Fair and Private Voting Rules

Project Report — CSCI 6968 — Trustworthy ML

Farhad Mohsin

Introduction

Voting rules are the most popular method of aggregating individual preferences to make group decisions. There are many existing voting rules and most satisfy the anonymity property, which means that every voter's vote is considered equal. But we think that sometimes when specific demographic groups are affected significantly differently by a voting outcome, it can be useful to consider group membership explicitly. So, we try to design a voting rule that is fair to groups. To circumvent the issue of privacy, we propose designing a differentially private voting rule. In this work, we analyze the trade-off between fairness and privacy of a simple Laplacian mechanism and a sampling based mechanism.

Background

Voting rules are probably the most popular method of making collective decisions from individual preferences. Given a set of alternatives, $\mathcal{A}$, let $\mathcal{L}(\mathcal{A})$ be the set of all linear orders over $\mathcal{A}$. A voting rule, $r : \mathcal{L}(\mathcal{A})^n \mapsto \mathcal{A}$ is defined as a mapping from a preference profile (a set of linear orders) to a single winner, i.e., a voting rule aggregates n preferences to find a winning alternative. Assume two groups of voters, with group size n_1 and n_2. So, we essentially have two preference profiles, $P_1 \in \mathcal{L}(\mathcal{A})^n_1, P_2 \in \mathcal{L}(\mathcal{A})^n_2$. For every voter, we can usually assume a utility function based on the outcome of the rule. For this work, we assume a simple ranked utility function, where an voter gets utility $m - k$ if there k-th ranked alternative wins. So, every voter has a different utility for an alternative, say $u_i(a)$ is the utility to some voter i for alternative a. We can thus define average group utilities.

$$W_1(a) = \frac{1}{n_1} \sum_{i \in P_1} u_i(a)$$

$$W_2(a) = \frac{1}{n_2} \sum_{i \in P_2} u_i(a)$$

Many popular voting rules (we refer to [1] for an introduction to voting rules) are defined to be *anonymous* (the rule treats every voter equally) and economically efficient (tries to maximize total utility to all voters). However, in presence of explicit group memberships among voters in terms of protected features (e.g., men and women, vegetarians and non-vegetarians etc.), anonymous and efficient voting rules can be unfair if one of the groups have majority in membership. This is motivated by the recent work in fair algorithmic decision-making(See [2] for an introduction to the topic). This may be undesirable in various scenarios, where being fair in terms of the protected features is important, even if it goes against the rule of the majority. So, designing new fair voting rules while considering the protected features is an important new problem.

Additionally, privacy regarding expressed preferences is also an important issue. There has been some recent work done in terms of private aggregation of preferences, which consider differential privacy [3, 4, 5, 6]. But they mostly focus on economically efficient aggregation as well. The relation between fairness and privacy is also an important consideration.

Related Work

To our knowledge, no work has considered the trade-off between group fairness, privacy and economic efficiency in voting other than our research group's past work [7]. The privacy mechanism used was randomized response whereas we had proposed two methods of designing fair voting rules: 1. a utility-constrained fairness maximization method and 2. a machine learning based framework. However, the results showed a three-way trade-off, in that high privacy requirement reduced both fairness and efficiency with fairness and efficiency also having an inverse relation.

Some work has analyzed the trade-off between economic efficiency (specifically utility) and privacy in voting before. Shang et al. [3] considered multiple methods, including the popular Laplacian mechanism. On the other hand, Lee [4] considers a sampling based method. We adapt both this approaches for designing fair and private rules in this work.

Algorithms for Fair and Private Voting Rules

As mentioned in the background, we assume only two groups of voters, with n_1 and n_2 members respectively. What we want is a voting rule that is fair and private.

First, we briefly discuss what it means to be fair in terms of groups. In group fairness literature in ML, we see that statistical parity is an important measure. Motivated by that, we say that an outcome is fairer, when the difference between utilities is smaller. That is, we want the utility difference, $\Delta W(a, P_1, P_2)$ to be small.

$$\Delta W(a, P_1, P_2) = \mid W_1(a) - W_2(a) \mid$$

Additionally, we require an $\epsilon - DP$ algorithm for voting. Based on this, we propose two algorithms.

