# AI driven adaptive scheduling for on-demand transportation in smart cities.

Veneta Markovska[1] Margarita Ruseva[2] and Stanimir Kabaivanov[3][0000-0002-8686-8112]

[1] University of Food Technologies, Plovdiv Bulgaria
[2] Plovdiv University "Paisii Hilendarski", Plovdiv, Bulgaria
[3] Plovdiv University "Paisii Hilendarski", Plovdiv, Bulgaria
stanimir.kabaivanov@uni-plovdiv.bg

**Abstract.** Artificial intelligence algorithms can be used to automate and improve various processes in public transportation. Using a combination of data sources like positioning devices, ticketing and sales notifications and video surveillance we can obtain details on the load and utilization of transportation network segments. These results can be used not only to improve marketing campaigns and increase quality of transportation services, but also to provide better on-demand transportation options with flexible schedules. In this paper we discuss one such automation system that benefits from delay analysis and real-time processing of video streams.

**Keywords:** adaptive transportation schedules, AI in transportation, real-time video processing, Monte Carlo simulation.

## 1 Theoretical background

### 1.1 Transportation scheduling

In order to provide efficient schedule for public transportation vehicles, one needs to answer several crucial questions:

- What exactly is the reasonable set of criteria in order to consider a solution to be optimal?
- How often can we assume that restrictions and parameters used during our optimization procedures will remain unchanged, or at least close to the values we have observed?
- How feasible and execution-friendly will be the solution?

The list of questions can be extended to include also special needs and restrictions, depending on the city and area where optimization is done. As a result, it is possible that some technically feasible and theoretically sound solutions, do not meet all acceptance criteria and fall short on providing sufficiently adequate answers to all entries in the list.

We can greatly benefit from existing studies that focus on optimization with regard to common factors like available transfers ( [1], [2], [3]), use of new IoT devices ( [4],

[5]) and contemporary traffic management systems ( [6], [7]). Yet we believe there is a substantial potential in providing better decisions, due to the fact that focusing solely on the optimization techniques and used devices can lead to skipping a very important part of the issue – customer satisfaction and convenience of transportation network. If we consider that public transportation also provides a social and welfare support, it is evident that in addition to technical and purely economic criteria we should also consider this special case. There are several directions in which more traditional optimization techniques can be improved, and we use them to suggest an AI-powered solution:

- Combine different sources of information and decision criteria.

Use of various information sources imposes a technical difficulty as that would require to integrate several systems and to be able to handle distinct inputs and data types. But the gains may more that outweigh the efforts, due to the fact that use of events and notifications from independent sources can improve the accuracy and on-time delivery of crucial inputs (like for example traffic management and notification system and accident reporting can help avoid bottlenecks and choose alternative route whenever possible).

- Real-time processing of information.

For a public transportation system that operates on a fixed schedule (or one that is updated at sufficiently large periods – typically few months) real-time processing of information is still beneficial, although not crucial. As a result, such a system would exhibit regular delays of different magnitude. Even if schedules are updated frequently, timing issues will still be present due to irregular traffic jams, changes in road and weather conditions.

- Use of consumer behavior models and preferences.

Consumer-centric approach in transportation schedule optimization is a key factor for generating results that are not only utilizing resources efficiently but also help improve citizen experience and wellbeing. Due to the fact that public transportation system impacts also economic development of individual cities and regions, accounting for consumer behavior and preferences can also boost economic growth and improve long-term development prospects.

- Efficient dissemination of information on changes.

Changes that aim for better transportation schedules need to be communicated in an efficient and convenient way. Otherwise they will simply not be considered by end users and their positive impact will be reduced. While contemporary means of communication and IT systems can spread information quickly, automating this process and development of proper visual representation can significantly improve the overall use experience and benefits from transportation optimization efforts.

Artificial intelligence algorithms can support each of these transformations due to their ability to adapt to changing environment conditions and provide flexible solutions. These advantages come at a cost, that typically AI systems are harder to understand – this being especially true for deep learning algorithms, and even harder to certify ([8]). To overcome this problem, we suggest several easy-to-understand and visualize key performance indicators that allow to benchmark the results of the AI optimization.

