Spatial data science for sustainable mobility

Martin Raubal

Institute of Cartography and Geoinformation, ETH Zurich, Switzerland

Abstract: The constant rise of urban mobility and transport has led to a dramatic increase in greenhouse gas emissions. In order to ensure livable environments for future generations and counteract climate change, it will be necessary to reduce our future CO_2 footprint. Spatial data science contributes to this effort in major ways, also fuelled by recent progress regarding the availability of spatial big data, computational methods, and geospatial technologies. This paper demonstrates important contributions from spatial data science to mobility pattern analysis and prediction, context integration, and the employment of geospatial technologies for changing people's mobility behavior. Among the interdisciplinary research challenges that lie ahead of us are an enhanced public availability of mobility studies and their data sets, improved privacy protection strategies, spatially-aware machine learning methods, and evaluating the potential for people's long-term behavior change towards sustainable mobility.

Keywords: spatial data science, sustainable mobility, human mobility analysis, spatial big data, behavior change, spatially-aware machine learning, geosmartness

1 Introduction

Our world is currently facing dramatic problems, such as major effects of climate change, spread of diseases, overconsumption of goods, a lack of universal access to quality education, and poverty. Although, contrary to common belief, data clearly demonstrates that humankind—through science, reason, and humanism—has been making enormous progress towards a better world over the course of its history, these ideals must be more than ever defended and applied in order to solve our present problems [19].

One of the key problems, which directly and indirectly impacts many facets of our lives, is the constant increase of urban mobility and transport. Although they have brought major

advances to our society, mobility and transport are also main contributors of greenhouse gas emissions. These, in turn, stand in the way of achieving several of the Sustainable Development Goals[1] as formulated by the United Nations Development Programme, such as *good health and well-being*, *sustainable cities and communities*, or *climate action*.

Reducing our future CO_2 footprint depends on both technological as well as behavioral research. Spatial data science, which encompasses space and spatial data at a wider spectrum of scales than geographic information science [20], can contribute to this endeavor in major ways. From a technological perspective, decarbonisation of transport will require improved vehicle efficiency, electric vehicles relying on renewable energy sources, and driverless cars. Vehicles can be equipped with a variety of (geo)sensors and evaluated for efficiency within spatial networks. Electric vehicles may utilize smart charging strategies, i.e., on the one hand consume fluctuating electricity production such as from photovoltaics, and on the other hand recharge the grid through vehicle-to-grid applications [18]. Optimizing such strategies involves sophisticated spatio-temporal mobility pattern analysis to evaluate real-time situations but also making predictions regarding future states [4]. The employment of autonomous vehicles can lead to an optimization of traffic and presents significant challenges regarding positional accuracy and geo-databases [10].

Long-term decarbonisation of transport will also require that people actively make an effort to contain demand and shift to lower emission transport modes [3]. Mobility pattern analysis [29] allows to evaluate people's transport behavior and mode choices—even in real time—and to predict future behavior and transport network states. In addition, it can be utilized to detect behavioral changes over time, such as people switching to more sustainable modes of transport. Location based services [11] and other mobile technologies support people in their spatio-temporal decision-making, potentially leading to increased sustainability and therefore curbing greenhouse gas emissions.

2   Contributions from spatial data science

Over the last decade, spatial data science has contributed in major ways towards the goal of making our society's mobility sustainable. The combination of novel spatial data sources and spatial big data [15, 22], computational methods, and geospatial technologies—also termed *geosmartness* [21]—has provided researchers with great opportunities to perform large-scale spatio-temporal analyses of mobility patterns as well as investigate people's decision-making behavior.

Digitalization has provided us with an unprecedented volume of movement and context data, derived from GPS sensors, social networks, or mobile phone usage. The goal is to mine these data for mobility patterns, both historically and in real-time. Evidence for a conserved quantity of human mobility has been demonstrated as well as a correlation between people's location capacity and the number of their social connections [1]. Such regularities and constraints help generating knowledge of individual and aggregated city-wide mobility behavior [28], which can further be utilized to make predictions regarding future states. This, in turn, facilitates the allocation of spatial resources, optimization of transportation networks and infrastructure.

When analysing mobility data, it is important to take the internal and external context in which people are moving into account [2, 25] because movement is always influenced

---

[1]https://www.undp.org/content/undp/en/home/sustainable-development-goals.html

by spatio-temporal phenomena, such as weather, transportation infrastructure, or personal time schedules. Nowadays, many different sources of context data are available at various spatio-temporal resolutions, and cities, such as London, Singapore, New York, or Zurich, publish them on open data platforms.[2] Volunteered geographic information [9] projects, e.g., OpenStreetMap[3], also contribute their share.

Due to the massive volume, variety, velocity, and veracity of spatial big data [15], traditional spatio-temporal analysis methods have been enhanced by various methods from machine learning. For example, graph convolutional neural networks can be utilized for traffic forecasting [6] and imputing human activity purposes from GPS trajectory data [16], by embedding a large variety of spatio-temporal information and structure in the graphs, and subsequently exploiting this structure for the tasks to be solved.

