Fair Facility Location for Socially Equitable Representation

Extended Abstract

Helen Sternbach
Hebrew University
Jerusalem, Israel
helen.sternbach@mail.huji.ac.il

Sara Cohen
Hebrew University
Jerusalem, Israel
sara@cs.huji.ac.il

ABSTRACT

We consider the problem of effectively finding a subset of reps (representatives) from a large group of agents belonging to a metric space while considering three distinct notions of fairness. First, each agent should be close to a rep (while precisely how close depends on population densities). Second, reps should satisfy a given social equity constraint specifying the number of representatives with each property value. Finally, reps should be similar, in their property value, to that of the community of agents whom they represent.

KEYWORDS

Facility Location; Fairness; Choosing Representatives

ACM Reference Format:

Helen Sternbach and Sara Cohen. 2023. Fair Facility Location for Socially Equitable Representation: 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

Given a large set of agents (or people, data items, etc.) a common problem is to find a subset of representatives (or *reps*, for short). Numerous techniques have been developed, for different settings of the agents, and a variety of requirements that the subset should satisfy. In this paper, we make two main assumptions about the set of agents. First, agents belong to a metric space, e.g., they may have a geographic location, or may be tuples in a database over which a standard distance metric is defined. In addition, agents have a (possibly sensitive) property, such as ethnicity, gender identity, category, price range. Our goal is to find a set of representatives, in a manner that is fair in terms of the distance of agents to their closest rep, and is socially equitable w.r.t. to the given property.

Example 1.1. Suppose that the state of California wishes to establish a citizens committee to discuss a pressing issue plaguing society (e.g., community violence, environmental hazards). The committee composition could be established via a voting mechanism or by making a call for volunteers. However, to well represent the entire population body, it is reasonable to require the committee to satisfy three distinct notions of fairness.

The first notion is *distance-based fairness*. Citizens in all areas should have a rep that is located close by, i.e., it is reasonable to expect citizens in densely populated areas to have a rep quite close by, while citizens in sparsely populated areas may have a rep somewhat farther off. Second, we may desire the satisfaction of a *social equity constraint*. Reps should not only represent all geographic areas, but also all social circles. Thus, we may require the committee to have a given number of reps of different ethnicities or gender identities. Finally, it is desirable that communities have a rep that is *similar socially* (e.g., ethnically) to their own composition. Thus, it is not sufficient for a social equity constraint to ensure the inclusion of Asian American citizens. Fairness requires that the rep chosen for the Chinatown neighborhood of Los Angeles should be Asian American (as Chinatown is over 70% Asian American). Finding a set of reps that achieves all three notions of fairness is challenging, raising new problems, and is studied in this paper. ◻

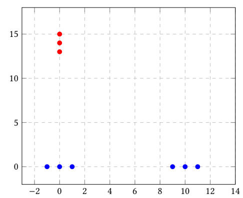
There has been extensive previous work on facility location [10, 19], including recent focus on strategy-proof mechanisms for the facility location problem, e.g., [8, 16, 17, 21, 23]. While traditionally facility location often seeks locations that minimize the maximum or sum of distances of agents to facilities, in this paper we focus specifically on fair facility location [13, 18], which considers the density or sparsity of areas surrounding agents when determining how close facilities must be to agents. Our work extends this line of research to consider social equity constraints.

Our main goal is to choose a set of reps from among a larger set of agents or items. This problem has been studied extensively in a variety of settings, such as multi-winner elections, and the axioms that election rules should satisfy [1, 7, 11], strategic voting [15, 20, 22], and balancing the load size of districts [5, 9]. Work on choosing diverse or representative committees, given candidates with sensitive properties [2, 4, 6] considers rankings of candidates and does not utilize a more general distance function. Previous work also does not attempt to ensure that reps chosen are similar in their property value to the communities they represent. Also related is previous work on fair clustering, e.g., [3, 12, 14]. This problem differs significantly from that on hand as our requirements are on the reps and not the set of agents they represent.

