Bus Congestion Project Proposal

Contributor:

Name	**Email address**
Chengjun Wu(Lead)	simonwu@bu.edu
Annan Miao	annan@bu.edu
Jing Li	jingli18@bu.edu

1. Background

With a huge amount of public school students taking school buses and the MBTA in Boston for daily commuting, the City of Boston is faced with the problem improving the traffic system to reduce disruption to student arrival times and keep students out of dangerous situations. Tackling this problem will not only help students arrive to school on time, but also lower down the risk of traffic accidents. By leveraging historical traffic data, our team aims to build a model to fit in this route system and utilize it to evaluate the variance of switching from one route to another.

2. Objective

The overall objective of this project is to optimize BPS bus routes in order to decrease the amount of time that students are on buses and therefore decrease the number of students arriving late to school. We will also explore the following questions to better understand and respond to the overall goal:

- Which streets have the greatest impact on bus schedules neighborhoods have the highest rate of traffic congestion
- Which route has the highest idling and hence emissions?
- Can we identify problem spots for bus routing (intersections, stretches of road)? Do these problem spots extend to traffic in general
- What are the characteristic attributes of the problem areas (neighborhood type, street geometry, nearby hubs) and what is the impact of time-of-day? Are there any creative solutions beyond rerouting?

3. Dataset Used

We will utilize the following datasets in our project in addition to BPS sensor data which comes from a private database and will not be disclosed in this report.

- BPS sensor data:
  Get the current buses status and analyze the traffic patterns.

- Google Map APIs
  (https://cloud.google.com/maps-platform/credits/application/edu)
  a. Get map information to analyze traffic status;
  b. Find alternative routes between two core stations;
  c. (maybe) get the current and historical congestion status of target routes.

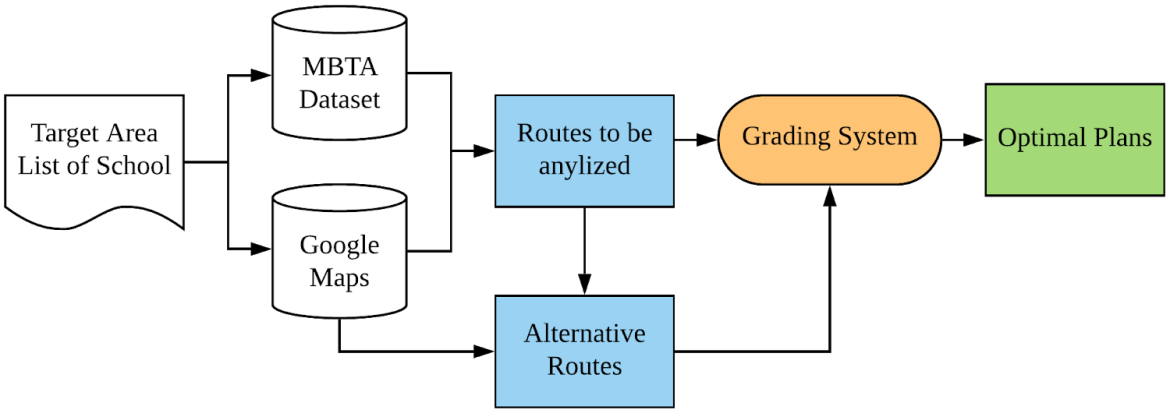
- PATI bus stops
  (https://geo-massdot.opendata.arcgis.com/datasets/pati-bus-stops-1)
  Get the detailed information of core stations that need to be analysed

- Historic MBTA dataset
  (https://mbtabackontrack.com/performance/index.html#/download)
  Get the historic information and status of MBTA routes to find congestion spots and learn the bus traffic patterns.

- Lists of school
  (https://www.bostonpublicschools.org/Page/628)
  Determine the target area and the end points of the routes that need to be analysed and optimized.

4. Solution Concept

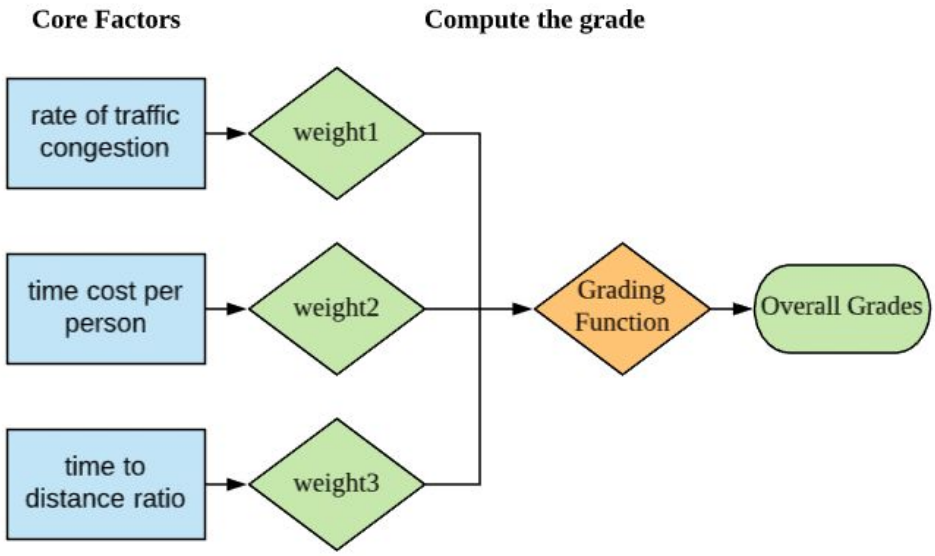
Overall Solution:
Based on researching on traffic systems we will build a grading system to analyze the routes, which will allow us to find alternative choices and provide optimization options.



