Algorithmic Optimizations of Boston's Public Bus Network

Adrian Law
Mark Bestavros
Tyrone Hou

MOTIVATION

- Inefficiency and unreliability are the biggest challenges to Boston's public bus system
- Smart transit control systems rely on big data techniques to dynamically allocate resources under variable conditions
- **Goal:** To take an algorithmic approach to bus route planning by:
  - Relocating bus stops according to urban population density
  - Optimizing bus allocation per route with machine learning

STOP PLACEMENT OPTIMIZATION

Data

- **Region of Interest (ROI)**
  - Urban regions: OpenData + ArcGIS
  - T/Commuter stops: MBTA API
- **Bus Route & Stop locations:**
  - MBTA API
- 202,434 locations
- 264 Routes & 2,447 Stops

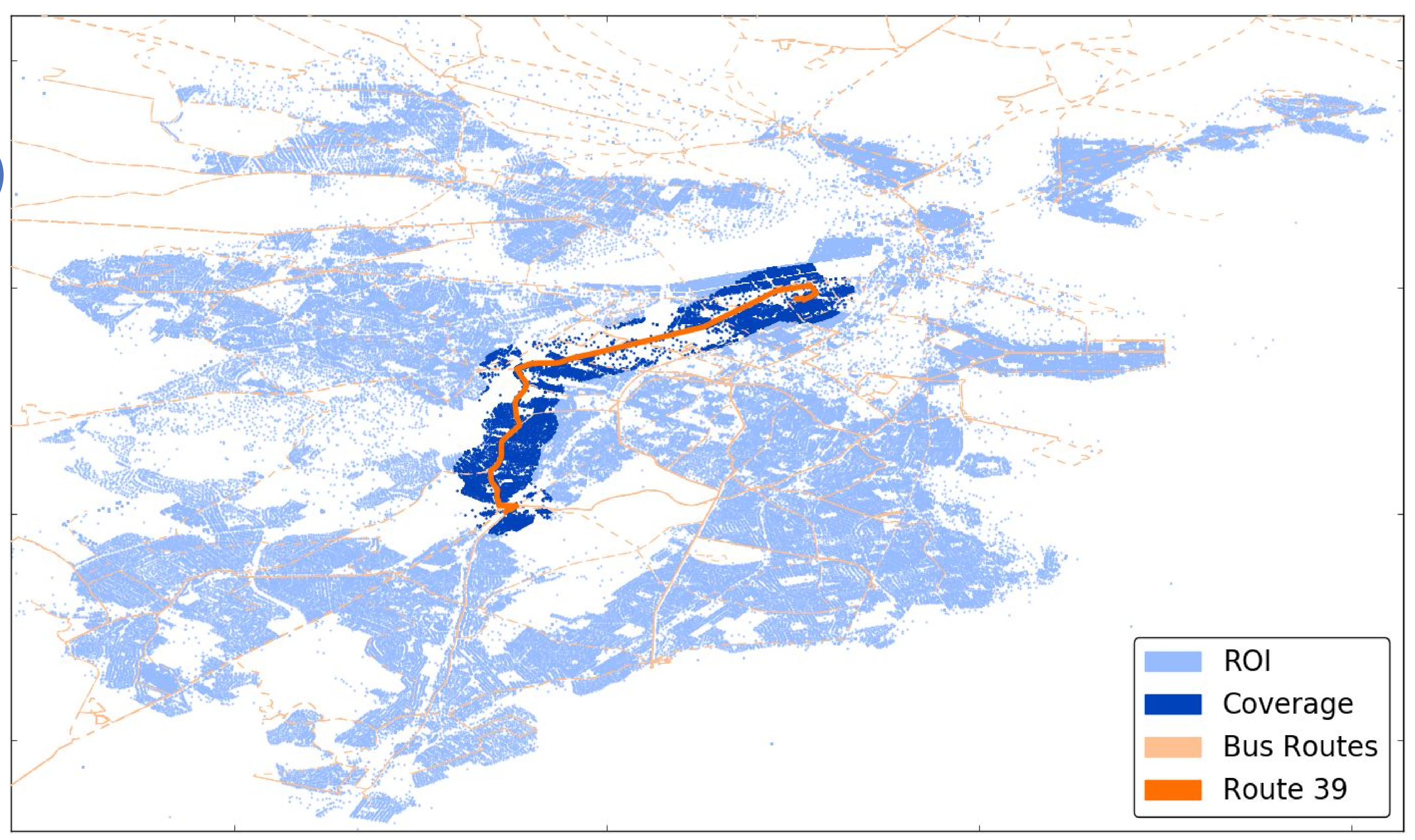

Fig. 1 - Map of Boston Bus Routes and Regions of Interest (ROI)
Bold orange line is Route 39 and bold blue points are regions of interest within 0.5 km

