URBANITE H2020 project
Algorithms and simulation techniques for decision - makers

Alina Machidon
alina.machidon@ijs.si
Jožef Stefan Institute
Jamova cesta 39
Ljubljana, Slovenia

Maj Smerkol
maj.smerkol@ijs.si
Jožef Stefan Institute
Jamova cesta 39
Ljubljana, Slovenia

Matjaž Gams
matjaz.gams@ijs.si
Jožef Stefan Institute
Jamova cesta 39
Ljubljana, Slovenia

ABSTRACT

URBANITE (Supporting the decision-making in URBAN transformation with the use of dIsruptive TEchnologies) is a H2020 project with the goal to provide an ecosystem model that articulates the expectations, trust and attitude from civil servants, citizens and other stakeholders in the use of disruptive technologies. This model will be supported with the provision of a data management platform and algorithms for data – driven decision – making in the field of urban transformation. One of the main output of the project will be a Decision-Support System including (AI based) predictive algorithms and simulation models for mobility that support the decision–making process by analyzing the current situation, the trends that occurred in a certain time frame and allowing to predict future situations, when changing one or more variables. URBANITE will analyze the impact, trust and attitudes of civil servants, citizens and other stakeholders with respect to the integration of disruptive technologies such as Artificial Intelligence (AI), Decision Support Systems (DSS), big data analytics and predictive algorithms in a data–driven decision-making process. The results of the project will be validated in four real use cases: Amsterdam, Bilbao, Helsinki and Messina. This paper overviews the current state of the project's progress.

KEYWORDS

AI, Big Data, DSS, disruptive technologies, URBANITE project

1 INTRODUCTION

In recent times, the cities and urban environments are facing a revolution in urban mobility, bringing up unforeseen consequences that public administrations need to manage. It is in this new context that public administrations and policy makers need means to help them understand this new scenario, supporting them in making policy–related decisions and predicting eventualities. The traditional technological solutions are no longer valid for this situation and therefore, disruptive technologies such as big data analytics, predictive algorithms as well as decision support systems profiting from artificial intelligence techniques to support policy – makers come into place.

The main technical objective of the URBANITE project is the development of advanced AI algorithms for analysis of big data on mobility. The developed methods and tools will provide substantial support for policy-makers to tackle complex policy problems on the mobility domain and will enable their validation on case-specific models. The goal of the activities will be to implement novel tools and services in order to enable policy-makers to use advanced data analysis and machine learning methods during the design of novel policies for a specific city

URBANITE will allow the analysis of the traffic flows that are currently happening and have happened up until that moment. In addition to the visualization of the traffic, usage of economy sharing vehicles and other aspects, URBANITE will analyse which are the bottlenecks and critical points, based on a set of parameters to be determined by the civil servants. Due to the fact that historic data is stored, trends can be determined by URBANITE by big data algorithms. These trend analyses can entail the understanding of, for instance, the use of a certain transportation system (e.g. bikes) in a certain neighbourhood of the municipality, or the peak hours in which a street is blocked. URBANITE will also provide means to simulate the effect of different situations such as opening a pedestrian street at certain times, location of electric charging stations, or bike sharing points through the implementation of artificial intelligence algorithms. To achieve that, URBANITE will build first generic models from the data across all the cities and then provide adaptation mechanisms to apply these models to the different use cases. From the data available, URBANITE will extract and formalize knowledge and then, through a combination of classification, regression, clustering, and frequent pattern mining algorithms, conclude into some decisions and actionable models that will enable city policy-makers to simulate and assess the outcomes and implications of new policies.

2 SYSTEM'S ARCHITECTURE

The URBANITE project will combine various data sources, algorithms, libraries and tools that provide the best solutions to the scope of the project. The technical "core" of the project has to fulfill the following objectives:

- Deploy tools for big data exploration with the active involvement of policy-makers.
- Design methods for the detection of important events that need to be addressed.

