JOURNEY-BASED CHARACTERIZATION OF MULTI-MODAL PUBLIC TRANSPORTATION NETWORKS

Cecilia Viggiano[1] · Haris N. Koutsopoulos[2] · Nigel H.M. Wilson[1] · John Attanucci[1]

July 2015

[1]Massachusetts Institute of Technology · [2]Northeastern University

MOTIVATION

- Improved understanding of passenger behavior in a multi-modal public transport network can inform planners making strategic decisions

- Smart card transactions produce a rich new data set that can be converted into new information

OBJECTIVES

- Understand roles of different modes in a multi-modal public transport network
  - Do bus services feed rail or are they used as a stand-alone service?
  - Do bus routes provide an alternative to rail for certain trips?

PREVIOUS RESEARCH

- Smart card analysis focused on observing typical usage patterns and user classification

  (Utsonomiya et al., 2006; Morency et al., 2007)

- Network structure characterization using metrics derived from graph theory assessed based on the physical network topology

  (Garrison and Marble, 1964; Vuchic and Musso, 1991; Derrible and Kennedy, 2010)

INPUTS

- Smart card and AVL data processed with ODX methodology to infer complete journey itineraries
  (Gordon et al., 2013)

- Stop and station locations

OUTPUTS

OD pairs assigned to seven categories according to mode or combination of modes used

METHODOLOGY OVERVIEW

1. Stops and stations are clustered based on location to assign journeys to zonal OD pairs according to their initial and final stops or stations

2. OD pairs are categorized based on the share of journeys by each mode or combination of modes

STOP AND STATION CLUSTERING

- Most existing zonal schemes (such as postcodes or census tracts) use roads as boundaries, resulting in stops or stations at the boundary of multiple zones

- Instead, stops and stations clustered according to geographic coordinates using the k-means algorithm

