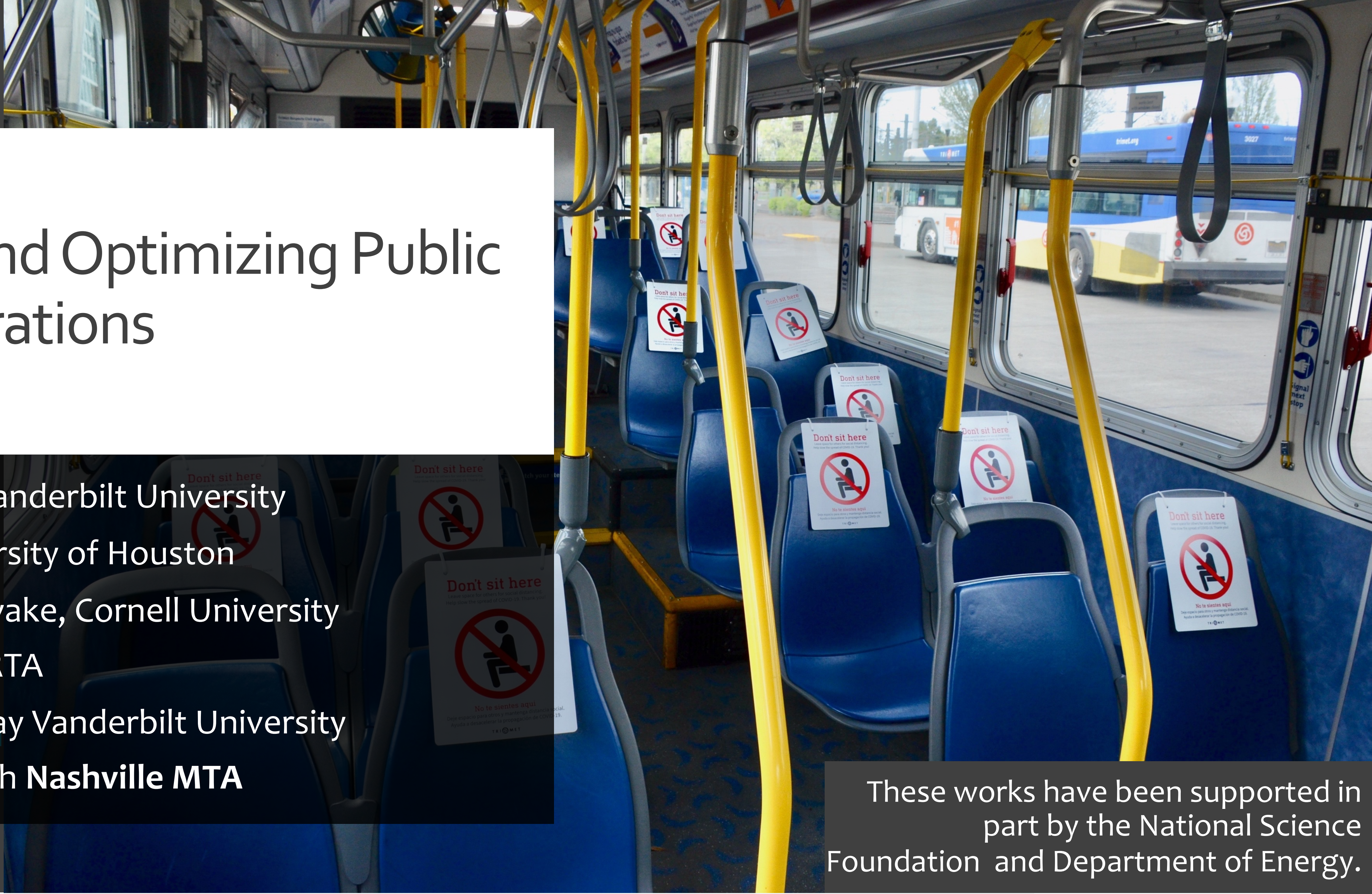
Analyzing and Optimizing Public Transit Operations

Abhishek Dubey, Vanderbilt University
Aron Laszka, University of Houston
Samitha Samaranayake, Cornell University
Philip Pugliese, CARTA
Ayan Mukhopadhyay Vanderbilt University
in collaboration with **Nashville MTA**

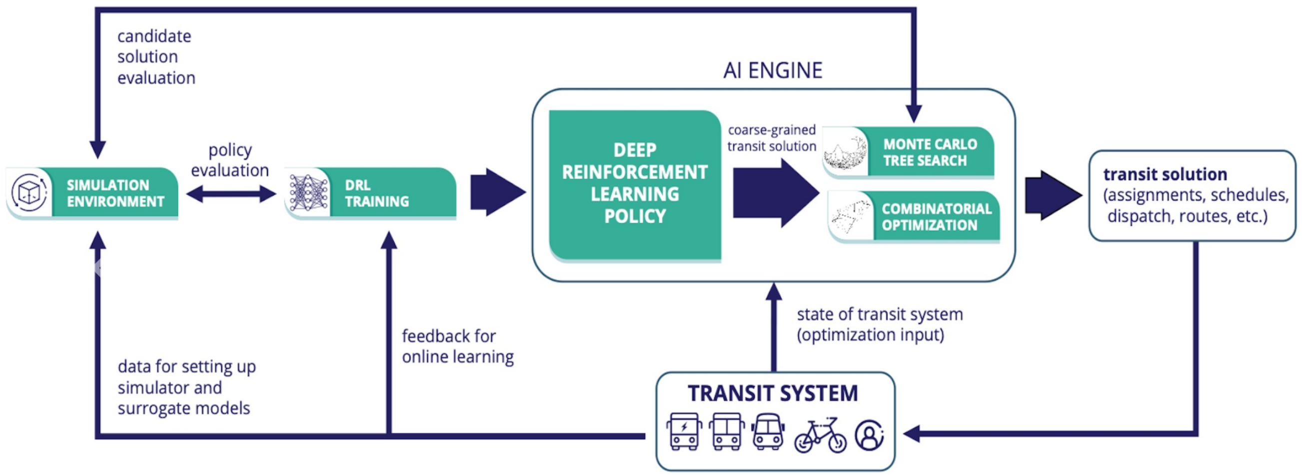


These works have been supported in part by the National Science Foundation and Department of Energy.

About us

Smart Transit Systems - Home

smarttransit.ai



Overview of AI Engine for optimization of integrated transit services.

About Us

We are a research team that collaborates with Chattanooga Area Regional Transportation Authority (CARTA) and Nashville WeGo to design efficient transit operation algorithms by using artificial intelligence and real-time data analysis at scale. This includes reinforcement learning, Monte-Carlo tree search, and operations-research based optimization for system-wide integrated scheduling and dispatch of transit operations. As part of this work, we are also developing models to estimate the load factors and real-time energy consumption of mixed-vehicle transit fleets and use those models to predict and optimize operations in order to lower overall energy impact while ensuring that system-wide capacity remains unaffected.

Abhishek Dubey

Philip Pugliese

Aron Laszka

Samitha Samaranayake

Lead Investigators

Full Investigator List

- Paul Speer, Vanderbilt University
- Ayan Mukhopdhyay, Vanderbilt University
- Himanshu Neema, Vanderbilt University
- Chandra Ward and Mina Sartipi, University of Tennessee at Chattanooga
- Siddhartha Banerjee, Cornell University
- Lillian Ratiliff, University of Washington
- Shashank Shekhar, Siemens, Corporation
- Malini Ghoshal, PNNL
- **Graduate Students and Post docs:** Jordan Jurinsky, Sayyed Vazirizade, Michael Wilbur, Amutheezan Sivaganam, Afiya Ayman, Geoffrey Pettet, Juan Martinez, Daniel Gui, Ruxiao Sun

Our approach

We use a **socio-technical** approach to address **community-scale transit problems improving accessibility and coverage** in an **equitable** and fair manner, while optimizing for energy efficiency and cost.

Research Overview

Machine Learning

- Predicting ridership using machine learning models
- Predicting energy consumption of the mixed transit fleet.
- Predicting travel time and delays considering weather, local events and likelihood of incidents.
- Analyzing equity and fairness metrics for public transit.
- Origin destination models for understanding demand shift using computer vision models.

Scheduling and Optimization of Operations

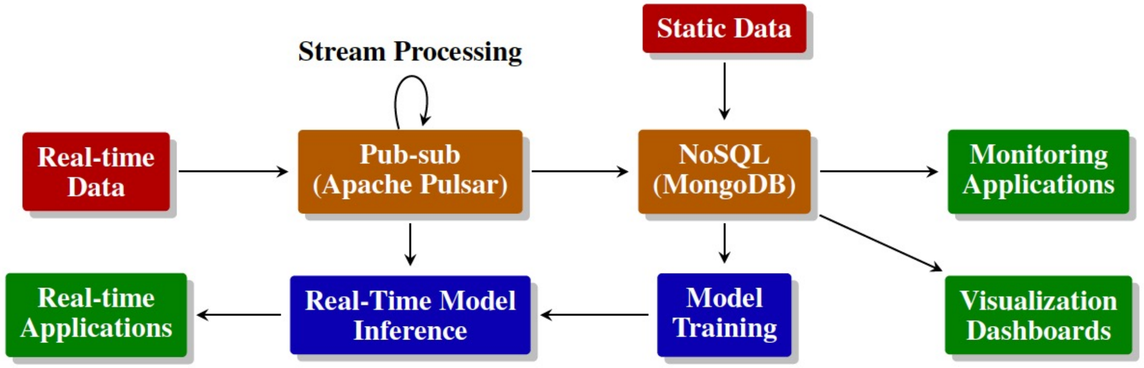
- Proactive optimization of fixed-route transit services
  - maximize transit accessibility while minimizing crowding
  - minimize energy usage by choosing optimal vehicle type assignments for each trip.
  - Considering exigent circumstances
- On-demand prioritization and dispatch for microtransit and paratransit services
  - Assign the calls to on-demand transit in anticipation of the fixed line schedule
  - Consider real-world constraints like driver hours and driver contracts on minimum and maximum hour of work
  - Integrate day ahead and real-time scheduling.
  - Provide pluggable service quality and optimization metrics.

Research Overview

Building Community Oriented and Agency Oriented Operational Software and Simulation Tools.

- Provide multi-modal travel planning software and apps considering predicted travel delays.
- Full city transportation and transit simulation for analyzing the impact of changes in service schedules to the community.
- Passenger guidance feed for estimating the likelihood of congestion in transit vehicles.
- Dashboards for visualizing occupancy, delays, travel times and energy usage for past trips.
- Big data architecture for managing real-time and historical telemetry of multimodal performance and vehicle level data for the fleet.

