

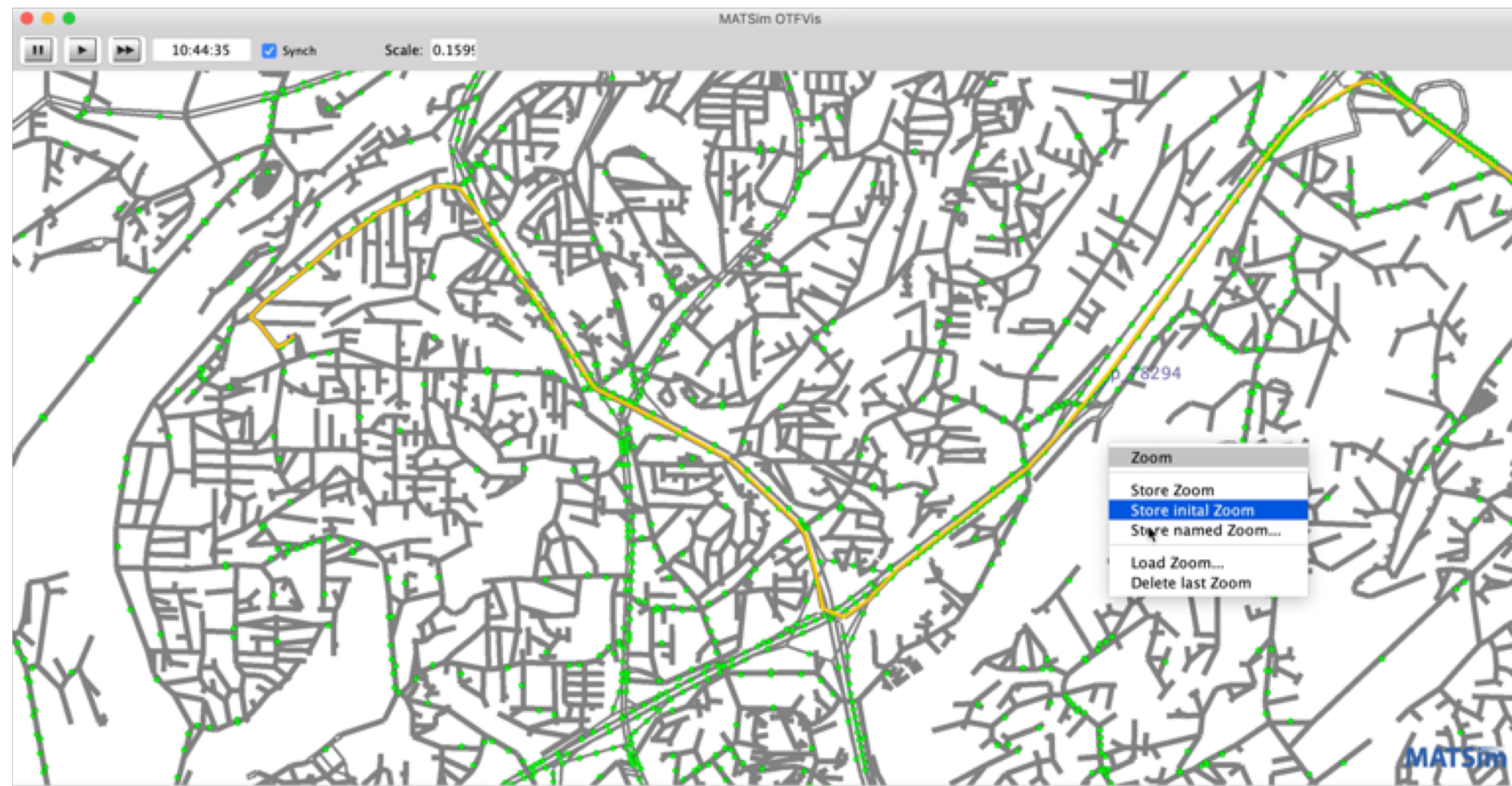
Realistic Transport Simulation with Open Data

