Traffic flow city index based on public transportation vehicles data

Georgi Yosifov
gkjosifov@fmi.uni-sofia.bg
Milen Petrov, PhD
milenp@fmi.uni-sofia.bg

Department of Software Engineering
Sofia University “St. Kliment Ohridski”
Sofia, Bulgaria

Contents

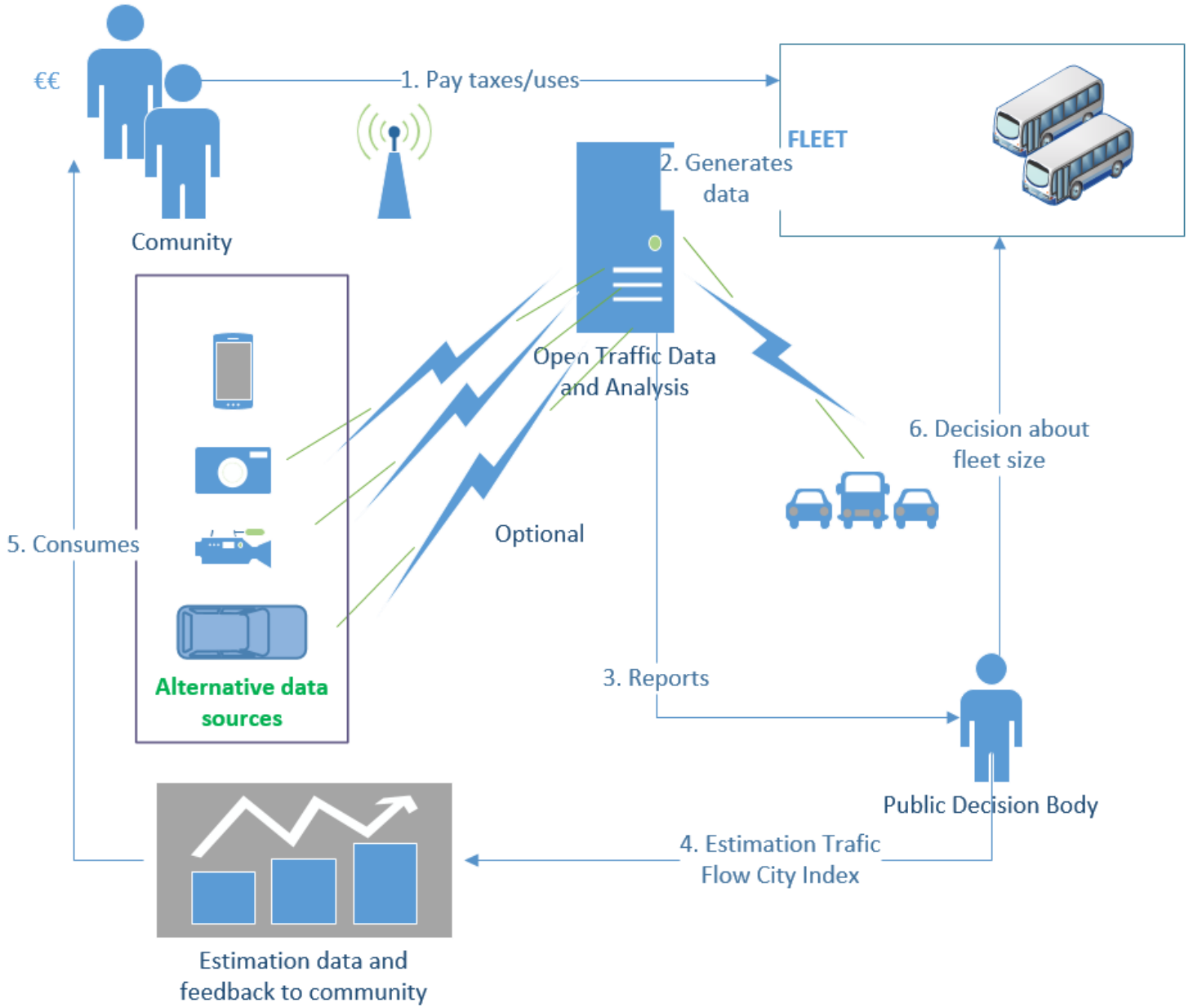
- Why traffic congestion data is important?
- Ways to identify traffic congestions
- Proposed algorithm using public transportation vehicles
- Experimentation
- Conclusion and future work

Why traffic congestion data is important?

