

Improved Approximation Algorithms for Individually Fair Clustering

Ali Vakilian & Mustafa Yalçiner

March 28, 2022

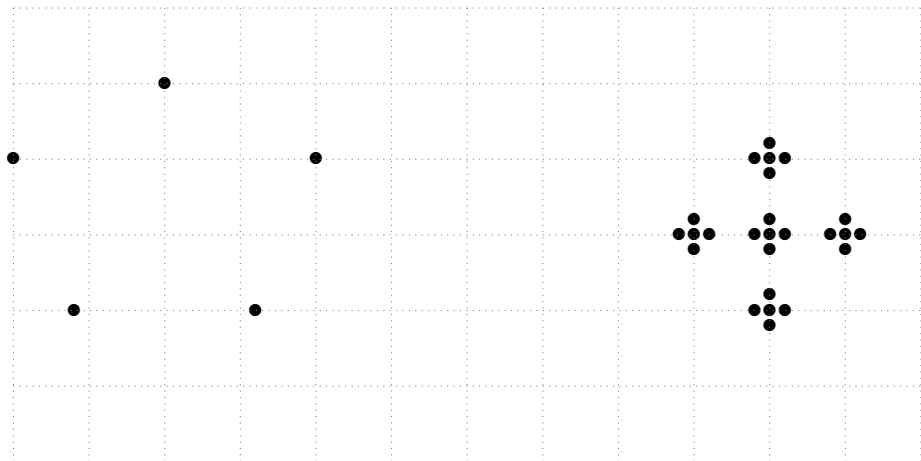


Ali Vakilian (TTIC)



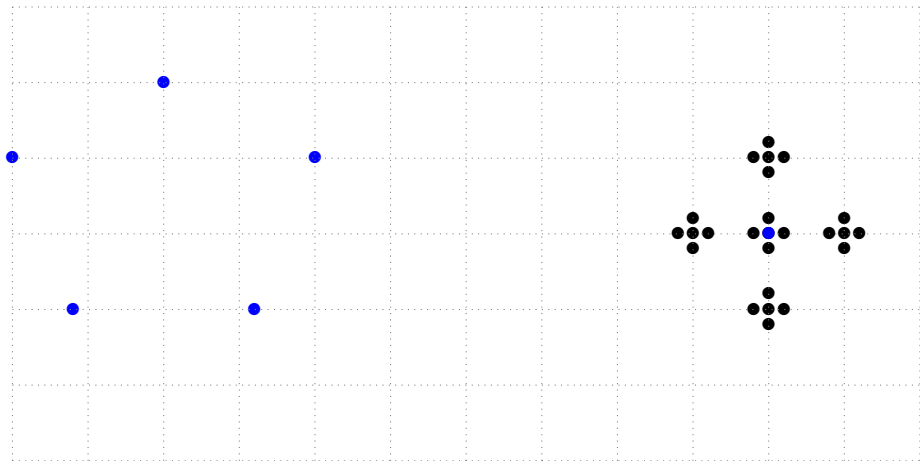
Mustafa Yalçiner (TU Dortmund)

Example - Input: Population Densities



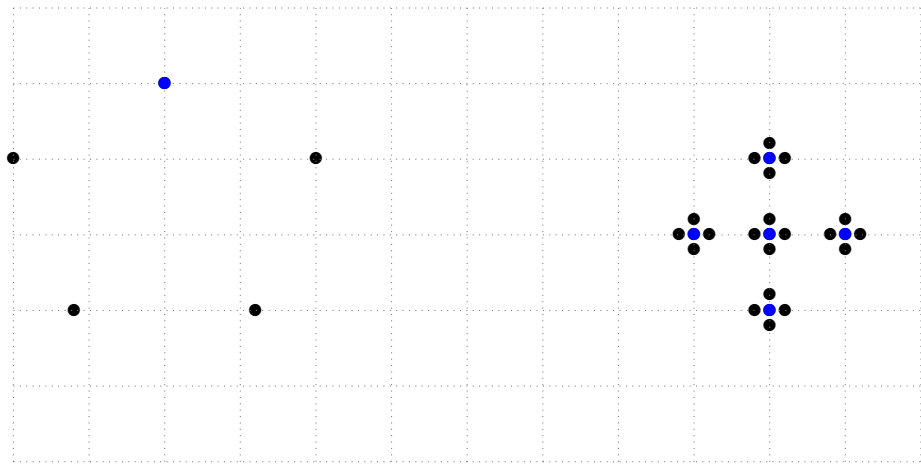
Population densities ($n = 30$)

Example - Output: Hospital Locations



(unfair) minimum cost clustering ($k = 6$ hospitals)

A Center in your Neighborhood [Jung et al.,2019]



