

Feature-based Individual Fairness in k-clustering

Extended Abstract

Debajyoti Kar*
IIT Kharagpur, India
debajyoti.apeejay@gmail.com

Sourav Medya
University of Illinois Chicago, USA
medya@uic.edu

Mert Kosan*
UC Santa Barbara, USA
mertkosan@ucsb.edu

Arlei Silva
Rice University, USA
arlei@rice.edu

Debmalya Mandal
MPI-SWS, Germany
dmandal@mpi-sws.org

Palash Dey
IIT Kharagpur, India
palash.dey@cse.iitkgp.ac.in

Swagato Sanyal
IIT Kharagpur, India
sanyalswagato@gmail.com

ABSTRACT

Ensuring fairness in machine learning algorithms is a challenging and essential task. We consider the problem of clustering a set of points while satisfying fairness constraints. While there have been several attempts to capture group fairness in the k -clustering problem, fairness at an individual level is not so well-studied. We introduce a new notion of individual fairness in k -clustering based on features not necessarily used for clustering. The problem is NP-hard and does not admit a constant factor approximation. Therefore, we design a randomized heuristic algorithm. Our experimental results against six competing baselines validate that our algorithm produces individually fairer clusters than the fairest baseline.

KEYWORDS

Individual Fairness, Randomized Algorithms, Clustering

ACM Reference Format:

Debajyoti Kar, Mert Kosan, Debmalya Mandal, Sourav Medya, Arlei Silva, Palash Dey, and Swagato Sanyal. 2023. Feature-based Individual Fairness in k -clustering: Extended Abstract. In *Proc. of the 22nd International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2023), London, United Kingdom, May 29 – June 2, 2023*, IFAAMAS, 3 pages.

1 INTRODUCTION

Machine learning systems are increasingly being used in various societal decision-making, including predicting recidivism [2, 8], deciding interest rates [10], and even allocating healthcare resources [17]. However, beginning with the report on bias in recidivism risk prediction [2], it has been known that such systems are often biased against certain groups of people. This paper focuses on fairness in unsupervised learning, particularly in clustering. There has been an increasing interest in designing clustering algorithms that are fair with respect to different subgroups [1, 3, 4, 7].

Compared to group fairness, individually fair clustering has received significantly less attention. Individually fair clustering is motivated by the *facility location problem* where the goal is to open

*Both authors contributed equally to this research.

k facilities while minimizing the total transportation cost between individuals and their nearest facility. If we choose k facilities (or centers) uniformly at random, then each point x could expect one of its nearest n/k neighbors to be one of such facilities. This led a few studies [6, 13, 15] to consider the following notion of individual fairness in clustering. For a point x , let $r(x)$ be the radius such that the ball of radius $r(x)$ centered at x has at least n/k points. An individually fair clustering guarantees that, for every x , a cluster center is chosen from the $r(x)$ -neighborhood of x .

Proposed Definition of Individual Fairness. Motivated by the original definition of individual fairness in supervised learning [9], we introduce a feature-based notion of individual fairness. We say that two individuals are similar if their features match significantly (parameterized by γ in Definition 1). For each individual v , let $C(v)$ denote the cluster v is assigned to. Then our feature-based individually fair clustering requires that $C(v)$ also contains at least m_v individuals that are similar to v . This guarantees that a point v is not isolated in its own cluster but that the cluster has a desired representation (or participation) from points similar to it. Note that, the features that are used to compute similarity for individual fairness might not necessarily be used for clustering. Our notion of individual fairness guarantees that feature-wise similar individuals often share similar clusters. If one converts the cluster centers into representations for downstream tasks, then similar individuals get similar representations and hence similar decisions.