## 2 Building a model

A high-level overview of the model is presented on **Fig. 1** and **Fig. 2**, where we have split the process into two steps:

- Data processing and fusion of different information sources (like traffic control and monitoring systems, ticketing).

The main function of this step is to combine information from different sources (vehicle location systems, ticketing sales and use, traffic control and video surveillance) into one stream of inputs. At the end of this step the system is able to provide intermediate estimates on the load of the public transportation (thus optimize the available in accordance with it) as well as short-term estimates on the demand.

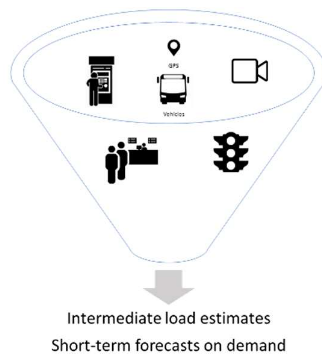


**Fig. 1.** Overview of data inputs and first stage of the AI-supported scheduling.

- Use of intermediate load estimates from the first processing steps to create dynamic schedules and monitor their execution.

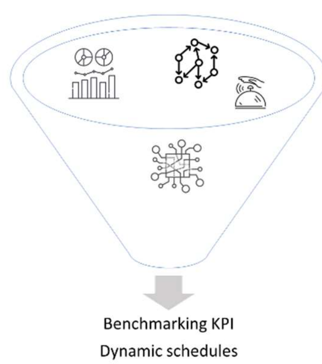


**Fig. 2.** Overview of the second state of the AI-supported scheduling.

Dynamic schedule generation and calculation of benchmarking KPIs is the second step of the model. Here artificial intelligence algorithms can support the process in speeding-up the optimization, accessing robustness ([9]) and simulating special conditions to facilitate proactive public transportation management.

By combining both stages, as demonstrated on **Fig. 3**, it is possible to create schedules that track user demand and traffic conditions. Unlike static allocation of resources, this makes it possible to react to changes in customer behavior, and when short-term forecasts are available to even pro-actively deploy available vehicles.

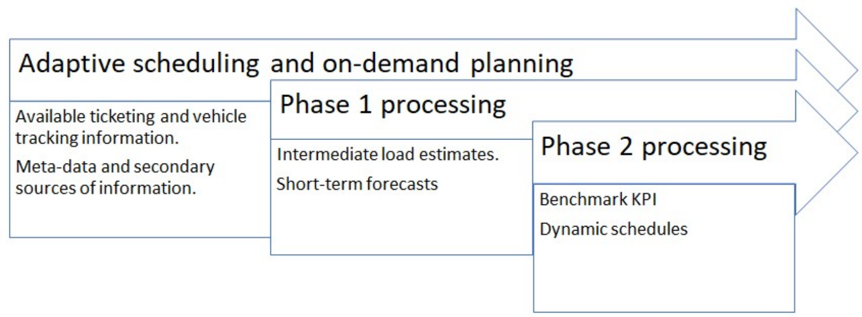


**Fig. 3.** Combination of different model stages to provide dynamic scheduling

Therefore, dynamic scheduling can help meet the demand with a smaller number of transportation vehicles. As a result, available funds can be re-allocated to acquiring more modern (thus environment-friendly) equipment and introduction of new means of transportation.

### 2.1 Traceability of passengers

As argued in [8], to be able to get the most of such systems it is necessary to extend the analysis of collected information. No matter how detailed the inputs are, if their processing is limited to simply calculating aggregated values (like for example means or deviations), the benefits will remain limited. To achieve a better segmentation and true understanding of public transportation needs, it is necessary to process in more intelligent way collected data and constantly monitor the changes in the results.

In our model, the ability to model and analyze consumer behavior and preferences is a key requirement for economically sound and efficient decisions. Regardless of the decision criteria, if the ultimate need is to achieve sustainable and usable transportation system then end user preferences should be studied at any given moment.