2 CONTRIBUTIONS

In the *fair facility location problem* [13] we are given a set P of n points in a metric space (X, d). For a point $x \in P$, and a value k, the *neighborhood radius* of x, denoted $NR_k(x)$, is the minimum radius r such that at least n/k points in P are within distance r of x. Let $S \subseteq P$ be a set of size k. We use $d(x, S)$ to denote the distance of x from the closest point in S. The *distance ratio* for S, P and k is defined as $\alpha_k(P, S) = \max_{x \in P} \frac{d(x,S)}{NR_k(x)}$. The *optimal distance ratio* $\alpha^*_k(P)$ is simply the value for which $\alpha_k(P, S)$ is minimal, when considering all $S \subseteq P$ of size k. A set S is an *optimal rep set* if $\alpha_k(P, S) = \alpha^*_k(P)$. Intuitively, an optimal rep set is a set in which the the worst ratio of distance from a point x to S over the neighborhood radius of x is minimized.

Proc. of the 22nd International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2023), A. Ricci, W. Yeoh, N. Agmon, B. An (eds.), May 29 – June 2, 2023, London, United Kingdom.



Consider the set of points P_0. For all $x \in P$, $NR_3(x) \leq 2$. Set S_0 containing the middle point from each cluster would be optimal for $k = 3$, and thus $\alpha_3^*(P) = \alpha_3(P, S) = 1/2$, as each point has a rep that is at distance at most $1/2$ the size of its radius.

Now, a facility location algorithm $\mathcal{A}$ over a metric space (X, d) is said to be *α-fair [13]* if, for all $P \subseteq X$ and for all $1 \leq k \leq |P|$, it holds that $\mathcal{A}(P, k)$ returns a set $S \subseteq P$ of size k such that $\alpha_k(P, S) \leq \alpha$.

We are interested in choosing reps that satisfy the same fairness principle introduced by the fair facility location problem. In addition, we assume that points are associated with a *property* (such as race, gender, socioeconomic level). Our goal is not only to solve the fair facility location problem, but to ensure that the resulting set S preserves a given constraint of social equity, i.e., has sufficiently many points with different property values.

Let C be a set of values, and let $v : P \longrightarrow C$ be a function that associates each point in P with a value. Let $\mathcal{E} : C \longrightarrow \mathbb{N}$ be a function associating each value $c \in C$ with a natural number, called an *equity constraint.* Intuitively, $\mathcal{E}(c)$ indicates the number of reps that should be chosen with value c. We use k to denote the total number of reps that $\mathcal{E}$ requires, i.e., $\sum_{c \in C} \mathcal{E}(c)$. We say that a set $S \subseteq P$ is *$\mathcal{E}$-equitable* (or simply *equitable* for short) if, for all $c \in C$ it holds that $|S_c| = \mathcal{E}(c)$, where S_c is the subset of points in S with value c. For example, S_0 is $\mathcal{E}$-equitable only if $\mathcal{E}$ requires one red and two blue points.

We extend the definition of a neighborhood radius to consider $\mathcal{E}$. The *f-feasible neighborhood radius* of a point $x \in P$, denoted $NR_{f,\mathcal{E}}(x)$, is the minimum radius r s.t. there are at least n/k points in P within distance r of x *and* there are at least f points with different values c (with $\mathcal{E}(c) > 0$) within distance r of x. Observe that $NR_{0,\mathcal{E}}(x) = NR_k(x)$. Also, $NR_{1,\mathcal{E}}(x) = NR_k(x)$ if $\mathcal{E}(c) > 0$ for all c, as any radius will satisfy the second requirement, since it will contain the point x itself.

Given the notion of an f-feasible neighborhood radius, we define the *f-feasible distance ratio* $\alpha_{f,\mathcal{E}}(P, S)$, and the *optimal f-feasible equitable distance ratio* $\alpha^*_{f,\mathcal{E}}(P)$ analogously to $\alpha_k(P, S)$ and $\alpha_k(P)$. Now, a facility location algorithm $\mathcal{A}$ over a metric space (X, d) is *(f, α)-fair* if, for all $P \subseteq X$ and for all equity constraints $\mathcal{E}$, it holds that $\mathcal{A}(P, \mathcal{E})$ returns an $\mathcal{E}$-equitable set $S \subseteq P$ such that $\alpha_{f,\mathcal{E}}(P, S) \leq \alpha$. We show the following results, which depend on the choice of f.