2 PRELIMINARIES

Let V be a set of n points. We denote the tuple of q features of the point i by $X_i = (X_i^\ell)_{\ell \in [q]}$. We write $C = (C_i)_{i \in [n]}$ to denote a clustering of the set V . Given a clustering C and a point v , $\phi(v, C)$ assigns the cluster center to the point v . We are also given a distance function $d : V \times V \rightarrow \mathbb{R}$. The clustering cost is defined as follows: $\text{Cost}(C, \phi) = \sum_{v \in V} d(v, \phi(C, v))$

Our definition of individual fairness builds upon the features of individual points that are not necessarily used for clustering. We convert the discrete features into one-hot encoding vectors and min-max normalize the continuous features. We convert the distance to similarity: $s(X_i, X_j) = e^{-d(X_i, X_j)}$ where s is the similarity between X_i and X_j . We say that X_i and X_j are γ -similar if $s > \gamma$.

DEFINITION 1 (γ -SIMILARITY). For a parameter $\gamma \in [0, 1]$, we say two points $i, j \in V, i \neq j$ are γ -similar if $s > \gamma$ where $s(X_i, X_j) = e^{-d(X_i, X_j)}$. We assume that a point is not γ -matched with itself.

DEFINITION 2 (INDIVIDUAL FAIRNESS IN CLUSTERING). Given a set V of n points with a q -length feature vector X_v for every point $v \in V$, a similarity parameter $\gamma \in [0, 1]$, an integer tuple $(m_v)_{v \in V}$, and an integer k , we say that a clustering $(C_i)_{i \in [\ell]}$ ($\ell \leq k$) is $(m_v)_{v \in V}$ -individually fair if it satisfies the following constraint for all $v \in V$:

$$|\{u : u \in \Gamma(v) \text{ and } \phi(u) = \phi(v)\}| \geq m_v \quad (1)$$

The fairness constraint (1) says that at least m_v points that are γ -similar to point v must belong to the cluster of v . Our main goal is to compute a clustering $(C_i)_{i \in [\ell]}$ of V into ℓ ($\leq k$) clusters and corresponding cluster centers (or facilities¹) $(c_i)_{i \in [\ell]}$ that is individually fair for every point and minimizes the clustering cost. Formally, we define our INDIVIDUALLY FAIR CLUSTERING problem:

DEFINITION 3 (INDIVIDUALLY FAIR CLUSTERING (IFC)). The input is a set V of n points with a q -length feature vector $X_v = (x_v^1, \dots, x_v^q)$ for each $v \in V$, a similarity parameter $\gamma \in [0, 1]$, an integer tuple $(m_v)_{v \in V}$, a set F of potential facilities. The objective is to open a subset $S \subseteq F$ of at most k facilities, and find an assignment $\phi : V \rightarrow S$ to minimize $\text{Cost}(C, \phi)$ satisfying the fairness constraints (Eq. 1).

3 HARDNESS AND METHOD

Hardness Results. For hardness results, we consider the decision version of the IFC problem. It is always possible to find a (trivial) individually fair clustering by one cluster containing all the points. However, the cost of such a fair clustering could be high, and we ask whether it is possible to beat the cost of such trivially fair clustering.

DEFINITION 4 (TRIVIAALLY FAIR CLUSTERING). Given a set V of n points with q -length feature vector X_v for every point $v \in V$, the trivially fair clustering puts all points in one cluster and picks the point as cluster center which minimizes the cost: $\min_{k \in F} \sum_{v \in V} d(v, k)$

We show that it is NP-complete to compute if there exists any clustering better than TRIVIAALLY FAIR CLUSTERING.

DEFINITION 5 (SATISFACTORY-PARTITION). Given a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ and an integer λ_v for every vertex $v \in \mathcal{V}$, compute if there exists a partition $(\mathcal{V}_1, \mathcal{V}_2)$ of \mathcal{V} such that (1) $\mathcal{V}_1, \mathcal{V}_2 \neq \emptyset$ and (2) For every $i \in [2]$ and every $v \in \mathcal{V}_i$, the number of neighbors of v in \mathcal{V}_i is at least λ_v . We denote an arbitrary instance of SATISFACTORY-PARTITION by $(\mathcal{G}, (\lambda_v)_{v \in \mathcal{G}})$.

Theorem 1. It is NP-complete to decide whether an instance of INDIVIDUALLY FAIR CLUSTERING admits a clustering of cost less than the TRIVIAALLY FAIR CLUSTERING even when there are only 2 facilities.

