Semantic-aware building and summarization of multiple aspect trajectories

Chiara Pugliese
ISTI-CNR, University of Pisa, Pisa, Italy
chiara.pugliese@isti.cnr.it
chiara.pugliese@phd.unipi.it

ABSTRACT

The proliferation of motion sensors has significantly contributed to the availability of mobility data. An important line of research focuses on augmenting these datasets with diverse semantic information, referred to as *aspects*, thereby yielding multiple aspect trajectories (MATs). However, a notable gap in the existing literature pertains to the absence of methodologies for obtaining MATs and the scarcity of real-world datasets. To address this gap, we introduce MAT-Builder, an innovative system designed to facilitate the customization of semantic enrichment of trajectories through the use of arbitrary aspects and external data sources. Notably, the richness of information endowed by MAT-Builder may introduce challenges in terms of data management and storage. Consequently, we propose MAT-Sum, an approach tailored to summarize trajectories while preserving their semantic information.

CCS CONCEPTS

• **Information systems** → **Location based services**; **Geographic information systems**.

KEYWORDS

Semantic trajectory, multiple aspect trajectory, summarized semantic trajectory, semantic enrichment

ACM Reference Format:
Chiara Pugliese. 2023. Semantic-aware building and summarization of multiple aspect trajectories. In *The 31st ACM International Conference on Advances in Geographic Information Systems (SIGSPATIAL '23), November 13–16, 2023, Hamburg, Germany*. ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3589132.3628363

1 INTRODUCTION AND MOTIVATIONS

With the proliferation of sensor devices, there has been a consistent expansion in the volume of available mobility data. Over recent years, researchers have been directed towards managing and analysing this movement data. These efforts often involve the enrichment of such data with contextual information (i.e., *semantic dimensions*), including the activities carried out by the moving entity, the Points of Interest (POIs) visited, the use of public transportation, and other relevant factors. Within the existing literature, the resulting combination of movement data and contextual semantics has been denoted as *semantic* or *multiple aspect* trajectories (MATs) [3, 6, 10, 11]. Incorporating semantic dimensions (i.e., aspects) into movement data holds potential value in various analytical contexts, including but not limited to the comprehension of human behavior patterns, the optimization of public transportation routes, and the assessment of how weather conditions might influence the dynamics of moving objects [7]. However, the availability of comprehensive datasets containing multiple aspect trajectories that accurately represent real-world scenarios remains limited often due to privacy and confidentiality issues. Furthermore, the integration of extensive semantic information has the potential to significantly increase the resources and computational capabilities to ensure effective analysis and meaningful interpretation. To overcome these challenges, researchers have focused on developing methods to reduce the volume of mobility data through summarization or simplification of trajectories [1, 2, 4].

Consequently, our study aims to address the following research questions:

RQ1 What methodologies can be employed to obtain datasets of multiple aspect trajectories?
RQ2 How can we effectively summarize multiple aspect trajectories while preserving their semantics?

To answer the first research question, which pertains to the absence of a standardized approach for the semantic enrichment of trajectories in current literature, we present MAT-Builder. MAT-Builder stands as a methodology and a system characterized by configurability, extensibility, and modularity, allowing users to enrich trajectories with semantic aspects derived from external data sources [5, 9]. To address the second research question, we propose a new approach, i.e., MAT-Sum. This approach focuses on enriching trajectories with semantic information while simultaneously condensing the mobility data without sacrificing its semantic integrity. By doing so, MAT-Sum aims to overcome the limitations encountered in MAT-Builder and effectively tackle the challenges previously delineated. We detail both MAT-Builder and MAT-Sum in the next section.

2 MAT-BUILDER AND MAT-SUM

As delineated in the previous section, the existing literature lacks an approach for generating MAT datasets that transcends constraints tied to specific semantic aspects, datasets, or application scenarios. However, the enrichment of the same dataset of movement data with heterogeneous semantic dimensions derived from diverse data sources holds the potential to unlock various applications and research analyses. For instance, this enriched dataset can serve to

SIGSPATIAL '23, November 13–16, 2023, Hamburg, Germany

ACM ISBN 979-8-4007-0168-9/23/11.
https://doi.org/10.1145/3589132.3628363

understand the behavior of tourists in urban landscapes or comprehend the use patterns of public transportation within distinct city zones. In light of this and to answer the **RQ1**, we introduce a methodology designed to empower users in creating MAT datasets by tailoring the entire semantic enrichment process. More precisely, our proposed MAT-Builder system (top part of Figure 1), which embodies the design principles of the above-mentioned methodology, offers a high degree of configurability, extensibility, and modularity. Therefore, it enables stakeholders to (1) define trajectory segments to enrich, (2) select data sources for semantic enrichment aligned with their preferences, (3) opt for optimal techniques to combine movement data with semantic aspects, and (4) ultimately represent the resulting dataset by means of Resource Description Framework (RDF) knowledge graphs. Such uniform representation of the final enriched trajectories facilitates querying and analysing enriched trajectory data across MAT datasets. Consistent with analogous methodologies, we evaluate the effectiveness of MAT-Builder via a comprehensive qualitative empirical assessment, emphasizing its effectiveness in creating and analysing semantically enriched datasets. In pursuit of this objective, we conduct two analyses on different representative scenario instances to exemplify the practicality and adaptability of our methodology. These instances, situated in the domains of tourism and urban mobility, illustrate the adeptness of MAT-Builder in effectively answering common queries.

