
BENJAMIN COOKSON, University of Toronto
Unifying Proportional Fairness in Centroid and Non-Centroid Clustering

Proportional fairness criteria inspired by democratic ideals of proportional representation have received growing attention in the clustering literature. Prior work has investigated them in two separate paradigms. Chen et al. (2019) study centroid clustering, in which each data point's loss is determined by its distance to a representative point (centroid) chosen in its cluster. Caragiannis et al. (2024) study non-centroid clustering, in which each data point's loss is determined by its maximum distance to any other data point in its cluster.

We generalize both paradigms to introduce semi-centroid clustering, in which each data point's loss is a combination of its centroid and non-centroid losses, and study two proportional fairness criteria, the core and, its relaxation, fully justified representation (FJR). Our main result is a novel algorithm which achieves a constant approximation to the core, in polynomial time, even when the distance metrics used for centroid and non-centroid loss measurements are different. We also derive improved results for more restricted loss functions and the weaker FJR criterion, and establish lower bounds in each case.

Based on joint work with Nisarg Shah and Ziqi Yu