Given this result, we explore the possibility of approximation for the INDIVIDUALLY FAIR CLUSTERING (IFC) problem. However, IFC is inapproximable within factor δ for any $\delta > 0$.

Theorem 2. Distinguishing between instances of the IFC problem having zero and non-zero optimal costs is NP-complete even when there are 2 facilities. Hence, for any computable function δ , there does not exist a δ -approximation algorithm for IFC unless $P=NP$.

¹We use cluster center and facility interchangeably.

Our Algorithm, LP-FAIR: Algorithm 1 describes our randomized approximation algorithm for IFA. The linear program (LP) described in Inequality (2) is a relaxation of the IFA problem. It has a variable x_{v, f_k} for each vertex v and facility f_k . In an (integral) "solution" the variable x_{v, f_k} takes value 1 if and only if the point v is assigned to the facility f_k . After LP, Algorithm 1 determines the assignment ϕ by assigning point v to f_k with probability x_{v, f_k}^* .

Algorithm 1 LP-FAIR, Algorithm for IFA

Input: $(V, (X_v)_{v \in V}, \gamma, (m_v)_{v \in V}, k)$, and δ .

- 1: **for** $t = 1, 2, \dots, T = \log_{1+\delta} n$ **do**
- 2: Solve the following LP to get solution x_t^* .

$$\begin{aligned} \min_x \quad & \sum_{v \in V} \sum_{f_k \in F} d(v, f_k)^p \cdot x_{v, f_k} \\ \text{s.t.} \quad & \sum_{u \in N(v)} x_{u, f_k} \geq m_v \cdot x_{v, f_k} \quad \forall v \in V, f_k \in F \\ & \sum_{f_k \in F} x_{v, f_k} = 1 \quad \forall v \in V \\ & x_{v, f_k} \geq 0 \quad \forall v \in V, f_k \in F \end{aligned} \quad (2)$$

- 3: **for each** $v \in V$ **do**
- 4: Set $\phi_t(v) = f_k$ with probability x_{t, v, f_k}^* .
- 5: **end for**
- 6: **end for**
- 7: **return** Assignment ϕ^* with the minimum cost.

4 EXPERIMENTS

Datasets and Baselines: We apply three commonly used datasets [4, 7, 15] from the UCI repository²: **ADULT** [14], **BANK** [16], and **DIABETES**. We evaluate the following seven algorithms: our LP-based approach (LP-FAIR), FairCenter [13], Alg-PP [5], Alg-AG [5], P-PoF-Alg [5], Hochbaum-Shmoys (H-S) [12], Gonzalez [11]. Our implementation³ is available online.

Performance measures: We use **Normalized Cost** which measures clustering cost normalized by the cost of trivially fair clustering (Def. 4). Our second metric is **Fairness**, which denotes the fraction of points that satisfy individual fairness.

Performance: Table 1 shows normalized cost and fairness results. LP-FAIR has a significantly lower cost than the baselines, with a 34.5% lower cost than the best baseline (H-S) on average. Moreover, LP-FAIR consistently clusters points fairer.

Table 1: Normalized cost and fairness results. The best (bold) and second-best (underlined) performances are emphasized.

| | Normalized Cost | | | Fairness | | |
|------------|-----------------|--------------|--------------|-------------|-------------|-------------|
| | ADULT | BANK | DIABETES | ADULT | BANK | DIABETES |
| FairCenter | 0.544 | 0.528 | 0.341 | 90.0 | 96.0 | <u>94.1</u> |
| Alg-PP | 0.625 | 0.516 | 0.422 | 86.8 | 91.3 | 88.1 |
| Alg-AG | 0.617 | 0.563 | 0.649 | 86.8 | 83.4 | 87.2 |
| P-PoF-Alg | 0.592 | 0.586 | 0.528 | 86.8 | 87.6 | 92.8 |
| H-S | <u>0.267</u> | <u>0.251</u> | 0.107 | <u>90.8</u> | <u>97.7</u> | 91.1 |
| Gonzalez | 0.331 | 0.327 | 0.088 | 88.6 | 90.4 | 89.9 |
| LP-FAIR | 0.194 | 0.176 | 0.057 | 92.3 | 96.3 | 97.9 |

