

# TTASC: Transport and Traffic Analytics in Smart Cities

## Special Session at DATA ANALYTICS 2016, Vencie, Oct 9-13, 2016

<http://www.iaria.org/conferences2016/DATAANALYTICS16.html>

Ivana Semanjski

Department of Telecommunications and Information Processing/Department of Intelligent Transportation Systems  
Ghent University/University of Zagreb  
Ghent, Belgium/Zagreb, Croatia  
e-mail: [ivana.semanjski@ugent.be](mailto:ivana.semanjski@ugent.be)

**Abstract**—Vast generation of high resolution spatial and temporal data, particularly in urban settings, started revolution in mobility and human behavior related research. However, after initial wave of first data oriented insights their integration into ongoing, and traditionally used, planning and decision making processes seems to be hindered by still opened challenges. These challenges suggest need for stronger integration between data analytics and dedicated domain knowledge. Special session on Transport and Traffic Analytics in Smart Cities tackles these challenges from transport planners' point of view. Collection of papers aims at identifying the existing gaps and bridging between related disciplines with aspiration to foster faster integration of data driven insights into smart cities' dedicated planning.

**Keywords**—data analytics; smart cities; big data; data-driven insights; transport; traffic; mobility.

### I. INTRODUCTION

Each of us has become a walking data generator as mobile phones, social networks, package tracking, bank card payments, Global Navigation Satellite Systems (GNSS), public transport ticketing etc. produce torrents of data as a by-product of their operations, generating daily a bit more than 1GB of content per person. This data production seems to be doubling every couple of years as more data cross the internet every second today than were stored in the entire internet just two decades ago [1]. However, increase of volume (data scale becomes increasingly big), variety (data comes in structured, semi-structured and unstructured forms) and velocity (swift generation) seem to be overshadowed by data value (huge information value with very low density concealed among data) as a main characteristic that differentiates between simply large data amounts and big data itself. For this reason, data collection-processing chain needs to be promptly and timely conducted to maximally utilize the potential value of big data. Scientific literature on this topic is exhibiting exponential growth over the last five years [2] as gleaning intelligence from data brings forward pioneering results. Applicability of these findings is evident as companies that have, at early stage, recognized and

invested in this potential demonstrate a strong role that data driven insights have in gaining business advantage. Many examples of this, like Amazon [3] or Google [4], are well known today. This first wave of business oriented big data applications is slowly flowing into other domains contributing to conscious efforts towards improving quality of lives and enabling the sustainable development of smart cities. However, such a complex task requires strong domain knowledge with interdisciplinary and cross cutting view. For this reason it is considered that current research lines on big data, which were so far mainly in the domain of computer science and physics, have reach its barrier regarding its applicability on transport and traffic behavior complexity [5]. Main arguments for this lie in the unbalanced focus between knowledge extraction from big data and its complementarity with conceptualization of frameworks and hypotheses of ideas that were long used by transport researchers. Human mobility and urban dynamics are pressing issue as urban population of the world has grown rapidly from 746 million in 1950 to 3.9 billion in 2014 [6] and its mobility contributes to 25% of the global CO<sub>2</sub> emission [7]. Furthermore, transport is the only major sector where emissions continue to grow and even though technological advances improved the energy efficiency in transport, nevertheless this has been outweighed by the increase in travel demand. Consequences of these trends have high societal and economic impacts as, for example, only congestion costs nearly 100 billion Euro, or 1% of the EU's annual GDP [8]. For this reasons, Transport and Traffic Analytics in Smart Cities are emerging area that aims to bridge existing gaps between different disciplines and ensure seamless applicability of big data based analytics on pressing smart city mobility issues.

### II. THE PAPERS AND BEYOND

There are three main topics in this special session, each of which tackles smart city's mobility oriented data analytics and based on the current state-of-the-art identifies existing gaps that hinder wider applicability of big data within transport and traffic domain. Papers presented here try to answer questions:

- How confident and applicable are big data for smart city's traffic and transport planning;
- Can big data be used to forecast future flows of people and goods;
- How can data analytics contribute in facilitating introduction of new mobility services?

A. *How confident and applicable are big data for smart city's traffic and transport planning?*

Introduction of location-enabled (or Global Positioning System, GPS-based) devices started an important revolution in mobility studies [9]-[12] as they allowed continuous tracking of movement locations and, this way, were able to fill some of the gaps that were present when collecting data using traditional methods, like travel diaries of household surveys. Lopez et al., in their paper entitled *Travelled Distance Estimation for GPS-based Round Trips: Car-Sharing Use Case*, take a deeper look into applicability of such data by examining how reliable are they when determining a car trip length. For this purpose they compare vehicles' Controller Area Network bus data, as a ground truth, with GPS recorded trip distances and investigate under which circumstances performance of location-enabled devices varies. Their findings show that particularly sensitive part of the car trip is its beginning as time-to-first-fix location highly impacts the observed trip length. However, poor quality of location precision after the first-fix is more pronounced in urban areas, than in rural ones.

Based on measurements from 170 participants over a period of one year, they observed that location enabled device will, in average, report on 9% shorter trips than they were actually. This effect seems strongly influenced by land use (was trip made in rural or urban area) but also the trip length as underreported distance were more pronounced for short trips than the longer ones.

Bernardino et al. follow the idea of location-enabled devices and their applicability for transport planning purposes. However, they focus more on non-motorized transport modes as walking and biking. Since these transport modes are considered to be more sustainable mobility option than others, they are particularly important for development of equitable and accessible transport services within smart cities.