Algorithm

- **Goal**: to minimize distances from ROI to given route through k-means
- ROI close to bus route stored in R tree (r<=0.5km)
- ROI are projected to closest point on bus route and mapped to a one dimensional space (D)
- We then perform a k-means clustering on D and map centers onto route as optimal bus stops

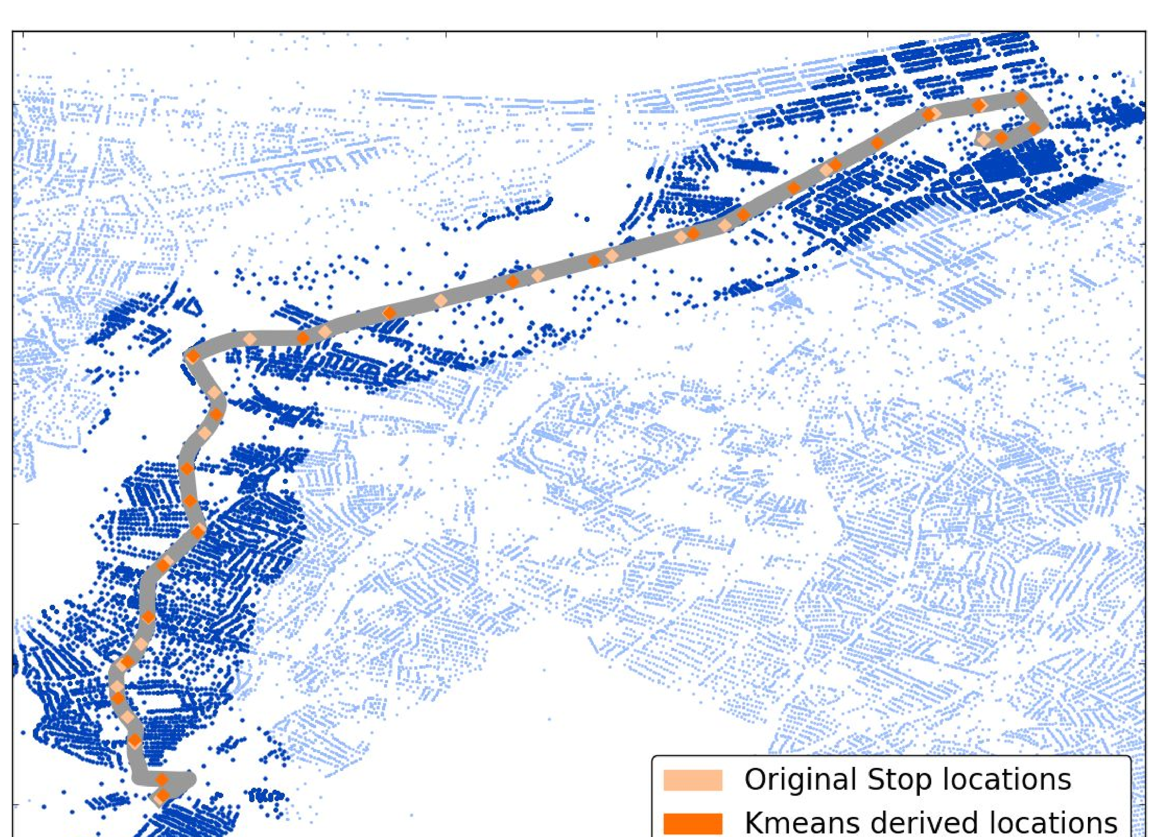

Fig. 2 - Closeup of old and new bus stops along Route 39.
Again, blue points are regions of interest, and bolded points are within 0.5km
Original stops are light orange, new stops are dark orange,

BUS ALLOCATION OPTIMIZATION

Data

- **Bus locations & velocities:**
  - NextBus API: Scraped for real time data every 5 mins
  - Calculated route completion time averages & variance per bus

$$Score = \frac{\mu_{FT}}{|K_{Buses}|} * n_r + \sum_{b \in K_{Buses}} \int_0^{\mu_{FT}} N(0, \sigma_b) - N(\mu_b, \sigma_b)$$

μ_{FT} = Average Route Finish Time
μ_b = Average Route Finish Time per Bus
n_r = Number of Stops
σ_b = Standard Deviation of Finish Time
K_{Buses} = Buses allocated to the Route