Overview diagram

Grading System:

We plan to define a set of criteria to help us quantify the degree of what is a "good" route. Such as the rate of traffic congestion on the route, the time cost per person, the time spent to distance ratio and so on. We will analyze the traffic datasets to learn the pattern and compute the weights of the core factors.



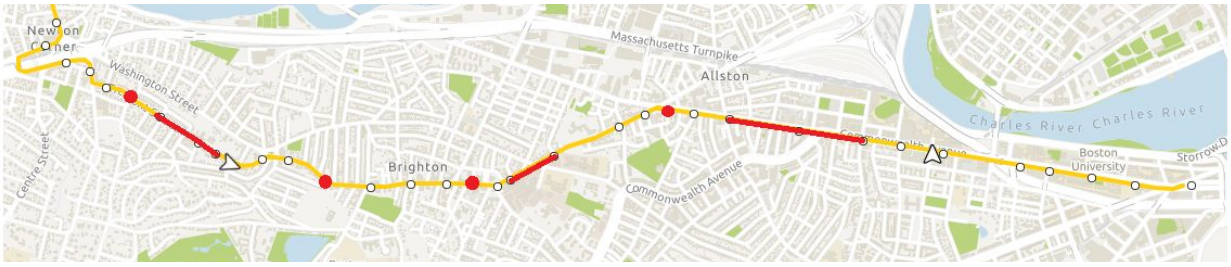
Grading System

1. Set up grading system based on strategies questions we had in project description

-- including two parts: grade route conditions and ridership for each stops during different time patterns.

-- e.g got -10 when met a congestion, got -5 when met an intersection and etc. We will analyze the route sections which have stations that have a large amount of ridership or lots of congestion. As shown in the image below, we will find the core stops(highlight in red spots) and

route sections (highlight in red lines) in the 57 Bus route and grade them to see there are better routes(with high score) to replace.



Core routes sections

```
In  [28]: bustop['features']

Out[28]: [{'attributes': {'OBJECTID': 1,
    'HastusId': 15660,
    'Name': 'Blue Hill Ave @ Devon St',
    'OnStreet': 'Blue Hill Ave',
    'AtStreet': 'Devon St',
    'Latitude': 42.310723,
    'Longitude': -71.081571,
    'Source': 'CTPS_2016',
    'Operator': 'MBTA',
    'SurveyId': '7B75AF51-3293-4537-A9D5-0009AF3F598B',
    'Town': 'Boston',
    'Score': 'Medium'},
   'geometry': {'x': -71.08160801231861, 'y': 42.31073924076076}}]
```

Core stations information from PATI dataset

Find the Alternative Routes:
With the help of several APIs and datasets we will find alternative routes. The main tool we will use is the Google Map Platform and PATI bus stops dataset. With the PATI dataset we can get the location of the stations that we want to optimize and we will find several alternative routes by Google Map Platform.

Give the optimized routes plans:
compare routes: only replace current route with a possible one with "much lower" score(threshold set up manually), or disseminate amount of passengers to several lower routes with closed scores to achieve load balance.

5. Out Of Scope

Although we are going to try best to cover all combinations of possible routes under current traffic system to explore the optimal ones, several limitations do exist within the strategy designed.

Our model only analyze on the data of BPS bus routes within MBTA bus service network. Because of technical constraints and time limitation, other factors that may bias our criteria are not discussed, which includes but not limited to

- The time taken for students walking from stations to schools
- The possibility of taking alternative vehicles(such as bike and ferry) to cut down the cost of time
- The exact number of students coming in and out at each station

Since our model gives simplified technical solution, some other constraints are not considered as well. For instance, the optimized routes plan may simulate disseminating student passengers to several different routes at a certain station based on the real-time traffic conditions in case of leading to a second traffic congestion to adding up the time spent. However our model will only handle such a scenario in a simple way.

The questions mentioned above will be kept opened and remained for future studying.

6. Delivery Planning

Sprint 1: 9/30 – 10/9:
Sprint Goals
In the first sprint, we are going to analyze project description, choose the datasets to work on and explore methods to solve the bus route optimization problem. Particularly, our team is going to determine what questions to be answered and what assumptions will be made under the context of Boston traffic system.
Delivery: Project Proposal

Sprint 2: 10/10 – 10/25:
Sprint Goals
Based on the outcome we had from the first sprint, our team is about to collect the target datasets online via any techniques, such as scraping. Furthermore, we'd like to study each of these datasets and write on how they will be used in our project.
Delivery: Data Source Report

Sprint 3: 10/26 – 11/11:
Sprint Goals
We aim to clean up and compile datasets from web portals and private database to generate pure data during this period, including but not limited to filter out all

uninterested fields and blocks of the data and replaces missing values with meaningful values. Finally we will transform them into client preferred format to deliver.
Delivery: Compiled and Cleaned Data

Sprint 4: 11/12 – 11/20:
Sprint Goals
The sprint is dedicated to dive deep into the solution of BPS bus route optimization problem. We plan to describe our solutions of the project problems in details and will cover all related questions raised in the project description.
Delivery: Draft Analysis

Sprint 5: 11/21 – 12/09:
Sprint Goals
The last sprint will focus on polishing the draft report as well as visualizing our results if possible. In addition to that, we will also host a poster session to better interact with clients and answer any kind of questions related to our project.
Delivery: Final Analysis and Poster