- Public transportation optimization
- Schedule improvements
- Public administration policies for infrastructure, resource allocation
- Open-data communities
- Business opportunities

Ways to identify traffic congestions

- Smart devices/ phones
- CCTV
- Car GPS
- Public transport positioning data
- … Road sensor stripes and other …


€€
1. Pay taxes/uses
FLEET
Comunity
2. Generates data
Open Traffic Data and Analysis
6. Decision about fleet size
5. Consumes
Optional
Alternative data sources
3. Reports
Public Decision Body
4. Estimation Trafic Flow City Index
Estimation data and feedback to community

Availability, Privacy, Cost

No	Approach	Availability	Privacy concerns	Cost
1.	Car GPS	Moderate	High	Moderate
2.	Smart devices/ phones	High	High	Low
3.	CCTV	High	High	High
4.	**Public transport positioning data**	**Moderate**	**Low**	**Low**

Public transportation positioning data

- Using them as probes in traffic
- Sufficient road presence
- Most main roads have public transportation

Classification of studies that use positioning data

- Type of traffic: **urban** / highway / combined
- Use of periodic public transportation: **yes** / no
- Defines a traffic index: **yes** / no
- Makes future traffic predictions: yes / **no**
- Type of data analysis: real-time / historic / **both**

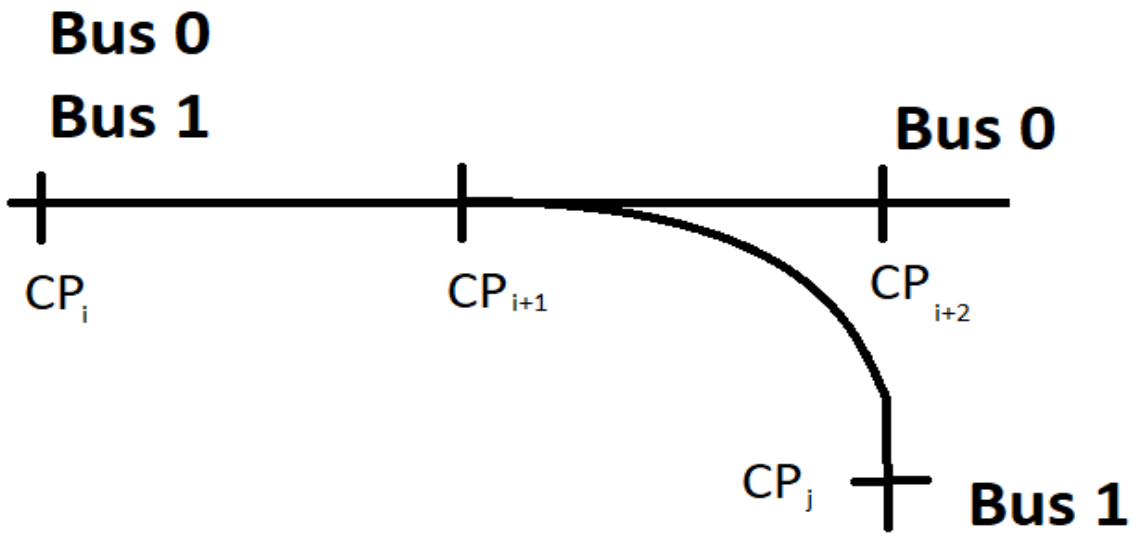
Goals

- Being able to measure traffic flow
  - For a road segment
  - For the whole city
- Classify the flow
- Visualize the results

Traffic index

Index value	Description
0	Free flow speed
1	Mild traffic
2	Mild congestion
3	Congestion