In order to provide the desired functionalities, several state-of-the-art technologies are currently examined and tested in order to be adapted, customized and integrated into the platform. A simplified preliminary architecture is presented in Figure 1.

2.1 Data analysis module

One of the first tasks involves the development of various methods for exploratory data analysis and user interaction. Multimodal methods, tools and services for big data on urban mobility will be implemented that will provide exploratory analysis capabilities and enable the policy-makers to actively search for causal relations in the data will be provided by the platform.

Information Society 2020, 5–9 October, 2020, Ljubljana, Slovenia

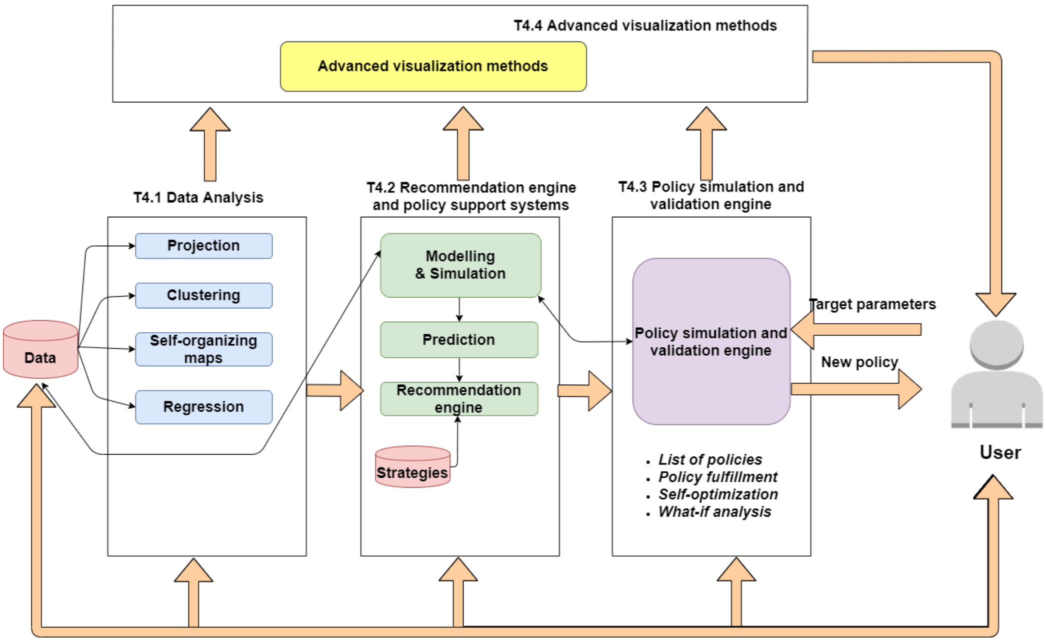

Figure 1: High level architecture of the URBANITE platform.

The methods to be included in the platform can be segmented in four main groups:

- clustering, where the main goal is to reduce the amount of data by grouping together similar instances. The implemented method will provide mechanisms to group instances based on GIS data or any subset of attributes that users will define. For example, platform users might choose to cluster all instances based on the type of transportation used (shared bikes, electric cars, etc.)
- projection methods that will be used to reduce the dimensionality of the data items. The goal of these methods is to represent the data in a lower dimensional space in such a way that the key relations of the data structures are preserved. The results of the methods can be used to more clearly visualize the data or use the transformed data in the next rounds of analysis
- self-organizing map involves the use of a type of artificial neural network, trained in an unsupervised manner. The method can at the same time reduce the amount of data (similar to clustering) and nonlinearly projects the data into lower dimensionalities
- prediction/regression methods, or classification models, that will allow to exploit the data

2.2 Recommendation engine

Recommendation engines (also known as recommender systems) are information filtering systems that deal with the problem of information overload [5] by filtering key information "chunks" out of large amount of dynamically generated information according to user's preferences, interest, or observed behavior about item [6][4]. Recommendation engines have the ability to predict whether a particular user would prefer an item or not based on the user's profile [4]. Recommendation engine is defined as a decision making strategy for users under complex information environments [3]. Recently, various approaches for building recommendation engines were developed, based on either collaborative filtering, content-based filtering or hybrid filtering [10], [9], [7].