K-MEANS

1. K data points selected as initial centroids (can be random, user-specified, or k-means++)

2. Assign all data points to closest centroids

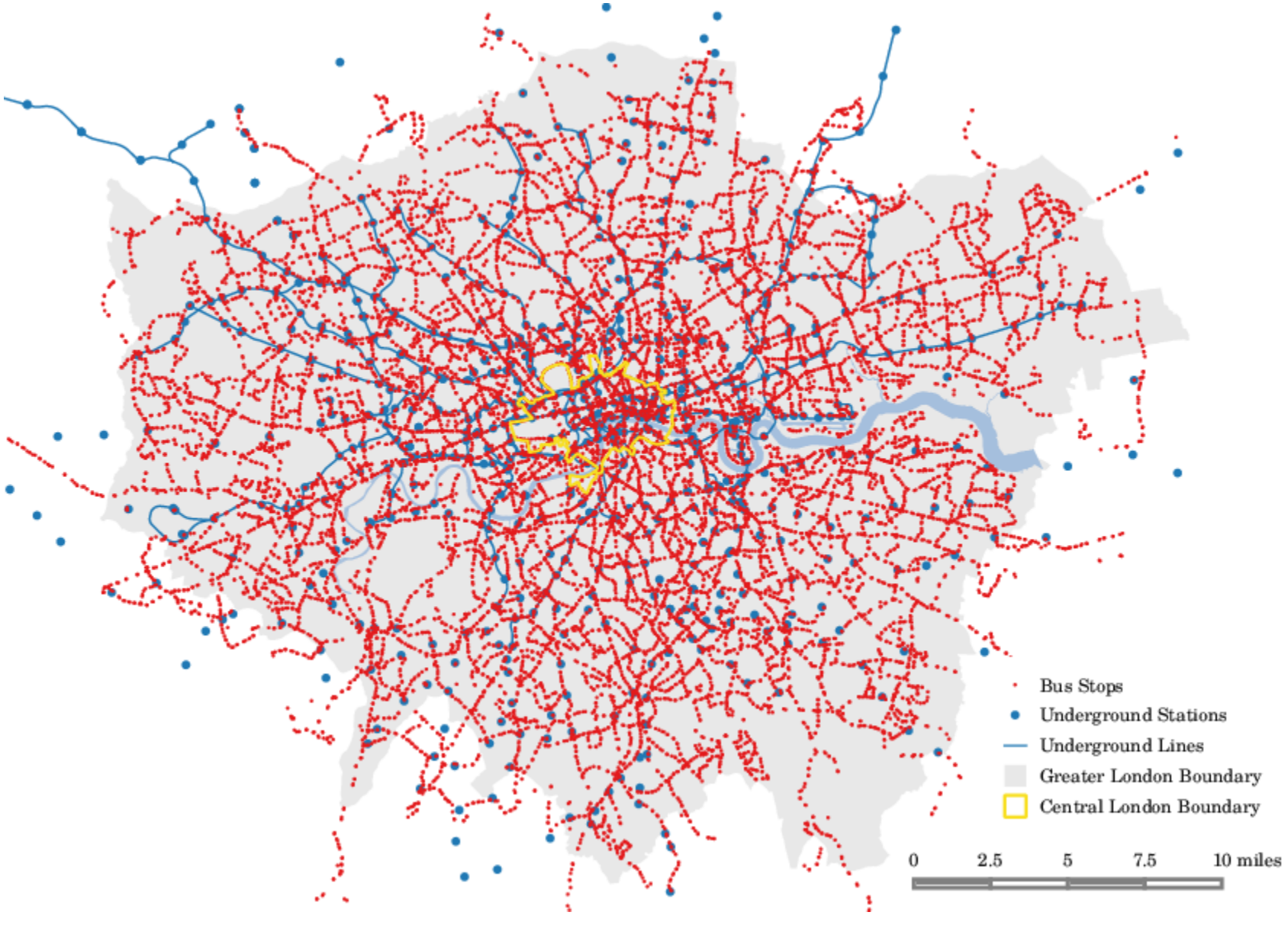
3. Recalculate centroids

   Iterate until centroid locations no longer change significantly