Equation 1- Scoring Formula for Allocation Algorithm

Algorithm

- **Goal:** to minimize waiting time at each stop and buses that "overtake" each other
- Scoring Function (Eq. 1) is minimized through varying allocation of buses for all routes & calculating the score for K allocation
- Score represents the balance of latency vs inefficiency of a route

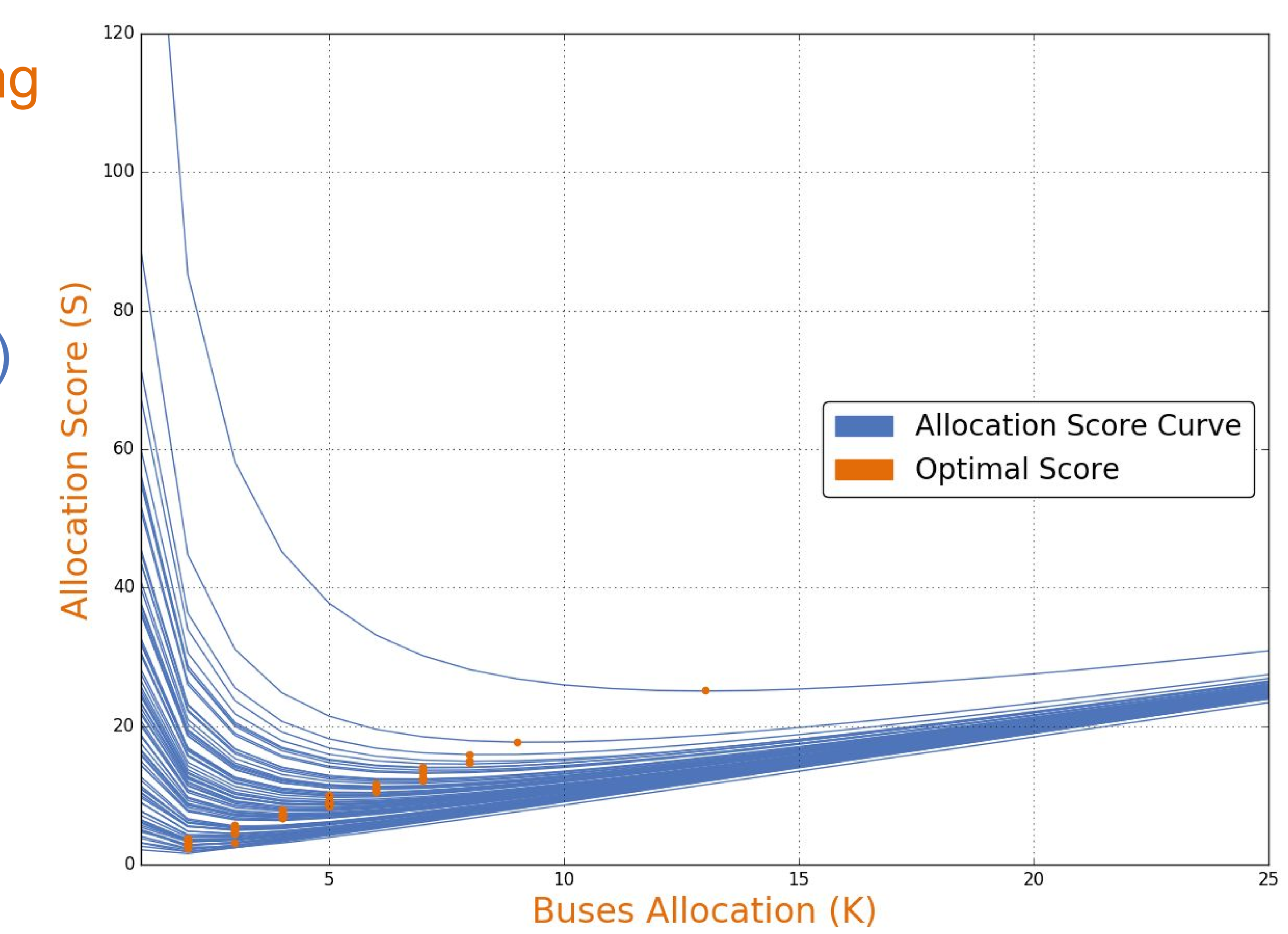

Figure 4 - Graph of the allocation scores for each bus route route & their minimum score as the optimal allocation

FUTURE WORK

- Expand ROI set to include commercial data for better resolution
- Allocation algorithm is a heuristic; optimum inconsistent with actual allocation for most routes
- Allocation is a more cost-effective task than stop relocation and has more potential for application
- Future Goal: Implement a central MBTA bus allocation platform using machine learning techniques