But to achieve traceability of passengers we have to define the fine balance between collecting data that can at the same time:

- Comply with regulations about privacy and data protection – e.g. does not violate the General Data Protection Regulation (GDPR) rules.
- Is easy to collect and does not require excessive costs for collecting and storing.

- Allows to identify groups of public transportation users that are of interest for the analysis.

While the list is quite specific and demanding, it is often the case that public transportation monitoring authorities and municipalities already have a lot of the necessary details. However, they are also often spread around and kept in different IT systems. As a result, the process of gaining traceability of passengers is actually the one of integrating different systems holding necessary data pieces.

Table 1 shows major sources of information as well as the inputs that could be used to retrieve them.

Table 1. Data types, sources and availability

| Data type and characteristics | Source / Availability |
|---|---|
| Routes and location of public transportation vehicles | Dynamic information with high frequency (often in the range of few seconds – 10 to 30) from GPS systems of the respective AVL devices. |
| Planned schedules and frequency of different types of public transport. Location of stops and links between different types of transport. | Information that is relatively constant and could be found in the respective schedules and municipality decisions/local legislation. |
| Preprocessed inputs which indicate the discrepancies between planned (scheduled) locations (resp. arrivals and departures) and actual times. | Depending on the feature set of the software system controlling ALVs such information may already be available, or may be easily calculated from the GPS inputs. |
| Ticketing and fare information | Summary of ticketing and fare information is typically available from the fare system database. |
| Passenger location and movement information | Such information cannot be obtained due to violation of GDPR and privacy protection. However mobile networks can give anonymized and aggregated information on movement of mobile devices to match against utilization of public transport network. |

To avoid violation of privacy regulations, all the information used should be anonymized and only aggregated (e.g. preprocessed) output should be made available to users of the traceability analysis.

### 2.2 Identification of fundamental needs

Traceability of passengers alone is not sufficient to select the best policies and deploy them to improve public transportation. As there are multiple ways to optimize the public transportation services, it is very important to define appropriate goals and translate them into optimization criteria. Our study has been inspired by Innovative demand responsive green public transportation for cleaner air in urban environment (INNOAIR) project, which aims to provide alternative transportation methods and improve air quality by reducing traffic pollution. Financing from EU Urban Innovative Actions initiative, makes it possible to test innovative and creative ways to address problems in modern cities.

Establishing on-demand transport, as one of the major targets for INNOAIR, needs to fit inside the existing public transportation framework. To guarantee this, goals must be in line with general regulations for the public transportation in Sofia, as well as to be measurable.

Metrics and goals presented in Table 2 are defined to be S.M.A.R.T. – e.g. they are specific, measurable, achievable, relevant and time-bound ([13]).

Table 2. Important goals and metrics used to identify needs for public transportation

| **Goals/Optimization criteria** | **Metrics involved** |
|---|---|
| Adaptive schedule that considers congestion and expected load (including traffic and number of passengers). | Time between consecutive trips (in hours/minutes) for different weekdays, weekends, and official holidays.<br>Delays and advances, compared to established schedules.<br>Load and number of tickets validated for a specific period. |
| Adaptive bus lanes, that consider passenger point of entry and destination. | Load and number of tickets validated for a specific period.<br>Time series of people entering or leaving a specific area. |
| Fixed lanes optimization and bus stop location optimization | Top destinations based on passenger location and movement over time (generalized and not individual data).<br>Top origin-destination pairs based on selectable scale (station, lane, residential area, city-level). |
| Traffic lights and special rules for public transportation vehicles | Delays and advances, compared to established schedules over time.<br>Time to stay on every stop and the number of people entering/leaving the vehicle.<br>Traffic light cyclogram data. |

Traffic data and bus location is available from the tracking systems that vehicles have and use to signal their location (as shown on **Fig. 4**). A typical application of GPS data is to check that all scheduled courses have been completed in time and without missing bus stops. While this information is very important, it does not reflect any patterns in passenger behavior. **Fig. 4** demonstrates how a usual report looks like and indicates that the information contained in it is purely technical.