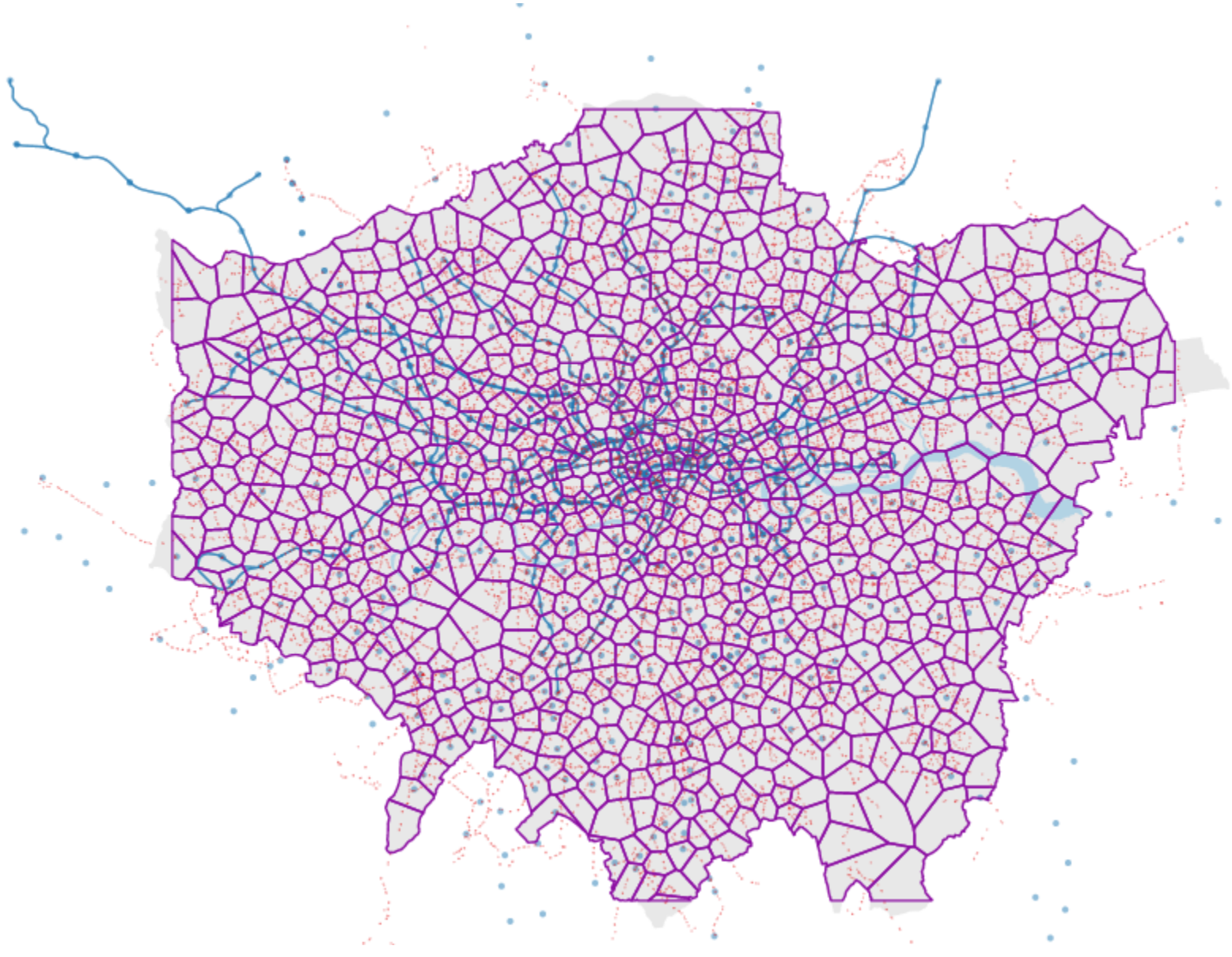
SELECTING THE NUMBER OF CLUSTERS

- Too many clusters can result in insufficient journeys per zonal pair
- Too few clusters can result in walking distances that are unrealistic

