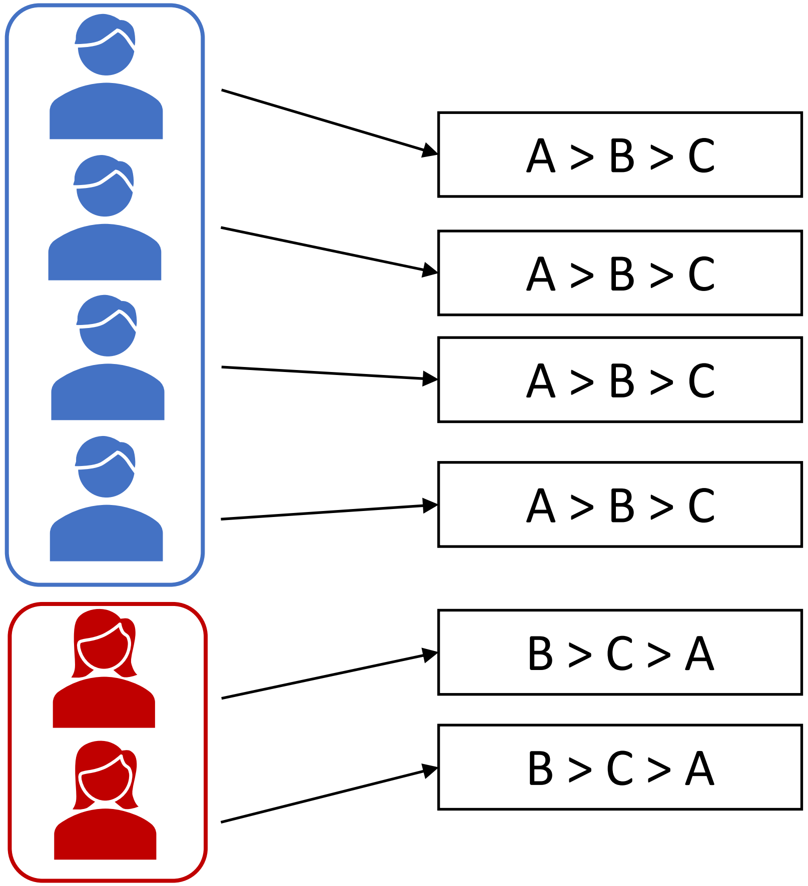
Designing Fair and Private Voting Rules

Farhad Mohsin (mohsif)

Course project for CSCI 6968 (Trustworthy ML)

Introduction



- Voters divided into groups by features, e.g., gender, race, age
- Traditional voting rules maximizes some measure of economic efficiency
  - Winner might be highly preferred to majority group, being *unfair* to minorities
- A fair voting rule looks at voter features and chooses an alternative with similar utility to both groups
- Considering voter features risk loss of privacy
- How to get a voting rule both fair and private?

Preliminaries

Voting scenario

- Two groups, n_1, n_2 voters in each.
- Set of m alternatives $\mathcal{A}$. Voters give a ranking over all m alternatives
- Assumption
    - Any voter receives utility $(m - j)$ if their j-th ranked alternative won
    - Example:

      Alternatives = $\{A, B, C, D\}$

      Voter i's ranking : $A \succ C \succ B \succ D$

      i gets 3 utility if A wins, 2 for C, 1 for B, 0 for D
- A collection of votes ≡ preference profile, we have two preference profiles P_1, P_2
- Average utility of group for alternative a is $W(a, P_1)$

[1]. Mohsin et al., *Learning to design fair and private voting rules* (2021)

Fair and Private Outcomes in Voting

Definition (Group Imbalance in Voting) [1]

In a voting scenario, if two groups have voting profile (collection of rankings) P_1, P_2, then the group imbalance for any alternative is

$$\Delta W(a, P_1, P_2) = | W(a, P_1) - W(a, P_2) |$$

where $W(a, P)$ is the average utility for voters in P if a wins.