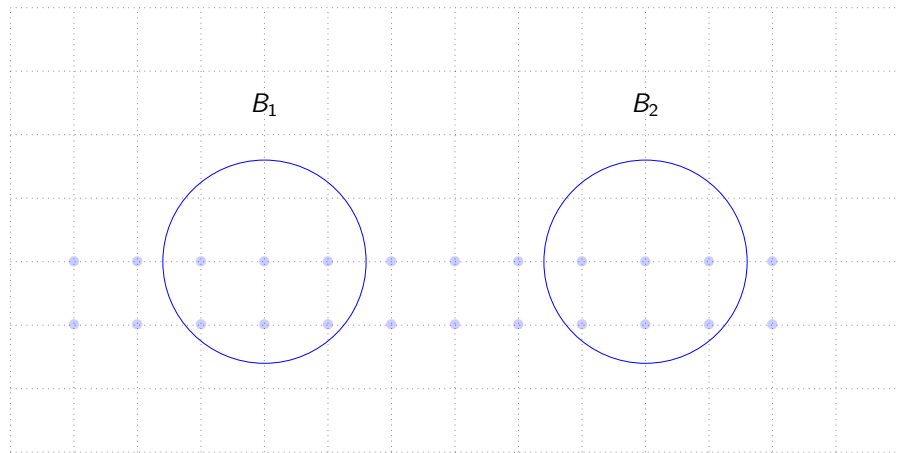
Every resident has a center among its $n/k = 5$ nearest neighbors

1. Urban residents *expect* shorter distances
2. Load balancing

Problem Definition

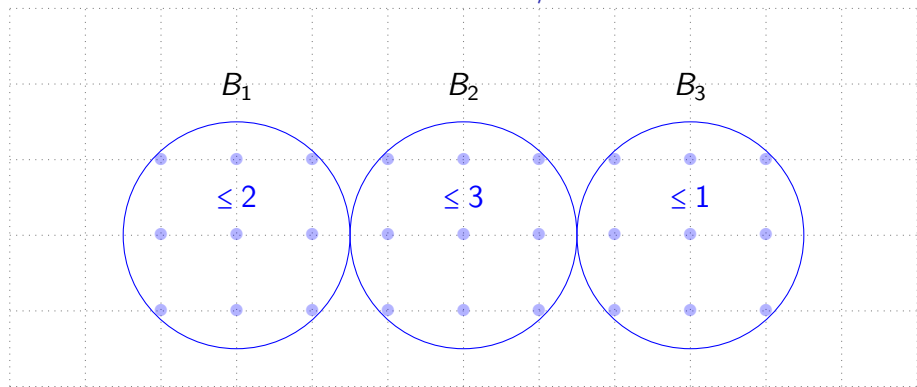
- **Input:** Points P of size n , number of centers k
- **Output:** Set of k centers $C \subseteq P$ s.t.:
 - 1 Every $x \in P$ has a center among its n/k nearest neighbors.
 - 2 C minimizes the ℓ_p -clustering cost: $\sum_{x \in P} \min_{c \in C} d(x, c)^p$.

Critical Regions - Achieve Fairness



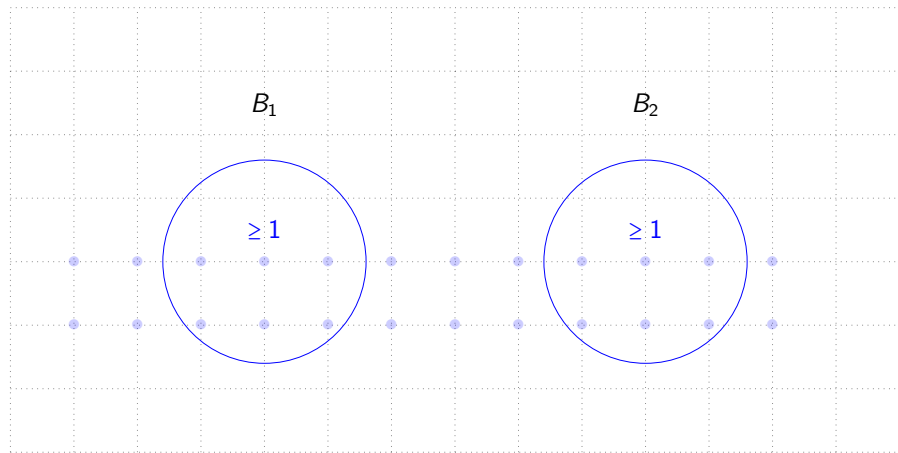
Solution is fair \iff Solution contains point(s) in each critical region