The URBANITE recommendation engine will identify and predict important or problematic events related to mobility and will provide suggestions to tackle the issue. The policy support system will provide support to the policy-makers for identifying possible policies that tackle events based on specific criteria. The inputs will have to be aggregated for effective decision-making using hierarchical multi-criteria decision models.

2.3 Policy simulation and validation engine

Simulation transparency is a vital feature of the decision making process when quantitative computer tools are used to justify some strategies [8]. Simulation predictions can play a catalytic role in the development of public policies, in the elaboration of safety procedures, and in establishing legal liability. Hence, given the impact that modelling and simulation predictions are known to have, the credibility of the computational results is of crucial importance to engineering designers and managers but also to public servants, and to all citizens affected by the decisions that are based on these predictions [8].

To create trust and increase the model's credibility and the simulation results delivered, it is crucial to deal with a validation strategy in which non-simulation-trained end-users could feel comfortable and trust the simulation model [8].

In the URBANITE project, the policy simulation and validation module will provide methods and tools to simulate the efficiency of specific policies in the target domain. Given a new policy, urban mobility model and the target parameters, the system can evaluate the performance of the new policy based on the observed parameters. The implementation of credible traffic simulations for the entire city has been addressed by various project; however, it is not yet adequately solved, due to its complexity. In URBANITE ,the constructed model will be used to predict and classify traffic flow changes based on the provided changes in the new policies. Policy-makers will select the defined KPI's that need to be evaluated by the validation engine and based on the scores the new policies achieve, policy-makers will be able to make an informed decision about which policies should be deployed in the city.

2.4 Advanced visualization methods

Another important task will be the implementation of advanced visualizations for mobility patterns, highlighting important events, and results of policy validations. The main visualization functionalities will present the information on a combination of map layers, describing where in the city specific events or a sequence of events occurred. Visualizations will involve the use of heat maps, traffic flow graphics, and other transportation clusters. Users will be able to change and interact with the visualization parameters. For example, select specific time ranges, zoom, highlight, display additional information, etc. Considering the variety and characteristics of the data, one concern is regarding the depicting multidimensional data in a human-perceivable manner. Several graphical methods are customarily used for a preliminary analysis of generic multivariate datasets [1]:

- Scatter plots
- Pie charts and bar plots
- Histograms
- Box plots
- Violin and bean plots
- Spider/radar/star/polar plots
- Biplots
- Conditioning plots/coplots
- Glyph plots
- Mosaic and spine plots
- Treemaps

Traffic datasets are generally high-dimensional or spatial-temporal [2], thus visualizing traffic data mostly employs information visualization and visual analytics.

Traffic data contain multiple variables, of which the most important ones are time and space. Several different types of visualisation are currently used for traffic data, among them:

- visualization of time
- visualization of spatial properties
- spatio-temporal visualization

Location is the main spatial property of traffic data. Based on the aggregation level of location information, visualization of spatial properties can be categorized into three classes: point-based visualization (no aggregation) , line-based visualization (first-order aggregation), and region-based visualization (second-order aggregation) [2].

Heatmaps are the most used visualisation tools to show the integrated quantity of a large scale of objects in a map.