Results

Without detailed proofs, we provide the following claims with a short discussion.

1. Algorithm 1 is $\epsilon - DP$. This is easy to see when we notice that for each $W_k(a)$, the sensitivity is $\frac{m-1}{n_k}$ for our defined utility function

Algorithm 1 Laplace Fair-Private Voting

```
1: Inputs: Group preference profiles $P_1, P_2$
2: Calculate $W_1(a)$ and $W_2(a)$ for each alternative, $a$
3: $\hat{W}_k(a) \sim W_k(a) + Lap(0, \frac{m(m-1)}{2n_k\epsilon})$ for all $a$ and $k = 1, 2$
4: Compute $\Delta\hat{W}(a) = \mid \hat{W}_1(a) - \hat{W}_2(a) \mid$
5: $a = \arg\min_a \Delta\hat{W}(a)$
6: Output: Fair and private outcome $a$
```

Algorithm 2 Sampling-based Fair-Private Voting

```
1: Inputs: Group preference profiles $P_1, P_2$, sampling parameter $n_s$, and $\gamma$
2: With probability $\gamma$, output any random alternative, otherwise go to line 3
3: Initialize $S_k(a) = 0$ for all alternatives $a$ and $k = 1, 2$
4: Sample $n_s$ voters uniformly from both $P_1, P_2$ to get sampled preference profile $\tilde{P}_1, \tilde{P}_2$
5: for Each alternative $a$ do
6:     for $k \in \{1, 2\}$ do
7:         $n_x \sim Binom(n_s, \frac{2}{m})$
8:         for $i = 1$ to $n_x$ do
9:             Sample a random voter $v$ from $P_k$
10:            Sample a random alternative $b \neq a$
11:            if $v$ prefers $a$ to $b$ then
12:                $S_k(a) += 1$
13:            end if
14:        end for
15:    end for
16: end for
17: $\tilde{W}_k(a) = S_k(a)/n_k$ for all $a$ and $k = 1, 2$
18: Compute $\Delta\tilde{W}(a) = \mid \tilde{W}_1(a) - \tilde{W}_2(a) \mid$
19: $a = \arg\min_a \Delta\tilde{W}(a)$
20: Output: Fair and private outcome $a$
```

and each W_k is essentially an m-dimensional vector. We can also easily show $\Delta\hat{W}_k$ to be an unbiased estimator for ΔW_k.

2. Algorithm 2 is $\epsilon{-}DP$ if $n_s = O(\frac{m}{\epsilon^2}log\frac{m}{\delta})$ and also $n_s \leq \frac{\min(n_1,n_2)}{\epsilon}$. So, effectively n_s defines the value of ϵ for the method. The proof of this is slightly more involved, but it is essentially a variant of the proofs of Lemma 1, 2 and 5 from [4]. We do want to note some significance of the conditions though. The number of samples is in fact independent of n_1, n_2 and only dependent on m. This allows for equal sample size to be equally effective for both groups. However, the sample complexity indicates, that for small preference profiles, the sampling method might not be that useful.

Experimental Results

We show some very simple experimental results. We generate random preference profile for each group from separate Mallow's models [8] which is a popular model for group preferences. So, for our experiments shown here, the two groups have distinctly different preferences. We limit ourselves to just elections with four alternatives. We change the group sizes between the simulations. The results are presented in Table 1.

n_1	n_2	Algorithm	ϵ	ΔW	W
1000	500	laplace	0.3	0.45	0.68
			0.6	0.45	0.68
			1	0.45	0.68
		sampling	0.3	N/A	0.74
			0.6	0.57	0.95
			1	0.68	1.06
2000	1000	laplace	0.3	0.45	0.68
			0.6	0.45	0.68
			1	0.45	0.68
		sampling	0.3	N/A	0.77
			0.6	0.55	0.94
			1	0.58	1.04
4000	2000	laplace	0.3	0.45	0.68
			0.6	0.45	0.68
			1	0.45	0.68
		sampling	0.3	0.47	0.77
			0.6	0.56	0.97
			1	0.65	1.05