²<https://archive.ics.uci.edu/ml/datasets>

³<https://github.com/mertkosan/lp-fair>

REFERENCES

- [1] Sara Ahmadian, Alessandro Epasto, Ravi Kumar, and Mohammad Mahdian. 2019. Clustering without over-representation. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. 267–275.
- [2] Julia Angwin, Jeff Larson, Surya Mattu, and Lauren Kirchner. 2016. Machine bias. *ProPublica*, May 23, 2016 (2016), 139–159.
- [3] Arturs Backurs, Piotr Indyk, Krzysztof Onak, Baruch Schieber, Ali Vakilian, and Tal Wagner. 2019. Scalable fair clustering. In *International Conference on Machine Learning*. PMLR, 405–413.
- [4] Suman Bera, Deeparnab Chakrabarty, Nicolas Flores, and Maryam Negahbani. 2019. Fair algorithms for clustering. In *Advances in Neural Information Processing Systems*. 4955–4966.
- [5] Darshan Chakrabarti, John P Dickerson, Seyed A Esmaeili, Aravind Srinivasan, and Leonidas Tsepenekas. 2022. A New Notion of Individually Fair Clustering: α -Equitable k -Center. In *International Conference on Artificial Intelligence and Statistics*. PMLR, 6387–6408.
- [6] Deeparnab Chakrabarty and Maryam Negahbani. 2021. Better Algorithms for Individually Fair k -Clustering. *NeurIPS* (2021).
- [7] Flavio Chierichetti, Ravi Kumar, Silvio Lattanzi, and Sergei Vassilvitskii. 2017. Fair clustering through fairlets. In *Proceedings of the 31st International Conference on Neural Information Processing Systems*. 5036–5044.
- [8] Alexandra Chouldechova. 2017. Fair prediction with disparate impact: A study of bias in recidivism prediction instruments. *Big data* 5, 2 (2017), 153–163.
- [9] Cynthia Dwork, Moritz Hardt, Toniann Pitassi, Omer Reingold, and Richard Zemel. 2012. Fairness through awareness. In *Proceedings of the 3rd innovations in theoretical computer science conference*. 214–226.
- [10] Andreas Fuster, Paul Goldsmith-Pinkham, Tarun Ramadorai, and Ansgar Walther. 2020. Predictably unequal? the effects of machine learning on credit markets. *The Effects of Machine Learning on Credit Markets (October 1, 2020)* (2020).
- [11] Teofilo F Gonzalez. 1985. Clustering to minimize the maximum intercluster distance. *Theoretical computer science* 38 (1985), 293–306.
- [12] Dorit S Hochbaum and David B Shmoys. 1985. A best possible heuristic for the k -center problem. *Mathematics of operations research* 10, 2 (1985), 180–184.
- [13] Christopher Jung, Sampath Kannan, and Neil Lutz. 2019. A center in your neighborhood: Fairness in facility location. *arXiv preprint arXiv:1908.09041* (2019).
- [14] Ron Kohavi et al. 1996. Scaling up the accuracy of naive-bayes classifiers: A decision-tree hybrid.. In *Kdd*, Vol. 96. 202–207.
- [15] Sepideh Mahabadi and Ali Vakilian. 2020. Individual fairness for k -clustering. In *International Conference on Machine Learning*. PMLR, 6586–6596.
- [16] Sérgio Moro, Paulo Cortez, and Paulo Rita. 2014. A data-driven approach to predict the success of bank telemarketing. *Decision Support Systems* 62 (2014), 22–31.
- [17] Ziad Obermeyer, Brian Powers, Christine Vogeli, and Sendhil Mullainathan. 2019. Dissecting racial bias in an algorithm used to manage the health of populations. *Science* 366, 6464 (2019), 447–453.