In line with this, they examine how tracking information can be integrated into smart city planning and policy processes. Based on the review of state-of-the-art practical examples they identify relevant stakeholders and the most beneficiary scenarios on how traditional travel data collection processes can be replaced by tracking data. Some of such scenarios involve implementation of Geographic Information System tools to produce quantified information about the state and performance of the cycling or walking network, allowing timely detection of bottlenecks,

monitoring effects of introduced policy measures or studying user preferences.

B. *Can big data be used to forecast future flows of people or goods?*

Petri et al., in their paper *Datamining and big freight transport database. Analysis and forecasting capabilities*, investigate how existing databases on freight transport can be used, in more efficient way, to support policy making and future investments activities.

For this reason, they focus on freight demand forecasting models that are still in early development stage. They contribute to the topic by further developing existing, mainly deterministic models, with hybrid approach that integrates Decision Tree and Bayesian Networks to forecast future road and rail freight flows. They demonstrate the applicability of the suggested approach on the official European Union's statistical data for the Sixth European Rail Freight Corridor.

Similarly, Lopez et al. in their paper *Forecasting travel behaviour from crowdsourced data with machine learning based model*, integrate hybrid approach, but this time to investigate in more details potential to forecast human, not freight, mobility. They focus on multimodal urban transport system and integrate mobile sensed data to forecast which transport mode one is more likely to use for his/her next trip.

Achieved results demonstrate potential of the suggested approach to be integrated into smart city mobility system management and planning processes, particularly in development of more advanced pre-travel information service.

C. *How can data analytics contribute in facilitating introduction of new mobility services?*

Following the idea of supporting the smart city's mobility planning with availability of big data, Vidan et al., take a deeper look into process of introducing the new service. Integration of new mobility service, in this case the first European commercial costal seaplane service, requires broad view on complementarity with already existing transport services, integration in existing legal frames and study on environmental impacts.

They exhibit the impact of data availability to support such processes, but also to learn from best practices worldwide. The applicability of such approach is demonstrated on an example that includes connectivity of 66 inhabited islands in Croatia with costal smart cities. The new service should aim to increase quality of live (both on islands and along the main coast) but also to facilitate integration of different sized communities into sustainable ensemble.

### III. CONCLUSION

Transport and Traffic Analytics in Smart Cities comprise a process of acquisition, synchronization, integration and analysis of big and heterogeneous data generated by a diversity of sources (e.g. wearable and fixed sensors, vehicles and buildings) to gain reliable insights into urban dynamics and interactions. This still emerging domain aims to bridge between traditional data collection techniques and advanced sensing approaches, big data storage capacity and processing challenges, sustainability, information and communication technologies, behavioral sciences, policy related studies, data visualization and knowledge extraction analytics. In this context, gleaning intelligence from big data aims towards high applicability of its findings, particularly regarding its complementarity with conceptualization of frameworks and hypotheses of ideas that were long used by transport researchers. However its sole purpose is not just in supporting traditionally used approaches but also in enriching them with innovative, data driven, insights. To achieve these goals transport and traffic analytics for smart cities still need to overcome existing challenges, particularly in domain of crowdsourced data representativeness of overall urban dynamics and higher integration between data analytics and transport research domains.

### REFERENCES

- [1] F. Provost and T. Fawcett, "Data Science and its Relationship to Big Data and Data-Driven Decision Making" *Data Science and Big Data*, 1(1), 2013, pp. 51-59.
- [2] I. Semanjski, R. Bellens, S. Gautama, and F. Witlox, "Integrating Big Data into Sustainable Mobility Policy 2.0 Planning Support Systems". *Sustainability*, 2016, In press.
- [3] G. Linden, B. Smith, and J. York, "Amazon.com Recommendations Item-to-Item Collaborative Filtering". *IEEE Internet Computing*, 7(1), 2013, pp. 76-80.
- [4] L. Hasan, A. Morris, and S. Proberts, "Using Google Analytics to Evaluate the Usability of E-Commerce Sites". In: *Human Centered Design*. Berlin Heidelberg: Springer, 2009, pp. 697-706.
- [5] C. Chen, et al., "The promises of big data and small data for travel behavior (aka human mobility) analysis". *Transportation Research Part C: Emerging Technologies*, Volume 68, 2016, p. 285-299.
- [6] United Nations, 2014. *World's population increasingly urban with more than half living in urban areas*, New York: UN.
- [7] IEA, 2014. *Key World Energy Statistics*, Paris: International energy agency.
- [8] European Commission, 2007. COM(2007) 551 Green paper: *Towards a new culture for urban mobility*, Brussel: European Commission.
- [9] I. Cavar, H. Markovic, and H. Gold, "GPS Vehicles Tracks Data Cleansing Methodology". 2006, *International Conference on Traffic Science*.
- [10] T. Feng and H. Timmermans, "Extracting Activity-travel Diaries from GPS Data: Towards Integrated Semi-automatic Imputation". *Procedia Environmental Sciences*, Volume 22, 2014, p. 178-185.
- [11] J. Wolf, S. Bricka, T. Ashby, and C. Gorugantua, "Advances in the Application of GPS to Household Travel Surveys" Washington, 2004, *National Household Travel Survey Conference*
- [12] J. Wolf, R. Guensler, and W. Bachman, "Elimination of the travel diary: Experiment to derive trip purpose from global positioning system travel data". *Transportation Research Record: Journal of the Transportation Research Board*, Volume 1768, 2001, pp. 125-134.