On Modeling and Analyzing Museum Visitor Movements with Semantic Trajectories

Alexandros Kontarinis

On Modeling and Analyzing Museum Visitor Movements with Semantic Trajectories

Alexandros Kontarinis[1,2]

[1] ETIS UMR 8051, Université Paris Seine, Université de Cergy-Pontoise, ENSEA, CNRS, F-95000, Cergy, France
[2] DAVID lab., Université Paris-Saclay, Université de Versailles Saint-Quentin-en-Yvelines, F-78035, Versailles, France
alexandros.kontarinis@ensea.fr

Abstract. Traditional museum audio guides have gradually been enhanced with multiple additional functionalities. The implementation of location-based services in particular, has enabled the collection of large volumes of spatiotemporal visitor movement data, from which individual visitor trajectories can be extracted. These trajectories can be studied to enable museums to better "know" their visitors, by means of trajectory mining. In addition, tapping into context data could make such analyses more expressive and revealing, by adding meaning and intention to trajectories. This work discusses important challenges in studying museum visitor movements with the help of context-aware indoor trajectory modeling and mining, and proposes work directions for each challenge.

Keywords: Data Mining · Indoor Trajectories · Trajectory Mining · Movement Patterns · Museum Experience

1 Motivation

Thanks to the advent of wireless technologies such as Bluetooth and WiFi, the location of museum visitors can be automatically acquired as they move through the exhibition spaces. The goal is to support location-based services offered via the museum's multimedia guides or mobile applications (e.g. automated audio content delivery). These services are ultimately aimed at improving the visiting experience. However, by collecting visitor movement data, museums can also study their visitors' mobility behavior, in an effort to gain a deeper understanding of their needs and expectations.

Useful actionable insight can be more readily extracted from movement data if those are first structured as individual trajectories, which may require some pre-processing steps or even be impossible, depending on the granularity and structure of the raw spatiotemporal data. A trajectory refers to the geometric aspect of the spatiotemporal path of a moving object. In simpler terms, it is a sequence of the object's positions in space and time: $T = (p_1, p_2, ..., p_n)$. Trajectory data-based applications are usually not primarily concerned with the physics of trajectories nor with the geometry of the moving objects. Indeed, museum visitors can be effectively modeled as moving points.

However, trajectories do not only have a geometric aspect. [5] considered trajectories from a semantic point of view, claiming that they should correspond to semantically meaningful travels. Hence, an object's whole movement lifespan consists of potentially many trajectories, each one meaningfully interpreted and defined by its own starting and ending time instants. A semantic segmentation of trajectories into application-specific sub-intervals of moves and stops was also introduced in [5]. But actually, trajectories can be semantically enhanced even beyond the "stops-moves" concept; any type of annotation which adds meaning to trajectories and their subdivisions, gives rise to so-called "semantic trajectories". Thus, a semantic trajectory can be seen as a sequence of spatiotemporal points complemented with annotations containing semantic values related to places, activities, transportation modes, or any other domain or environmental knowledge: $ST = ((p_1, \mathcal{A}_1), (p_2, \mathcal{A}_2), ..., (p_n, \mathcal{A}_n))$.

2 Identifying the Challenges: Current Approaches and Limitations in Trajectory Modeling and Mining

A major challenge in building an indoor visitor trajectory analytics system is designing a formal trajectory data model, which accounts for the complexities of indoor environments. For example, architectural elements can considerably affect movement. In museums, long and narrow hallways often dictate the path taken by the visitors, similar to how transportation networks restrict vehicle movement in outdoor settings. Also, accessibility constraints are far more dynamic (e.g. rooms closed for restoration purposes). Current trajectory models for the most part ignore such intricacies of indoor environments and need to be extended to indoor settings in non-trivial ways. This is partly due to the fact that existing navigation-oriented indoor space modeling standards have so far seen limited application [2]. Equally importantly, the trajectory data model should also account for varying degrees of data quality, because unlike outdoor trajectories which are based on GPS data, indoor trajectories are obtained through a wide variety of positioning technologies and techniques [4]. This leads to a whole range of location perceptions, each of different precision and quality.