$$\text{Fair outcome} \equiv \operatorname*{argmin}_{a} \Delta W(a, P_1, P_2))$$

Definition (ϵ −Differential Privacy in Voting) [2,3]

A randomized single-winner voting rule r satisfies $\epsilon - DP$ if for preference profiles P, P' differing only in one vote

$$\Pr(r(P) \in S) \leq \exp(\epsilon) \Pr(r(P') \in S)$$

For any subset of alternatives S

[1]. Mohsin et al., *Learning to design fair and private voting rules* (2021)
[2]. Hay et al., *Differentially Private Rank Aggregation* (2017)
[3]. Lee, Efficient, Private, and eps-Strategyproof Elicitation of Tournament Voting Rules (2013)

Fair and Private Voting Rules

- Problem
  - Input: Preference profile of two groups P_1, P_2
  - Goal: Design a voting rule that is $\epsilon - DP$ and approximately fair
  - Approximate fairness:
    - If winning alternative is a, it satisfies α-approximate fairness if
$$\Delta W(a, P_1, P_2) \leq (1 + \alpha) \min_{a'} \Delta W(a', P_1, P_2)$$
i.e, not too imbalanced compared to least imbalance

Baseline Algorithm

- Laplace mechanism (baseline)
  1. Add Laplace noise to group utility values $W(a, P_k)$ for all a and $k = 1{,}2$(both groups) to get noisy estimates for all alternatives $\widehat{W}(a, P_k)$
  2. Compute Imbalance and final outcome in terms of $\widehat{W}(a, P_k)$

- Theoretical guarantees
  - If added noise is $Lap(\frac{m(m-1)}{2n\epsilon})$, then Laplace mechanism is $\epsilon - DP$
  - Estimate for group utility is an unbiased estimate
  - Thus, estimate for utility difference is also unbiased

Sampling Algorithm

- Sampling algorithm
  1. With probability $\delta/2$, return a random winner
  2. With probability $1-\gamma$, follow steps 3-5
  3. Fix some $n_s \leq \min(n_1, n_2)$ as sampling parameter
  4. For each alternative $a \in \mathcal{A}$, each group $k = 1{,}2$
     - Sample $\sim n_s$ pairwise comparisons from the voters in group k
       - e.g. from ranking $A \succ B \succ C$, sample $A \succ B$
     - Assign $\overline{W}(a, P_k)$ = number of pairwise comparisons where a wins
  5. Compute Imbalance and final fair outcome in terms of $\overline{W}(a, P_k)$

- Theoretical guarantees
  - Differentially private when samples from each group, $n_s = O(\frac{m}{\epsilon^2}\ln\frac{m}{\delta})$
  - Value for fairness gives $\epsilon - approximate$ fairness with probability $\geq 1-\delta$
  - Only works when number of samples from a group $n_s \leq \epsilon n$

Experimental Results

Experimental setup

- $m = 4$ alternatives
- Rankings for two groups sampled from different Mallow's distributions
- Results averaged over 1000 samples

				ϵ	
n1	n2	Algorithm	0.3	0.6	1
1000	500	laplace	**0.45**	**0.45**	**0.45**
		sampling	0.47	0.55	0.62
2000	1000	laplace	**0.45**	**0.45**	**0.45**
		sampling	0.46	0.58	0.65
4000	2000	laplace	**0.45**	**0.45**	**0.45**
		sampling	0.46	0.55	0.69

Table: Difference of utilities for the two fair-private voting algorithms at different levels of privacy

Conclusion and Future Work

- Results
  - Laplace mechanism is sufficient enough for fairness
  - Sampling mechanism suffers loss due to sampling and loss of information

- Future work
  - Look at the three-way fairness-utility-privacy trade-off
  - Local-DP version of sampling mechanism possible. Compare with regular local-DP method (e.g., randomized response)