Several short- and long-term studies based on mobility data have been performed but many of the data sets are not publicly available due to privacy issues [13], which also hinders their reproducibility [26]. These studies have provided general insights into statistical properties and spatio-temporal regularities of human movement [8], or more specifically investigated the spatial distribution of different demographic groups in a city [28].

Geospatial technologies, in combination with spatio-temporal data, are also important tools to be utilized for changing people's mobility behavior. When employed as persuasive technologies, they can nudge users to travel more sustainably [27]. Based on the detection and identification of activities and transport modes [17], changes in behavior over time can be analysed and evaluated. Feedback, gamification methods, and computationally generated suggestions for alternative and more sustainable travel options can support people in their spatio-temporal decision-making, eventually resulting in decreased CO_2 emissions. Different studies have demonstrated the positive effects of persuasive information and communication technologies on mobility choices and behavior [5, 24].

3 Research challenges

Spatial data science has contributed broadly to solving the challenges of sustainable mobility, from movement analysis and prediction based on spatial big data and spatio-temporal analysis methods, to geospatial technologies such as location based services for supporting behavior change. Still, important research challenges, which require collaboration with other scientific fields, are ahead of us.

It will be important to compare and benchmark both the results of mobility studies and their data sets. This is a challenge due to the variety of distinctive tracking styles, sampling rates, spatio-temporal distributions, and sampling biases [14]. The resulting differences in accuracy and privacy issues pose additional problems. Public availability of mobility datasets, such as GeoLife [30], is still limited and more comprehensive benchmark datasets including variable context information are needed. This requires further research regarding strategies for privacy protection of individuals [7,12,13]. Further challenges regarding data concern real-time availability of different sources—including tracking, context, and social media data—and efficient filtering processes due to the sheer volume of data. Whereas in previous times, one could not get enough data, the flood of today often makes it difficult

---

[2]https://data.london.gov.uk/, https://data.gov.sg/, https://opendata.cityofnewyork.us/, https://data.stadt-zuerich.ch/.

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

to find the right data for solving a particular problem. Making progress on these issues, and also on problems of data integration, requires interdisciplinary efforts including data science, geoinformatics, and social science.

From the perspective of analysing mobility data, the integration of machine learning methods into spatio-temporal analysis has already commenced [22,23]. But truly spatially-aware machine learning methods are for the large part still missing. Most of the machine learning models, which are applied to spatial data do not account for spatial autocorrelation and structure, and can therefore either not fully take advantage of spatial constraints, or in the worst-case lead to incorrect results.

From a cognitive perspective, we still do not know how long-term behavior change towards sustainable mobility can be achieved. Although some outcomes of longer-term studies are encouraging, most effects resulting from the use of geospatial technologies such as location based services have only been demonstrated within short-term studies with small samples of users, strongly biased participant groups, and often a lack of control groups [5]. Results can therefore not be generalized to society as a whole. Technological development and ubiquitous access to smartphone apps will hopefully provide us with opportunities in the future to more comprehensively evaluate the potential of such devices to contribute to the vision of a truly sustainable mobile society.

References

[1] ALESSANDRETTI, L., SAPIEZYNSKI, P., SEKARA, V., LEHMANN, S., AND BARONCHELLI, A. Evidence for a conserved quantity in human mobility. *Nature Human Behaviour 2*, 7 (2018), 485–491.

[2] ANDRIENKO, G., ANDRIENKO, N., AND HEURICH, M. An event-based conceptual model for context-aware movement analysis. *International Journal of Geographical Information Science 25*, 9 (2011), 1347–1370.

[3] BOULOUCHOS, K., CELLINA, F., COX, B., GEORGES, G., HIRSCHBERG, S., HOPPE, M., JONIETZ, D., KANNAN, R., KOVACS, N., KÜNG, L., MICHL, T., RAUBAL, M., RUDEL, R., SCHENLER, W., AND CIARI, F. Towards an energy efficient and climate compatible future Swiss transportation system. Tech. rep., SCCER Mobility (Swiss Competence Center for Energy Research—Efficient Technologies and Systems for Mobility), 2017.

[4] BUFFAT, R., BUCHER, D., AND RAUBAL, M. Using locally produced photovoltaic energy to charge electric vehicles. *Computer Science-Research and Development 33*, 1-2 (2018), 37–47.

[5] CELLINA, F., BUCHER, D., MANGILI, F., VEIGA SIMÃO, J., RUDEL, R., AND RAUBAL, M. A large scale, app-based behaviour change experiment persuading sustainable mobility patterns: methods, results and lessons learnt. *Sustainability 11*, 9 (2019), 2674.

[6] CUI, Z., HENRICKSON, K., KE, R., AND WANG, Y. Traffic graph convolutional recurrent neural network: A deep learning framework for network-scale traffic learning and forecasting. *IEEE Transactions on Intelligent Transportation Systems* (2019).

[7] DUCKHAM, M., KULIK, L., AND BIRTLEY, A. A spatiotemporal model of strategies and counter strategies for location privacy protection. In *International Conference on Geographic Information Science* (2006), Springer, pp. 47–64.