Guimu Guo*, Jalal Khalil, Da Yan, Virginia Sisiopiku

The University of Alabama at Birmingham, Contact: guimuguo@uab.edu

The Problem

MATSim: agent-based transport simulation



User survey: 451 people with their day plans

```
In [20]: print_traj(1)
***** 014 *****
(51759.26, 3694245.51), (03:20:00, 03:40:00), Work, Car
(51846.15, 3694138.06), (12:30:00, 12:50:00), Pickup passenger, Car
(52108.24, 3714180.15), (12:58:00, 01:35:00), Shopping-Retail, Car
(52939.47, 3722913.29), (19:00:00, 19:10:00), Home, Car

***** 015 *****
(513811.97, 3710229.8), (12:30:00, 12:50:00), Services (e.g. Bank, post office), Car
(516433.42, 3698821.97), (14:00:00, 14:20:00), Home, Car
(519219.47, 3700995.9), (18:10:00, 18:10:00), Shopping-Grocery, Car
(516433.42, 3698821.97), (19:00:00, 19:10:00), Home, Car

***** 016 *****
(517956.27, 3708248.89), (17:10:00, 17:20:00), Shopping-Retail, Uber/Lyft
(517478.44, 3708718.93), (17:18:00, 17:44:00), Home, Uber/Lyft
(52132.09, 3715671.58), (16:42:00, 17:00:00), Nightlife/Ent, Car rental
(517478.44, 3708718.93), (19:00:00, 07:18:00), Home, Car rental

In [154]: getPerson('005').print_traj()
(516227.31, 3695236.73), Home (Initial Location), Home, Car
(520497.7, 3703029.62), (05:45:00, 06:00:00), Home, Car
(520843.16, 3702785.6), (06:30:00, 07:00:00), Work, Car
(519515.05, 3707745.67), (07:00:00, 07:15:00), Work, Car
(520497.7, 3703029.62), (08:00:00, 08:20:00), Work, Car
(519659.7, 3701123.61), (09:00:00, 09:20:00), Work, Car
(520497.7, 3703029.62), (09:20:00, 09:30:00), Work, Car
(516227.31, 3695236.73), (10:00:00, 10:15:00), Home, Car
(520497.7, 3703029.62), (12:00:00, 12:20:00), Work, Car
(520129.24, 3703687.88), (13:00:00, 13:10:00), Work, Car
(520497.7, 3703029.62), (13:20:00, 13:20:00), Work, Car
(516227.31, 3695236.73), (15:00:00, 15:20:00), Home, Car
(517820.51, 3696054.02), (16:00:00, 16:08:00), Shopping-Grocery, Car
(516227.31, 3695236.73), (16:15:00, 16:23:00), Home, Car
```

Persons with 4 Legs

Legs of a Uber Driver

The Problem: to generate a realistic population from the survey as an input to MATSim for simulation

The simulation model can be used to study traffic problems such as road congestions, and impact of Uber traffic flow

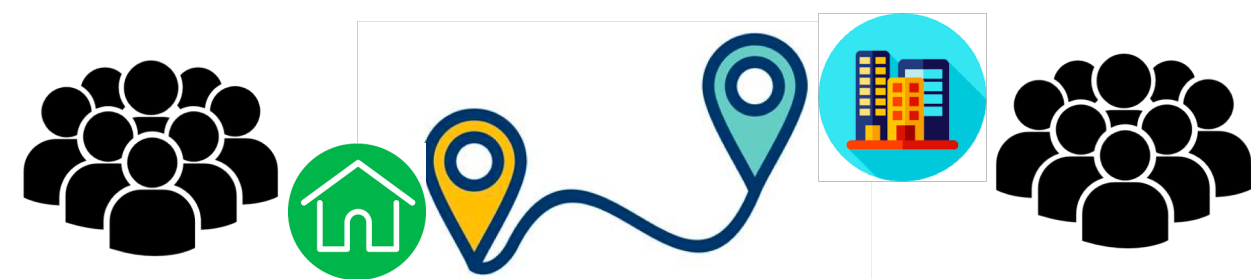
The Challenges

Conventional Approach: Iterative Proportional Fitting (IPF) according to agents' attribute distribution in the US Census

Challenges: the small data problem (tackled with open data)

IPF only generates # of people in each region; missing details like locations and activities

Sampling day plans?

