ComputerScience

Samantha Karam | Sarah Medeiros
Boston University

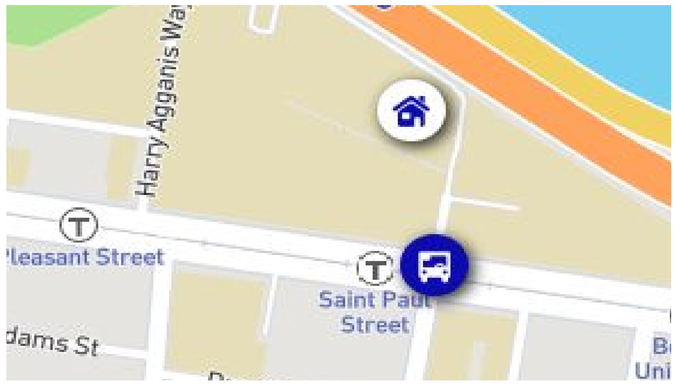
# Optimizing Bus Routes

## Introduction

Transportation costs for Boston Public Schools currently account for 11% of the district's budget. Our goal was to optimize the location of corner bus stops so that buses could pick up the maximum number of "corner stop" students (students that did not need to be picked up at their doorstep) at the least number of stops while also taking into account the distance a child is allowed to walk to reach their stop.

## Data

We chose two datasets issued by the City of Boston: the students-simulated.geoson dataset and the school-real.geojson dataset.The students-simulated.geojson dataset included generated student data including their school, address, max walking distance allowed, ect. The school-real.geojson dataset is a collection of all of the Boston Public Schools. The last dataset we used was the Boston, MA geojson (OSM2PGSQL)





## Data Visualization

We wanted to let a user add new students and have our algoirthms show them where that student's bus stop would be. After getting the student's address and school, we run k-means on the set of students and we map the new stops found from k-means to their closest corners. We used Leaflet to show the address and new stop and we use MapShakers MapKey Icons to change the image and color of the Leaflet markers.

## Methods

**K-Means:** We employed the k-means algorithm to choose preliminary bus stop locations. We grouped students by school and ran k-means on these groups

**R-Trees:** The city of Boston requires that children must be picked up at a corner. We mapped our results from k-means to the closest possible street corner. To do this, we followed the theory behind r-trees. We split up the Boston, MA geojson dataset into searchable subsets for each school so that we could minimize the amount of intersections that we had to check for possible bus stop placement.

## Conclusion

This problem is extremely complex and our solution is far from complete. Not only could we further optmize our run time, but there are many constraints we have not addressed such as school start time, school end time, bus start locations, and bus capacity.