[8] GONZALEZ, M. C., HIDALGO, C. A., AND BARABASI, A.-L. Understanding individual human mobility patterns. *nature 453*, 7196 (2008), 779–782.

[9] GOODCHILD, M. F. Citizens as sensors: the world of volunteered geography. *GeoJournal 69*, 4 (2007), 211–221.

[10] GOODCHILD, M. F. GIScience for a driverless age. *International Journal of Geographical Information Science 32*, 5 (2018), 849–855.

[11] HUANG, H., GARTNER, G., KRISP, J. M., RAUBAL, M., AND VAN DE WEGHE, N. Location based services: ongoing evolution and research agenda. *Journal of Location Based Services 12*, 2 (2018), 63–93.

[12] KACHORE, V. A., LAKSHMI, J., AND NANDY, S. Location obfuscation for location data privacy. In *2015 IEEE World Congress on Services* (2015), IEEE, pp. 213–220.

[13] KESSLER, C., AND MCKENZIE, G. A geoprivacy manifesto. *Transactions in GIS 22*, 1 (2018), 3–19.

[14] LAUBE, P. *Computational movement analysis*. Springer, 2014.

[15] LI, S., DRAGICEVIC, S., CASTRO, F. A., SESTER, M., WINTER, S., COLTEKIN, A., PETTIT, C., JIANG, B., HAWORTH, J., STEIN, A., AND CHENG, T. Geospatial big data handling theory and methods: A review and research challenges. *ISPRS journal of Photogrammetry and Remote Sensing 115* (2016), 119–133.

[16] MARTIN, H., BUCHER, D., SUEL, E., ZHAO, P., PEREZ-CRUZ, F., AND RAUBAL, M. Graph convolutional neural networks for human activity purpose imputation. In *NIPS spatiotemporal workshop at the 32nd Annual conference on neural information processing systems (NIPS 2018)* (2018).

[17] MONTINI, L., RIESER-SCHÜSSLER, N., HORNI, A., AND AXHAUSEN, K. W. Trip purpose identification from GPS tracks. *Transportation Research Record 2405*, 1 (2014), 16–23.

[18] MWASILU, F., JUSTO, J. J., KIM, E.-K., DO, T. D., AND JUNG, J.-W. Electric vehicles and smart grid interaction: A review on vehicle to grid and renewable energy sources integration. *Renewable and sustainable energy reviews 34* (2014), 501–516.

[19] PINKER, S. *Enlightenment now: The case for reason, science, humanism, and progress*. Penguin, 2018.

[20] RAUBAL, M. It's the spatial data science, stupid! In *Spatial Data Science Symposium "Setting the Spatial Data Science Agenda"* (2019), Center for Spatial Studies at the University of California, Santa Barbara.

[21] RAUBAL, M., BUCHER, D., AND MARTIN, H. Geosmartness for personalized and sustainable future urban mobility. In *Urban Informatics* (2020), W. Shi, M. F. Goodchild, M. Batty, and M.-P. Kwan, Eds., Springer.

[22] Raubal, M., Wang, S., Guo, M., Jonietz, D., and Kiefer, P. *Spatial Big Data and Machine Learning in GIScience, Workshop at GIScience 2018, Melbourne, Australia.* 2018.

[23] Sabek, I., and Mokbel, M. F. Machine learning meets big spatial data. In *2020 IEEE 36th International Conference on Data Engineering (ICDE)* (2020), IEEE, pp. 1782–1785.

[24] Semanjski, I., Lopez Aguirre, A. J., De Mol, J., and Gautama, S. Policy 2.0 platform for mobile sensing and incentivized targeted shifts in mobility behavior. *Sensors 16*, 7 (2016), 1035.

[25] Sharif, M., and Alesheikh, A. A. Context-aware movement analytics: implications, taxonomy, and design framework. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery 8*, 1 (2018), e1233.

[26] Stodden, V., McNutt, M., Bailey, D. H., Deelman, E., Gil, Y., Hanson, B., Heroux, M. A., Ioannidis, J. P., and Taufer, M. Enhancing reproducibility for computational methods. *Science 354*, 6317 (2016), 1240–1241.

[27] Weiser, P., Scheider, S., Bucher, D., Kiefer, P., and Raubal, M. Towards sustainable mobility behavior: Research challenges for location-aware information and communication technology. *GeoInformatica 20*, 2 (2016), 213–239.

[28] Yuan, Y., and Raubal, M. Analyzing the distribution of human activity space from mobile phone usage: an individual and urban-oriented study. *International Journal of Geographical Information Science 30*, 8 (2016), 1594–1621.

[29] Zheng, Y. Trajectory data mining: an overview. *ACM Transactions on Intelligent Systems and Technology (TIST) 6*, 3 (2015), 1–41.

[30] Zheng, Y., Wang, L., Zhang, R., Xie, X., and Ma, W.-Y. Geolife: Managing and understanding your past life over maps. In *The Ninth International Conference on Mobile Data Management (mdm 2008)* (2008), IEEE, pp. 211–212.