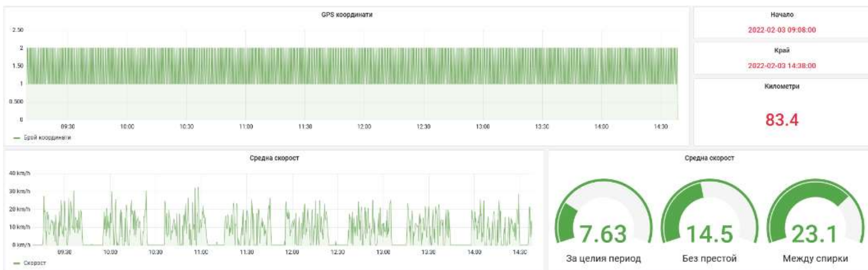


**Fig. 4.** Example technical output, showing distance travelled, GPS coordinate quality and average speed between different stops.

### 2.3 Measuring the impact of changes

Measuring changes in accessibility in specific areas can be useful for local planning or for promoting project-based coordination between different public agencies. Transportation projects, by themselves, cannot create denser, mixed-use, and active neighborhoods, but they can be catalysts for redevelopment and create conditions for improved economic development.

Due to the stochastic nature of process, the impacts on individual passengers vary, however in an aggregated way passenger mainly experience the following three effects:

1) Impacts on duration of travel time components, being in-vehicle time and waiting time, which lead to arriving early or late.
2) Impacts on passenger perception of the public transport mode depending on the variability of travel time components, being departure time, arrival time, in-vehicle time and waiting time, which lead to uncertainty of the actual travel time.
3) Impact on the probability of finding a seat and of crowding, affecting the level of comfort of the journey.

Reliability is important for operators and passengers alike. For operators, unreliable services cause difficulties in timetabling and resource planning. Also, unreliable services are typically more unevenly loaded, causing issues of passenger overloading and possible breaching of loading licenses.

For passengers, unreliable services cause adjustments in an individual's desired trip making behavior to account for the possibility of a service not operating 'as normal'. Variable departure times force the traveler to arrive earlier at the service and create uncertainty and anxiety about whether the service has arrived. Variable arrival times cause travelers to arrive at their destination late and force them to take an earlier service. In-vehicle time variability causes the traveler to experience uncertainty and anxiety about how long they will have to spend in the service.

Valuations of reliability can be estimated using revealed and stated preference data. However, most valuations are undertaken using stated preference techniques, where a survey asks respondents about hypothetical situations.

### 2.4 Benchmarking and key indicators

The first and most obvious performance area for public transport relates to the portion of travelers using the services. Although it is not a direct measure of the quality of a public transport system it is a definite indicator of its popularity or in some cases the patron's dependency on it for essential travel.

Performance measures/indicators are used in both performance measurement and benchmarking. The performance measures are normally a quantitative measure or index that numerically expresses a specific activity. In the context of this study, reference is made to key performance indicators (KPIs), as the aim is not to measure a complete set of performance measures, but rather focus on some key ones that will provide a sufficient understanding of relative comparison in the process. The challenge in defining KPIs is to select the appropriate ones that will give a sufficient understanding of overall performance. The KPIs should also be practical in terms of data availability and understandable to the audience. Useful KPIs can normally be associated with the S.M.A.R.T principle:

- Specific – A KPI must cover concisely one aspect of the activity.
- Measurable – KPIs must be quantifiable as subjective measures, e.g., a rating scale, could lead to distorted comparisons.
- Achievable – Available data and common items normally measured must be used for KPIs. It would not be useful to develop sophisticated KPIs for which data are unobtainable.
- Relevant – The KPI must be relevant to the activity being considered.
- Sometimes a different KPI is used to indicate or estimate a different activity. For example, one can use fuel consumption as a surrogate of CO2 emission if no actual emission data exist and Timebound – KPIs of similar timeframes need to be used to be an effective comparison tool for benchmarking.