Considering that MATs might be massive, complex, and contain redundant content, we propose a perspective shift. This shift entails moving away from a trajectory-centric enrichment approach towards one centred around locations. This transition leads to a more compact representation of trajectories, particularly those like MATs. The objective of this novel methodology, referred to as MAT-Sum, is to condense mobility data while retaining the inherent semantic information it encompasses (**RQ2**). The core concept of MAT-Sum revolves around considering the geographical region through which trajectories pass. This geographical context is used for both semantic enrichment and summarization of trajectories. To achieve this, MAT-Sum divides the geographic area of interest into cells, each of which is then enriched with semantic aspects. Subsequently, these enriched cells are used by MAT-Sum to generate condensed semantic summarized trajectories from the original ones. Specifically, every trajectory – whether it has undergone prior semantic enrichment from a trajectory-centric approach or not – is discretized into the set of enriched cells it intersects. This discretization serves a dual function: it not only enriches the trajectory with semantics from a location-centric perspective but also aids in compacting information about its movements and any associated semantic elements it might carry. The bottom part of Figure 1 shows an overview of how MAT-Sum works.

In the initial phase of our experimental evaluation, we assess the effectiveness of MAT-Sum from two distinct yet complementary perspectives: reduction in information within the summarized semantic trajectories and preservation of adequate semantic quality. This evaluation involves employing two distinct datasets across various scenarios. Subsequently, we compare the performance of our approach with two baseline methods, namely, RLE and Seqscan-D [2], in terms of its ability to generate high-quality summarized semantic trajectories. The evaluation is based on two key metrics: summarization rate and MUITAS [8], the latter being a similarity

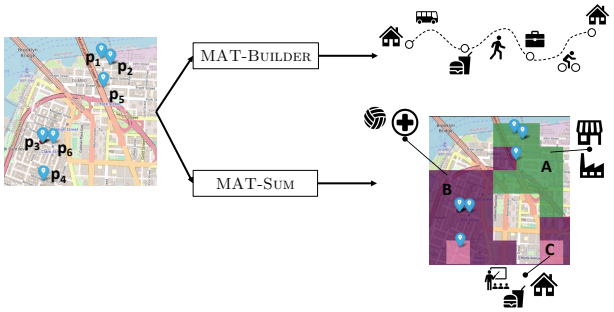

Figure 1: Overview of MAT-Builder and MAT-Sum.

measure designed for MATs. The results indicate that MAT-Sum effectively achieves its objectives according to these evaluation criteria.

3 CONCLUSIONS AND FUTURE WORKS

We think the trajectories obtained through MAT-Sum may open the way for new analyses, as the context of the traversed locations is added, alongside semantics augmenting trajectories using semantic enrichment methods such as MAT-Builder. Furthermore, our objective is to determine whether comparable analytical results can be achieved with summarized trajectories as with the original trajectories. Should such comparability be established, it could potentially reveal a pathway to expedite the analysis of trajectories without compromising the quality of the semantics that enrich them.

REFERENCES

[1] D. Amigo, D. S. Pedroche, J. García, and J. M. Molina. Review and classification of trajectory summarisation algorithms: From compression to segmentation. *Int. J. of Distributed Sensor Networks*, 17(10):15501477211050729, 2021.
[2] M. L. Damiani, F. Hachem, C. Quadri, M. Rossini, and S. Gaito. On location relevance and diversity in human mobility data. *ACM Transactions on Spatial Algorithms and Systems (TSAS)*, 7(2):1–38, 2020.
[3] R. Fileto, C. May, C. Renso, N. Pelekis, D. Klein, and Y. Theodoridis. The baquara^2 knowledge-based framework for semantic enrichment and analysis of movement data. *Data Knowl. Eng.*, 98:104–122, 2015.
[4] C. Fu, H. Huang, and R. Weibel. Adaptive simplification of GPS trajectories with geographic context - a quadtree-based approach. *Int. J. Geogr. Inf. Sci.*, 35(4):661–688, 2021.
[5] F. Lettich, C. Pugliese, C. Renso, and F. Pinelli. General methodology for building multiple aspect trajectories. In *The 38th ACM/SIGAPP Symposium On Applied Computing, ACM SAC 2023, Tallin, Estonia, March 27-31, 2023, Proceedings*, 2023.
[6] R. d. S. Mello, V. Bogorny, L. O. Alvares, L. H. Z. Santana, C. A. Ferrero, A. A. Frozza, G. A. Schreiner, and C. Renso. MASTER: A multiple aspect view on trajectories. *Transactions in GIS*, 23(4):805–822, 2019.
[7] C. Parent, S. Spaccapietra, C. Renso, G. Andrienko, N. Andrienko, V. Bogorny, M. L. Damiani, A. Gkoulalas-Divanis, J. Macedo, N. Pelekis, et al. Semantic trajectories modeling and analysis. *ACM Computing Surveys (CSUR)*, 45(4):1–32, 2013.
[8] L. M. Petry, C. A. Ferrero, L. O. Alvares, C. Renso, and V. Bogorny. Towards semantic-aware multiple-aspect trajectory similarity measuring. *Transactions in GIS*, 23(5):960–975, 2019.
[9] C. Pugliese, F. Lettich, C. Renso, and F. Pinelli. MAT-builder: a system to build semantically enriched trajectories. In *MDM 2022*, pages 274–277, 2022.
[10] L. Ruback, M. A. Casanova, A. Raffaetà, C. Renso, and V. M. P. Vidal. Enriching mobility data with linked open data. In *IDEAS 2016*, pages 173–182. ACM, 2016.
[11] S. Spaccapietra, C. Parent, M. L. Damiani, J. A. de Macedo, F. Porto, and C. Vangenot. A conceptual view on trajectories. *Data & Knowledge Engineering*, 65(1):126–146, 2008.