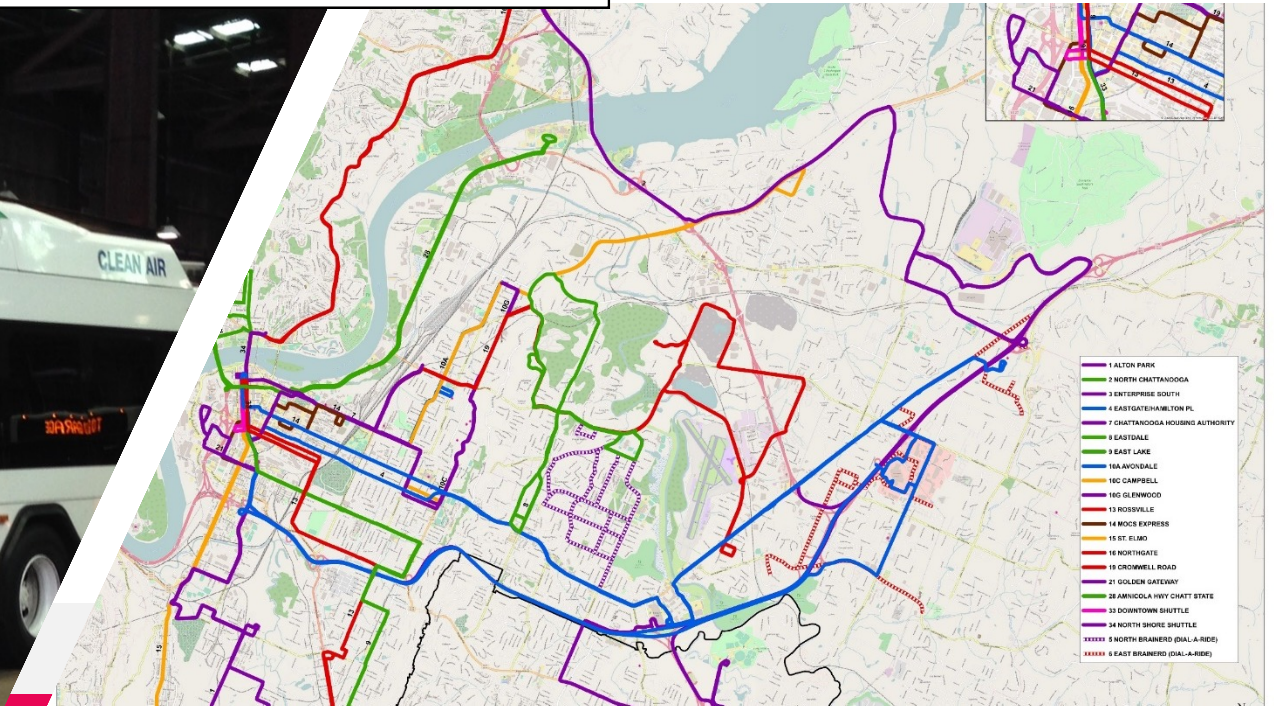
Data	Source	Frequency	Scope	Features	Schema/Format
Diesel vehicles	ViriCiti and Clever Devices	1 Hz	50 vehicles	GPS, fuel-level, fuel rate, odometer, trip ID, driver ID	Viriciti SDK and Clever API
Electric vehicles	ViriCiti and Clever Devices	1 Hz	3 vehicles	GPS, charging status, battery current, voltage, state of charge, odometer	Viriciti SDK and Clever API
Hybrid vehicles	Viriciti and Clever Devices	1 Hz	7 vehicles	GPS, fuel-level, fuel rate, odometer, trip ID, driver ID	Viriciti SDK and Clever API
Traffic	HERE and INRIX	1 Hz	Chattanooga region	TMC ID, free-flow speed, current speed, jam factor, confidence	Traffic Message Channel (TMC)
Road network	OpenStreetMap	Static	Chattanooga region	Road network map, network graph	OpenStreetMap (OSM)
Weather	DarkSky	0.1 Hz	Chattanooga region	Temperature, wind speed, precipitation, humidity, visibility	Darksky API
Elevation	Tennessee GIC	Static	Chattanooga region	Location, elevation	GIS - Digital Elevation Models
Fixed-line transit schedules	CARTA	Static	Chattanooga region	Scheduled trips and trip times, routes, stops	General Transit Feed Specification (GTFS)
Video Feeds	CARTA	30 Frames/Second	All fixed-line vehicles	Video frames	Image
APC Ridership	CARTA	1 Hz	All fixed-line vehicles	Passenger boarding count per stop	Transit authority specific



Chattanooga Area Regional Transportation Authority

CARTA is a representative sample of a mid size city transit agency

Chattanooga Area Regional Transportation Authority (CARTA)



https://smarttransit.ai


THERMO KING
All Electric Zero Emission-Clean Air
753
HAVE A
MAIN POWER DISCONNECT SWITCH
Wireless
Inductive
Charging


751
DRIVER IN
ELECTRIC VEHICLE
RESEARCH CENTER
CARTA
ELECTRIC


206
CARTA
CARTA
206


CHAMPION


CHOO CHOO
Healthy Relationships Start Here™
firstthings.org
726
FIRST things FIRST
firstthings.org


CHATTANOOGA

CHATTANOOGA
PARKING
AUTHORITY

CARTA GO

Mobility on Demand

bike. bus. drive. park.

Machine Learning and Data Analytics

Integrating Sensors in the Vehicles

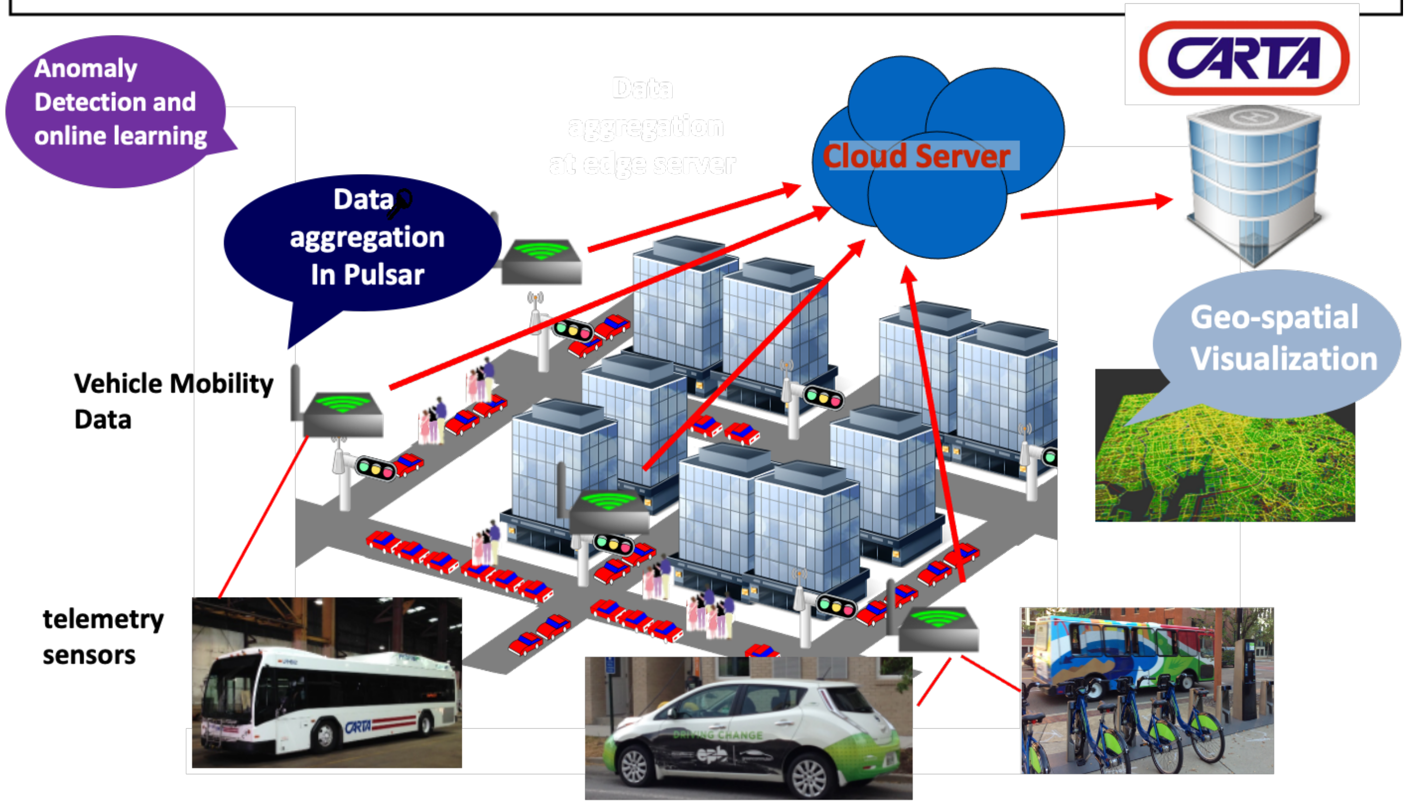
VIRICITI

This is the DataHub

The **DataHub** is ViriCiti's own **plug & play** on board solution. The DataHub is small, energy efficient and easy to install.

Due to its unique firmware, the DataHub is highly flexible and allows you to run your own programs on its **AppLayer**.

High resolution sensor data aggregation from all transit vehicles.
Anomaly detection and data store for supporting high integrity, velocity, and volume
Micro (Vehicle Specific), Macro (Elevation, Weather and Traffic) Energy Prediction for Mixed Fleet
Operational Guidance for Mixed Fleet Operations and City-wide geo-spatial visualization.

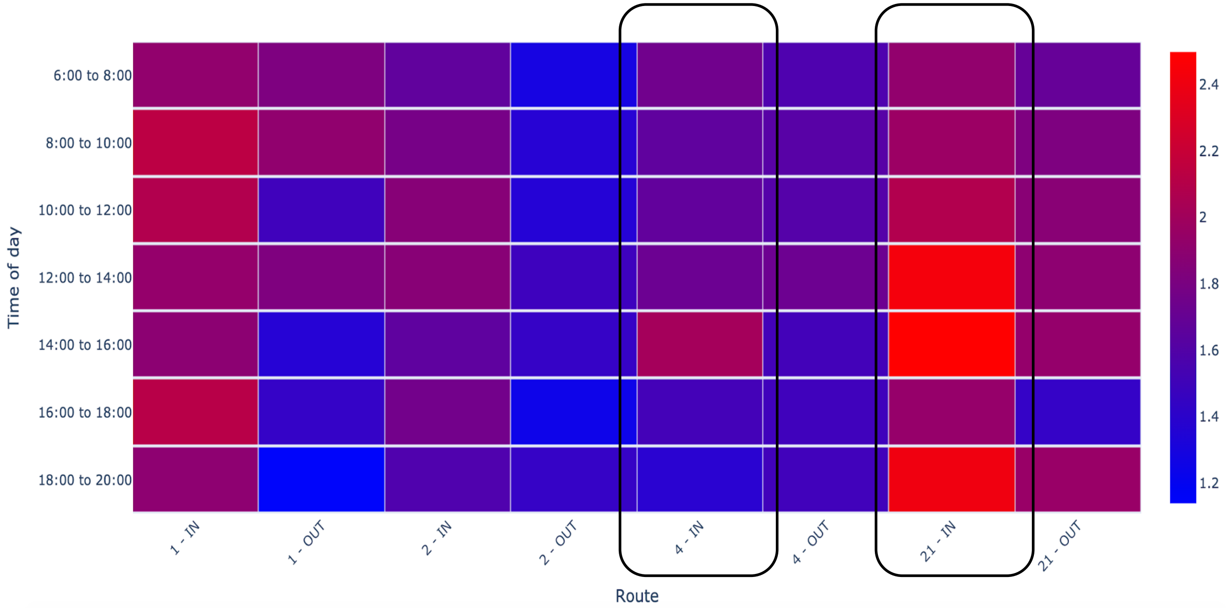


System Architecture

Understanding Energy Consumption

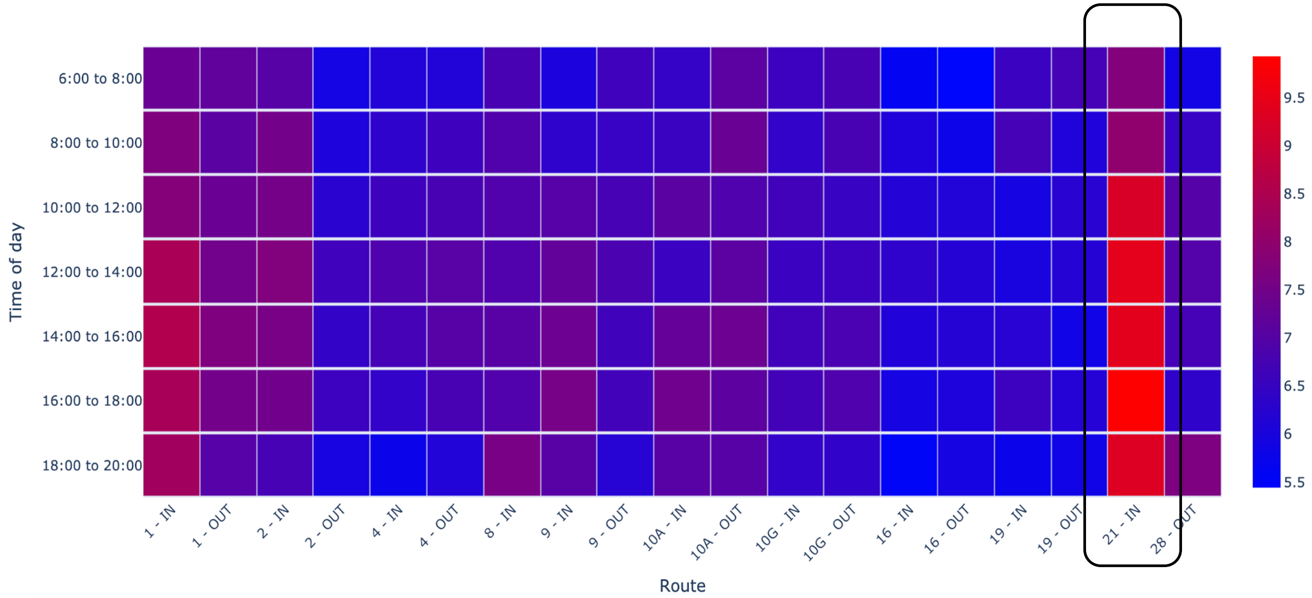
Understanding Energy Costs

Energy (kWh/mile) per Route – BYD Electric Vehicles



Route 21 – has more stops and hilly terrain

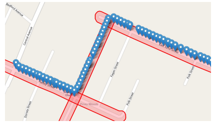
Energy (kWh/mile) per Route Gillig Diesel Vehicles



