Algorithms

Algorithm 1 FINDPAIRS algorithm

Require: a dataset of points $\mathcal{P}$, a number of partitions p and a distance threshold ε.

1: **function** FINDPAIRS ($\mathcal{P}$, p, ε)
2: Partition $\mathcal{P}$ using a Quadtree and p partitions ▷ Using Algorithm 1 in next page.
3: Create a circle of radius ε for each point in $\mathcal{P}$ and store them in $\mathcal{Q}$ ▷ keep same id
4: Partition $\mathcal{Q}$ using the same partitioner of $\mathcal{P}$
5: Build local index in $\mathcal{P}$ ▷ Using operations provided by GeoSpark
6: Execute a distance join query in $\mathcal{P}$ and $\mathcal{Q}$ using ε as distance ▷ Using Algorithm 4 in following pages.
7: Filter those pairs where $p_1.id < p_2.id$
8: **end function**