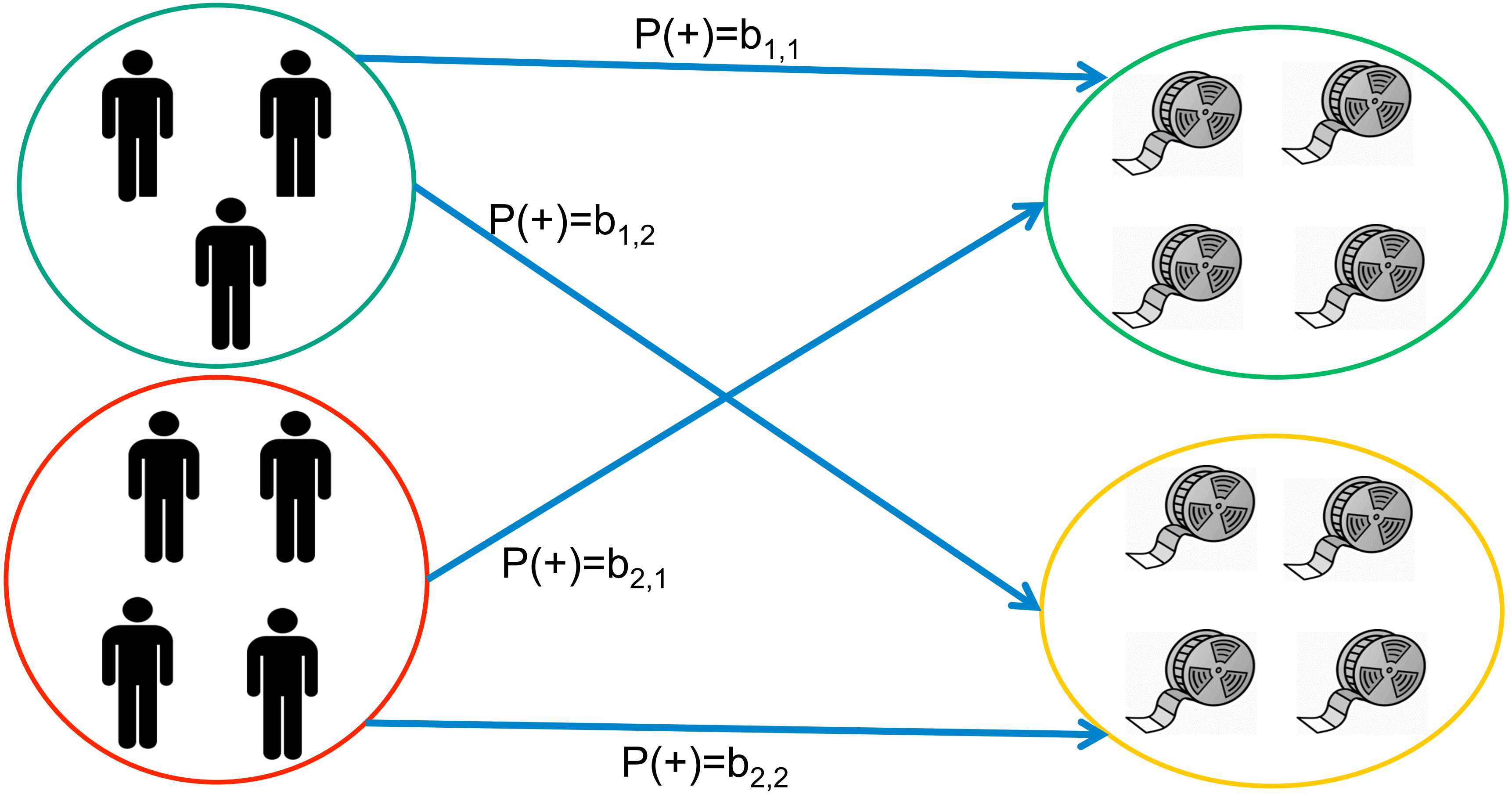
A Simple Generative Model: The "Stochastic Block Model" [Holland et al. 83]



- ❑ Each user belongs to one of K user classes
- ❑ Each movie belongs to one of L movie classes
- ❑ The rating of a user for a movie depends only on the user & movie classes



A Simple Generative Model: The “Stochastic Block Model”



Minimal requirement for recommendation:

learn movie clusters

→ Can tell what "Users who liked this have also liked"

→ Can reveal clusters and let users decide on their own their affinity to distinct clusters

Challenge: how to do so while respecting users' privacy? Without them trusting you?