Route 21 – consumption is ~ 4 times more than electric

- Diesel vehicles are more affected by time of day than electric vehicles.
  This supports our thoughts that electric vehicles perform better in high traffic.
- The scales of the heatmap are different because of the difference in energy consumption magnitude between electric and diesel vehicles

Macroscopic Energy Prediction



Vehicle make and model

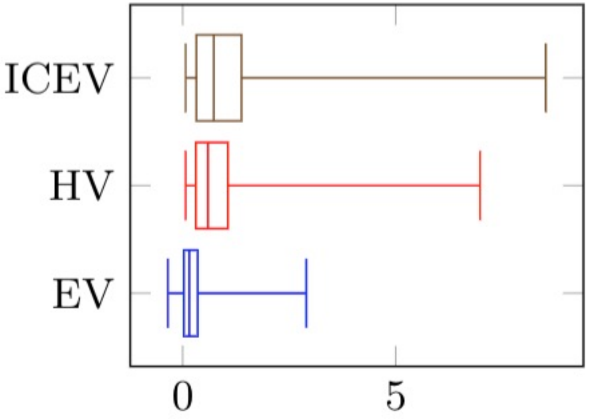
Route (series of locations)

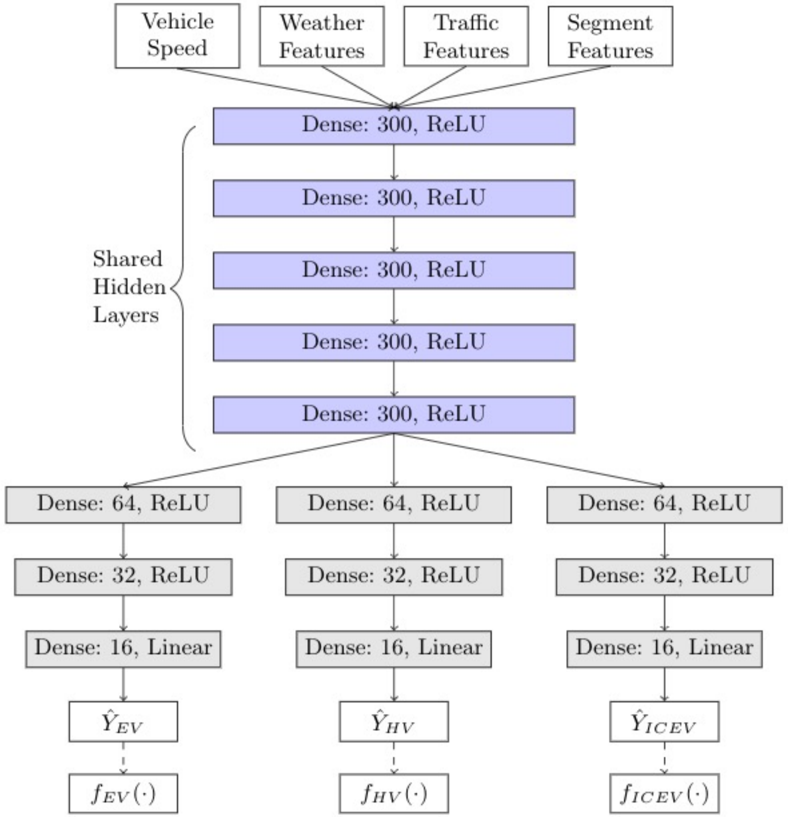
Day of week, time of day

Expected Weather, Expected Traffic

Neural network

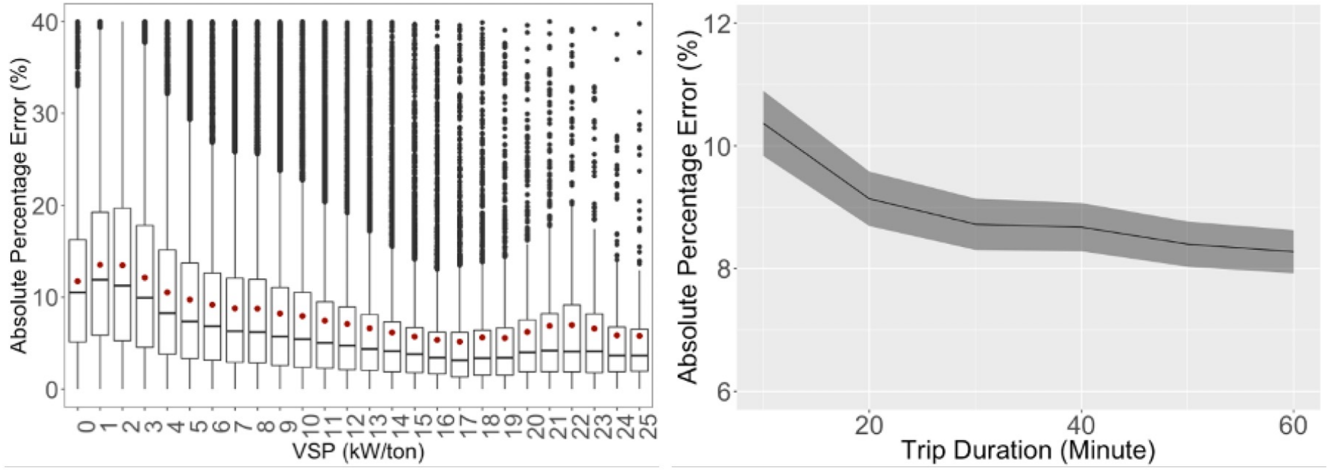
Energy use (gallons or Watt-seconds)





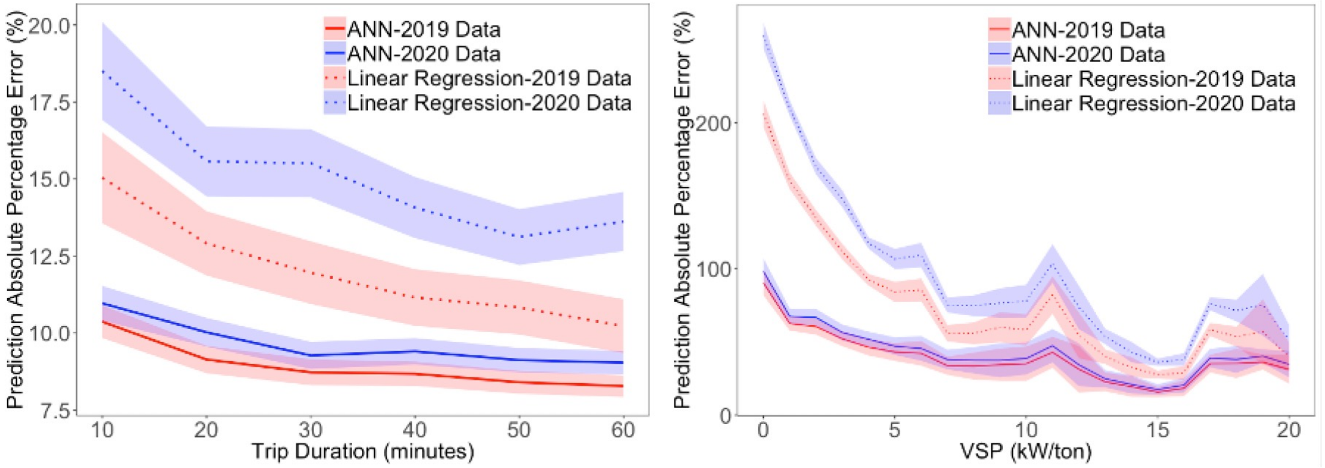
Microscopic Energy Prediction Models

Predicting Vehicle Specific Statistics along the trip



(a) (b)

Figure 8: (a) Boxplot of absolute percentage error (%) for the microscopic energy consumption as a function of vehicle specific power. (b) Mean and 95% confidence interval of absolute percentage errors of microscopic energy prediction at different trip duration



(a) (b)

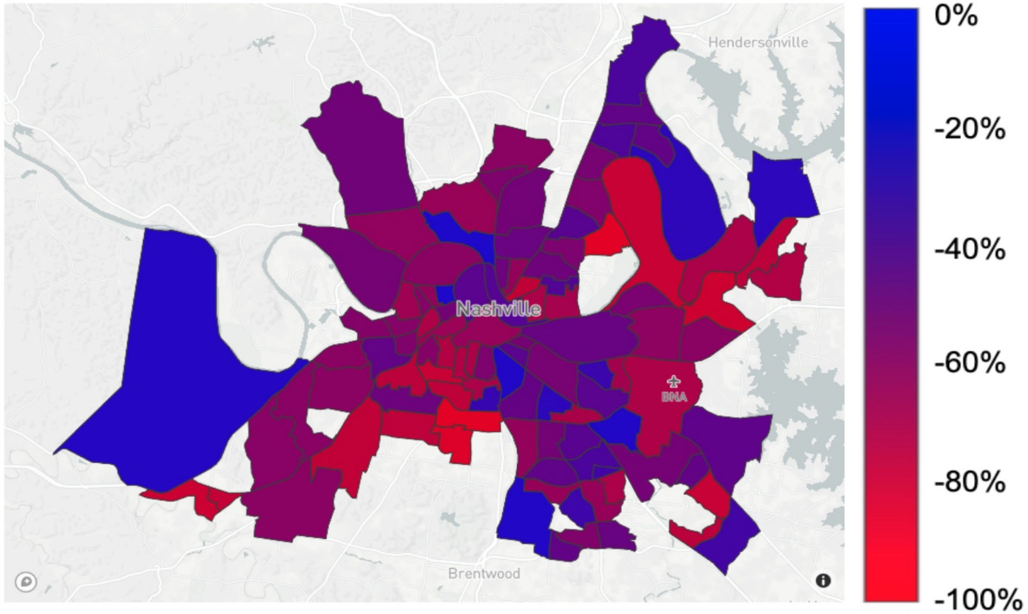
Figure 9: (a) Mean and 95% confidence interval of absolute percentage errors of microscopic energy prediction with trip duration for different models. (b) Mean and 95% confidence interval of absolute percentage errors of microscopic energy prediction at with respect to the vehicle specific power.