Choosing the space and the trajectory model has a direct effect on the types of analysis that are implementable on top of them. For example, indoor distances can not be calculated using the typical 2D euclidean metric, primarily due to walls and multiple floors (vertical movement is also more frequent). This has many implications on the trajectory mining methods, such as rendering conventional trajectory similarity measures ineffective. Similarly, semantic trajectories also raise requirements in the types of trajectory distance measurement (e.g. two visitor trajectories which do not share any spatial characteristics, could still be considered semantically similar if they both start with stops on ancient Greek statues, then Italian paintings, and then finish at an exit).

Finally, semantic trajectories are so far mostly defined at a conceptual level and therefore most existing trajectory data mining methods and techniques deal exclusively with the spatiotemporal dimensions of trajectories. These methods include clustering, classification and mobility pattern extraction (frequent patterns, sequential patterns, association rules, group movement identification), etc., but with few exceptions, they ignore the semantic aspect of trajectories. As a matter of fact, there is still no clear consensus on which types of trajectory semantics would better describe human mobility behavior in general, let alone domain-specific or application-specific behavior.

3 Research Directions

One approach at accounting for the specificities of indoor environments is to represent indoor spaces by graph-based or set-based models consisting of symbolic locations (e.g. human-readable hall identifiers). Ideally, an indoor trajectory model should combine both symbolic and coordinate trajectory representations. Also, to deal with positioning data quality issues, it should incorporate hierarchical elements that enable the representation of trajectories at multiple levels (e.g. as sequences of exact points, sequences of regions, sequences of rooms). More generally, a separation between a trajectory's abstract perception and its physical encoding is needed. Abstractly, a trajectory can be viewed as a continuous mapping function, defining a position in an indoor space for a visitor and time instant, while physically it can be described by a sequence of discrete predefined spatial cells (in the spirit of how the IndoorGML standard [2] represents indoor space) and temporal intervals, along with movement attributes (e.g. speed).

There exist few algorithms and data structures to support the process of semantically enriching trajectories, and therefore expressiveness and consistency issues in modeling semantic trajectories merit further investigation. Semantic analysis at an arbitrary number of different levels of detail (e.g. stops at different collections of a big museum each consisting of stops and moves at the room level and in turn at the exhibit level) is achievable by exploring enrichment processes based on the hierarchic subdivisions of movement (such as in [1]). Methods for direct domain-specific semantic enrichment of trajectories should also be investigated, as existing works have neglected domain specific datasets (in favor of DBpedia, Open Street Map, Open Weather Map and other popular sources) [3].

With regards to trajectory mining, one direction is to try to extend the existing methods and movement pattern definitions to indoor spaces, or alternatively investigate new movement patterns by considering the context as well, proposing new features of typical context-aware indoor trajectories, and analyzing them to capture visitor behavior and intention. Finally, on-line trajectory mining methods should be explored more, given that most works opt for historical data rather than trajectory data streams.

4 Conclusion

Museums are starting to consider the use of computational data analytics to study the moving patterns of their visitors. In this work, we identify the most important challenges to be resolved, in order to enable an advanced type of museum visitor movement analytics, and provide some general research directions for overcoming them.

References

1. Renato Fileto, Alessandra Raffaet, Alessandro Roncato, Juarez A.P. Sacenti, Cleto May, and Douglas Klein. A Semantic Model for Movement Data Warehouses. In *Proceedings of the 17th International Workshop on Data Warehousing and OLAP*, DOLAP '14, pages 47–56. ACM, 2014.
2. Hae-Kyong Kang and Ki-Joune Li. A Standard Indoor Spatial Data ModelOGC IndoorGML and Implementation Approaches. *ISPRS International Journal of Geo-Information*, 6(4):116, 2017.
3. Luiz Andr P. Paes Leme, Chiara Renso, Bernardo P. Nunes, Giseli Rabello Lopes, Marco A. Casanova, and Vnia P. Vidal. Searching for Data Sources for the Semantic Enrichment of Trajectories. In *Web Information Systems Engineering WISE 2016*, pages 238–246. Springer, Cham, November 2016.
4. Rainer Mautz. *Indoor Positioning Technologies*. PhD thesis, ETH Zurich, 2012.
5. Stefano Spaccapietra, Christine Parent, Maria Luisa Damiani, Jose Antonio de Macedo, Fabio Porto, and Christelle Vangenot. A Conceptual View on Trajectories. *Data Knowl. Eng.*, 65(1):126–146, 2008.