Roadmap

- Recommendation as Learning
- "Local" Differential Privacy
- Query Complexity Bounds
  - Mutual Information and Fano's Inequality
  - Information-Rich Regime: Optimal Complexity via Spectral Clustering
  - Information-Scarce Regime: Complexity Gap and Optimality of "MaxSense"

Formal definition: Differential Privacy [Dwork 06]

- Input (private) data: X

→x, x’: any two possible values differing in just one user’s input

- Output (public) data Y

→y: any possible value

Definition

$$P(Y=y\mid X=x)\leq e^{\varepsilon}P(Y=y\mid X=x')$$

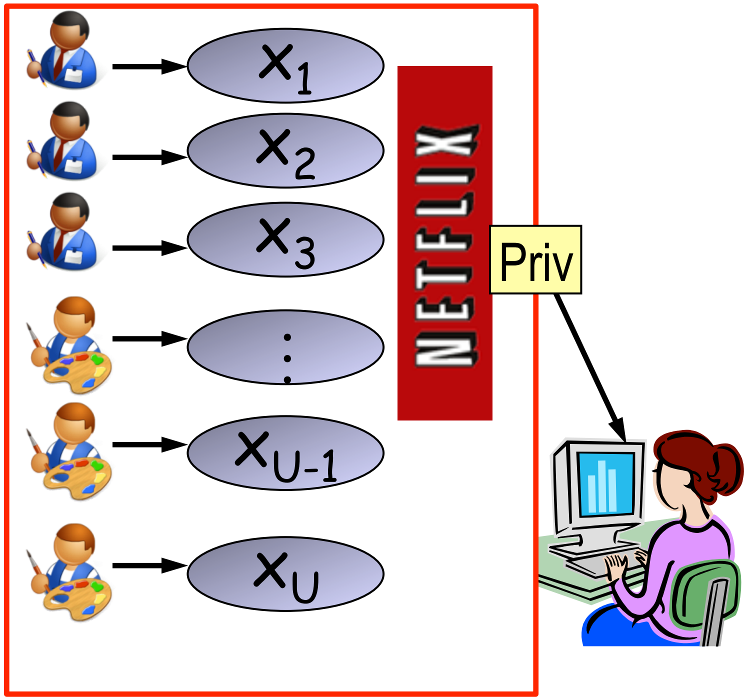
Key property: attacker holding **any** side information S trying to know whether user u has **any** property A. Then public data does not help:

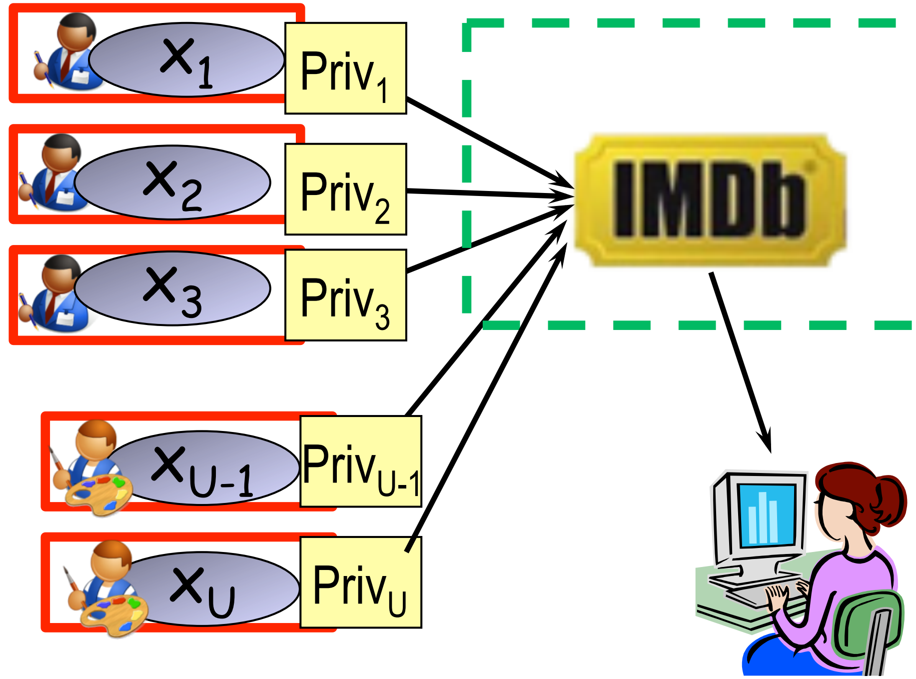
$$e^{-\varepsilon}\leq\frac{P(\text{user } u \text{ has A}\mid S \text{ and } Y)}{P(\text{user } u \text{ has A}\mid S)}\leq e^{\varepsilon}$$

