Title
Progress in computational movement analysis - towards movement data science

Permalink
https://escholarship.org/uc/item/0nb4q1cq

Authors
Dodge, Somayeh
Gao, Song
Tomko, Martin
et al.

Publication Date
2020-06-25

DOI
10.1080/13658816.2020.1784425

Peer reviewed
Progress in computational movement analysis – towards movement data science

Editorial

Somayeh Dodge,
University of California Santa Barbara, Santa Barbara, USA (sdodge@ucsb.edu)

Song Gao,
University of Wisconsin-Madison, Madison, USA (song.gao@wisc.edu)

Martin Tomko,
University of Melbourne, Australia (tomkom@unimelb.edu.au)

Robert Weibel,
University of Zurich, Switzerland (robert.weibel@geo.uzh.ch)

Introduction
There hasn't been a time in the history of GIScience when movement analytics and mobility insights have played such an important role in policymaking as in today’s global responses to the COVID-19 crisis (Oliver et al. 2020). Non-pharmaceutical interventions such as physical distancing measures and stay-at-home orders have slowed down mobility across the globe and disrupted normal flows of travelers, commuters, trades, etc. Today more than ever we recognize the significance of movement data and movement analytics in urban planning, crisis mitigation, public health (Wang and Taylor 2016, Li et al. 2019, Kraemer et al. 2020). While our community has played a major role in the progress of movement data analytics and advancing its methods and applications over the past two decades (Demšar et al. 2015, Dodge et al. 2016, Long et al. 2018, Dodge 2019, Miller et al. 2019), we still have to further advance developing methods that enable the large-scale characterization of mobility patterns and knowledge discovery in large mobility data (Scherrer et al. 2018). With ever increasing access to large repositories of raw trajectory data contributed voluntarily or often involuntarily through the widespread use of cellphones (Huang et al. 2019), location-aware apps and mobile services, the technology industry has gained an unprecedented data advantage to develop new computational movement analysis methods. Academic researchers often struggle to overcome barriers of data access to develop new mobility
indices (e.g. estimates of mile travelled, percentage/duration of time at-home or at-work, foot traffic information, measures for movement regularity) as well as real-time mobility analytic methods to explore daily changes in mobility patterns and for digital contact tracing starting to proliferate recently (Kitchin 2020, Naghizade et al. 2020, Oliver et al. 2020). Yet, access to detailed data at fine granularity triggers major concerns regarding ethics and privacy (McKenzie et al. 2019), as well as the representativeness and adequacy of mobility indicators currently used for major policymaking and intervention planning (see Fillekes et al. 2019 for an example in the health sciences). We, as the scientific community, should stand at the forefront of research on advancing a ‘responsible data science’ of movement (Dodge 2019, Yang et al. 2020).

This special section further builds on previous efforts by the editorial team and others from the GIScience community and beyond to advance the body of knowledge in Computational Movement Analysis (CMA) (Laube 2014). CMA generally refers to a series methods and analytical approaches to process, structure, visualize, and analyze tracking data and movement patterns to facilitate knowledge discovery and modeling of movement. Specifically, this special section was proposed as part of a pre-conference workshop on Analysis of Movement Data (AMD’2018) at the GIScience 2018 meeting, 28 August 2018, Melbourne, Australia. The focus of this special section is on three aspects of CMA: (1) representation and modeling of movement (Buchin et al. 2019, Graser et al. 2020); (2) urban mobility analytics (Qiang and Xu 2019, Li et al. 2020), and (3) movement analytics using social media data (Ma et al. 2020, Xin and MacEachren 2020). With the papers presented in the special section we highlight recent advancements in CMA with the development of methods and techniques for big movement data analytics and utilization of trajectories constructed using user-generated crowdsourced contents such as geo-tagged social media posts. Traditional CMA methods were often developed and evaluated using a smaller set of movement data involving smaller numbers of individuals and contextual variables. As the momentum to generate more geo-enriched movement data at large volumes, high frequencies, and for longer durations continues, this is a timely and significant achievement towards movement data science (Dodge, 2019). As the papers of this special section illustrate, movement data science leverages the advancements in big data analytics, cyberinfrastructure, parallel computing, and data fusion to enhance the analysis of large, multi-faceted, and multi-sourced movement data (Dodge, 2019). Below we summarize the six original papers presented in this special section.
Representation and modeling of movement

Graser et al. (in this issue) propose a new model named Massive Movement Model (M^3) to summarize patterns and variations in movement parameters (e.g. speed, direction, and their derivatives). M^3 supports exploratory data analysis and hypothesis generation using a scalable and distributed framework for the representation of massive trajectory dataset using a grid-based approach. The novelty of M^3 lies in the use of multiple data-driven prototypes per grid cell instead of simply aggregating movement parameters in a single value. The authors demonstrate the performance of their proposed model using two large tracking data sets of ship and vehicle movements including 3.9 billion GPS records.

Group diagrams, proposed by Buchin et al. (in this issue) offers a new approach for representing and summarizing groups of trajectories while maintaining the spatiotemporal structure of the groups’ movement patterns. The proposed approach is a generic algorithm and takes three inputs: a distance threshold, a similarity measure, and a minimality criterion. This article analyzes the algorithmic complexity to compute the group diagrams with different similarity measures including the equal-time distance (when objects move simultaneously), the similar-time distance (when objects move within a given time shift), and the Fréchet distance (when objects move independently of time constraints). A case study to demonstrate the proposed group diagrams is conducted using the GPS tracks of one goose family on their spring migration route.