Table 1: Difference in utility and total average utility under various levels of privacy for the sampling and Laplace voting algorithms

Each value in the table represents the average of 1000 sample elections. For each sample, we compute the outcomes of the noisy elections. For this noisy outcome, we compute both the difference of utilities (ΔW) and also the overall average utility (W, average utility for the whole population, not just a single group). For the fairness requirement, we want the difference of utilities to be low. We notice that the fairness values are roughly equal for all ϵ values of the Laplace mechanism.

The results for the sampling mechanism is more interesting though. First, we notice that we cannot get enough samples for a differentially private mechanism for a couple of scenarios (e.g., $n_1 = 1000, n_2 = 500, \epsilon = 0.3$). Then, in a regular noise adding mechanism, usually the low noise (and thus low privacy versions) behave more closely to the ideal version. We surprisingly see that the high privacy version actually has better fairness (lower utility difference). This is because of the sample complexity's dependence on a $\frac{1}{\epsilon^2}$ term, the high privacy version actually uses more samples. Our theoretical analysis indicates that this was actually done to get something that is both close to the original ranking in terms of utility and privacy, and we should be able to tweak the sampling algorithm to better represent expected trade-off behavior.

Conclusion

Going into the project, we had hoped to find the sampling algorithm of more interest and the Laplace sampling as a simple baseline. We notice that the Laplace algorithm actually has better performance in terms of fairness. And the sampling algorithm has unexpected behavior in terms of the fairness-privacy trade-off. Additionally, in Table 1, the utility values (W column) indicate that looking into the utility-fairness-privacy trade-off together might give better direction for designing a better sampling based algorithm.

So, in conclusion, the project has left us with more questions than answers regarding the sampling algorithm.

Acknowledgment

I have previously worked on using ML techniques to design fair and private voting rules as part of an ongoing project [7] with my advisor, Prof Lirong Xia and IBM collaborators. Work for this project would be completely new work that I would do by myself. However, I may discuss the project with my advisor and collaborators from RPI and IBM Research. I also thank the Professors for their guidance throughout the semester. In particular, the multiple deadlines and the discussions prior to those helped. Without them, I feel like I might have procrastinated and tried to do everything in the last week.

References

[1] Felix Brandt, Vincent Conitzer, Ulle Endriss, Jérôme Lang, and Ariel D Procaccia. *Handbook of computational social choice*. Cambridge University Press, 2016.

[2] Alexandra Chouldechova and Aaron Roth. A snapshot of the frontiers of fairness in machine learning. *Communications of the ACM*, 63(5), 2020.

[3] Shang Shang, Tiance Wang, Paul Cuff, and Sanjeev Kulkarni. The application of differential privacy for rank aggregation: Privacy and accuracy. In *17th International Conference on Information Fusion*, 2014.

[4] David T Lee. Efficient, private, and e-strategy proof elicitation of tournament voting rules. *Proceedings of the Twenty-Fourth International Joint Conference on Artificial Intelligence*, 2015.

[5] Shaowei Wang, Jiachun Du, Wei Yang, Xinrong Diao, Zichun Liu, Yiwen Nie, Liusheng Huang, and Hongli Xu. Aggregating votes with local differential privacy: Usefulness, soundness vs. indistinguishability. *arXiv preprint arXiv:1908.04920*, 2019.

[6] Ziqi Yan, Gang Li, and Jiqiang Liu. Private rank aggregation under local differential privacy. *International Journal of Intelligent Systems*, 35 (10):1492–1519, 2020.

[7] Farhad Mohsin, Ao Liu, Pin-Yu Chen, Francesca Rossi, and Lirong Xia. Learning to design fair and private voting rules. In *AI for Social Good workshop at IJCAI-2021*, 2021.

[8] Colin L. Mallows. Non-null ranking model. *Biometrika*, 44(1/2):114–130, 1957.