Theorem 2.1. *Let $n = |\{c \mid \mathcal{E}(c) > 0\}|$ and $0 < f < n$:*

(1) *determining whether $\alpha^*_{0,\mathcal{E}}(P) \leq 2$ is NP-complete;*
(2) *$\alpha^*_{0,\mathcal{E}}(P)$ is not bound from above;*
(3) *there exists an $(n, 3)$-fair polynomial algorithm;*
(4) *there exist P and $\mathcal{E}$ for which $\alpha^*_{n,\mathcal{E}}(P) \geq 3$;*
(5) *determining whether $\alpha^*_{f,\mathcal{E}}(P) \leq 3$ is NP-complete.*

Note that Item 4 implies that the algorithm of Item 3 is the best that can be defined. Although Item 5 states hardness for $0 < f < n$, a polynomial algorithm can be devised that returns a set of reps which is, in practice, significantly closer to the points that the algorithm of Item 3. Details are omitted due space considerations.

Until now, we have assumed that we have a single set P, out of which we must choose a set of $\mathcal{E}$-equitable reps. We now consider a different setting, in which we have multiple disjoint sets $P^1, \ldots, P^m$ (which can be thought of as communities), and we must choose one rep from each set such that *(1)* the set of reps chosen satisfy an equity constraint *and (2)* the rep s_i for P^i *well represents* P^i, i.e., has a popular property value from P^i. Note that a solution to this problem can be used as a secondary step after finding an $\mathcal{E}$-equitable set of reps S for P. Such reps can be seen as dividing P into *communities* P^i, where a point is in P^i if it is closest to rep s_i. We are then interested in determining how similar a rep is to its community, and to indeed replace reps with other community points in a manner that increases representativeness while preserving equity.

To demonstrate this issue, consider sets of points P^1, P^2 where: P^1 has $n/2$ points, all but one with property value c_r, and one point x_b has value c_b *and* P^2 has $n/2$ points, all but one with property value c_b, and one point x_r has value c_r. Consider the equity constraint $\mathcal{E}(c_r) = \mathcal{E}(c_b) = 1$. Now the set of reps $S = \{x_b, x_r\}$ is $\mathcal{E}$-equitable. However, every other $\mathcal{E}$-equitable set has reps that are much more representative of the points in their communities.

We now formalize the notion of representativeness for a set of reps. Let $x \in P^i$ be a point. Let k^i_x be the number of points in P^i with property value $v(x)$ and let k^i_* be the number of points in P^i with the most common property value, i.e., $k^i_x = |P^i_{v(x)}|$, $k^i_* = \max\{|P^i_c| \mid c \in C\}$. The *representativeness index* of x, denoted $RI(x)$, is simply k^i_x / k^i_*.

Now, given $P^1, \ldots, P^m$ and and equity constraint $\mathcal{E}$ s.t. we have $\sum_{c \in C} \mathcal{E}(c) = m$, we say that a set of reps S is an *$\mathcal{E}$-equitable set of community reps* if S is $\mathcal{E}$-equitable and contains precisely one point from each P^i. We say that S is *max-min optimally representative* (resp. *max-sum optimally representative*) if it is an $\mathcal{E}$-equitable set of community reps that maximizes $\min_{s_i \in S} RI(s_i)$ (resp. maximizes $\sum_{s_i \in S} RI(s_i)$).

Theorem 2.2. *A max-min (max-sum) optimally representative $\mathcal{E}$-equitable set of community reps can be found in polynomial time.*

As discussed earlier, a solution to the problem at hand can be used as a second step after first solving the fair and equitable facility location problem. In such a case, it is possible to find an optimal set S which is guaranteed to be $(n, 4)$-fair. Thus, the cost of a more representative set is a possibly somewhat larger ratio of neighborhood size to distance from a rep.

3 CONCLUSION

We presented a new method to choose reps in a fair and socially equitable manner by leveraging and extending previous work on fair facility location. Our results are applicable when agents are in a metric space and have special properties. Our theoretical results carefully analyze the cost of equity. Due to space considerations, we have omitted experimentation, however, the cost is reasonable in practice. As future work, we intend to consider agents with multiple properties, as well as a broader class of equity constraints. We will also study strategy-proof mechanisms, when agents may lie about their location and/or their property value. Finally, we will search for ways in which load balance can be ensured, while still deriving a fair and equitable result.