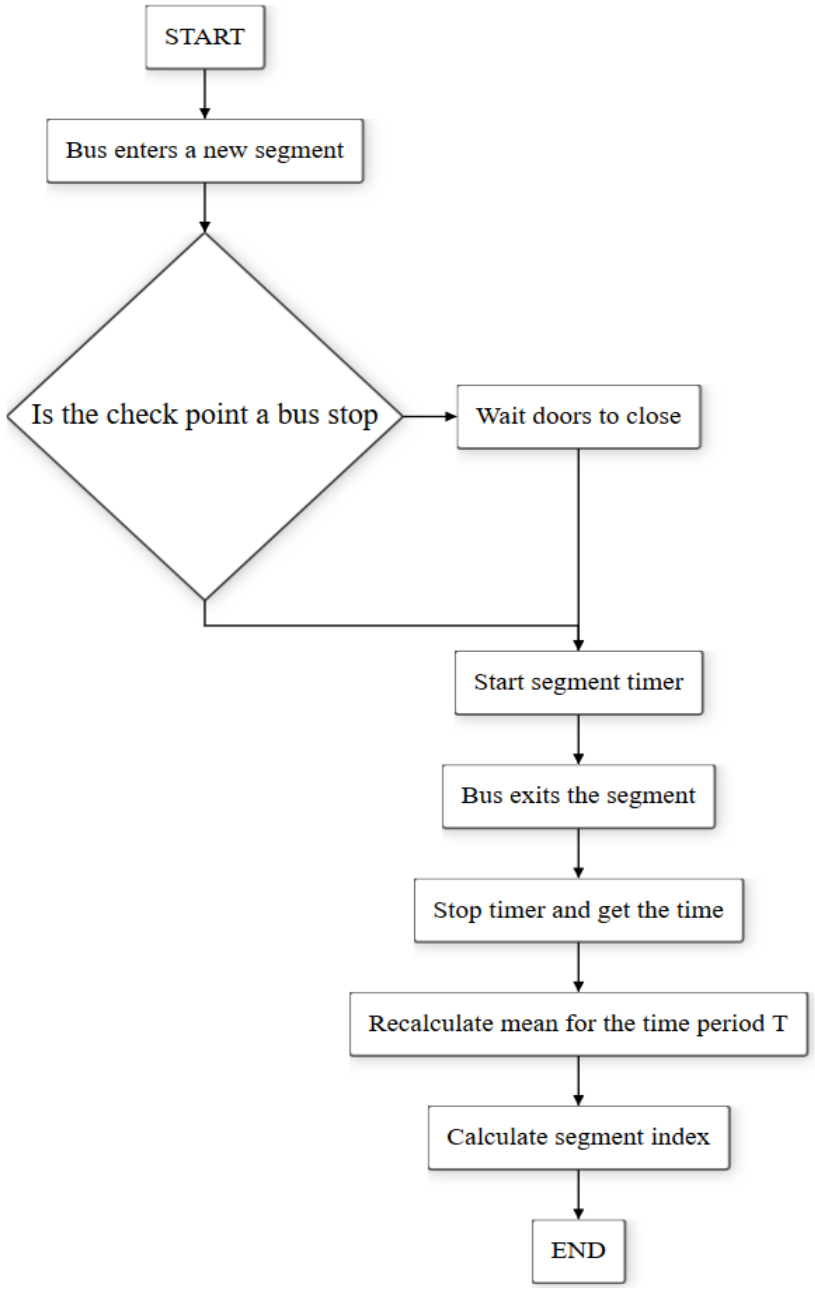
Checkpoints and Segments



Time intervals

- Divide the day to 30 minutes intervals
- On each vehicle pass we recalculate the arithmetic mean time for the segment
- When the time interval is over, we take the arithmetic mean time for it and store it in the historical dataset (HDS)
- We use this data to calculate the traffic index for that segment

