

Smoking and Activity Space

Akhil Kumar Pabbathi
Yumin Zhang
Mentor: Tianyuan Liu
Faculty Sponsor:
Joe McClearnon
Jason Oliver
Matt Hallyburton

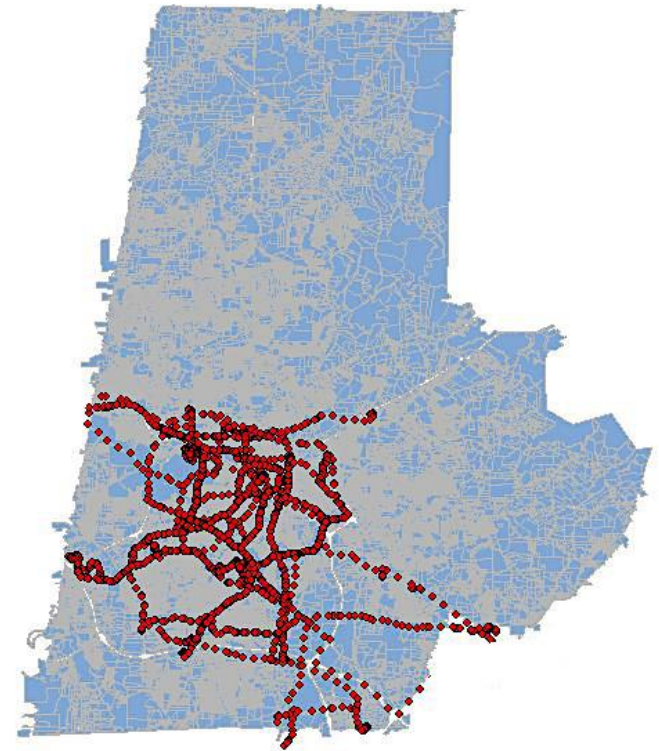


Data:

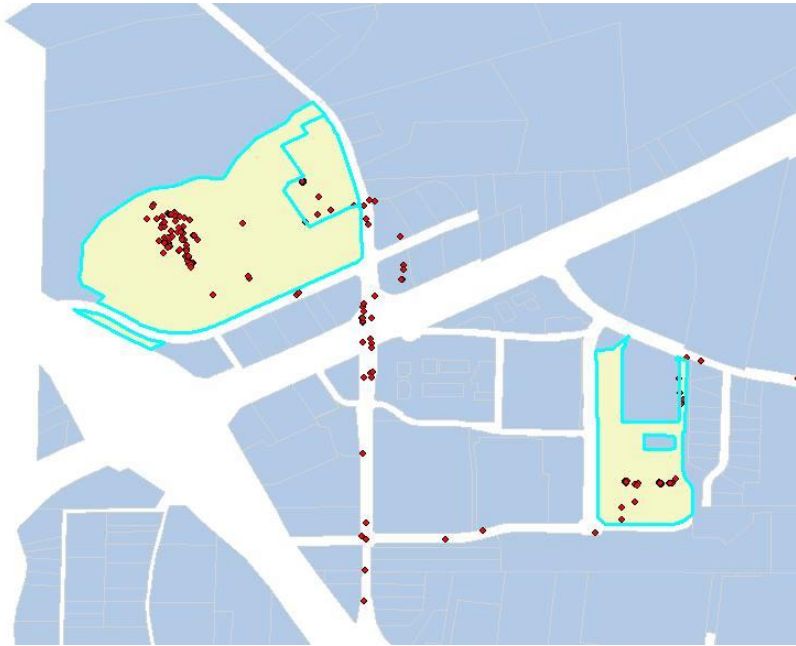
The dataset consists of about two-week activity tracks acquired from smokers and nonsmokers working and living in Durham County, NC.

Objectives:

1. Data cleaning
2. Identify the activity space (clustering)
3. Analysis the activity space in the graph



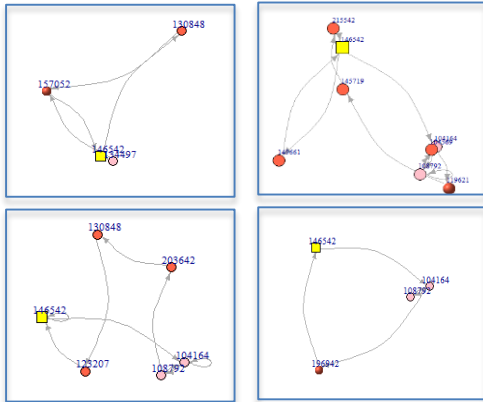
Method and Result of Clustering



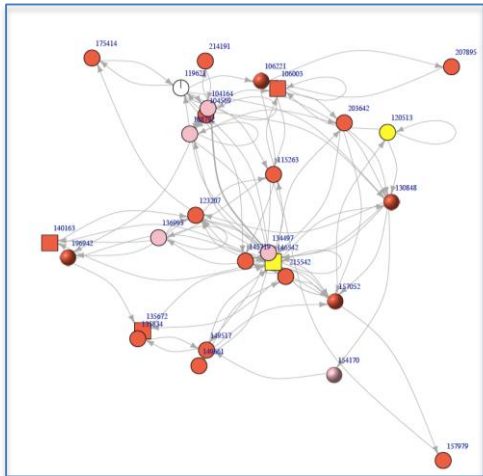
- In order to analyze the activity space, the staying place of the participant's activity tracks is the point of interest. The staying places are identified as clusters in the geo-spatial dataset.
- The GPS coordinates are joined with the parcel, which is a small unit of land with a unified function, to form the basic structure of the clusters. The figure above shows how the parcel spatial joins with the points falling onto it.
- Because some clusters might span over multiple parcels, some parcels are combined based on whether two parcels are within a certain distance and have the same land use type. Each cluster is represented by an unique parcel ID associated with the GPS coordinates.
- The final clusters are verified in a data-driven-produced map to ensure the thresholds for combining the parcels will produce the optimal result.



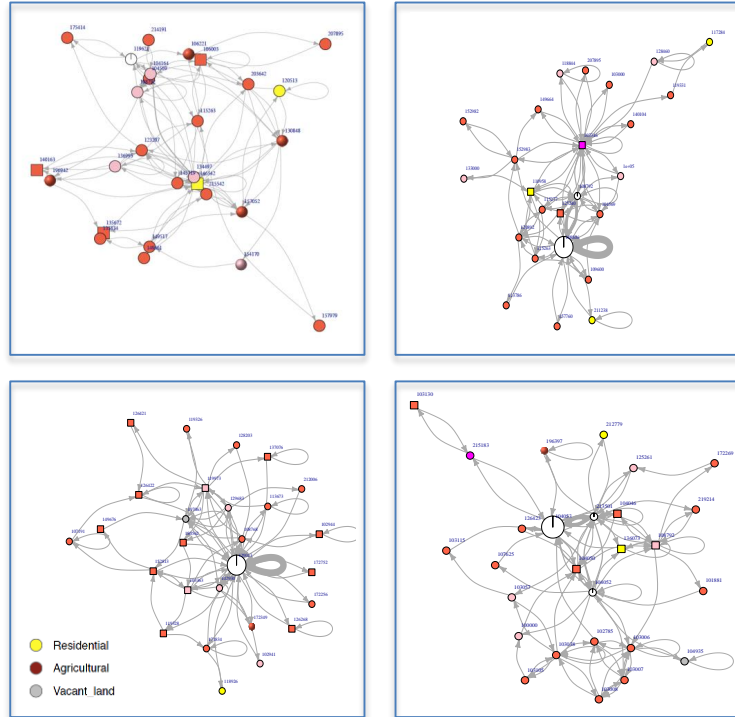
Network Analysis of Activity Space



Comparing daily networks of one participant



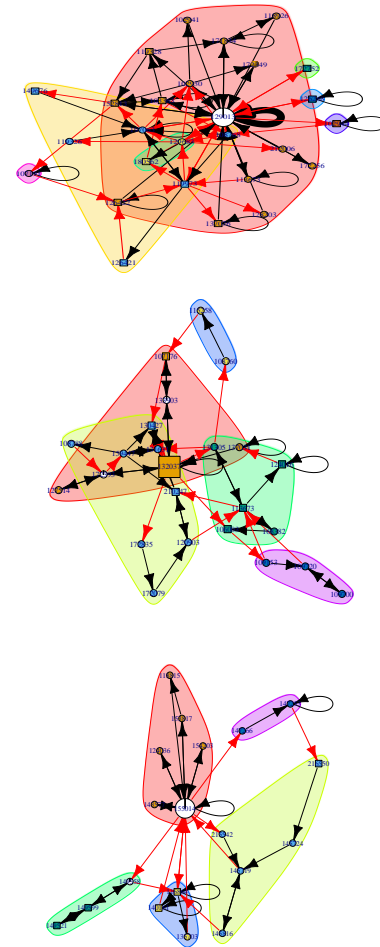
Accumulative network for 16 days



Compare aggregated graphs of multiple participants

	edge_density	reciprocity	dyad_census	transitivity	diameter	centralization
1	0.208	0.522	0.560	0.512	5	0.240
2	0.117	0.389	0.463	0.281	3	0.618
3	0.119	0.452	0.489	0.307	5	0.433
4	0.108	0.516	0.545	0.139	4	0.478
5	0.108	0.485	0.541	0.156	6	0.443
6	0.132	0.648	0.684	0.268	4	0.470
7	0.128	0.561	0.600	0.303	4	0.286
8	0.092	0.375	0.392	0.216	6	0.176

Compare networks on different indices



Analyze activity space as communities