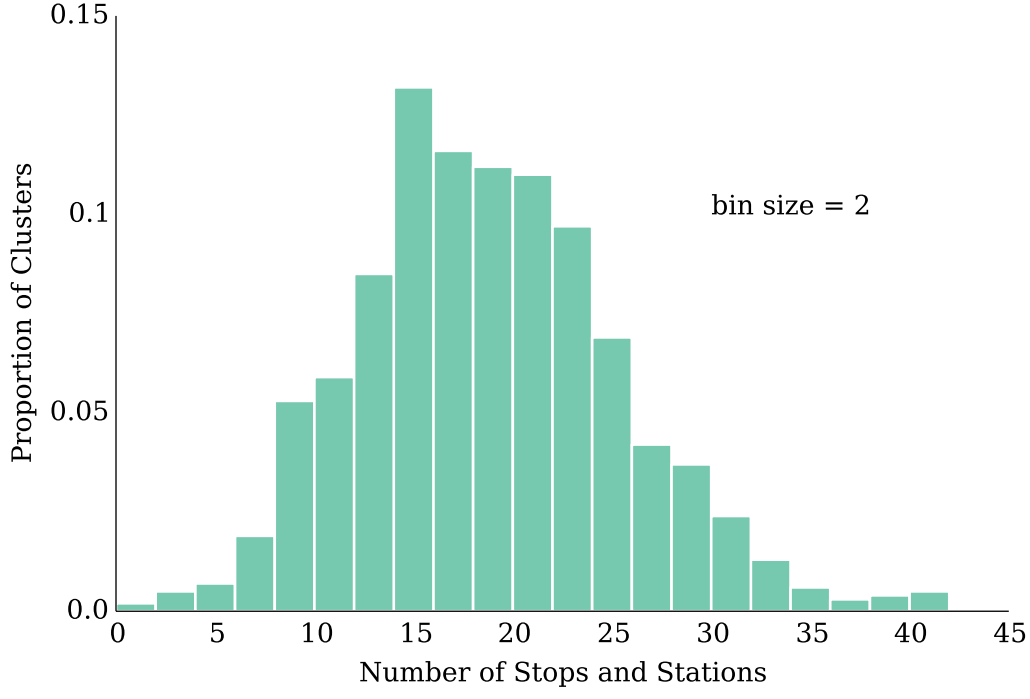
LONDON'S PUBLIC TRANSPORT NETWORK



LONDON'S 1,000 ZONES



STOPS AND STATIONS PER ZONE

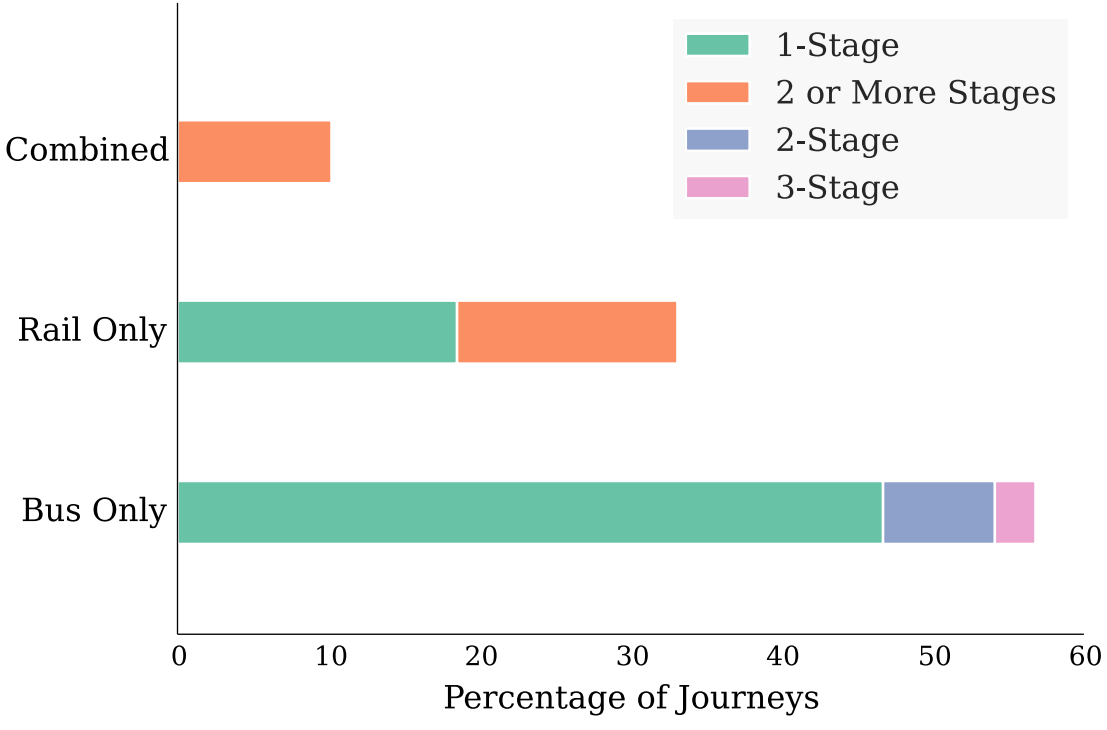


MODAL DEFINITIONS

Journeys categorized as:

- bus
- rail
- combined

JOURNEYS BY MODE

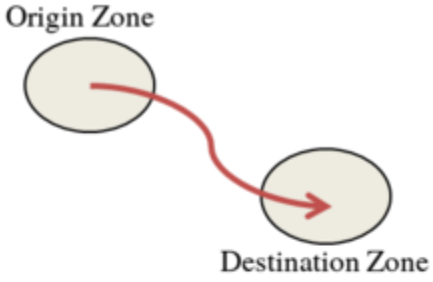


ZONAL PAIRS CATEGORIZED AS:

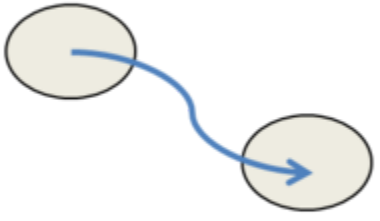
- primarily bus
- primarily rail
- primarily combined
- bus and rail
- bus and combined
- rail and combined
- bus, rail, and combined