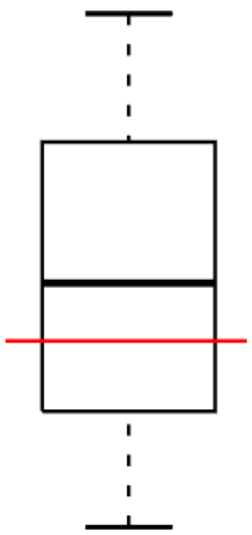
Time measuring algorithm

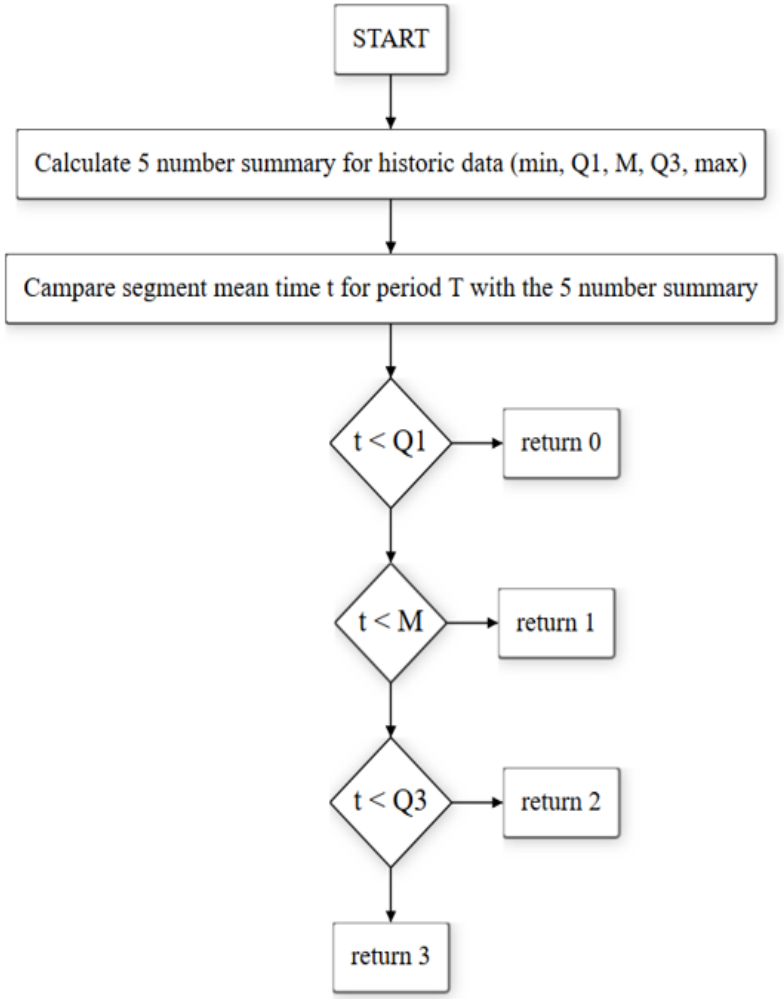


Historic dataset

- The historic dataset (HDS) consists of pods for each segment.
- In each pod we store the arithmetic mean times for each time interval, that has passed.
- We compare each segment only by itself. This is done due to the specific nature of each segment – how many pedestrian crosswalks there are, how busy they are, etc.
- When a new time interval begins, we create a new set for each segment. In that set we automatically add the arithmetic mean time for the previous interval in that segment.

Calculating the index for segment





City Index

$$CI = \sum_{i=0}^{n} \frac{\mathrm{MT_jS_i}}{n}$$

Limitations

- There should be public transportation vehicle routes that use the road
- There should not be a designated fast(bus) lane for the public transportation on the road
- The method assumes that each segment of the road has cycles of traffic – always busy, always free
- Another thing to consider is that for a smaller vehicle such as a car the segment times might differ

Ways to collect data

- GPS/GLONASS/Galileo
- Radio-frequency identification (RFID)