Understanding Ridership

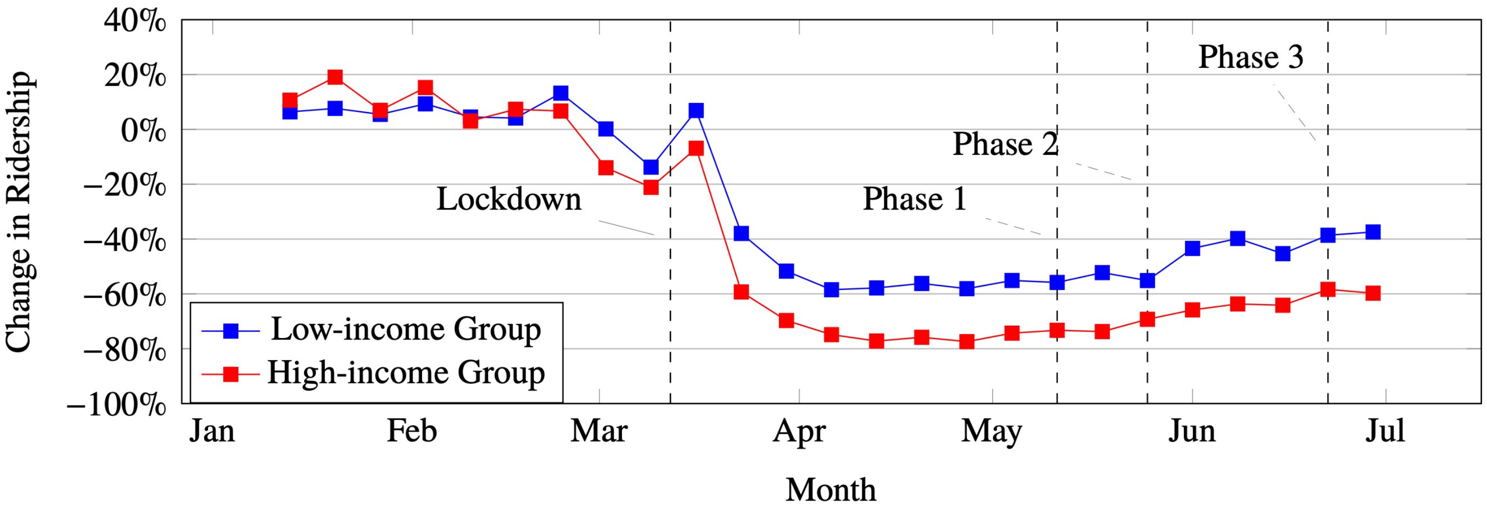
Ridership Trends

Spatial and Socio-Economic Analysis

- Significant decrease in ridership in retail and shopping areas
- Decrease in transit ridership is more significant in high-income neighborhoods than in low-income ones
- Other significant factors include housing value and rent



Metric	Pearson Correlation
Median housing value	0.35
Median income	0.21
Median rent	0.15
% White	0.01
% African American	-0.02
% Hispanic	-0.19



Publication: Wilbur et al., "Impact of COVID-19 on Public Transit Accessibility and Ridership," 2021 Annual Meeting of the Transportation Research Board.

Demand Estimation (boarding events, origin-destination events)

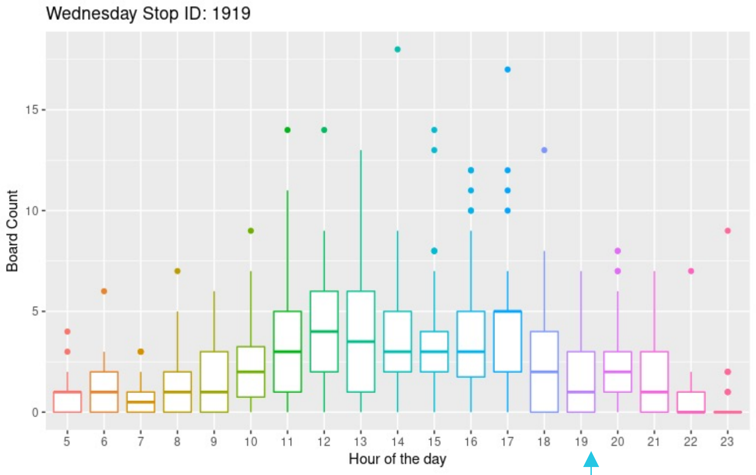
Approach: Use the automated passenger counter data, fare box data and camera data to create models for learning the distribution of commuters across bus stops and develop statistical models for prediction the future demand

Occupancy is a composition of two random processes: **boarding** and **alighting**.

- **Board counts:** $\gamma_t(s_i) \sim Po(\lambda_b^{(t)})$
- **Alight counts:** $\alpha_t(s_i) \sim Po(\lambda_a^{(t)})$

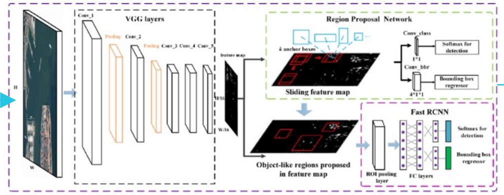
We need to learn distributions: $F_b(\gamma_t(s_i)|w)$ and $F_a(\alpha_t(s_i)|w)$. These can be used to seed a generative model that can be used to predict the likely demand at any bus stop at any time in the future given the nearby events, weather and information about that day.

Models for predicting demand per route, per stop



- Automated Passenger Counter (APC Data)
- Farebox Data
- Travel Demand Model Data

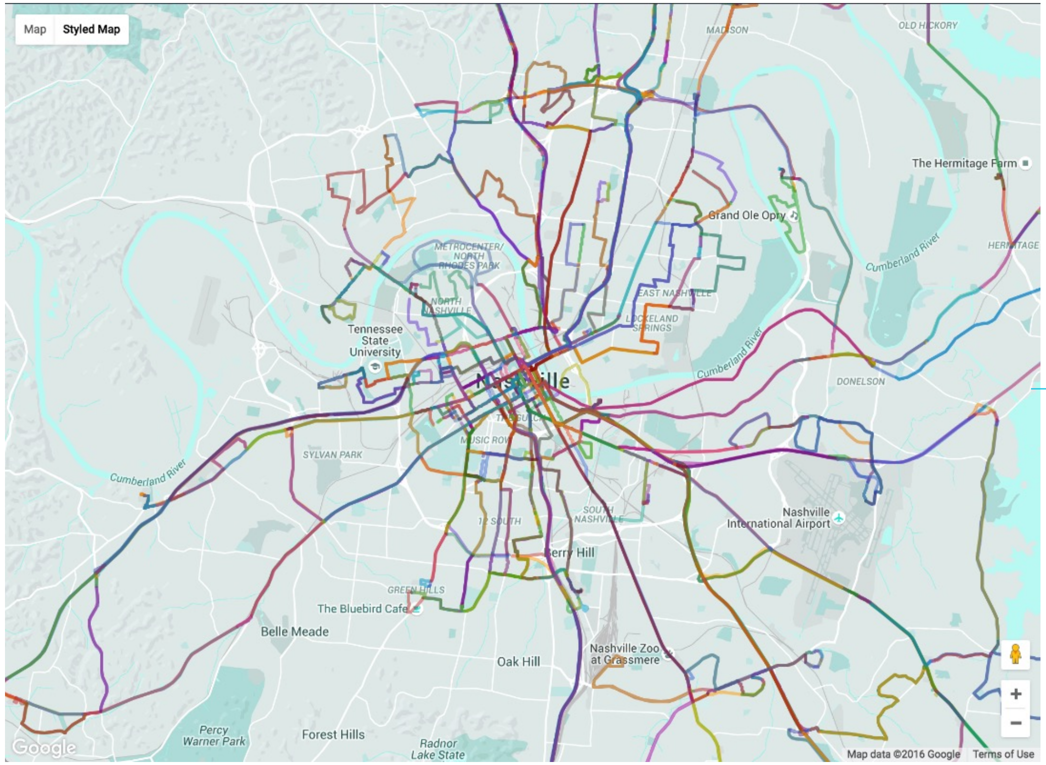
- Correction Factor



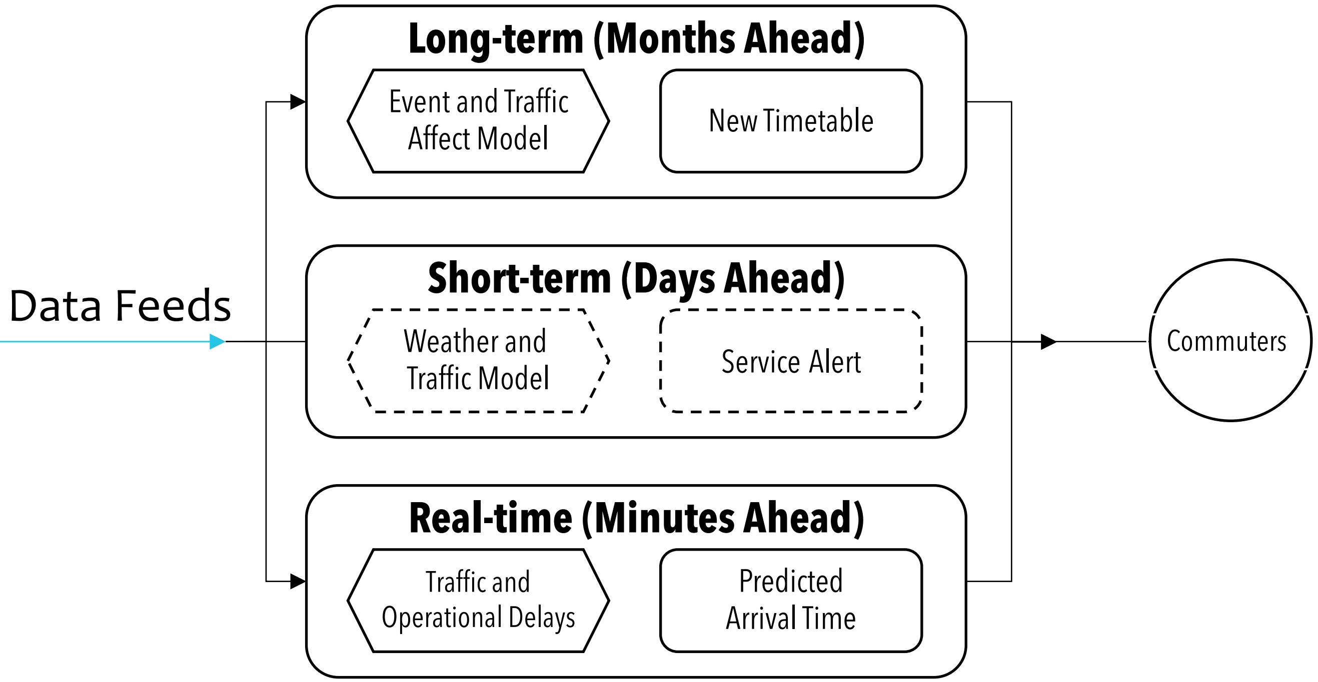
Challenges: privacy, robustness of prediction, understanding and responding to distribution shifts