The measures used are related to the factors affecting access, safety, efficiency, and affordability to public transportation system:

1) Comfort and safety - overall experience; safety; security; walking infrastructure; public transport infrastructure; operational performance; impact of traffic on walkability.
2) Service demand - daily trips.
3) Connecting destinations - access to public transport stops; access to jobs and services.
4) Support and encouragement – information, affordability; incentives.

## 3 Algorithm usage

Artificial intelligence algorithms can be applied in many different forms and contexts. Not all of them are suitable for managing on-demand transportation and dynamic scheduling, as some are too demanding to implement or too complicated to explain and calibrate on the available data. If we consider that on-demand transportation is often implemented as pilot projects with a limited scale, there is also little data available to train large deep learning systems. Therefore, we have decided to use the simplest possible solutions as a start, in order to simplify initial implementation and reduce the required amount of information for training.

Table 3. Sample application of intelligent algorithms to support data analysis

| **Algorithm** | **Usage information** |
|---|---|
| *Step 1* | |
| Hybrid neural network for anomaly detection | Based on research in [12], anomalies and outliers in data sets are detected to generate special events and warnings on unusual conditions, that can help in maintenance and provide for on-time reaction of potential issues. |
| Data profiling for high frequency inputs with deep learning neural network. | This pre-processing step helps to reduce the dimension of data and speed-up the analysis. |
| Clustering and classification of inputs on customer satisfaction with K-Means clustering with Mahalanobis distance ([13]). | This step helps to reduce significantly the amount of data stored for processing in the next step. Since classification is done with pre-defined groups, the only input passed on is the group identifier, instead of all initial consumer details. |
| *Step 2* | |
| Short term time series forecasting with neural networks | Multilayer neural networks provide good accuracy and efficiency when analyzing time series. Since necessary forecasts are limited in short-term, we can also benefit from avoiding issues like vanishing gradient problem. |
| Jump-diffusion process calibration with adaptive re-calibration in case of outlier detection. | We have used jump-diffusion processes to model bus delays ([14]), but when combined with anomaly detection, process parameters can be re-estimated in case of outlier detection or when enough arguments pointing toward environment conditions change are collected. |
| Geospatial analysis of on-demand requests identifies use frequency and patterns in requests. | In addition to request analysis, geospatial data can be used to cluster routes and improve efficient use of available transportation fleet. |

The list presented in Table 3 can be extended to include domain- and city-specific solutions, depending on requirements and problems to be solved.

Once the results of both steps are available, it is possible to continue analyzing the schedules under different simulated conditions. Depending on the context, different parameters can be modified in order to estimate how generated schedules will behave and under what scenarios they will become obsolete. Due to the modular and flexible structure of the suggested steps, it is possible to perform such simulations locally (for example modifying delays at single point – like individual bus stops) or globally, by introducing simultaneous delays or special events across the city area. The latter option is only limited by the available computational resources and time that can be spent on Monto Carlo simulations.

Local simulations on the other hand, can be applied to critical parts of the public transportation network. The main benefit of such restricted options is that they require much less computational power and can bring results faster.