CATEGORY DEFINITIONS

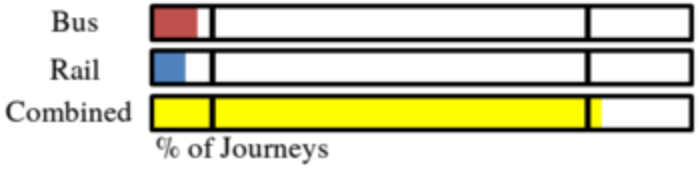
1. Primarily Bus

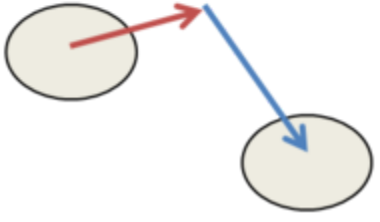


2. Primarily Rail



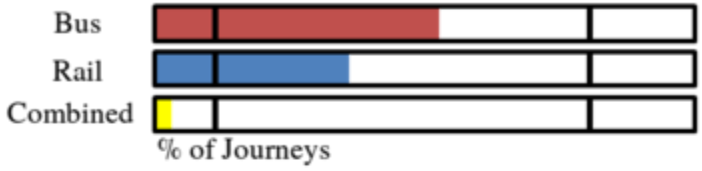
3. Primarily Combined

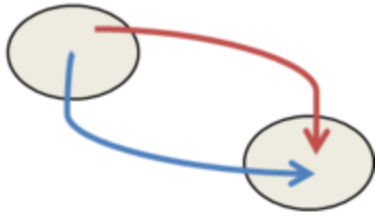




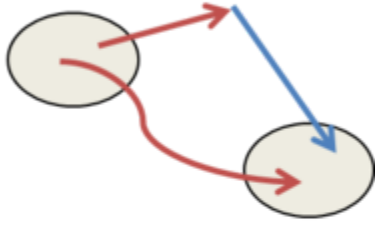