Understanding Delay and Exigent Circumstances

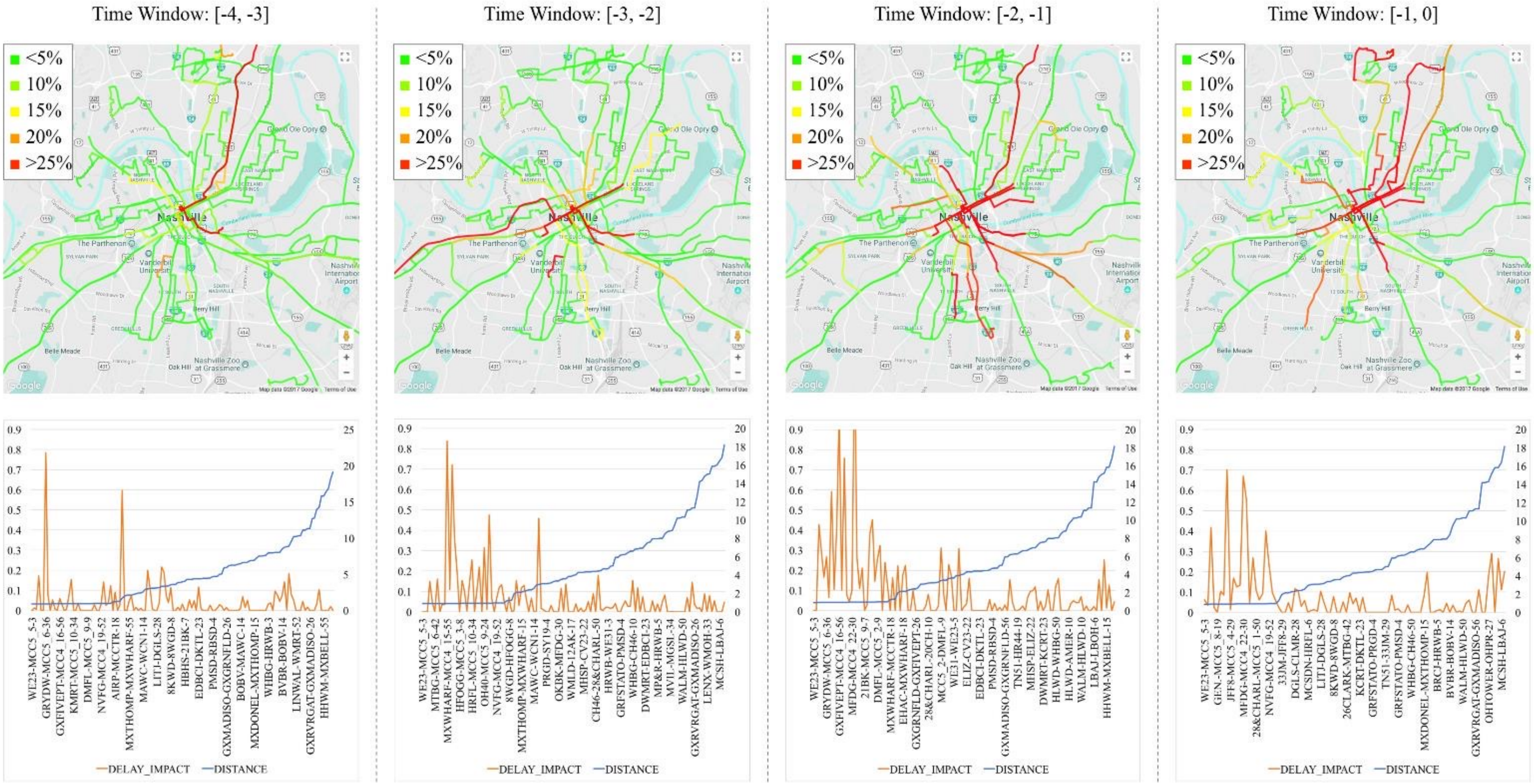
Computing Bottleneck Information



Shared Route Segment Network

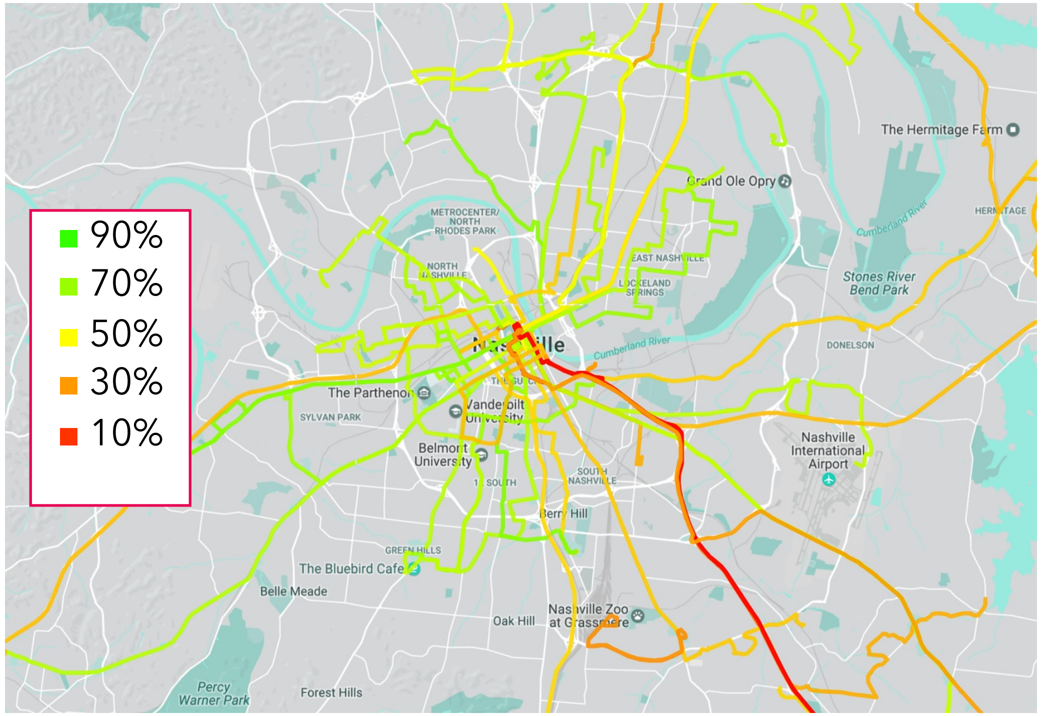


Short-Term Delay Non-Recurring Events: NFL Games

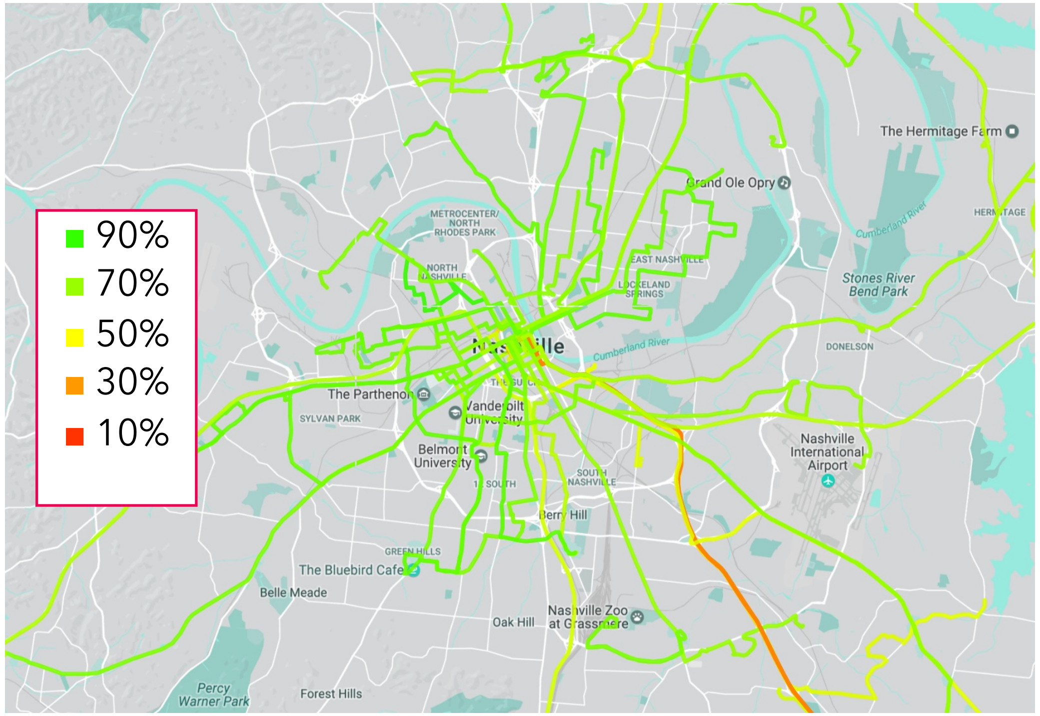


Schedule and Operations Optimization

Optimizing for delays



Original



Optimized

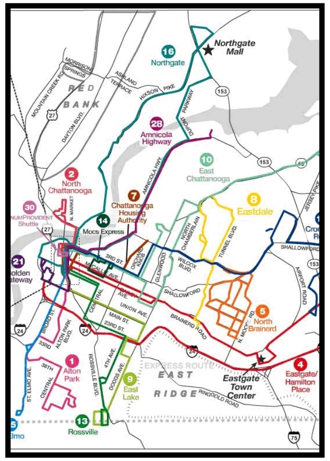
We scheduled using heuristics algorithms adjusting for seasonal delays. The result is an optimized GTFS for the next season.

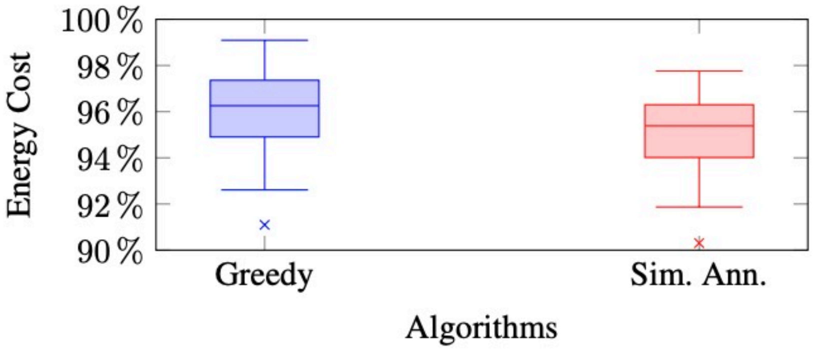
Optimizing for Energy Consumption

Transit Trips ⟷ Buses (EVs and ICEVs) ⟷ Charging Slots

Assign Transit Trips to Buses

Assign EVs to Charging Slots

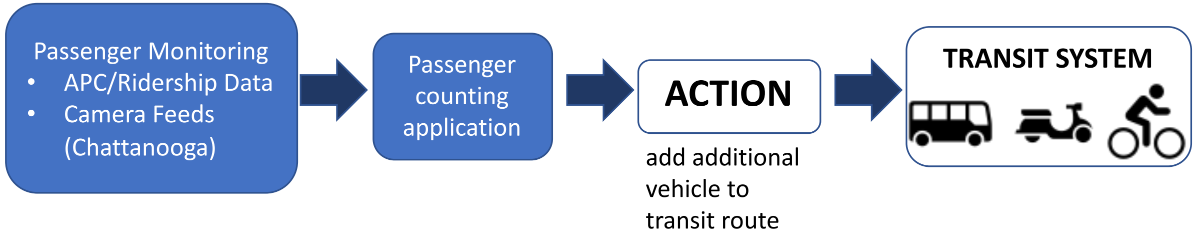




- Daily : saves \$399 of Energy Cost and reduces 1.58 metric tonnes of CO_2
- Annually: saves \$145k of Energy Cost and reduces 576.7 metric tonnes of CO_2