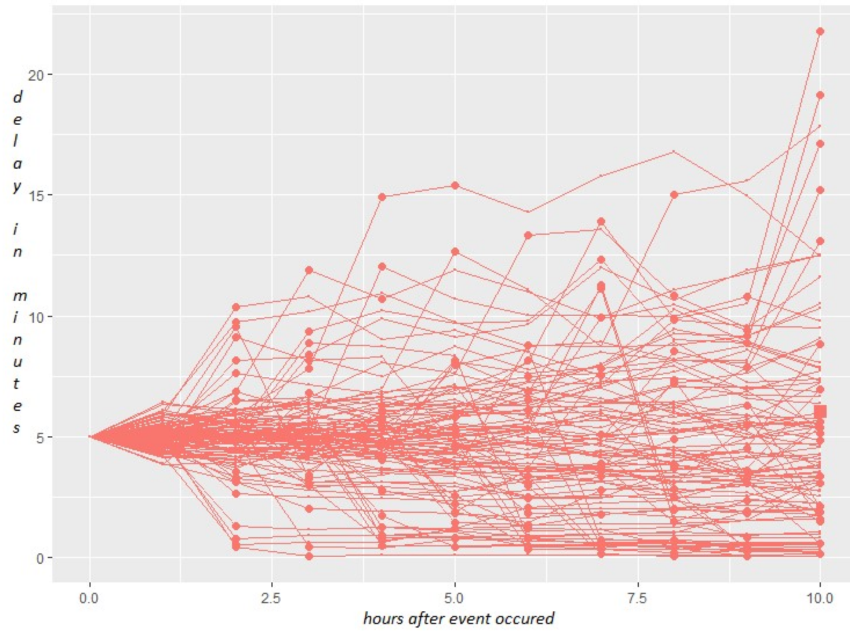


**Fig. 5.** Simulated delays with jumps in case of event that causes traffic jam

**Fig. 5** shows one such simulation, where a traffic jam event could cause different delays, because on reaction, number of vehicles and incident severity. In this particular scenario random behavior of delays have been assumed with potential for jumps, but other patterns can be applied as well. This provides additional flexibility and can improve proactive management of transportation networks.

## Conclusions

Transportation services are very important factor for driving economic development of individual cities or entire regions. They require careful planning, long-term investments and complex maintenance processes. Due to the fact that transportation services are important for both economic development and supporting specific social groups, there are various ways to analyze, manage and improve them. Based on experience gained in INNOAIR project, we have suggested a model where artificial intelligence algorithms can be used to improve existing transportation planning and scheduling algorithms, by considering different sources of information, real-time data processing and pay special attention to consumer preferences.

There are several key areas, where AI-supported data analysis can make a significant difference and add value to traditional transportation scheduling methods:

1) Pre-processing of visual records and information on traffic flows, jams and potential delays.
2) Analysis of passenger behavior and adaptive profiling of different customer groups, their needs and demands.
3) Delay analysis and proactive maintenance of fleet, assets and transportation support systems.

In combination with simple dashboards and carefully selected key indicators, it is also possible to overcome one of the significant drawbacks of AI-supported systems – their complexity and lack of transparency. People involved in the decision-making process do not need to dig into the details of the algorithms but only to make sure that performance indicators are adequate to the quality of service provided by the transportation system.

## Acknowledgement

This research was supported by UIA05-202 "INNOAIR - Innovative demand responsive green public transportation for cleaner air in urban environment", funded by the European Union initiative - Urban Innovative Actions. (UIA)."

## References

[1] A. Adamski, "Transfer optimization in public transport," in *Computer-Aided Transit Scheduling*, Berlin, Heidelberg, 1995.

[2] L. N. Jansen, M. B. Pdersen and O. A. Nielsen, "Minimizing passenger transfer times in public transport timetables," in *7th Conference of the Hong Kong Society for Transportation Studies, Transportation in the information age*, Hong Kong, 2002.

[3] A. Ceder, "Urban mobility and public transport: Future perspectives and review," *International Journal of Urban Sciences,* vol. 25, no. 4, pp. 455-479, 2021.

[4] P. Dharti, Z. Narmawala, S. Tanwar and S. P. Kumar, "A systematic review on scheduling public transport using IoT as tool," *Smart innovations in communication and computational sciences,* pp. 39-48, 2019.

[5] P. Kuppusamy, R. Kalpana and V. R. P. V., "Optimized traffic control and data processing using IoT," *Cluster Computing,* vol. 22, no. 1, pp. 2169-2178, 2019.

[6] N. Bhouri, F. J. Mayorano, P. A. Lotito, H. H. Salem and J. P. Lebacque, "Public transport priority for Multimodal urban traffic control," *Cybernetics And Information Technologies,* vol. 15, no. 5, p. 766, 2015.

