OPTIMUM Project: Geospatial data analysis for sustainable mobility

Luka Bradeško*
Jožef Stefan Institute
Ljubljana, Slovenia
luka.bradesko@ijs.si

Zala Herga
Jožef Stefan Institute
Jožef Stefan Postgraduate School
Ljubljana, Slovenia
zala.herga@ijs.si

Matej Senožetnik
Jožef Stefan Institute
Jožef Stefan Postgraduate School
Ljubljana, Slovenia
matej.senozetnik@ijs.si

Tine Šubic
Jožef Stefan Institute
Ljubljana, Slovenia
tine.subic@ijs.si

Jasna Urbančič
Jožef Stefan Institute
Ljubljana, Slovenia
jasna.urbancic@ijs.si

CCS CONCEPTS

• **Computing methodologies → Machine learning approaches**;
• **Applied computing** → *Transportation*;

KEYWORDS

geospatial data analysis, machine learning, sustainable mobility

ACM Reference Format:
Luka Bradeško, Zala Herga, Matej Senožetnik, Tine Šubic, and Jasna Urbančič. 2018. OPTIMUM Project: Geospatial data analysis for sustainable mobility. In *Proceedings of KDD 2018 Conference (KDD2018).* ACM, New York, NY, USA, Article 4, 2 pages. https://doi.org/10.475/123_4

1 INTRODUCTION

With the increase of sensor devices used in everyday life (phones, watches, cars etc.), there are some low hanging fruits in terms of "simply" collecting and analyzing the data, and using the results towards a practical goal. This applies to health, marketing, traffic, urbanism, ecology and other domains. Project OPTIMUM aims to embrace this potential in the traffic domain in order to push participants towards a more sustainable mobility. To this end, a platform integrating and aligning multiple data sources was established, enabling innovative data mining methods for solving specific predefined problems. The main innovation that we are going to present is NextPin, an online service for extracting mobility patterns from geo-spatial data.

2 PROJECT DESCRIPTION

OPTIMUM project[1] strives to improve transit, freight transportation and traffic connectivity throughout Europe. It uses scalable architecture for management and processing data from various sensors, systems, crowdsourcing and service providers to enable proactive decisions and actions in semi-automated way. Project benefits from big data fusion in order to design a distributed architecture that supports observations of multiple levels of the transport system in real time. Additionally, it acknowledges sustainable transportation behaviors through system-aware optimization that integrates adaptive charging models and real-time multi modal routing algorithms. The latter are supported by personalization to encourage proactive decisions on individual user or group levels. Finally, there are three real-life use cases to test the proposed solutions.

Use cases for pilot testing include:

(1) *Dynamic charging pilot*, where partners Infraestruturas de Portugal SA (national authority in Portugal, operates and maintains road network) and Luis Simões SA (freight company in Portugal) use model for calculating variable toll fees to motivate a shift of traffic from the overused toll-free road infrastructure to the underused toll infrastructure to decrease travel times and safety risks;
(2) *Car2X pilot*, where Adria Mobil (the third largest European caravan, motorhome and leisure vehicle manufacturing company from Slovenia) provides intelligent assistance to deal with the complexities of multi-sensor environments in motorhomes for long-distance and cross-border trips;
(3) *Multimodal routing application pilot* is deployed in 3 European cities, where local partners, including Ljubljana Public Transport and Birmingham City Council, use mobile phone application[2] developed within the consortium to enable trip optimisation, and to influence travel choices through recommendations and reward crediting scheme.

Technical partners, namely Austrian Institute of Technology GmbH, Fluidtime Data Services GmbH (provider of software solutions for mobility information and traffic data management from Austria), Institute for the Development of New Technologies UNINOVA (Portugal), Institute of Communications and Computer Systems (Greece), INTRASOFT International SA (European IT solutions and services group), Jožef Stefan Institute (Slovenia), Kapsch TrafficCom AG (provider of intelligent transportation systems from Austria), Nissatech Innovation Centre (software developer from

*Work done while working at Jožef Stefan Institute
[1]Funded from the EU H2020, GA No 636160-2 (http://www.optimumproject.eu/)

[2]Available on Google play as OPTIMUM Intelligent Mobility

KDD2018, August 2018, London, United Kingdom

ACM ISBN 123-4567-24-567/08/06.
https://doi.org/10.475/123_4

Serbia), Regional Environmental Center for Central and Eastern Europe (nonprofit organisation from Hungary), University of the Aegean (Greece) and University of Wolverhampton (United Kingdom), developed the architecture and support pilot partners during the testing.

3 PRESENTED INNOVATION

In Multi modal routing application pilot we utilized a reward crediting scheme to persuade users towards more sustainable travel choices. Specifically, individual user receives reward points each time they choose to walk, cycle or use public transport to their destination. The longer the duration of these trips, the more points they collect. At the end of the pilot trial each user will get a reward (credit, vouchers or free bus tickets), which depends on the number of collected points - the larger the number of collected points, the higher the reward.