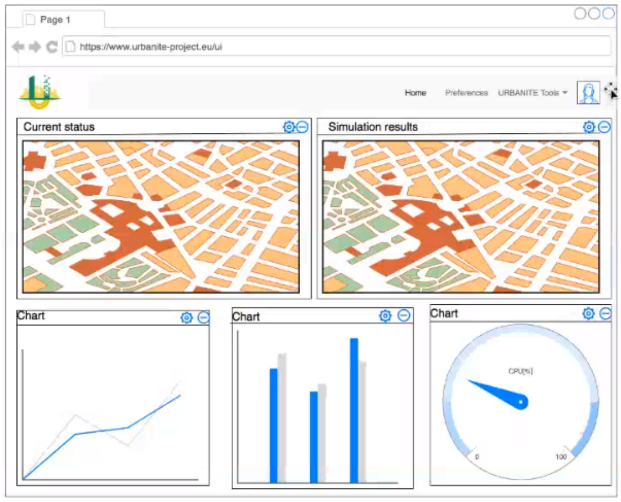

Figure 2: User interface mock up of the URBANITE platform.

3 DATA SOURCES

There are several collection procedures of the traffic related data and they range from sensor readings to airborne imagery and social media data [11]. The involvement of the municipalities of Bilbao, Helsinki, Amsterdam and Messina will provide a wide range of data sources related to the urban mobility, along with the public, open-source ones.

Several types of data sources were identified for the URBANITE project:

- geospatial data, e.g. maps (Open Street Maps [1], but also proprietary maps of the cities)
- additional info such as: car and lorry registration, information on parking lots, dynamic parking data, cadastre information, commercial register, care services, tourism accommodation
- demographics: statistical information on the number of inhabitants of different city districts, the number of households, population's age brackets, city boundaries, etc.
- public transportation: tram and metro lines, static and dynamic information about the public bus transport service, the GPS position of the buses
- traffic data: the count of car traffic and speeds, traffic status in real time, vehicle counts on the ring roads, etc.
- bicycle information: bike counters, bicycle collection points, calculated number of bikes in specific road segments, CityBikes [2]
- pedestrian: manual counts of pedestrians
- electric charging stations
- taxi stops available
- harbour transport data, ferry traffic statistics
- geographic airport information
- air quality (OpenAQ [3])
- noise maps
- wheather data (OpenWeatherMap [4]))

[1]https://www.openstreetmap.org/
[2]https://api.citybik.es/v2/
[3]https://openaq.org/
[4]https://openweathermap.org/

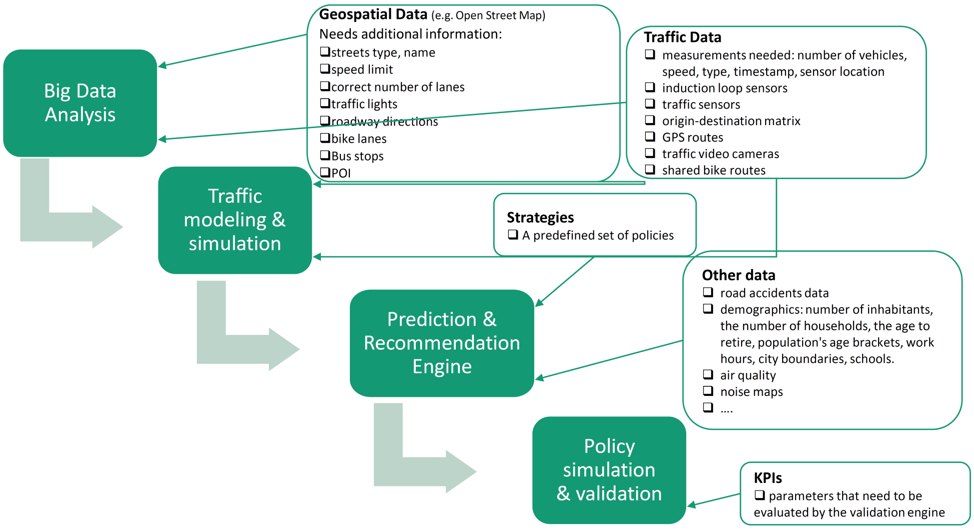

Figure 3: Data sources for the URBANITE platform.