4 ACKNOWLEDGEMENTS

The authors were partially supported by the ISF (Israel Science Foundation, Grant 359/21).

REFERENCES

[1] Haris Aziz, Markus Brill, Vincent Conitzer, Edith Elkind, Rupert Freeman, and Toby Walsh. 2017. Justified representation in approval-based committee voting. *Social Choice and Welfare* 48, 2 (2017), 461–485.

[2] Rachel Behar and Sara Cohen. 2022. Representative Query Results by Voting. In *SIGMOD '22: International Conference on Management of Data, June 12 - 17, 2022*, Zachary Ives, Angela Bonifati, and Amr El Abbadi (Eds.). ACM, Philadelphia, PA, USA, 1741–1754. https://doi.org/10.1145/3514221.3517858

[3] Suman Kalyan Bera, Deeparnab Chakrabarty, Nicolas Flores, and Maryam Negahbani. 2019. Fair Algorithms for Clustering. In *Advances in Neural Information Processing Systems 32: Annual Conference on Neural Information Processing Systems 2019, NeurIPS 2019, December 8-14, 2019*, Hanna M. Wallach, Hugo Larochelle, Alina Beygelzimer, Florence d'Alché-Buc, Emily B. Fox, and Roman Garnett (Eds.). Vancouver, BC, Canada, 4955–4966. https://proceedings.neurips.cc/paper/2019/hash/fc192b0c0d270dbf41870a63a8c76c2f-Abstract.html

[4] Robert Bredereck, Piotr Faliszewski, Ayumi Igarashi, Martin Lackner, and Piotr Skowron. 2018. Multiwinner Elections With Diversity Constraints. In *Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence, (AAAI-18), the 30th innovative Applications of Artificial Intelligence (IAAI-18), and the 8th AAAI Symposium on Educational Advances in Artificial Intelligence (EAAI-18), New Orleans, Louisiana, USA, February 2-7, 2018*, Sheila A. McIlraith and Kilian Q. Weinberger (Eds.). AAAI Press, 933–940. https://www.aaai.org/ocs/index.php/AAAI/AAAI18/paper/view/16769

[5] Markus Brill, Piotr Faliszewski, Frank Sommer, and Nimrod Talmon. 2019. Approximation algorithms for balancedCC multiwinner rules. In *Proceedings of the 18th International Conference on Autonomous Agents and MultiAgent Systems.* 494–502.

[6] L Elisa Celis, Lingxiao Huang, and Nisheeth K Vishnoi. 2017. Multiwinner voting with fairness constraints. *arXiv preprint arXiv:1710.10057* (2017).

[7] Edith Elkind, Piotr Faliszewski, Piotr Skowron, and Arkadii Slinko. 2017. Properties of multiwinner voting rules. *Social Choice and Welfare* 48, 3 (2017), 599–632.

[8] Edith Elkind, Minming Li, and Houyu Zhou. 2022. Facility Location With Approval Preferences: Strategyproofness and Fairness. In *Proceedings of the 21st International Conference on Autonomous Agents and Multiagent Systems.* 391–399.

[9] Piotr Faliszewski and Nimrod Talmon. 2018. Between proportionality and diversity: Balancing district sizes under the Chamberlin-Courant rule. In *Proceedings of the 17th International Conference on Autonomous Agents and MultiAgent Systems.* 14–22.

[10] Reza Zanjirani Farahani, Maryam SteadieSeifi, and Nasrin Asgari. 2010. Multiple criteria facility location problems: A survey. *Applied mathematical modelling* 34, 7 (2010), 1689–1709.