- Transmitting using 3G module

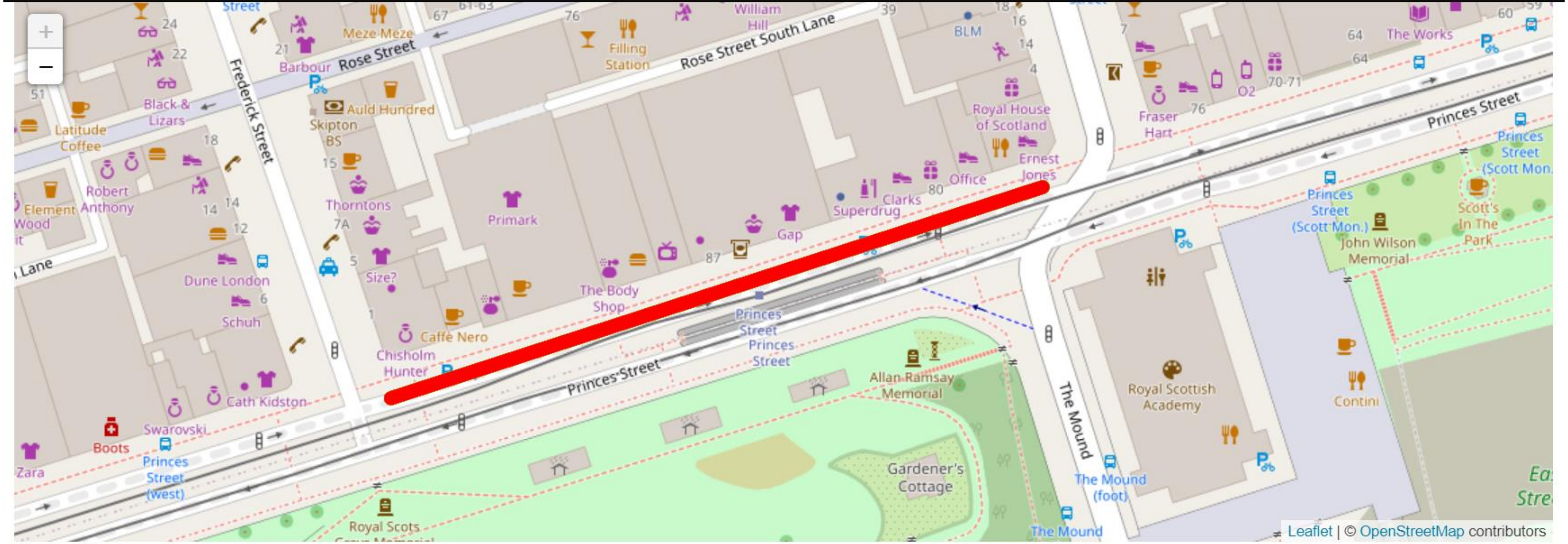
Data presentation

Index value	Description	Color
0	Free flow speed	
1	Mild traffic	
2	Mild congestion	
3	Congestion	
-	Not enough data / Unable to calculate	

Experimental setup and analysis

- GPS coordinates for Edinburgh, Scotland. The set contains data for total of 9 days
- Text files with vehicle id, timestamp, coordinates
- Street segment to run the algorithm was chosen to be in a central location with multiple bus lines passing thru it
- OpenStreetMap and Leaflet