[7] Y. Z. D. Wang, L. Hu, Y. Yang and L. H. Lee, "A data-driven and optimal bus scheduling model with time-dependent traffic and demand," *IEEE Transactions on Intelligent Transportation Systems,* vol. 18, no. 9, pp. 2443-2452, 2017.

[8] F. Tambon, G. Laberge, L. An, A. Nikanjam, P. S. Nouwou Mindom, Y. Pequignot, F. Khomh, G. Antoniol, E. Merlo and F. Laviolette, "How to certify machine learning based safety-critical systems? A systematic literature review," *Automated Software Engineering,* vol. 29, no. 2, pp. 1-74, 2022.

[9] M. Müller-Hannemann, R. Rückert, A. Schiewe and A. Schöbel, "Estimating the robustness of public transport schedules using machine learning," *Transportation Research Part C: Emerging Technologies,* vol. 137, no. 103566, 2022.

[10] C. Iliopoulou and K. Kepaptsoglou, "Combining ITS and optimization in public transportation planning: state of the art and future research paths," *European Transport Research Review,* vol. 11, no. 27, 2019.

[11] R. Bogue, "Use S.M.A.R.T. goals to launch management by objectives plan," TechRepublic, 2018.

[12] S. Kabaivanov and V. Markovska, "Hybrid deep-learning analysis for cyber anomaly detection," *IOP Conference Series Materials Science and Engineering,* vol. 878, no. 1, p. 012029, 2020.

[13] S. Kabaivanov, K. Roberts and S. Kovacheva, "Machine learning assisted social system analysis: Youth transitions in five south and east Mediterranean countries," *AIP Conference Proceedings,* vol. 2333, p. 030002 , 2021.

[14] S. Kabaivanov and V. Markovska, "Data driven public transportation delay modelling," INNOAIR Project, Sofia, 2021.

[15] Apache Software Foundation, "Apache NiFi," 11 11 2021. [Online]. Available: https://nifi.apache.org/. [Accessed 11 11 2021].

[16] influxdata, "InfluxDB Telegraf," 15 11 2021. [Online]. Available: https://www.influxdata.com/time-series-platform/telegraf/. [Accessed 15 11 2021].

[17] Apache Software Foundation, "Apache Airflow," 15 11 2021. [Online]. Available: https://airflow.apache.org/. [Accessed 15 11 2021].

[18] Apache Software Foundation, "Apache Spark MLib," 18 11 2021. [Online]. Available: https://spark.apache.org/mllib/. [Accessed 18 11 2021].

[19] Grafana Labs, "Grafana," 22 11 2021. [Online]. Available: https://grafana.com/. [Accessed 22 11 2021].

[20] G. Beirao and J. A. S. Cabral, "Understanding attitudes towards public transport and private car: A qualitative study," *Transport Policy,* vol. 14, no. 6, pp. 478-489, 2007.

[21] L. Steg, "Car use: lust and must. Instrumental, symbolic and affective motives for car use," *Transportation Research: Part A: Policy and Practice 39,* vol. 2, no. 3, pp. 147-162, 2005.

[22] M. R. Solomon, Consumer Behavior: Buying, Having, and Being, Pearson Prentice Hall, 2004.

[23] Y. Gaoa, S. Rasoulib, H. Timmermansb and Y. Wang, "Trip stage satisfaction of public transport users: A reference-based model incorporating trip attributes, perceived service quality, psychological disposition and difference tolerance," *Transportation Research Part A,* vol. 118, p. 773, 2018.

[24] J. De Vos, P. L. Mokhtarian, T. Schwanen, V. Van Acker and F. Witlox, "“Travel mode choice and travel satisfaction: bridging the gap between decision utility and experienced utility," *Transportation,* vol. 43, no. 5, p. 771–796, 2016.

[25] A. Carrel, R. G. Mishalani, R. Sengupta and J. L. Walker, "In pursuit of the happy transit rider: dissecting satisfaction using daily surveys and tracking data," *J. Intell. Transp. Syst.,* vol. 20, no. 4, p. 345–362, 2016.