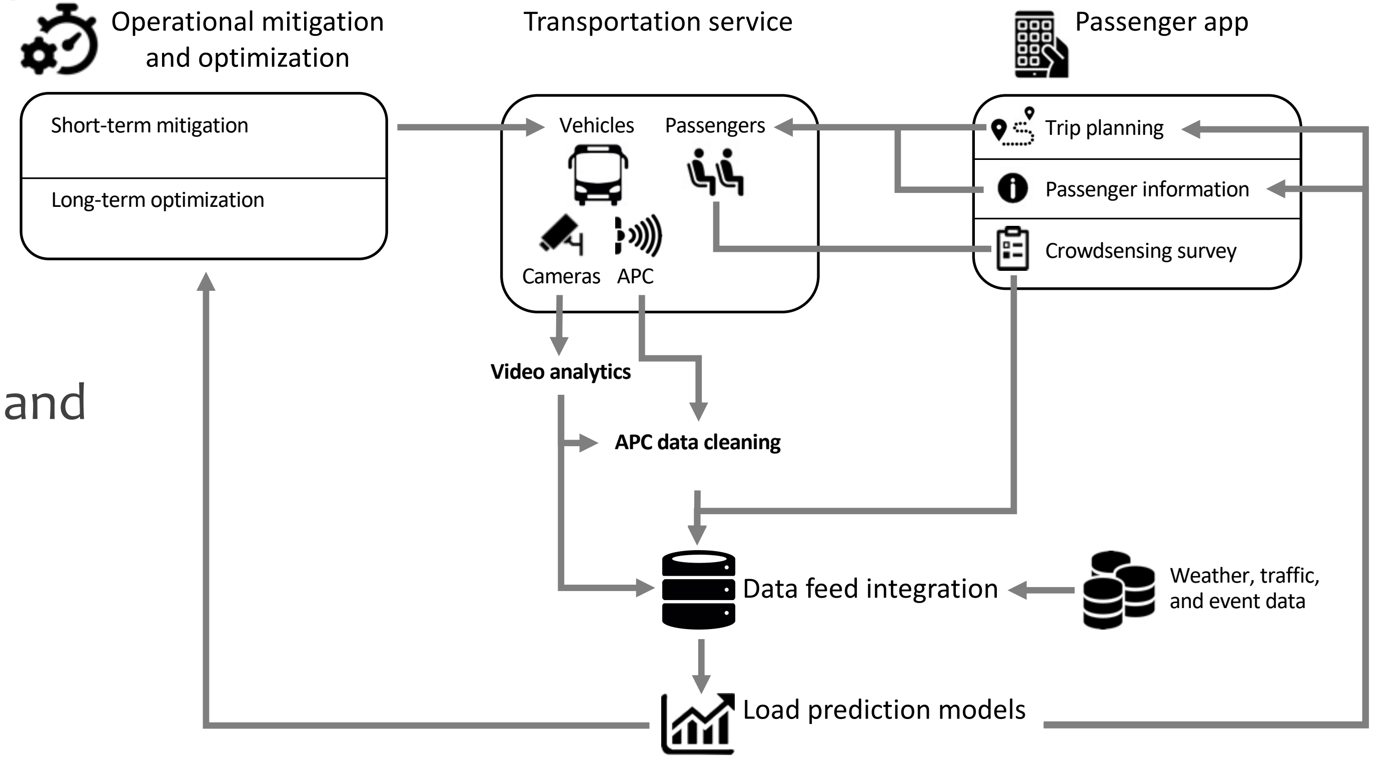
Optimizing for Social Distancing

- Affordable public transit services are the backbones of many communities, providing diverse groups of people with access to employment, education, and other public services
- COVID-19 has disrupted the operations of public transit agencies and created exigent challenges for them



- Resource Challenges

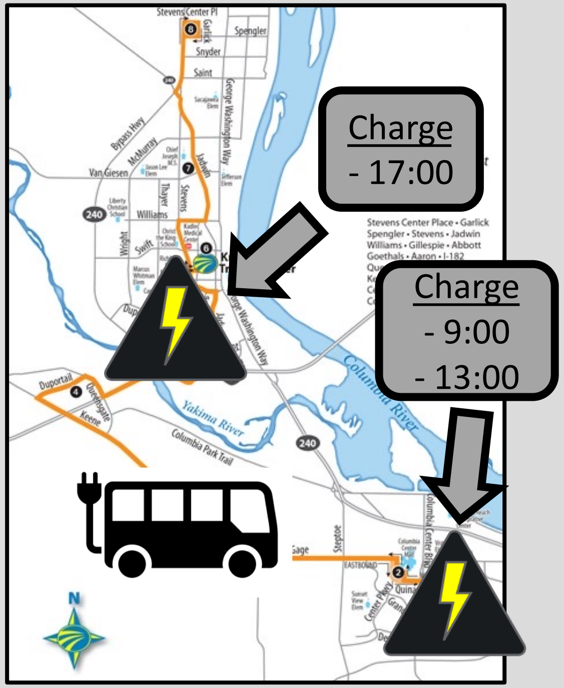
- **Reduced vehicle capacities** due to social-distancing requirements
- **Reduced driver availability** due to isolation and quarantine requirements
  - reduced service (*e.g.*, weekend schedule on weekdays)
- Disinfection requirements



Handling Large Scale EV Integration

Simulation

Replicates the transit system to estimate the impact of potential charging schedules



Traffic Simulator

Models ***travel times*** and ***battery discharge*** under varying traffic conditions

Traffic network →

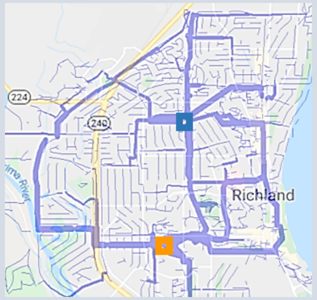
OpenStreetMaps
- Roadway Network
- Transit Schedules

Simulation of Urban MObility
- Micro traffic simulator
- Built in EV bus models

Grid Impact Model

Captures the impact charging actions have on the power grid

Impact metric derived from…
- Line losses
- Power phase balancing
- Etc.



Case study's feeder network

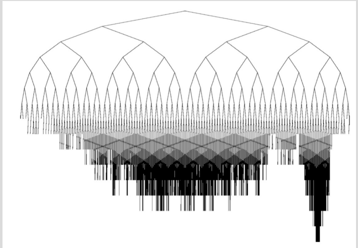
State updates, Estimated rewards →

← Charging actions to evaluate

Decision Agent

Evaluates potential charging schedules by estimating their long-term effects

Monte Carlo Tree Search
- Represents control process as game tree
- Asymmetrically grows tree, balancing exploration and exploitation
- Estimates values of actions using surrogate models
- Online algorithm; no training needed (unlike reinforcement learning). adaptable to dynamic environments



Notable application: world-champion defeating Go program[1]

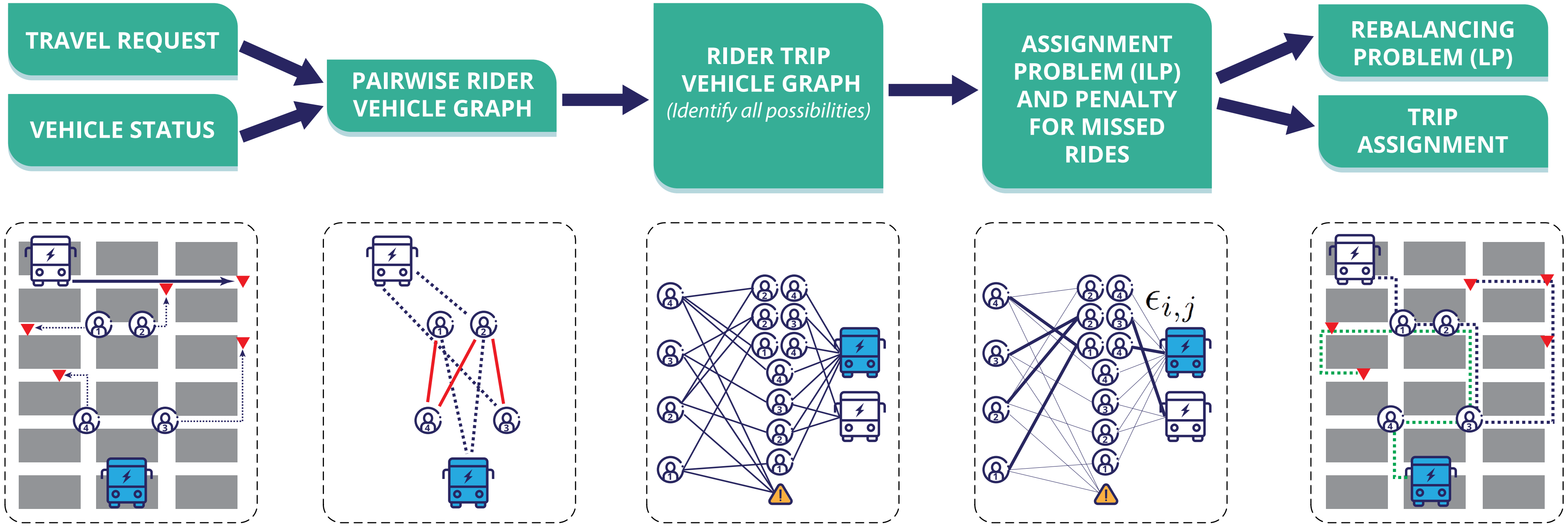
Reward Function

$$r_c = -e + \beta g + \varphi n_f$$

- energy costs (e)
- power grid impact (g)
- Number of failed buses (n_f)
- Tradeoff parameter (β)

Optimizing for Paratransit and Microtransit

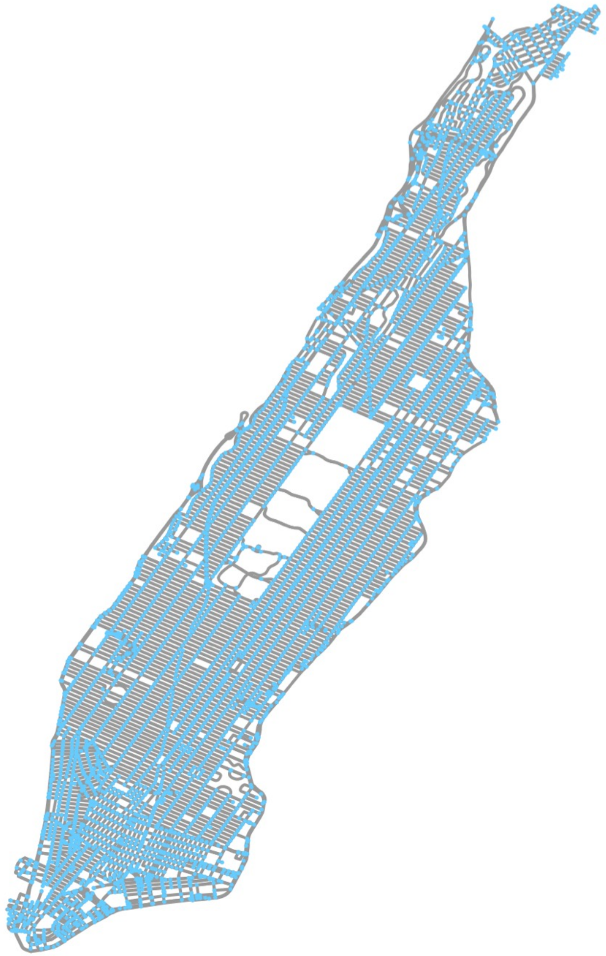
- **Approach: A generalized Modular framework.** Modular system allows easy integration of features (e.g., walk and ride, demand prediction, EV fleets)



Challenges: computation complexity, operational uncertainty and real-time requests.