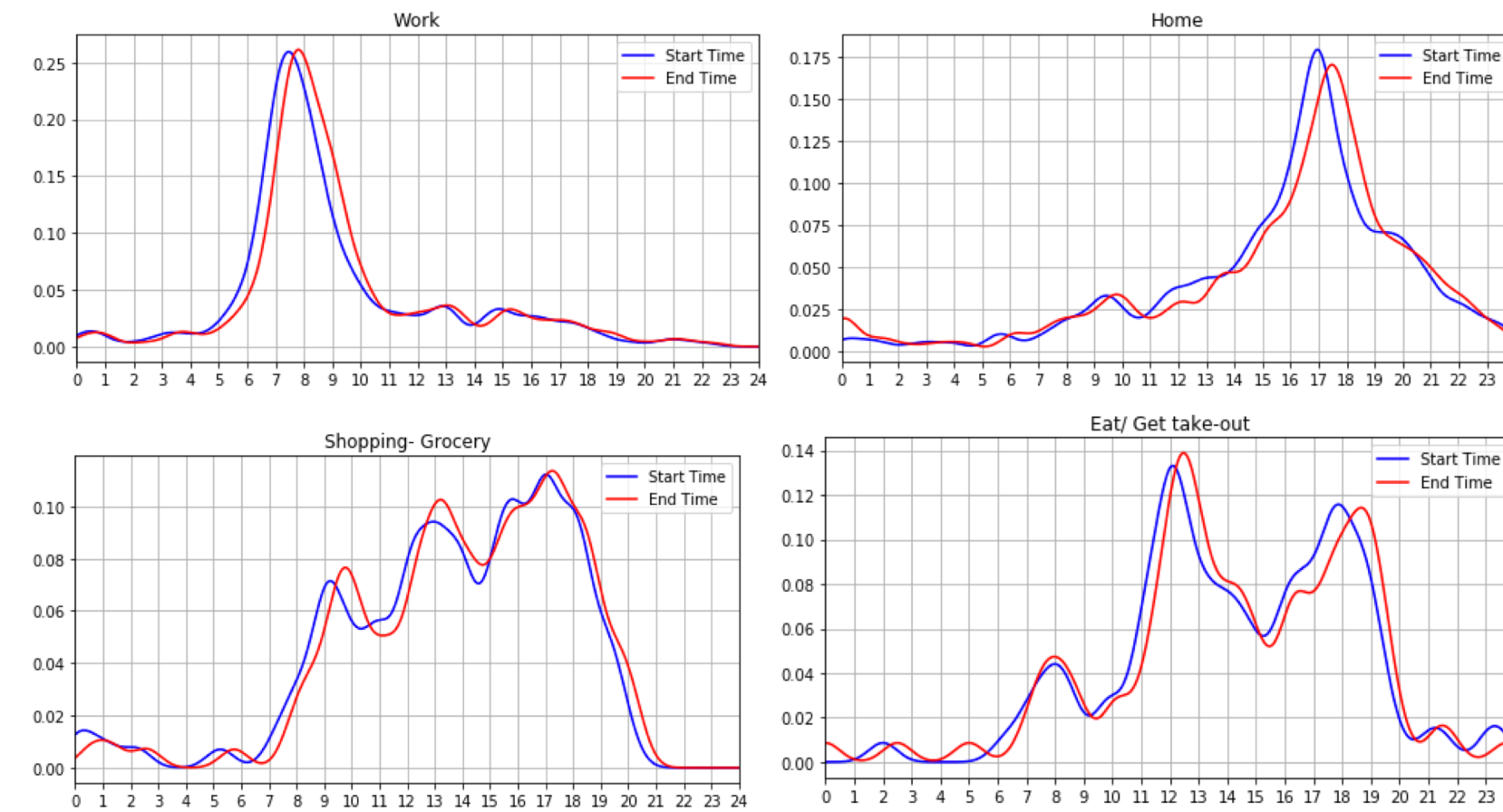


Our Solution

Data-Driven Approach: modeling (1) time, (2) location, (3) activities and (4) mode using generative machine learning models, with the help of open data.