CATEGORY DEFINITIONS

4. Bus and Rail

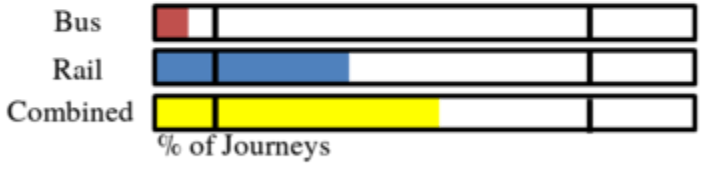


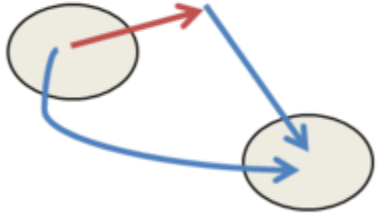


5. Bus and Combined

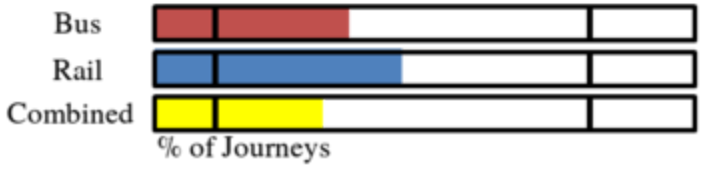


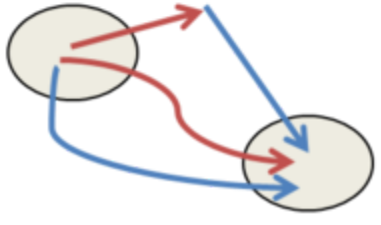
6. Rail and Combined



