Fair Resource Allocation

October 2019

We are developing an algorithm whose objective is to model the distribution of a resource across a community in order to maximize equitable access to the resource. Cities are often segregated (for example, by race and income) and therefore distributing public resources and infrastructure in a way that enables equal access for people within various subgroups becomes an important consideration; however, to date, most resource allocation algorithms have optimized for maximum usage and not most equitable usage of a resource. The proposed approach builds on a modified k-means clustering algorithm and takes into consideration the distribution of the underlying demographics in a community for planning where to place a resource.

1 Definitions & Preliminaries

Symbols	Description
$i \in I$, (x_i, y_i)	spatially distributed individuals, locations (x, y)
$c \in C$, $\{x_c, y_c\}$	spatially distributed centroids, locations (x, y)
M, $|X_m|$	discrete attributes and values
$s \in S$	subgroups based on binary attributes
L, $l_{i,c}$, l_s	loss, individual loss, subgroup loss
$d(i, c)$	Euclidean distance between individual, centroid
$\text{div}(l_s, l_{s'})$	divergence in loss between subgroups

- **Individual** The set of individuals, I, represent the relevant sample of a community: the subset of the population that is likely to use a resource (for example: potential cyclists, those exposed to covid-19). Each individual $i \in I$ has a location: $i = \{x_i, y_i\}$ where x_i is the x-coordinate and y_i is the y-coordinate of their residential location.

- **Centroid** The set of k centroids, $k = |C|$, denote the point locations of bicycle hubs. The location of a centroid $c \in C$ is defined as $\{x_c, y_c\}$ where x_c represents the x-coordinate, and y_c represents the y-coordinate, of the location.

- **Attribute** Each individual $i \in I$ is associated with M discrete attributes where each has $|X_m|$ values.

- **Subgroup** A subgroup s is defined by the M-dimension Cartesian set product of attribute values [**?**]:

$$|S| = \prod_{m=1}^{M} |X_m| \tag{1}$$

- **Loss** $l \in [0,1]$ denotes the generic cost of distance to a centroid.

- **Individual Loss** The individuals' loss is a function of the distance between their location and the closest centroid: $l_{i,c} = f(distance)$ where $l_{i,c} \in [0,1]$. To begin, this function may resemble a sigmoid function, for example:

$$l_{i,c} = \frac{1}{1 + e^{d(i,c)}} \tag{2}$$

where $d(i,c)$ is the Euclidean distance between the position of the individual (i) and the centroid (c):

$$d(i,c) = \sqrt{(x_i - x_c)^2 + (y_i - y_c)^2} \tag{3}$$

Eventually, we may be able to infer this relationship from data and/or use graph distance in place of Euclidean distance.

- **Subgroup Loss** l_s is the loss experienced by members of a subgroup s with probability $p \in [0,1]$ of the cumulative distribution function.

- **Subgroup Divergence** The divergence between subgroups is defined as the maximum difference in loss between two subgroups' cumulative distribution function at a given probability of experiencing some level of loss, p [**?**]:

$$\mathrm{div}(l_s, l_{s'}) = \max|l_{s,p} - l_{s',p}| \tag{4}$$

where $l_{s,p}$ is the loss experienced by the subgroup with the maximum loss at p, $l_{s,p} = \arg\max_s l_p$, and $l_{s',p}$ is the loss experienced by the subgroup with the minimum loss at p, $l_{s',p} = \arg\min_s l_p$.

- **Equitable Resource Allocation** Given:

  1. A set of spatially distributed individuals, I
  2. Some number of centroids, k
  3. Some set of discretized attributes, S
  4. Some upper-bound for the divergence in loss across subgroups, X
  5. Some notion of a loss function (e.g. the logistic map)

Identify the set of centroid locations, C, that minimize individual loss while also minimizing the divergence among subgroups.

$$\text{argmin} \left(\sum_{c \in C} \sum_{i \in I} \frac{l_{i,c}}{|I|} + \frac{\text{div}(l_s, l_{s'})}{|S|} \right) \tag{5}$$

Or, the problem can be defined in another way, like most existing work on constraint-based fair clustering [?, ?], where the fairness is put as a constraint that needs to meet, rather than a part of the objective function.

$$\text{argmin} \sum_{c \in C} \sum_{i \in I} \frac{l_{i,c}}{|I|}, w.r.t. \frac{\text{div}(l_s, l_{s'})}{|S|} < L \tag{6}$$

where L is a given loss threshold.

First draft of a procedure:

1. Initiate k random centroids

2. Assign each point to closest centroid using Euclidean distance

3. Refine location of centroid: "So that just leaves the centroid movement step. For that there are several possibilities, which all basically boil down to trying to move centroids in directions that improve the overall loss function (not just the first term). Perhaps the simplest way (probably not the best) is to propose moving a single center a small distance in a random direction (and reassigning every point to its closest center), but only make that move if it improves the score. Better would be figuring out how to move centers in a direction which is likely to improve the objective, e.g., prefer moves toward the cluster's center of mass and in directions that reduce disparities. Even better would be computing the negative gradient of the loss function and moving along that direction."

4. Repeat Assignment - Refinement steps until centroid converges.

2 Notes

$$\sum_{i \in I} \min_{c \in C} \frac{\text{loss}(i, c)}{I_c} + \sum_{a \in A} \text{Div}(F_{I_a}, F_{I'_a}) \tag{7}$$

$$\text{Div}(F_{I_a}, F_{I'_a}) = max|loss(p)_{F_{I_\alpha}} - loss(p)_{F_{I_\beta}}| \tag{8}$$

- F_{I_α} and F_{I_β} are the CDFs of minloss for individuals $i \in I_\alpha$ and $i \in I_\beta$ respectively.

- p is the proportion of people with loss $\leq$ some level of loss; so $p \in [0, 1]$.

DBN comments: let's use C for the set of cluster centers and I for the set of individuals. Then for term (1), we have $\sum_{i \in I} \text{minloss}(i) = \sum_{i \in I} \min_{c \in C} \text{loss}(i, c)$. I'm not sure what the point of term (2) is, since subgroups will be defined deterministically in terms of observable characteristics. We should discuss. For term (3), imagine that we have a set of binary attributes A, where each $a \in A$ partitions I into two groups I_a and I'_a. Then we could write (3) as $\sum_{a \in A} \text{Div}(F_{I_a}, F_{I'_a})$, where Div is a measure of divergence between two cumulative distribution functions (CDFs), and F_{I_a} and $F_{I'_a}$ are the CDFs of minloss for individuals $i \in I_a$ and $i \in I'_a$ respectively.