Urban mobility analytics

Li et al. (in this issue) use dockless bike sharing data as a new form of mobility for spatial, temporal, and statistical analysis of intra-urban human movement. Using data of recorded bike trips, the study quantifies short-trip transportation patterns in Shanghai, China. This work provides a new framework to integrate multiple sources of contextual and movement data such as the road and transit networks (including bike paths and public transit), information on road characteristics and urban land-use to explore space-time biking patterns at both the individual level and aggregate level in an urban area.

Qiang and Xu (in this issue) use crowdsource traffic data in Google Maps to evaluate road network resilience during natural hazards. A methodological framework of urban road network resilience measurements concerning both the reduction of transportation accessibility due to disturbance and the speed of recovering to normal performance is proposed. The performance of the proposed
framework is validated using a winter storm case study in the metropolitan area of Cleveland, Ohio, USA. The results show that low-accessibility communities in normal condition may also have lower road network resilience in an emergency, which can inform disaster management and transportation planning for reducing natural hazard impacts on urban mobility.

**Movement analytics using social media data**

Ma et al. (in this issue) analyze the heterogeneous human mobility patterns detected from geo-tagged tweets along road networks in two metropolitan areas: London and Tokyo. Specifically, the heterogeneity of urban human movements is analyzed by using a number of measures, including the number of tweets per trajectory, trajectory lengths and durations, the size of detected spatial hotspots, and observed movement flows along the road networks. The power-law function fits all these metric distributions well (e.g., a minority of road segments account for a majority of urban traffic). The study then conducts a two-level agent-based simulation that incorporates the scaling characteristics of urban space and human mobility using each of three distance types—metric, angular and combined. The correlation between street-based observed movement flow and simulated flow within the selected urban hotspots is found to be highest when using the metric-angular combined strategy.

Xin and MacEachren (in this issue) address the need to characterize the travel patterns of attendees of large-scale events. Such travel patterns facilitate understanding the dynamics of social events in cities and may well be relevant for applications in the COVID-19 pandemic as highlighted in the introduction of this editorial. The authors note that the labor-intensive data collection from travel diaries has formed a major impediment to the scientific study of event travel patterns. To overcome this data bottleneck, the authors apply topic modeling to geotagged data from social networks (i.e. Twitter). As such, they enrich the spatiotemporal data of time and location with thematic information and develop a data-mining approach to identify atypical behaviors of event attendees.

**Conclusion: towards a movement data science**

The work presented in this special section responds to previously identified gaps in the CMA literature (Dodge et al. 2016, Long et al. 2018, Dodge 2019, Miller et al. 2019) on computationally intensive movement data analytics and visualization (Graser et al., in this issue), representation of collective movement and interactions in groups of trajectories (Buchin et al., in this issue), sensor
fusion and data integration to contextualize movement (Li et al. and Ma et al., in this issue), as well as movement pattern analysis using crowdsourced data (Xin and MacEachren; Qiang and Xu, in this issue). These studies and the research presented in a preceding IJGIS special section on “Big Spatiotemporal Data Analytics” (Yang et al. 2020) highlight important achievements in data-driven approaches to modeling, representation, and analytics of movement using ‘big’ mobility and crowdsourced data, forming one of the pillars of the “data science framework for movement” (Dodge 2019) to advance the knowledge and understanding of movement processes and individuals’ behavior.

In 2020, with the COVID-19 pandemic we witness a change in the way large mobility data are shared and how access to geo-enriched mobility data is streamlined though data dashboards (Gao et al. 2020), and ‘open data’ and ‘data for good’\(^1\) initiatives supported by commercial companies, such as Facebook, Cuebiq, SafeGraph, Google, Apple, and Descartes Labs. With the shift in data access and rapid advancement in computational approaches, it is now time to further the development of real-time analysis, data-driven and theory-driven predictive movement models as another pillar of movement data science (Dodge 2019). Modeling of movement is key to the prediction of critical changes in real-time dynamic problems such as disease outbreak, climate-induced migrations, or human-wildlife conflicts.

In contrast to our previous special issues in the area of movement data analysis (Dodge et al. 2016, Long et al. 2018), which mainly used animal tracking data to demonstrate the applicability of CMA methods, the majority of papers in this issue focused on human mobility, with the exception of the work presented in Buchin et al. This also reflects the increasing access to human mobility data and the availability of crowdsourced data at higher temporal resolution. However, further research is required to bridge the gap between CMA for human and animal movement towards an integrated science of movement (Miller et al. 2019). Furthermore, the fusion of Eulerian and Lagrangian movement perspectives has been understudied in CMA (Laube 2014, Long et al. 2018, Dodge 2019). In addition, visualization still remains a key challenge in dissemination and communication of patterns and knowledge of movement, especially when dealing with large and long-term movement data sets (Dodge 2019, Miller et al. 2019). In this issue, Graser et al. and Buchin et al.

\(^1\) [https://www.cuebiq.com/about/data-for-good/](https://www.cuebiq.com/about/data-for-good/)
responded to this challenge by featuring different representation techniques to summarize patterns of movement in both network and Euclidean spaces.

Moving forward, we see a clear need for more reproducible research in CMA, following a growing mega trend in data-driven sciences. Data quality and privacy challenges as well as uncertainty in data, analytics, and modeling have been largely overlooked in the CMA literature so far. For a more responsible movement data science, careful considerations should be given to the quality, uncertainty, and representativeness of ‘large’ mobility data that are being used for generating important mobility insights for policymaking. Lastly, with the recent exciting developments in data access, as a community we should think about leveraging this advantage to make movement data science more relevant to real world problems for the mitigation of societal and environmental challenges such as disease outbreaks, population mobility, natural hazards, and human-wildlife conflicts.

References


distributed incrementally updatable solution for big movement data exploration.

*International Journal of Geographical Information Science*, 00 (00), in press.