High capacity sharing at the scale of NYC in real-time

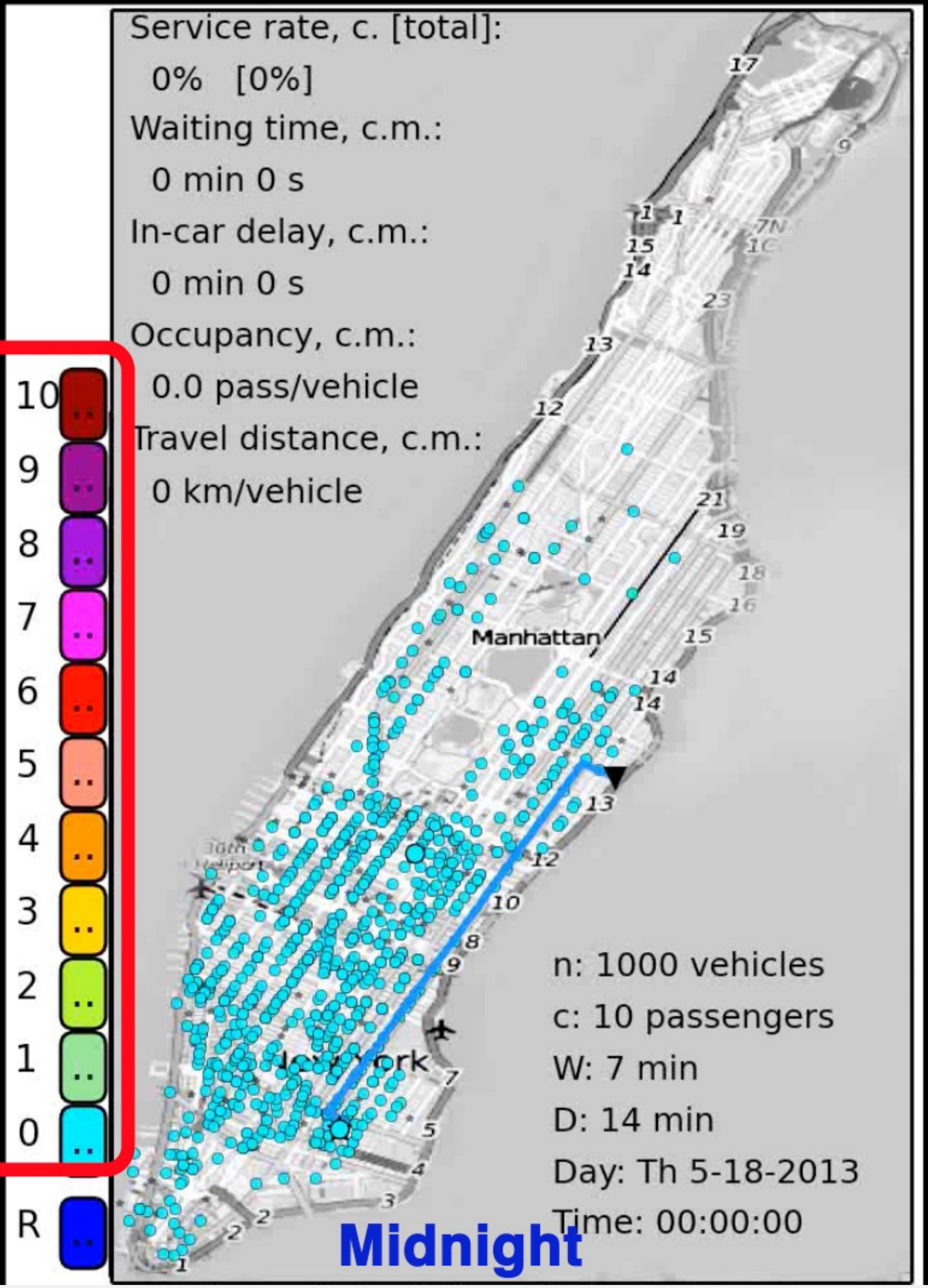
Sample week:
- May 5 - 11, 2013
- 380k (Sun) – 460k (Fri) trips/day
- 2000 active trips at anytime
- Served by 13,580 taxis



NYC Network: 4,092 nodes, 9,453 edges

The color of each vehicle (circle) shows the number of passengers, from light blue (empty) to dark red (full).

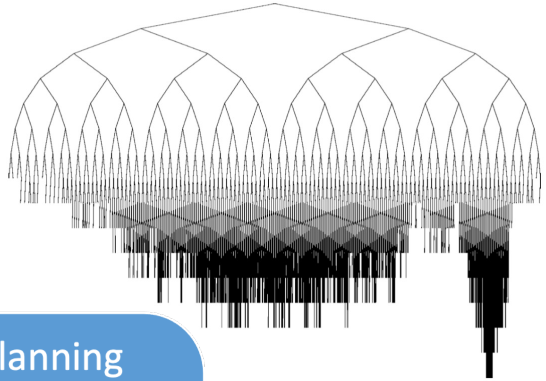
The vehicles move to serve incoming requests



Optimizing for Paratransit and Microtransit

Considerations: Multi-modal system. Incorporating real-time transit schedules (moving deadlines). Transfer deadlines need to be enforced (travel-time uncertainty in both shuttle and transit). Walking to pickup points (Group TSP problem)

Monte Carlo Tree Search: Game theoretic tree representation of process.
Nodes ➞ states , Edges ➞ actions. The tree grows asymmetrically and uses fast (online), simulated playouts to estimate value of node



Pre-processing

1. Each request is marked as a unique origin-destination pair.
2. A graph of requests and possible service runs is created.

➞

Fix Routes

1. Integer Program to optimize day-ahead requests.
2. Minimize energy and labor costs.
3. Maximize service.

➞

Real-Time Demand

1. Decision-Theoretic Approach
2. Anytime algorithm to modify existing routes.
3. Fast Monte-Carlo Search for tractability.

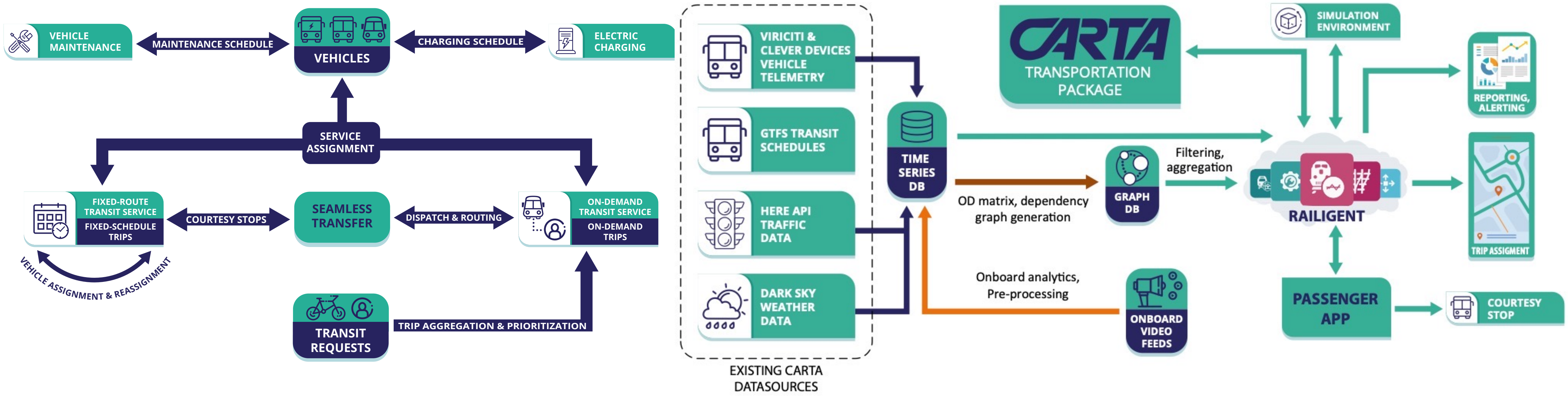
➞

Two-Stage Planning

1. Fixed routes generated by IP. Relatively time-consuming.
2. Once routes are decided, real-time requests are considered by an online optimizer.

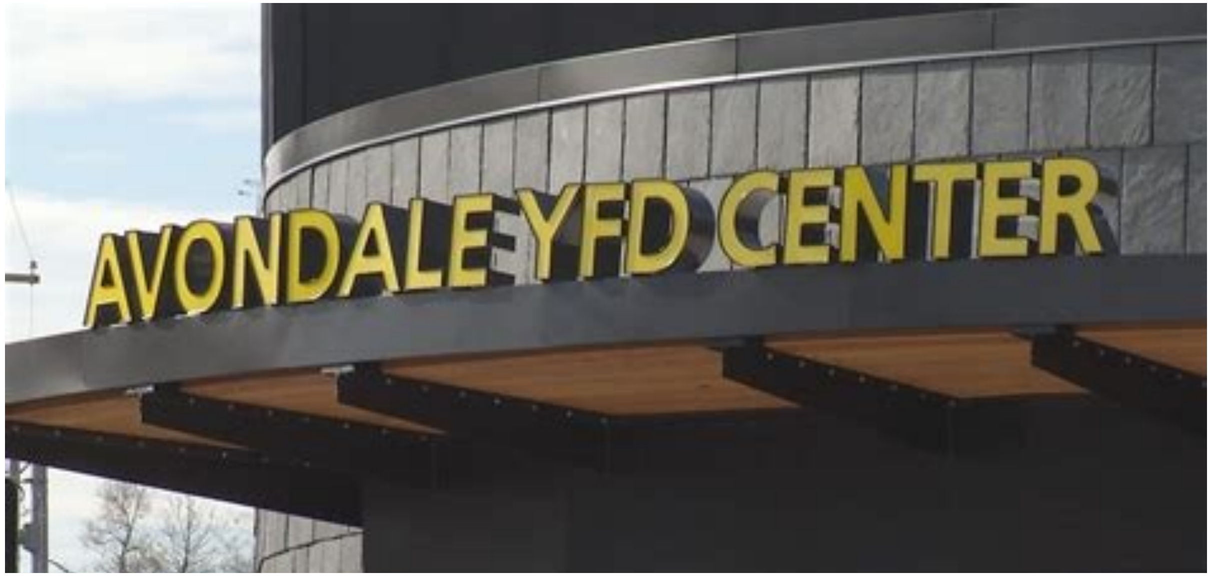
- **Challenges:** Transit centric online problems also lead to larger problem instances than in ridepooling. Real-world QoS constraints. Larger capacity vehicles. Advanced bookings

Integrated Vision



Our approach is to design a microtransit system which can serve some routes on-demand and integrate it with the fixed line system of the city. Designing this system, we must consider the operating region and think about the areas where we can transfer between the on-demand trip and the fixed line trips

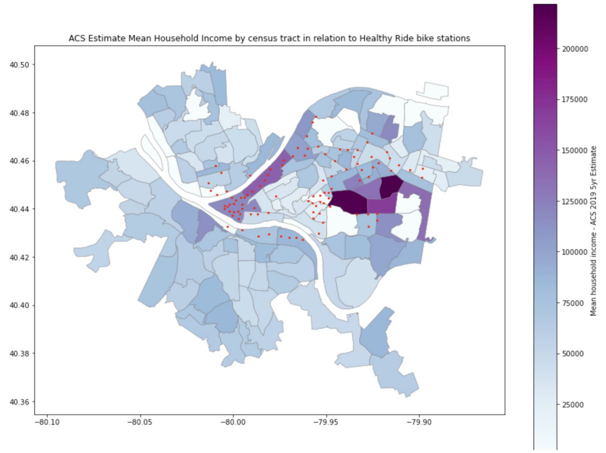
Community Outreach



Community Engagement

Identification of local community institutions composed of existing relationships. Settings to likely include:

- Community Centers
- Schools
- Signal Centers
- Public Housing
- Local Community College



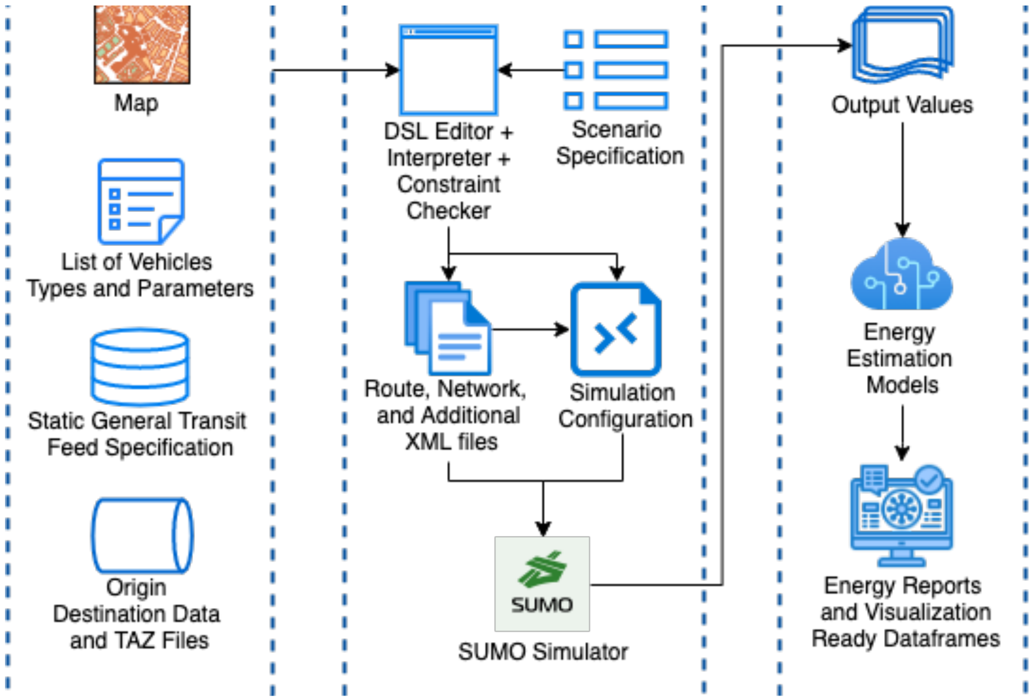
Bike stations in Pittsburgh are more accessible to high-income neighborhoods.

