GeoInformatica
https://doi.org/10.1007/s10707-024-00526-8

Special issue on the 5th International Workshop on Big Mobility Data Analytics (BMDA'23)

Christos Doulkeridis[1] · Alessandra Raffaetà[2] · Esteban Zimányi[3]

The aim of this special issue is to capture recent advances in data analytics for mobility data, foster the exchange of new ideas on multidisciplinary real-world problems, discuss proposals about innovative solutions, and identify emerging opportunities for further research in the area of big mobility data analytics, such as deep learning on mobility data, edge computing, and visual analytics.

Mobility analytics is a timely topic due to the ever-increasing number of diverse, real-life applications, including the movement of humans, animals, vehicles, vessels, aircraft, which produce massive amounts of streaming spatio-temporal data, whose acquisition, cleaning, representation, aggregation, processing, and analysis pose new challenges for the data management community.

The papers published in this issue constitute revised and extended versions of the accepted papers of the 5th International Workshop on Big Mobility Data Analytics (BMDA'23), held in conjunction with the 26th International Conference on Extending Database Technology (EDBT), in Ioannina, Greece, on March 28, 2023.

BMDA received eleven (11) submissions, each of which was carefully reviewed by three different reviewers, in order to select nine (9) papers to be presented during the workshop in Ioannina. Out of these papers, six (6) papers were invited to submit an extended version to this special issue of Geoinformatica. After a second round of reviews, five (5) papers have been accepted to be published in this special issue.

As such, this special issue consists of five papers covering a wide range of important topics related to big mobility data analytics for spatio-temporal and social data.

The first paper by Mandalis et al. present a distributed Unified Approach for Vessel Traffic Flow Forecasting, called dUA-VTFF. It takes as input historical timestamped

✉ Esteban Zimányi
esteban.zimanyi@ulb.be

1 Department of Digital Systems, University of Piraeus, Piraeus, Greece

2 Dipartimento di Scienze Ambientali, Informatica e Statistica, Università Ca Foscari, Venezia, Italy

3 Department of Computer & Decision Engineering (CoDE), Université libre de Bruxelles, Brussels, Belgium

Published online: 08 August 2024

vessel locations along with future vessel positions generated by a Vessel Route Forecasting model to create traffic flows. These data are arranged into a spatiotemporal grid and then each grid cell is allocated to a computing node through the Apache Spark big data distributed processing framework. Finally, Transformer models are employed for predicting vessel traffic flows in a time horizon of up to 30 min. Experimental evaluations highlight the advantages in terms of accuracy and execution times due to the incorporation of Big Data technology and of the Transformer architecture.

The second paper by Duarte and Sakr surveys both the state of the art in outlier detection algorithms and the available libraries that offer trajectory outlier detection and correction. The focus is on techniques for identifying points which are outliers inside a trajectory. The authors classify the approaches in literature into five categories, statistical-based methods, sliding-window algorithms, clustering-based methods, graph-based methods, and heuristic-based techniques providing the distinguishing features and the main related approaches. Then they describe ten libraries and evaluate their performance for outlier removal in trajectory data by experimenting them in four different datasets ranging in various domains, including maritime, aviation, and urban mobility.

The third paper by Oakley et al. presents a system, called Foresight Plus, for real-time forecasting over streaming traffic data in an urban context. The system relies on data about dynamic urban events and exploits a road sensor network for traffic data acquisition. In terms of forecasting, the system adopts a deep learning based approach and includes short-term (1 h ahead) but also long-term forecasting (up to 24 h ahead), which is more challenging. Foresight Plus is provided as a cloud-based architecture and has been extended to provide serverless inference, which can be more cost-effective for sporadic workloads.

The fourth paper by Abboud et al. proposes a methodology to enrich the maps generated by air pollution fixed stations with opportunistic mobile monitoring data to better estimate air pollution levels in unmonitored locations. This provides a more comprehensive understanding of air pollution levels than individual monitoring approaches, and it can be used to identify pollution hotspots and sources. To estimate air pollution levels in uncovered spots, deep learning methods are used and the efficacy of the approach is validated through experiments conducted on real-life datasets from two cities, Versailles (France) and Chicago (USA).

The last paper by Graser et al. provides an overview of deep learning techniques applied on trajectory data. Apart from classifying existing approaches according to (i) use case category, (ii) neural network architecture, and (iii) granularity of trajectory data, their work clarifies the use of deep learning versus traditional machine learning approaches. Furthermore, they explain the available benchmarks and metrics used for the evaluation of deep learning approaches over trajectory data. As a result, the paper offers a comprehensive overview of deep learning for trajectory data, making it a valuable read both for young and more experienced researchers.

The problems covered by these five papers present several problems and technical challenges that are currently being investigated, where mobility data are exploited to extract relevant information for delivering added-value services. Moreover, they introduce novel techniques and approaches that target emerging applications related to mobility analytics.

We would like to thank the authors for the excellent contributions to this issue; the reviewers for providing thoughtful and detailed manuscript reviews; as well as the editorial office staff at Springer for their assistance during the review process.