Reduction to Matroid Median/Center



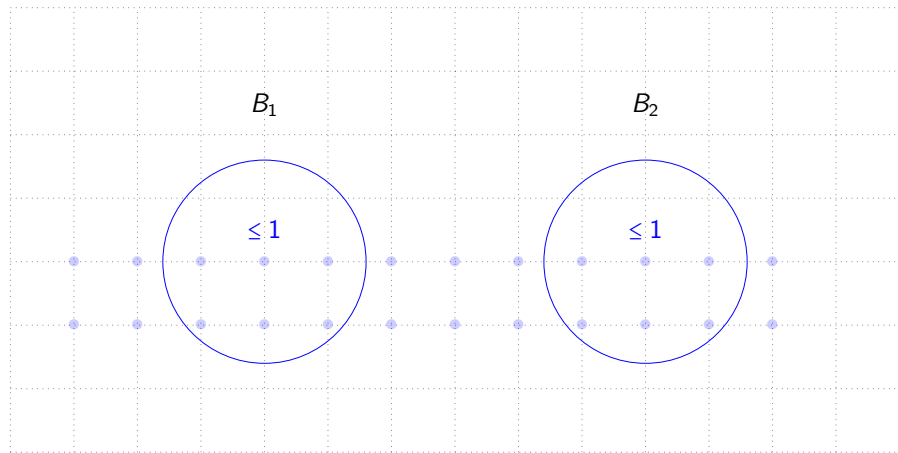
- Blue points P , partition matroid \mathcal{M}
- Find capacity-respecting (i.e. *independent*) set $C \subseteq P$ minimizing
 - $\sum_{x \in P} d(x, C)$ Median
 - $\max_{x \in P} d(x, C)$ Center

The Reduction



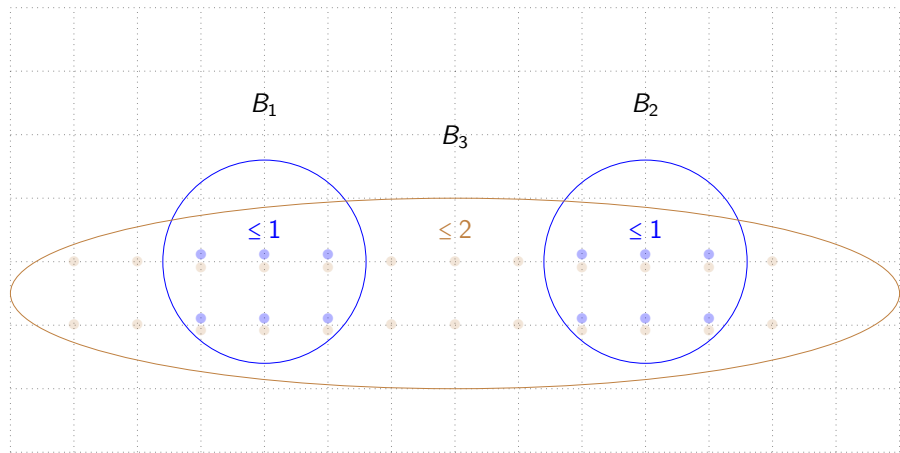
Lower bound requirement for critical regions ($k = 4$)

The Reduction



Upper bound constraint in the partition matroid ($k = 4$)

The Reduction








Upper bound constraint in the partition matroid ($k = 4$)

Results


	<i>k</i> -median		<i>k</i> -means		<i>k</i> -center	
	Cost	Fairness	Cost	Fairness	Cost	Fairness
[Mahabadi,Vakilian.,2020]	84	7	$O(1)$	7	$O(\log(n))$	7
[Chakrabarty,Negahbani.,2021]	8	8	4	8	$2 + \epsilon$	8
Ours	$7.081 + \epsilon$	3	$16 + \epsilon$	3	$3 + \epsilon$	3


- Generalize LP-Rounding framework for facility location to ℓ_p -cost function. [Charikar et al.,2002], [Swamy,2014]


References


-  Jung, Christopher and Kannan, Sampath and Lutz, Neil
A center in your neighborhood: Fairness in facility location
2019 *FORC*
-  Mahabadi, Sepideh and Vakilian, Ali
(Individual) Fairness for k -Clustering
2020 *ICML*
-  Plesník, Ján
A heuristic for the p -center problems in graphs
1987 *Discrete Applied Mathematics*
-  Deeparnab Chakrabarty and Maryam Negahbani
Better Algorithms for Individually Fair k -Clustering
2021 *NeurIPS*
-  Vakilian, Ali and Yalçiner, Mustafa
Improved Approximation Algorithms for Individually Fair Clustering
2021 *arXiv preprint arXiv:2106.14043*

References

 Jones, Matthew and Nguyen, Huy and Nguyen, Thy
Fair k-centers via maximum matching
2020 *ICML*

 Krishnaswamy, Ravishankar, Shi Li, and Sai Sandeep
Constant approximation for k-median and k-means with outliers via iterative rounding
2018 *STOC*

 Charikar, Moses and Guha, Sudipto and Tardos, Éva and Shmoys, David B
A constant-factor approximation algorithm for the k-median problem
2002 *Journal of Computer and System Sciences*

 Chaitanya Swamy
Improved Approximation Algorithms for Matroid and Knapsack Median Problems and Applications
2014 *APPROX*

References



Kay, Matthew, Cynthia Matuszek, and Sean A. Munson

Unequal representation and gender stereotypes in image search results for occupations

2015 *CHI*



Kleindessner, Matthäus, Pranjal Awasthi, and Jamie Morgenstern

Fair k-center clustering for data summarization

2019 *PMLR*