technicolor


Circle of trust
YOU

Differential Privacy: Centralized versus Local





Centralized model

- Trusted DataBase aggregates Users' private data
- DP applied at egress of DB

→learning is not affected by DP

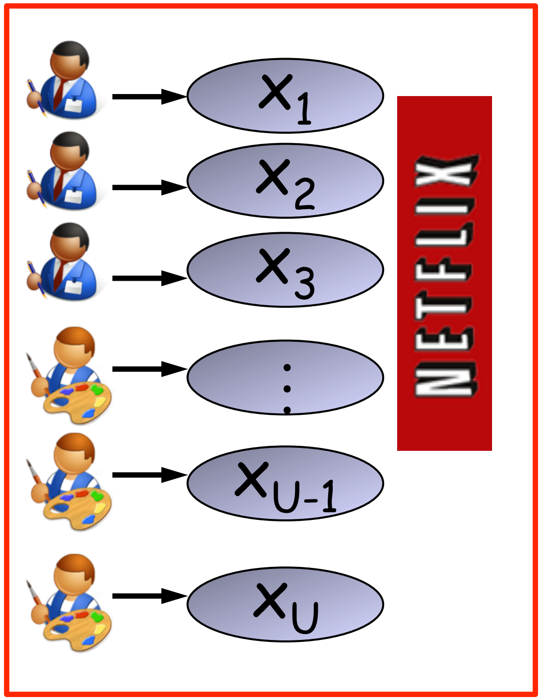
Local model

- No trusted DataBase
- DP applied locally at user end

→learning **is** affected by DP

technicolor

Example mechanisms: Laplacian noise and bit flipping

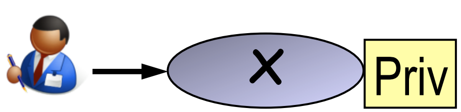


$$S = \sum_i X_i$$

$$N:\ P(N = n) = \frac{\varepsilon}{2} e^{-\varepsilon |n|}$$

$$\oplus\ \boxed{\text{Priv}}$$

$$S' = S + N$$



$$X' = X \quad \text{with prob } \frac{e^{\varepsilon}}{1+e^{\varepsilon}},$$
$$= 1 - X \text{ with prob } \frac{1}{1+e^{\varepsilon}}$$

Local DP- historical perspective

Aka “Randomized response technique” [Warner 1965]:

Used to conduct polls about embarrassing questions

“Do you understand the impact of euro-bonds on Europe’s future?”

Answer truthfully only if score >2

→Specific answers are deniable

→Empirical sums are still valid **for learning few parameters**

Inadequate for learning many parameters: with k distinct ε-private sketch releases, overall privacy guarantee becomes k ε

Roadmap

- Recommendation as Learning
- "Local" Differential Privacy
- Query Complexity Bounds
  - Mutual Information and Fano's Inequality
  - Information-Rich Regime: Optimal Complexity via Spectral Clustering
  - Information-Scarce Regime: Complexity Gap and Optimality of "MaxSense"

Learning, Mutual Information and DP

Want to learn hypothesis H from M distinct possibilities

(e.g. clustering of N movies into L clusters: $M \approx L^N$ options),

Having observed G (e.g., DP inputs of U distinct users)