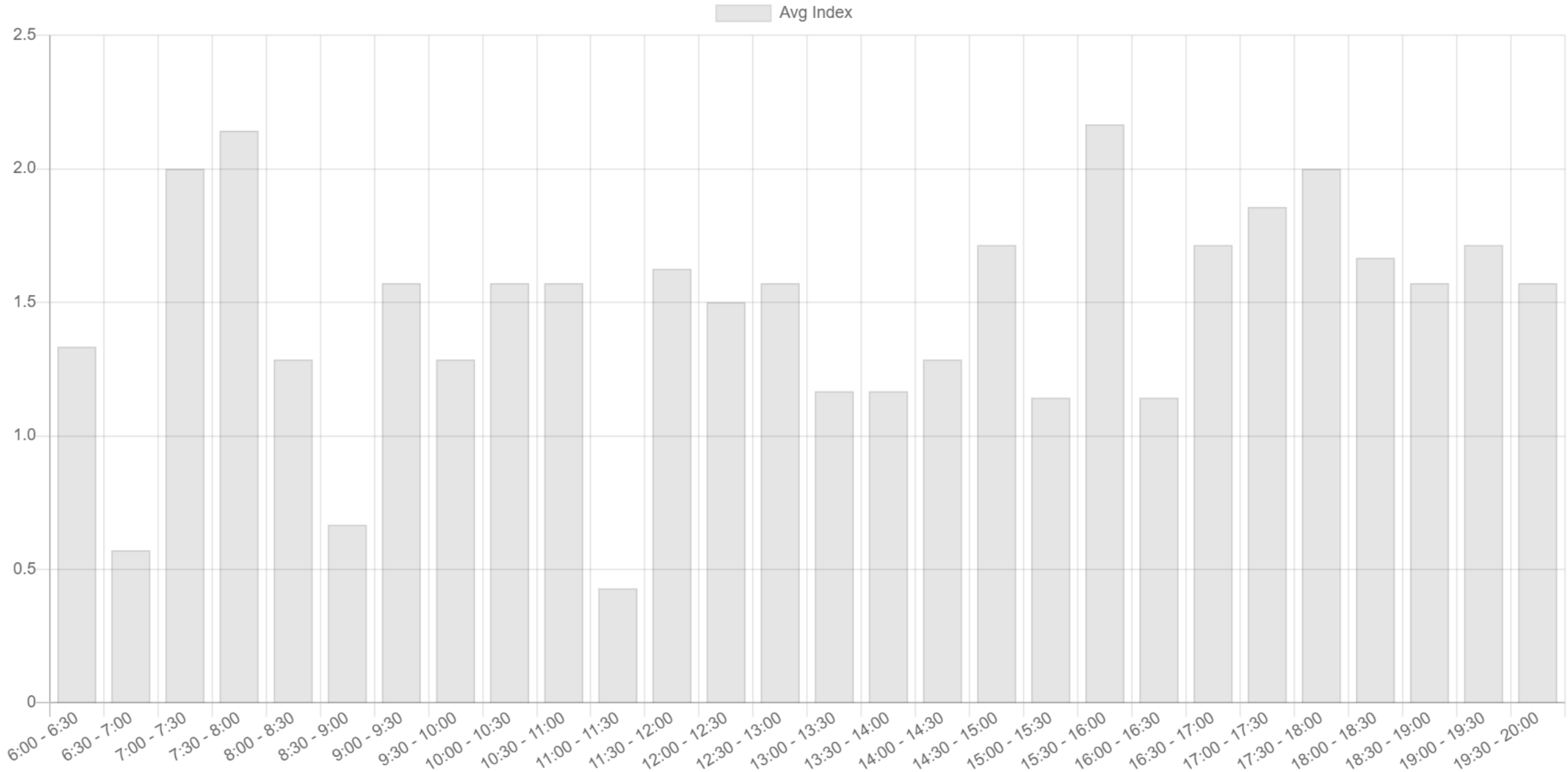
Data presentation



Choose time range:

2014-1-31 17:30 - 18:00

Index value: 3


Avg Index
2.5
2.0
1.5
1.0
0.5
0
6:00 - 6:30
6:30 - 7:00
7:00 - 7:30
7:30 - 8:00
8:00 - 8:30
8:30 - 9:00
9:00 - 9:30
9:30 - 10:00
10:00 - 10:30
10:30 - 11:00
11:00 - 11:30
11:30 - 12:00
12:00 - 12:30
12:30 - 13:00
13:00 - 13:30
13:30 - 14:00
14:00 - 14:30
14:30 - 15:00
15:00 - 15:30
15:30 - 16:00
16:00 - 16:30
16:30 - 17:00
17:00 - 17:30
17:30 - 18:00
18:00 - 18:30
18:30 - 19:00
19:00 - 19:30
19:30 - 20:00

Conclusions

- In this paper we present an algorithm for calculating the traffic flow index for a whole city (CI) as well as for a road segment, using public transportation vehicle data.
- This data can be used by the public administration to make infrastructure investment decisions, or for alternative path finding in real-time by routing algorithms.
- We also provide a mechanism for map visualization of the state of the city traffic both historically and in real-time.
- The model has been tested in the city of Edinburgh and despite limited experimental data yielding promising results.

Future work

- Predicting traffic congestions
- Ways to calculate roads that are not covered by public transportation using an association method
- Optimizing the calculations

ACKNOWLEDGMENTS

Research presented in this paper was partially supported by the FNI-SU-80-10-184/27.04.2020, FNI project of Sofia University “St. Kliment Ohridski” (Bulgaria) “Challenges of developing advanced software systems and tools for big data in cloud environment (DB2BD-3)” and project “An innovative software platform for big data learning and gaming analytics for an user-centric adaptation of technology enhanced learning (APTITUDE)” - research projects on the societal challenges – 2018 by Bulgarian National Science Fund with contract №: KP-06-OPR/1 from 13.12.2018.