Time Modeling using Kernel Density Estimation (KDE)

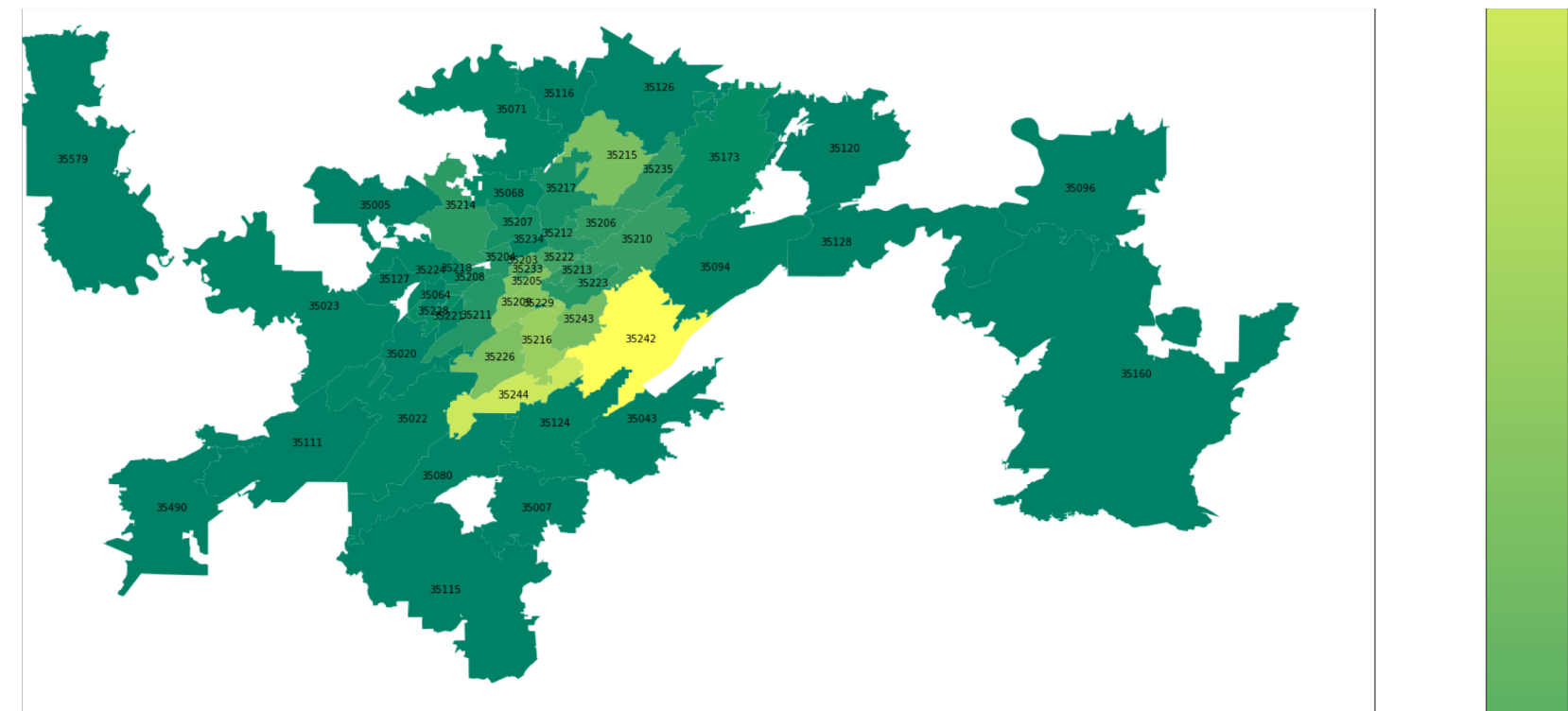
Activity start and end time distributions:



We can model the travel time using KDE similarly

Location Modeling with Open Data

Zip Code Tabulation Areas (ZCTA) from TigerWeb



Real addresses from OpenAddresses

NAME	DP03_0018E	zip code tabulation area
0	ZCTA5 35401	11636 35401
1	ZCTA5 35071	7141 35071
2	ZCTA5 35005	2777 35005
3	ZCTA5 35213	6409 35213
4	ZCTA5 35215	20017 35215
5	ZCTA5 35111	7809 35111
6	ZCTA5 35124	12974 35124
7	ZCTA5 35203	947 35203
8	ZCTA5 35206	6862 35206
9	ZCTA5 35209	16154 35209
10	ZCTA5 35218	2583 35218

IPF Source ZCTA marginal:

From American Community Survey (ACS), Variable "P03_0018E: COMMUTING TO WORK"

Destination ZCTA marginal: Birmingham Business Alliance

EMPLOYER NAME	EMPLOYER	ESTIMATED NUMBER OF EMPLOYEES	PRODUCTS OR SERVICES	COUNTY	METROLOCALITY
1	University of Alabama at Birmingham	23,000	Education and health care services	Jefferson	Birmingham
2	Regions Financial Corporation	9,000	Financial services, banking, corporate headquarters	Jefferson	Birmingham
3	St. Vincent's Health System	5,900	Health care services, hospital network serving matrix	Jefferson	Birmingham
4	Children of Alabama	5,000	Health care services, regional specialized health care	Jefferson	Birmingham
5	AT&T	4,517	Telecommunications, regional operations	Jefferson	Birmingham
6	Honda Manufacturing of Alabama	4,500	Manufacturing, vehicle assembly plant	Tallapoosa	Lincoln
7	Birmingham-Bartlett Health	4,459	Health care services, management	Jefferson	Birmingham
8	Jefferson County Board of Education	4,400	Government, public education	Jefferson	Birmingham
9	City of Birmingham	4,200	Government, city administration	Jefferson	Birmingham
10	Mercedes-Benz U.S. International, Inc.	3,600	Manufacturing, vehicle assembly plant	Tallapoosa	Venice
11	Ball-Cotter Ball-Bedell of Alabama	3,500	Financial services, insurance, corporate headquarters	Jefferson	Hyatt
12	Alabama Press Company	3,092	Utilities services, electrical, corporate headquarters	Jefferson	Birmingham
13	Birmingham Board of Education	2,721	Government, public education	Jefferson	Birmingham
14	Jefferson County Commission	2,500	Government, county administration	Jefferson	Birmingham

BBA Top Employer List

IPF Output Matrix

Generating agents' day plans using the IPF output matrix

Evaluation

Comparison with real traffic data

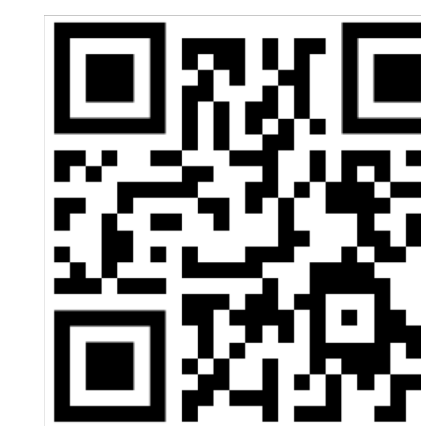
MATSim ID	Interstate	Monitored Speed	Simulation Speed
50779678_0	I-65 N	59	58.17957611
50779764_0	I-65 N	57	55.39014001
50780374_1	I-20 W/ I-59 S	67	67.23911964
119869822_0	I-20 W/ I-59 S	63	68.11513108

Future Work

OpenStreetMap POI

```
In [222]: root.print_maxDepth(3)
+-- ROOT, 443
+-- Home, 422
+-- Home, 23
+-- Home, 23
+-- Work, 170
+-- Home, 80
+-- Shopping-Grocery, 16
+-- Eat/Get take-out, 10
+-- Work, 10
+-- Pick-up passenger, 9
+-- Shopping-Retail, 8
+-- Services (e.g. Bank, post office), 9
+-- Drop-off passenger, 4
+-- Shopping-Grocery, 49
+-- Home, 23
+-- Shopping-Grocery, 8
+-- Services (e.g. Bank, post office), 43
+-- Home, 12
+-- Shopping-Grocery, 6
```

Video Demo



<https://youtu.be/Z1m0WsmKB4E>