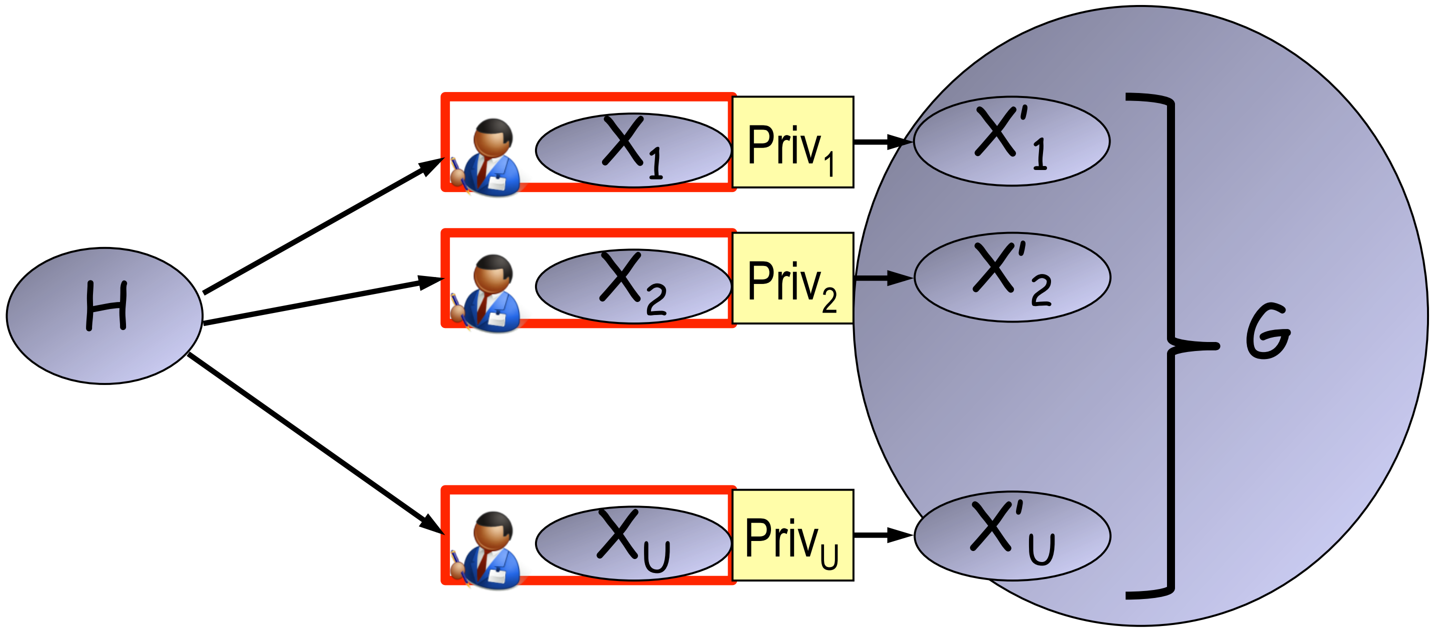
Fano's inequality: Learning will fail with high probability,

unless mutual information I(H;G) close to log(M)

Mutual information: $$I(H;G) = \sum_{h,g} P(H = h, G = g)\log\left(\frac{P(H = h, G = g)}{P(H = h)P(G = g)}\right)$$

Learning, Mutual Information and DP

Result: DP-sketch X' based on private data X verifies for any side information S: $\boxed{I(X;X'|S)\leq \varepsilon}$

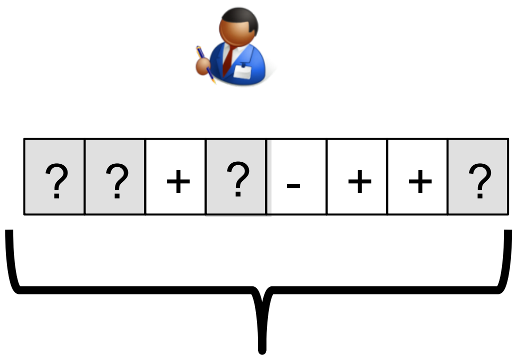


→ Mutual information I(H;G): at most U*ε

→ "Query complexity": need at least N/ε users' private inputs to recover hidden clusters

technicolor

The Information-Rich and the Information-Scarce Regimes



Out of N items in total, users rate W movies

(assumed picked uniformly at random)

