Differentially Private Decomposable Submodular Maximization
Anamay Chaturvedi†, Huy Lê Nguyên†, Lydia Zakynthinou†
†Northeastern University

Problem
We consider the problem of differentially private decomposable submodular maximization.
• Submodular functions \( f : 2^I \to \mathbb{R} \) have diminishing returns:
  \( S \subseteq T, u \notin T \Rightarrow f(S \cup \{u\}) - f(S) \geq f(T \cup \{u\}) - f(T) \).
• Decomposable submodular:
  \[ f_D(S) = \sum_{\text{agents } i \in D} f_i(S) \]
• Want to privately maximize decomposable submodular functions subject to a matroid constraint.
• Central model of differential privacy.
• Applications: – exemplar-based clustering – image summarization – recommender systems – document and corpus summarization

Approach
Continuous greedy methods of Vondrák 2008 and Feldman, Naor, and Schwartz 2011:
• Maximise multilinear relaxation of \( f_D \):
  \[ F_D(x) = \sum_{x \in S} f_D(S) \prod_{x \in S} (1 - x_i) \].
• \( T \) rounds. Iteratively pick feasible \( i \) maximising \( F(x) \) on increasing \( x_i \) by a \( 1/T \) step (monotone \( f_D \)); \( (1 - x_i)/T \) step (non-monotone \( f_D \)).
• \( x \) in convex hull of feasible sets. Swap-rounding of Chekuri, Vondrák, and Zenklusen 2010 returns feasible solution with good utility.

Highlights
• Greedy picks via exponential mechanism - following Gupta et al. 2010 get loss in privacy independent of number of rounds.
• Estimate \( f_D \) by sampling and sharing randomness between rounds - this avoids additional utility loss in each round.
• Directly replacing each round of continuous greedy by the private greedy does not work.
• Additive error \( \sim O \left( \frac{\epsilon}{r} \log nr \cdot \log \frac{1}{\delta} \right) \) close to known lower bound of \( O \left( \frac{\epsilon}{r} \log n/r \right) \).

Results
• Monotone rank \( r \) - matroid-constrained case we are \((\epsilon, \delta)\)-private using \( T \) rounds with expected utility
  \[ (1 - 1/e - O(1/T))f(0^T) - O \left( \frac{T}{\epsilon} \log nr \cdot \log \frac{1}{\delta} \right) \]
• Analogous non-monotone case:
  \[ (1/e - O(1/T))f(0^T) - O \left( \frac{T}{\epsilon} \log nr \cdot \log \frac{1}{\delta} \right) \]
Related work:
• Work by Gupta et al. 2010 and Mitrovic et al. 2017 used a discrete greedy algorithm. By adapting continuous methods improve multiplicative factor from \( 1/2 \) (Mitrovic et al. 2017) to \( (1 - 1/e - O(1/T)) \) in the monotone case.
• Rafiey and Yoshida 2020 also adapt continuous greedy methods but obtain significantly higher additive error of \( nr \cdot \log n/e^3 \).

Experimental results
We replicate the Uber location selection experiment of Mitrovic et al. 2017.
• Given a set of pick-up locations in Manhattan, the goal is to pick locations close to pick-ups while private with respect to pick-ups.
• Scaled \( r \) - distance between location \( l \) and pick-up \( p \):
  \[ M(l, p) = \frac{|l - p|}{\epsilon} \leq 1 \]
• Utility of locations \( S \) evaluated on pick-ups \( D \):
  \[ f_D(S) = \sum_{l,p \in D} \left( 1 - \min_{p \in D} M(l, p) \right) = |D| \cdot \sum_{l,p \in D} \min_{p \in D} M(l, p) \] (1)
\( f_D \) is monotone decomposable submodular function. We conduct two experiments:
• a rank constrained location selection for \( 100 \) agents at a time. Comparison with more general algorithm of Mitrovic et al. 2017 that uses the composition laws of privacy instead of the Gupta privacy analysis.
• simple 3-element partition matroid instance measuring per-capita utility versus dataset size. Comparison with discrete method for matroids of Mitrovic et al. 2017

Utility versus rank, \( \epsilon = 0.1 \)
Utility versus number of data, \( \epsilon = 0.1 \)