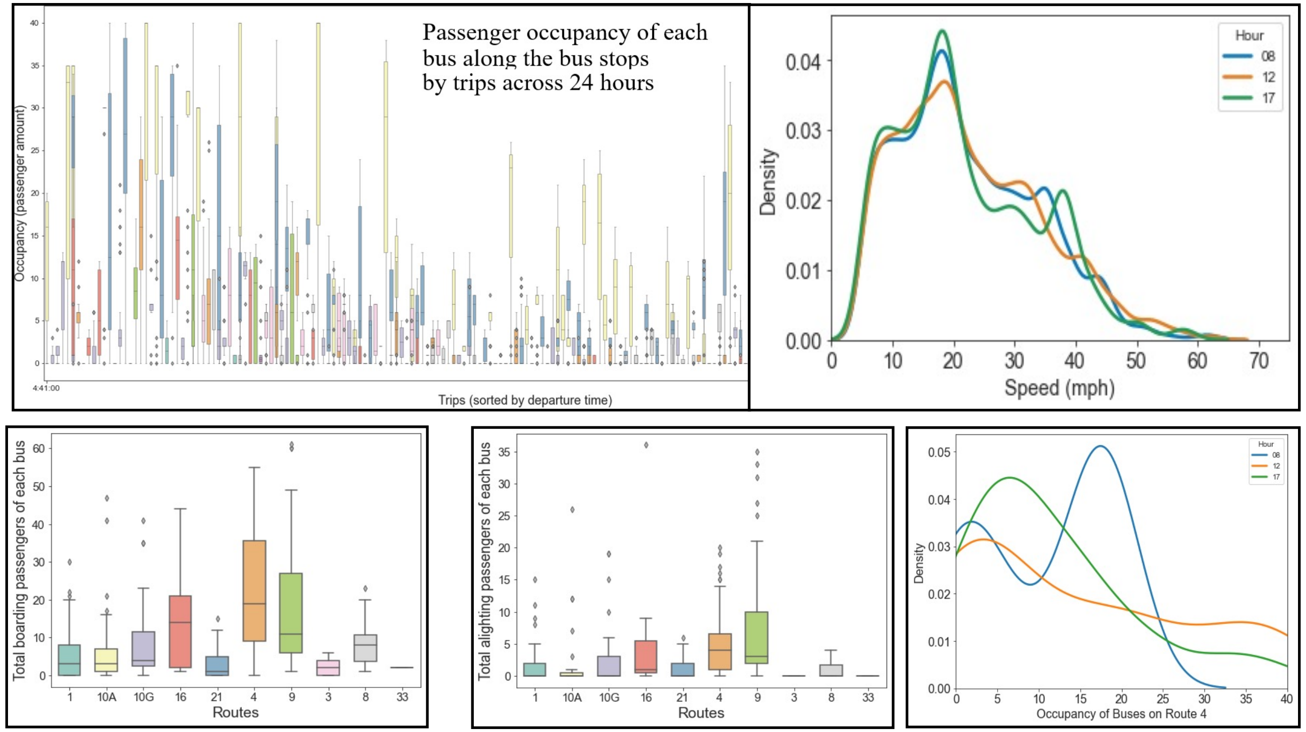
Less bike stations are available to low-income neighborhoods who have short commute times to work.

Answering questions and providing information to the community about equity

Evaluating What if Scenarios

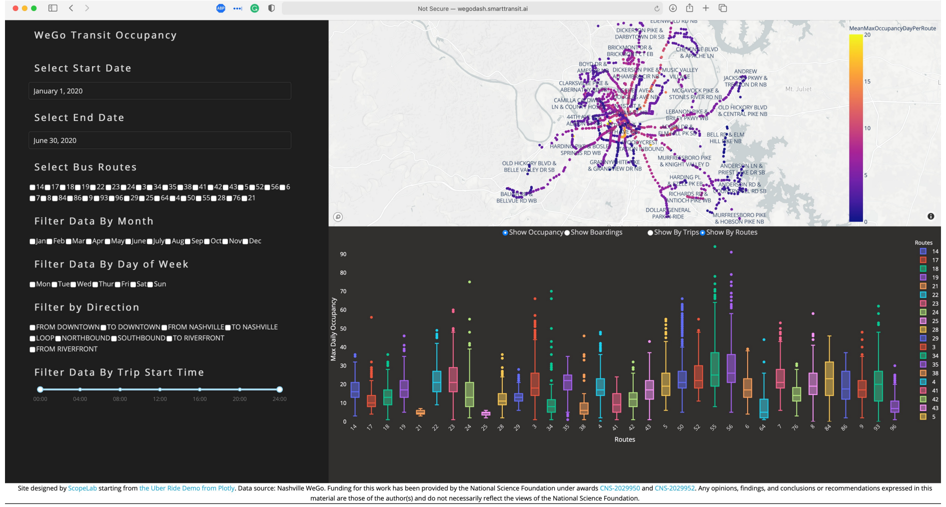
We are developing the ability to evaluate different scenarios and test the algorithms against changing demand, vehicle mix (electric vs ICEV vs Hybrid), weather and traffic patterns.

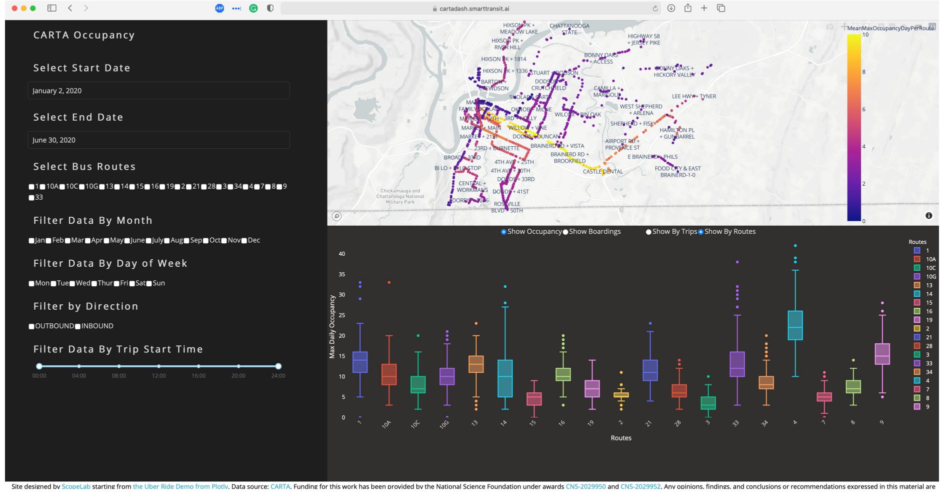




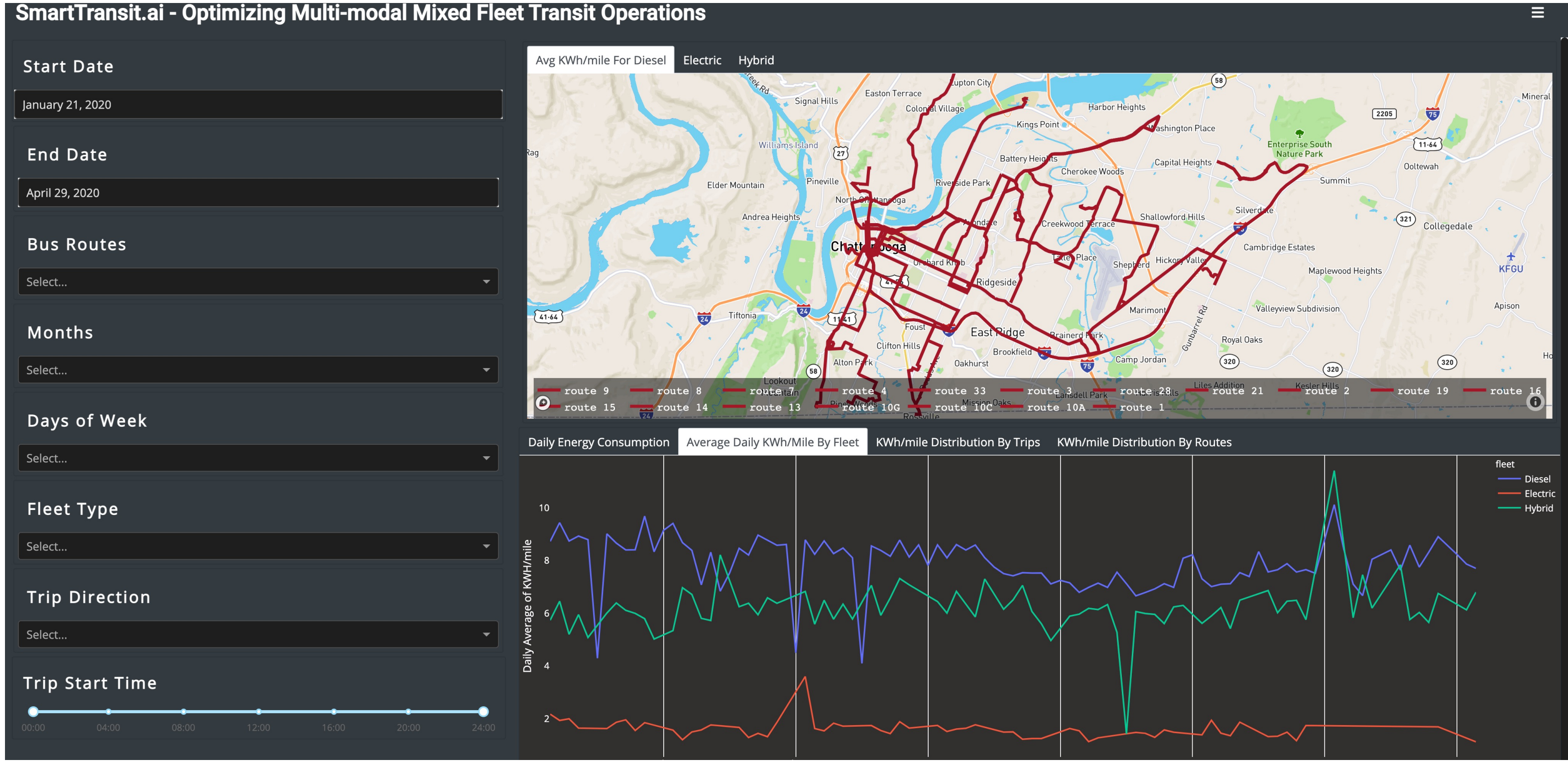
Challenges: multi-scale simulation, scenario specification, calibration of simulation models.

Real-time dashboards





We are building live community dashboards to make the service performance indicators readily available for analysis and introspection



Thank You

https://smarttransit.ai/publications/