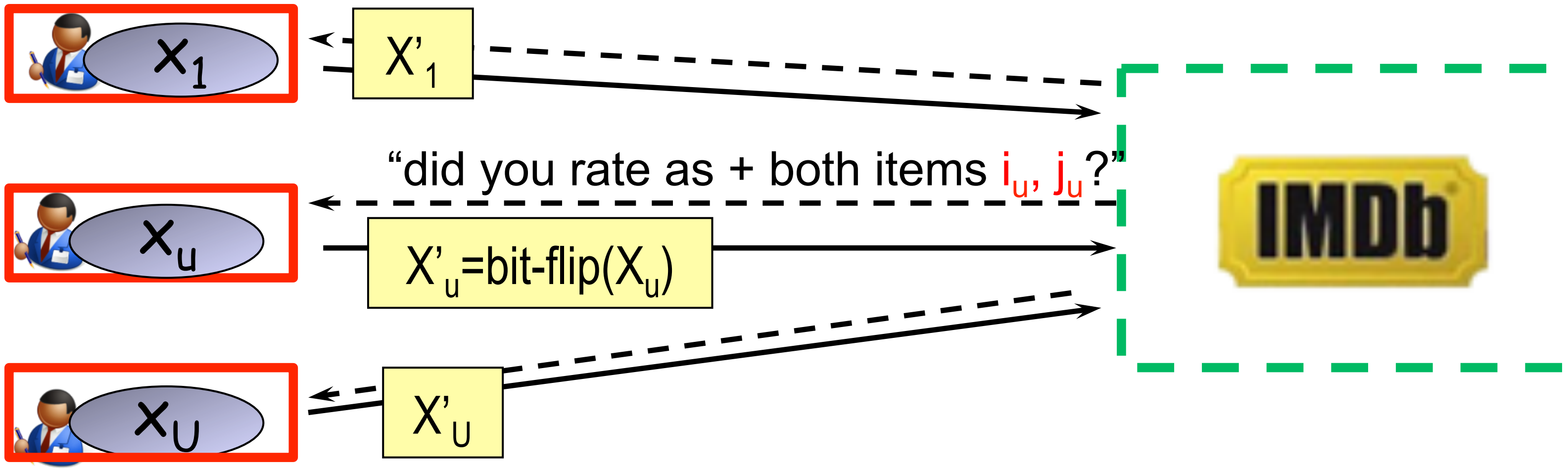
→ Information-rich regime: $W=\Omega(N)$

