Conclusions

Constantinos Antoniou*, Loukas Dimitriou† and Francisco Pereira‡
Chair of Transportation Systems Engineering, Department of Civil, Geo and Environmental Engineering, Technical University of Munich, Munich, Germany, †Lab. for Transport Engineering, Department of Civil and Environmental Engineering, University of Cyprus, Nicosia, Cyprus, ‡Denmark Technical University, Kongens Lyngby, Denmark

As our world faces grander challenges, fast technology development, together with societal evolution, feeds a stream of new visions, concepts, and opportunities, at a pace that is at times overwhelming to analyze. The concept of *Big Data* certainly is one that is timely born out of the confluence of new technologies, like sensors, faster hardware, and new more powerful software, together with their increasing integration in multiple systems of our life, generating a continuous deluge of data.

As a buzzword, it is likely that *Big Data* may be diluted in the upcoming stream of new concepts and hype, but certainly not what it stands for, particularly for Transportation. Data sets look nowhere close to be reduced, even with new constraints like the General Data Protection Regulation (GDPR), nor does the dramatic need to improve our transportation systems. Current best known data-driven algorithms tend more and more to rely on very large data sets, to avoid overfitting, while keeping high data dimensionality and model complexity.

In other words, *Big data in Transportation* is certainly here to stay. This book aimed to give a comprehensive, and as much in-depth as possible, account of current methodologies, tools, and practices of the area. We started by covering the methodological foundations, many of which cover much beyond the particular area of big data, including statistics, machine learning, visualization, and theory-driven methods. Still, other relevant foundations did not find space in this volume, but yet deserve particular mention and encouragement for readers to seek further, such as Network Sciences and Social Network Analysis (Alhazzani et al., 2016; Chodrow et al., 2016), Causal Inference and Counterfactual reasoning (Pearl, 2003; Pearl and Mackenzie, 2018), Complex Event Processing (Terroso-Sáenz et al., 2012), Ultra-high dimensionality reduction (Wang et al., 2017), or graphics-processing unit (GPU) computing (Jiang et al., 2015).

We also provided a number of practical applications, to provide the reader with the empirical realizations of those methodologies, and good practices in Big Data exploitation. Obviously, we had to leave out a very wide number of applications, that the reader might be interested in, such as traffic estimation from FCD (Fusco et al., 2016), traffic control with reinforcement learning (Wu et al., 2017), Big Data in Railways (Ghofrani et al., 2018), Logistics (Mehmood et al., 2017), Airlines (Badea et al., 2018), Maritime (Vouros et al., 2018), and many more, just to mention the most recent ones.

As a general conclusion, it is our belief that, despite the sheer vastness of research and practice already existing today in Big Data, its role in Transportation is still underexplored, particularly if we consider the needs and available opportunities. For example, prediction models are well-known to be limited in nonrecurrent conditions, such as traffic incidents (Fusco et al., 2016); many solutions do not really explore multimodality (e.g., in traffic control, operations management, traffic information); while the limitations of data-driven and model-based (e.g., simulation) approaches are well-known, little work exists in linking them, to get the best of both worlds; some major challenges in large-scale models remain open, such as transferability and model calibration, that deserve further attention including using Big Data; new hardware opportunities, such as GPU or quantum computing are also both promising and underexplored in Transportation research and practice.

This book was put forth with the expectation that, more than contributing with specific models, methods, tools, or examples, it would engage more researchers and practitioners in Big Data for Transportation. As we say, there are plenty of opportunities, and it is our belief that this book is a good basis to start, and to return to every so often, in the coming years.

REFERENCES

Alhazzani, M., Alhasoun, F., Alawwad, Z., González, M.C., 2016. Urban attractors: Discovering patterns in regions of attraction in cities. arXiv preprint arXiv., 1701.08696.

Badea, V.E., Zamfiroiu, A., Boncea, R., 2018. Big data in the aerospace industry. Inform. Econ. 22 (1), 17–24.

Chodrow, P.S., Al-Awwad, Z., Jiang, S., González, M.C., 2016. Demand and congestion in multiplex transportation networks. PLoS One. 11(9).

Fusco, G., Colombaroni, C., Isaenko, N., 2016. Short-term speed predictions exploiting big data on large urban road networks. Transp. Res. Part C: Emerg. Technol. 73, 183–201.

Ghofrani, F., He, Q., Goverde, R.M., Liu, X., 2018. Recent applications of big data analytics in railway transportation systems: a survey. Transp. Res. Part C: Emerg. Technol. 90, 226–246.

Jiang, H., Chen, Y., Qiao, Z., Weng, T.H., Li, K.C., 2015. Scaling up mapreduce-based big data processing on multi-GPU systems. Clust. Comput. 18 (1), 369–383.

Mehmood, R., Meriton, R., Graham, G., Hennelly, P., Kumar, M., 2017. Exploring the influence of big data on city transport operations: a Markovian approach. Int. J. Oper. Prod. Manag. 37 (1), 75–104.

Pearl, J., 2003. Causality: models, reasoning, and inference. Econ. Theory 19 (675–685), 46.

Pearl, J., Mackenzie, D., 2018. The Book of Why: The New Science of Cause and Effect. Basic Books.

Terroso-Sáenz, F., Valdés-Vela, M., Sotomayor-Martinez, C., Toledo-Moreo, R., Gómez-Skarmeta, A.F., 2012. A cooperative approach to traffic congestion detection with complex event processing and VANET. IEEE Trans. Intell. Trans. Syst. 13 (2), 914–929.

Vouros, G.A., Doulkeridis, C., Santipantakis, G., Vlachou, A., 2018. Taming big maritime data to support analytics. In: Information Fusion and Intelligent Geographic Information Systems (IF&IGIS'17). Springer, Cham, pp. 15–27.

Wang, Y., Shrivastava, A., Ryu, J., 2017. FLASH: Randomized Algorithms Accelerated over CPU-GPU for Ultra-High Dimensional Similarity Search. arXiv preprint arXiv, 1709.01190.

Wu, C., Parvate, K., Kheterpal, N., Dickstein, L., Mehta, A., Vinitsky, E., Bayen, A.M., 2017. In: Framework for control and deep reinforcement learning in traffic.Intelligent transportation systems (ITSC), 2017 IEEE 20th International Conference on, pp. 1–8. IEEE.