The format of this datasets varies from JSON, XML, CSV, XLSX, WMS , GEOJSO or GML. The main issue with the mobility related data sources it is related to the high level of heterogeneity, both in terms of data format and data availability. Most of the cities involved on the project have some data related to the traffic in city, for example, but the format of the data, the level of granularity (how often is the data updated) and the availability of historical data (for how long does the city stores their data) varies greatly from one case to another.

4 CONCLUSIONS

The technical core in the URBANITE project focuses on the development of advanced AI algorithms for analysis of big data on mobility. The developed methods and tools will provide substantial support for policy-makers to tackle complex policy problems on the mobility domain and will enable their validation on case-specific models. The goal of the activities is to implement novel tools and services in order to enable policy-makers to use advanced data analysis and machine learning methods during the design of novel policies for a specific city.

One underlining factor in URBANITE is the adaptation of everything that it is produced to civil servants, citizens and interesting parties that may or not be digitally literate. The use of big data techniques and artificial intelligence algorithms, up till now, is not a common skill among public servants and this is one of the reasons the data analysis processes and user interaction mechanisms described in this work are developed with the abilities of the non-experts in mind too.

ACKNOWLEDGMENTS

This paper is supported by European Union's Horizon 2020 Research and Innovation Programme, URBANITE project under Grant Agreement No.870338.

REFERENCES

[1] Sunith Bandaru, Amos HC Ng, and Kalyanmoy Deb. 2017. Data mining methods for knowledge discovery in multi-objective optimization: part a-survey. *Expert Systems with Applications*, 70, 139–159.

[2] Wei Chen, Fangzhou Guo, and Fei-Yue Wang. 2015. A survey of traffic data visualization. *IEEE Transactions on Intelligent Transportation Systems*, 16, 6, 2970–2984.

[3] Ricardo Colomo-Palacios, Francisco José García-Peñalvo, Vladimir Stantchev, and Sanjay Misra. 2017. Towards a social and context-aware mobile recommendation system for tourism. *Pervasive and Mobile Computing*, 38, 505–515.

[4] F.O. Isinkaye, Y.O. Folajimi, and B.A. Ojokoh. 2015. Recommendation systems: principles, methods and evaluation. *Egyptian Informatics Journal*, 16, 3, 261 –273. ISSN: 1110-8665. DOI: https://doi.org/10.1016/j.eij.2015.06.005.

[5] Joseph A Konstan and John Riedl. 2012. Recommender systems: from algorithms to user experience. *User modeling and user-adapted interaction*, 22, 1-2, 101–123.

[6] Chenguang Pan and Wenxin Li. 2010. Research paper recommendation with topic analysis. In *2010 International Conference On Computer Design and Applications*. Volume 4. IEEE, V4–264.

[7] Nymphia Pereira and Satishkumar L Varma. 2019. Financial planning recommendation system using content-based collaborative and demographic filtering. In *Smart Innovations in Communication and Computational Sciences*. Springer, 141–151.

[8] Miquel Angel Piera, Roman Buil, and Egils Ginters. 2013. Validation of agent-based urban policy models by means of state space analysis. In *2013 8th EUROSIM Congress on Modelling and Simulation*. IEEE, 403–408.

[9] Tomasz Rutkowski, Jakub Romanowski, Piotr Woldan, Paweł Staszewski, Radosław Nielek, and Leszek Rutkowski. 2018. A content-based recommendation system using neuro-fuzzy approach. In *2018 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*. IEEE, 1–8.

[10] Diego Sánchez-Moreno, Ana B Gil González, M Dolores Muñoz Vicente, Vivian F López Batista, and María N Moreno García. 2016. A collaborative filtering method for music recommendation using playing coefficients for artists and users. *Expert Systems with Applications*, 66, 234–244.

[11] G. Zhou, J. Lu, C.-Y. Wan, M. D. Yarvis, and J. A. Stankovic. 2008. *Body Sensor Networks*. MIT Press, Cambridge, MA.