[11] Rupert Freeman, Markus Brill, and Vincent Conitzer. 2014. On the Axiomatic Characterization of Runoff Voting Rules. In *Proceedings of the Twenty-Eighth AAAI Conference on Artificial Intelligence, July 27 -31, 2014, Québec City, Québec, Canada*, Carla E. Brodley and Peter Stone (Eds.). AAAI Press, 675–681. http://www.aaai.org/ocs/index.php/AAAI/AAAI14/paper/view/8561

[12] Lingxiao Huang, Shaofeng H.-C. Jiang, and Nisheeth K. Vishnoi. 2019. Coresets for Clustering with Fairness Constraints. In *Advances in Neural Information Processing Systems 32: Annual Conference on Neural Information Processing Systems 2019, NeurIPS 2019, December 8-14, 2019, Vancouver, BC, Canada*, Hanna M. Wallach, Hugo Larochelle, Alina Beygelzimer, Florence d'Alché-Buc, Emily B. Fox, and Roman Garnett (Eds.). 7587–7598. https://proceedings.neurips.cc/paper/2019/hash/810dfbbebb17302018ae903e9cb7a483-Abstract.html

[13] Christopher Jung, Sampath Kannan, and Neil Lutz. 2020. Service in Your Neighborhood: Fairness in Center Location. In *1st Symposium on Foundations of Responsible Computing (FORC 2020) (Leibniz International Proceedings in Informatics (LIPIcs), Vol. 156)*, Aaron Roth (Ed.). Schloss Dagstuhl–Leibniz-Zentrum für Informatik, Dagstuhl, Germany, 5:1–5:15. https://doi.org/10.4230/LIPIcs.FORC.2020.5

[14] Matthäus Kleindessner, Pranjal Awasthi, and Jamie Morgenstern. 2019. Fair k-center clustering for data summarization. In *International Conference on Machine Learning*. PMLR, 3448–3457.

[15] Martin Lackner and Piotr Skowron. 2018. Approval-Based Multi-Winner Rules and Strategic Voting.. In *IJCAI.* 340–346.

[16] Alexander Lam. 2021. Balancing Fairness, Efficiency and Strategy-Proofness in Voting and Facility Location Problems. In *AAMAS '21: 20th International Conference on Autonomous Agents and Multiagent Systems, Virtual Event, United Kingdom, May 3-7, 2021*, Frank Dignum, Alessio Lomuscio, Ulle Endriss, and Ann Nowé (Eds.). ACM, 1818–1819. https://doi.org/10.5555/3463952.3464250

[17] Minming Li, Chenhao Wang, and Mengqi Zhang. 2021. Budgeted facility location games with strategic facilities. In *Proceedings of the Twenty-Ninth International Conference on International Joint Conferences on Artificial Intelligence.* 400–406.

[18] Sepideh Mahabadi and Ali Vakilian. 2020. Individual fairness for k-clustering. In *International Conference on Machine Learning*. PMLR, 6586–6596.

[19] Susan Hesse Owen and Mark S Daskin. 1998. Strategic facility location: A review. *European journal of operational research* 111, 3 (1998), 423–447.

[20] Jaelle Scheuerman, Jason L. Harman, Nicholas Mattei, and Kristen Brent Venable. 2019. Heuristics in Multi-Winner Approval Voting. *CoRR* abs/1905.12104 (2019). arXiv:1905.12104 http://arxiv.org/abs/1905.12104

[21] Xin Sui and Craig Boutilier. 2015. Approximately Strategy-proof Mechanisms for (Constrained) Facility Location. In *Proceedings of the 2015 International Conference on Autonomous Agents and Multiagent Systems, AAMAS 2015, Istanbul, Turkey, May 4-8, 2015*, Gerhard Weiss, Pinar Yolum, Rafael H. Bordini, and Edith Elkind (Eds.). ACM, 605–613. http://dl.acm.org/citation.cfm?id=2773233

[22] Yongjie Yang. 2020. On the complexity of destructive bribery in approval-based multi-winner voting. *arXiv preprint arXiv:2002.00836* (2020).

[23] Houyu Zhou, Minming Li, and Hau Chan. 2022. Strategyproof Mechanisms for Group-Fair Facility Location Problems. In *Proceedings of the Thirty-First International Joint Conference on Artificial Intelligence, IJCAI 2022, Vienna, Austria, 23-29 July 2022*, Luc De Raedt (Ed.). ijcai.org, 613–619. https://doi.org/10.24963/ijcai.2022/87