To support this use case we used Mobility patterns detection and prediction service (NextPin[3]) and the associated mobile phone library, both developed at Jožef Stefan Institute. Mobile phone library transmits location and sensor data to the NextPin server and can be either integrated in any mobile application or used as a standalone one[4]. NextPin provides analysis of visited locations and routes on individual, group and global (all users) level. At its core, the service utilizes the Staypoint Detection Algorithm [1] (SPD), which is a clustering approach applied to geospatial data. The input to the algorithm is a stream of coordinates, which are clustered into *staypoints* (locations) and *paths* (GPS trajectories between two consecutive staypoints). Here, we focus on stream of data collected from mobile phones, but it can be also collected from other sources (e.g. company vehicles) as long as the data contains timestamp, latitude and longitude. A staypoint will be detected if the coordinate stream stays at a location limited by some radius for long enough time; both radius and time thresholds are adjustable parameters of the algorithm. There is a trade-off when setting both parameters: radius needs to be big enough to be insensitive to the noise in collected GPS data, but if it's too big the algorithm might not reliably differentiate between two locations that are nearby. Similarly, if the threshold for time that needs to be spent at location is too low, SPD might detect e.g. stop at a traffic light, but if it's too high it might miss completely an actual quick stop.

Staypoints are time dependent, meaning that two visits of the same location at different times map to two different staypoints. Additionally, we cluster staypoints that refer to the same location into *global locations* which are independent of time. Address (obtained from Open Street Maps[5]) and name (obtained from Foursquare API[6]) are attached to each global location. Each user can also add personalized names (e.g. home, work) to their locations; personalized naming data is stored separately and is not visible to other users. Analogous to global locations, we defined *routes* as GPS trajectories between two global locations.

NextPin can also make next location predictions for individual user based on the historical data. For that purpose it implements a first order Markov Chain model. More advanced methods with

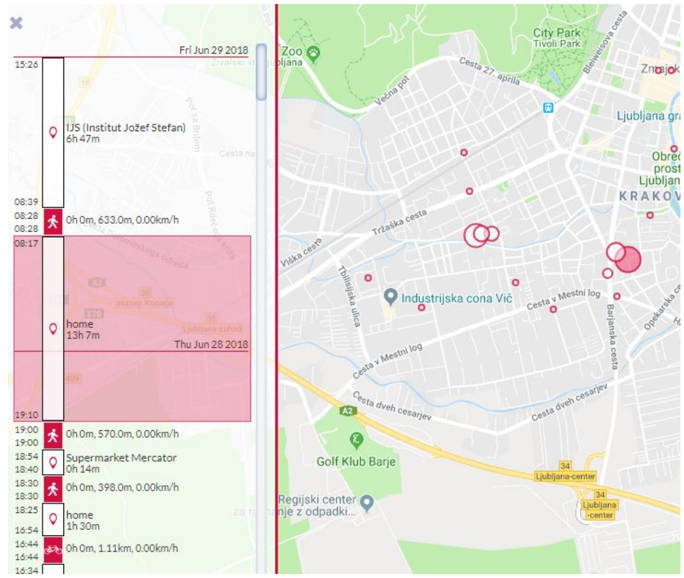

Figure 1: NextPin: user's timeline on the left side and global locations displayed on map.

arrival profiles and Monte Carlo simulations were also explored [2]. The results (recall and F1 measure) highlight the increased prediction power of these methods. However, their computational complexity is too high to include them in real-time analytics.

For the reward crediting scheme use case, we trained an online travel mode detection model which supplements the predictions and place/route detections with travel mode data. Currently supported travel modes are walking, running, cycling, traveling by car, traveling by bus and traveling by train. To detect non-motorized means of transportation on mobile phones, native OS APIs are used. When these APIs detect traveling in a vehicle, user's phone sends a short accelerometer sample to the NextPin server. The accelerometer sample is pre-processed to extract features for classification. Three one-against-all SVM classifiers with RBF kernel are used to decide whether the user is traveling by car, bus, or train.

In our demonstration we will show the mobile phone application and NextPin web interface with analysis tools it provides. We're going to present the technical challenges encountered while developing the system and the approaches to overcome them.

REFERENCES

[1] Jong Hee Kang, William Welbourne, Benjamin Stewart, and Gaetano Borriello. 2004. Extracting places from traces of locations. In *Proceedings of the 2nd ACM international workshop on Wireless mobile applications and services on WLAN hotspots*. ACM, 110–118.

[2] Blaž Kažič, Jan Rupnik, Primož Škraba, Luka Bradeško, and Dunja Mladenič. 2017. Predicting UsersâĂŹ Mobility Using Monte Carlo Simulations. *IEEE Access* 5 (2017), 27400–27420.

[3] http://traffic.ijs.si/NextPin/?user=zalah

[4] Available on Google Play as Mobility Patterns

[5] https://www.openstreetmap.org

[6] https://developer.foursquare.com/