→ Information-scarce regime: $W=o(N)$

Users' "information wealth" will affect optimal query complexity

technicolor

The information-rich regime: Pairwise-preference algorithm

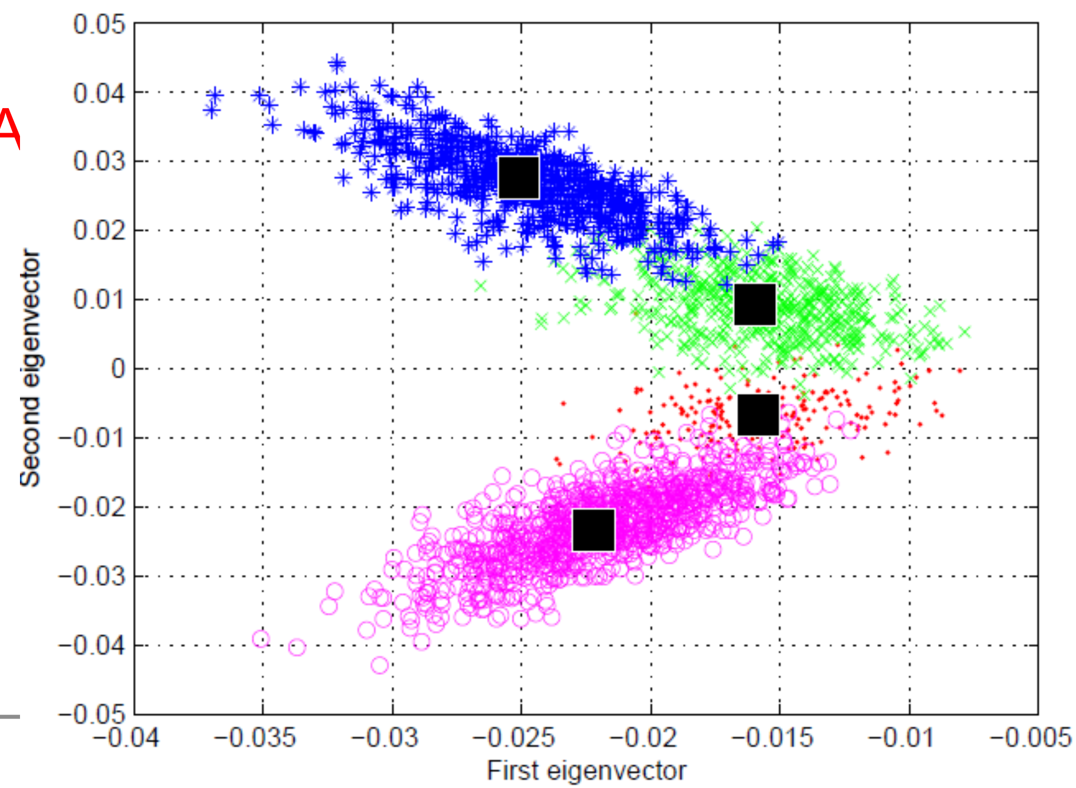


Construct item affinity matrix A

$$A_{ij} = \mathrm{Min}\left(1, \sum_{u=1}^{U} X'_u 1_{(i_u j_u)=(ij)}\right)$$

Spectral clustering of items

based on A



The information-rich regime: Pairwise-preference algorithm

Result: Algorithm finds hidden clusters w.h.p. if $U = \Omega(N \log N)$

under "block distinguishability" conditions on underlying model

→optimal, up to logarithmic factor

Proof elements: matrix A: adjacency of ER-like graph, with

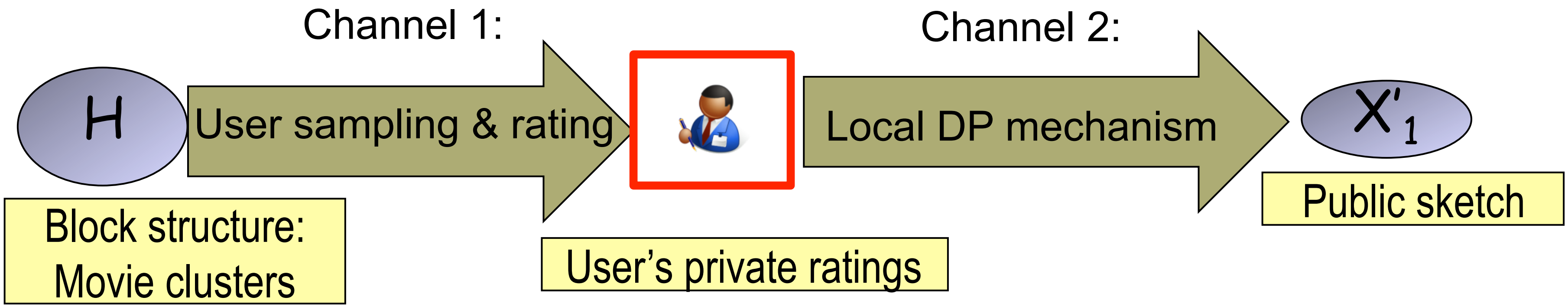
$$E(A_{ij}) = \underbrace{2\frac{U}{N(N-1)}\frac{W(W-1)}{N(N-1)}}\sum_k \pi_k\left[(1-2\varepsilon)b_{k\ell(i)}b_{k\ell(j)}+\varepsilon\right]$$

When prefactor is $\Omega(\log N/N)$, top eigenvectors determine underlying block structure

[Feige-Ofek 2005; Tomozei-M 2011]

technicolor

The information-scarce regime: lower bounds



Channel mismatch will make end-to-end mutual information much lower than minimum of each mutual information

Intuition: to question "did you rate item i with a +?", user's answer will be informative only with chance W/N

→ Information in public sketch is "diluted" by factor W/N

The information-scarce regime: lower bounds

Result: Assume two item clusters, and each user u observes true type Z_i of W randomly picked items i

Then: a user's DP sketch X' verifies $I(H;X')=O(W/N)$

Corollary: to learn hidden clustering of N items from parallel queries to U users needs $U=\Omega(N^2/W)$

e.g. $N=10^4$, $W=100$ needs $U= \Omega(10^6)$

$N=10^6$, $W=100$ needs $U= \Omega(10^{10})$

→ need to query non-humans!