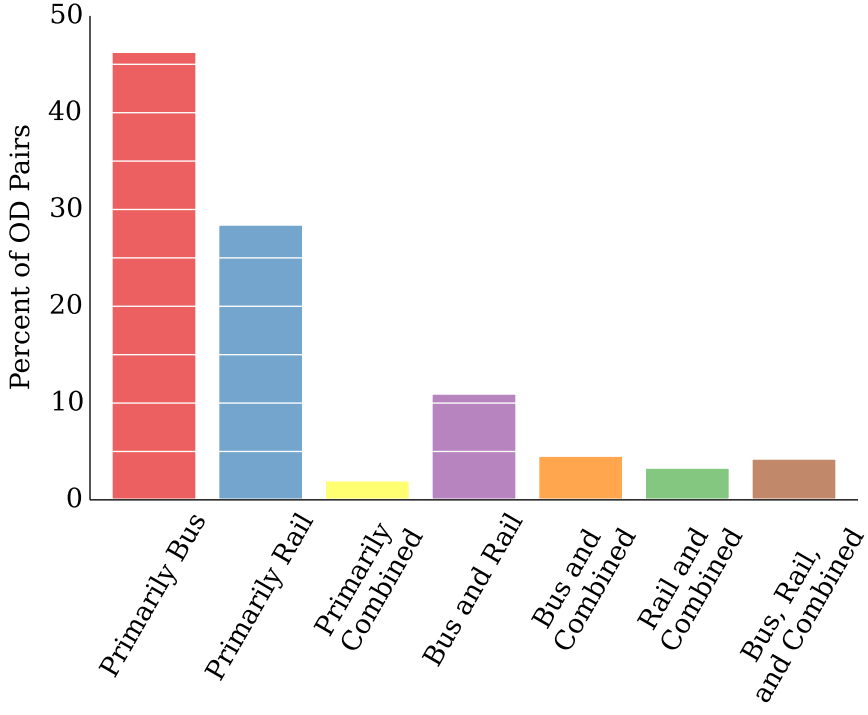


7. Bus, Rail, and Combined

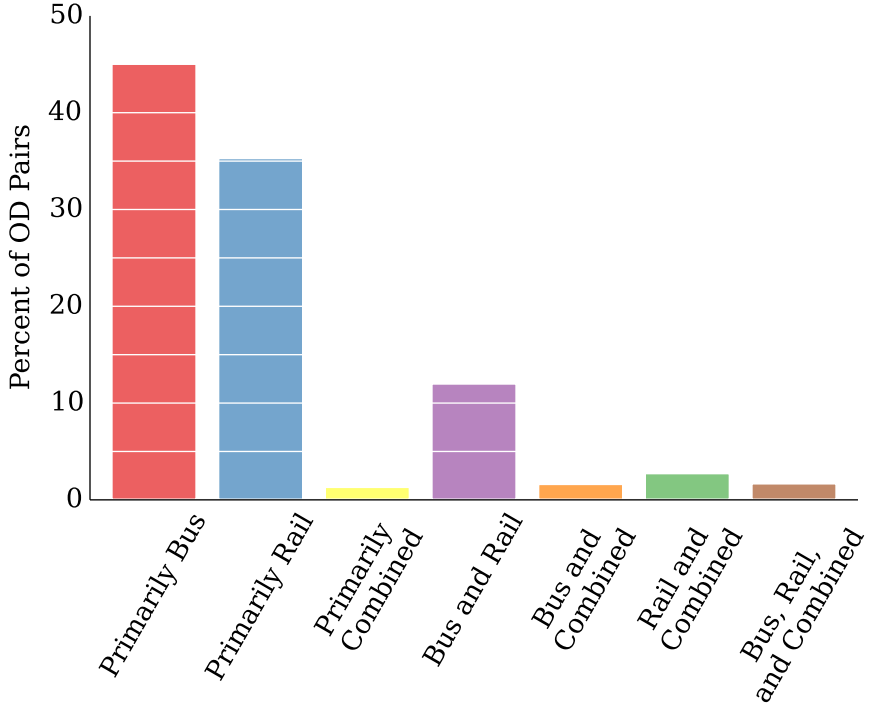




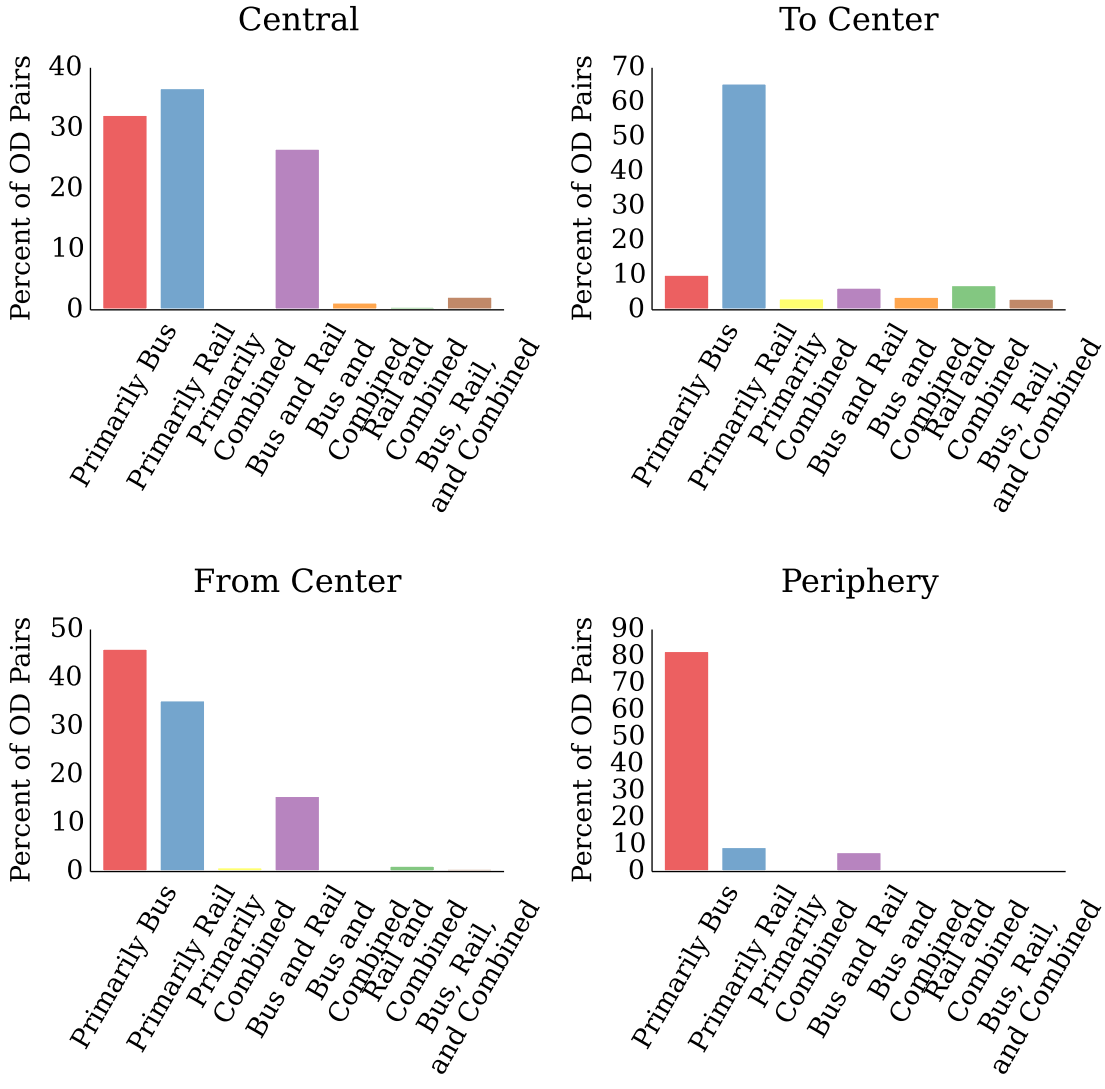
LONDON CATEGORIZATION



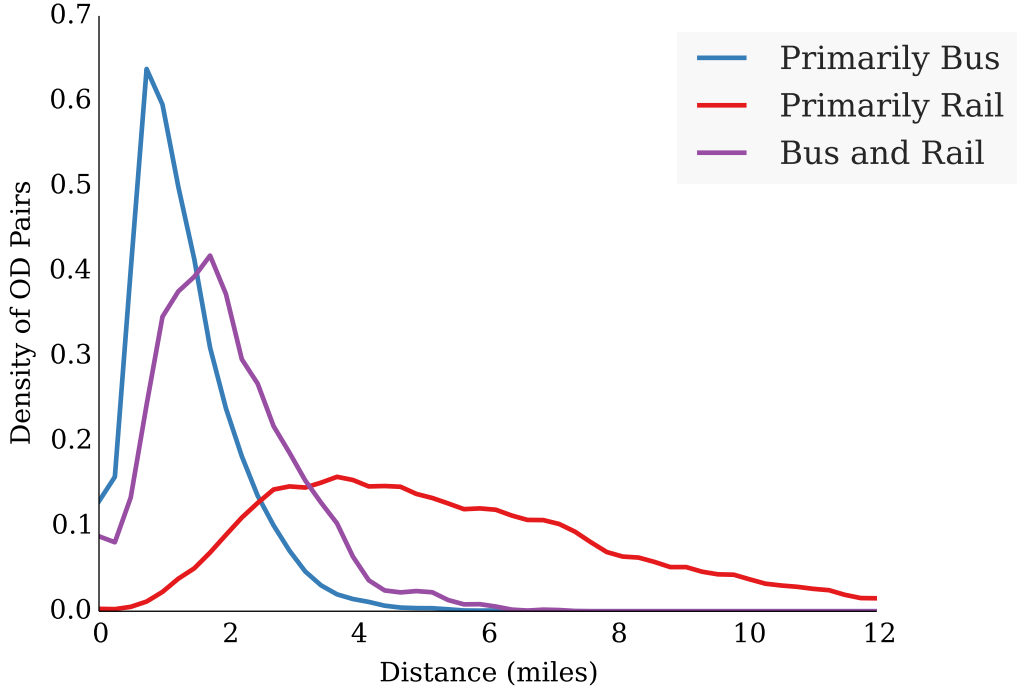
AM PEAK CATEGORIZATION



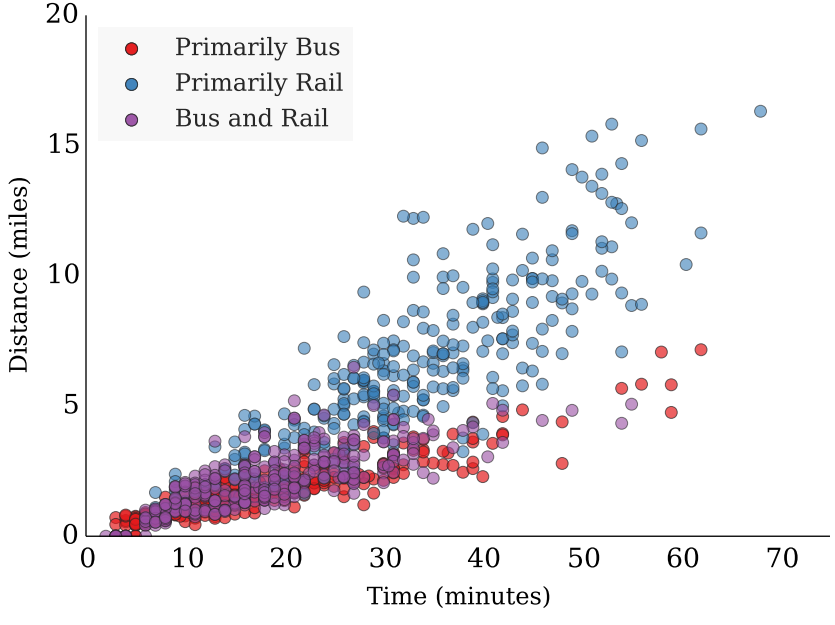
GEOGRAPHIC VARIATION



VARIATION IN JOURNEY LENGTH



VARIATION IN SPEED



CONCLUSIONS

- Categorization allows for journey data to be interpreted in a meaningful way

- Zonal and categorization schemes also provide a foundation for network evaluation methods

QUESTIONS?