Proof elements

1) Bound on mutual information

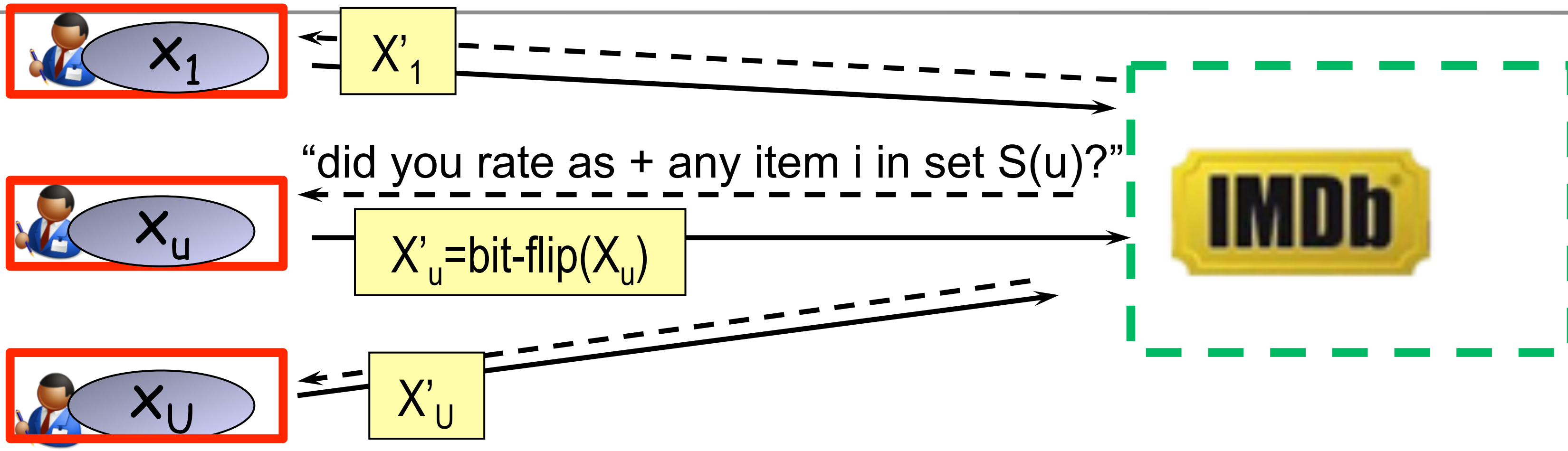
$$\mathcal{I}(\mathbf{Z}; S) \leq \mathbb{E}_S \left[\mathbb{E}_{(I_1,Z_1)|S \perp\!\!\!\perp (I_2,Z_2)|S} \left[2^{|I_1 \cap I_2|} \mathbb{1}_{\{Z_1 \equiv Z_2\}} - 1 \right] \right]$$

→ A convex quadratic form of the kernels p(I,Z | S)

2) Identification of extremal kernels

3) Some Euclidean geometry…

Information-scarce regime: Max-Sense algorithm



User query: Sense random set S(u) of size N/W

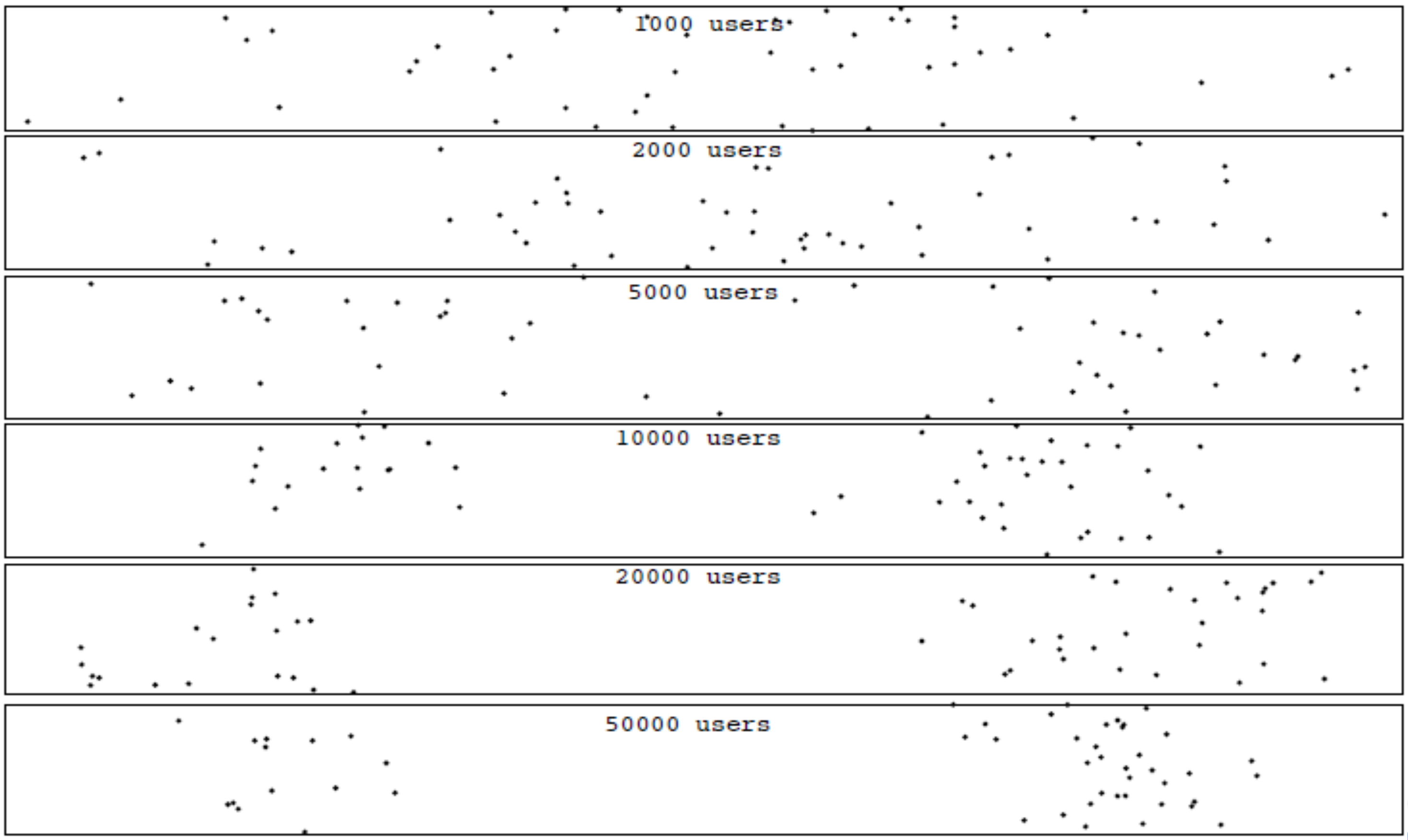
Item representative: $$\boxed{T(i) = \sum_{u=1}^{U} X'_u 1_{i \in S(u)}}$$

technicolor

Information-scarce regime: Max-Sense algorithm

Result: under separability assumption, k-means clustering of item representatives find hidden clusters w.h.p. if $U=\Omega(N^2\log(N)/W)$

→Optimal scaling, up to logarithmic factor



Conclusions and Outlook

- Mutual Information adequate to characterize learning complexity under local DP constraints
- Accurate Clustering, Local Differential Privacy, Low (linear) Query Complexity: leave one out!
- MaxSense achieves optimal complexity for parallel queries
- Can one beat its complexity with adaptive queries?
- Alternatives to Differential Privacy?

Questions?

Lower bounds for adaptive queries

Can one improve complexity by adapting queries based on previous user answers?

Result: for W=1, arbitrary side information S

Then user's DP sketch X'_u verifies $I(X'_u ; H \mid S) \leq O\left(\frac{1}{N}\right) \mathrm{Max}(1, I(H;S))$

→Adaptive query complexity at least $\Omega(N \log(N))$

Larger than initial lower bound by logarithmic factor

CONJECTURE: Query complexity lower bound of N^2/W still holds with adaptive queries