International Cartographic Association
Association Cartographique Internationale

A Novel Indoor Trajectory Pattern Mining Framework

Hassam Ali [a,*], Wangshu Wang [a]

[a] *Research Unit Cartography, Department of Geodesy and Geoinformation, Vienna University of Technology, hassam.ms20@gmail.com, Wangshu.Wang@geo.tuwien.ac.at*

* Corresponding author

Keywords: Trajectory, Similarity Measures, Indoor LBS, Pattern Mining, Framework

Abstract:

The recent shift in research from outdoor to indoor environments with reference to trajectory pattern mining was significantly influenced by the fact that humans spend most of their time indoors. Furthermore, the development of many indoor positioning technologies also paved the way to study indoor movement trajectory. But, because of the movement restrictions imposed by the characteristics of the indoor environment, indoor trajectories are different from outdoor trajectories. In order to locate similar indoor movement trajectories that may be aggregated together to find patterns, a variety of distance functions and similarity metrics are proposed. These similarity measures include Indoor Semantic Trajectory Similarity Measure (ISTSM) (Zhu et al., 2021), Weighted Edit-distance (Cheng et al., 2021), revised Longest Common Sub-Sequence (LCSS) and R-tree (Wang et al., 2022).

However, due to the paucity of indoor trajectory data sets, privacy issues, and the unreasonably high cost of technology and infrastructure for continuously capturing indoor movements mostly semantic component remains the primary focus of indoor similarity measures and pattern mining studies. Some research studies attempted to also include the spatial component in indoor trajectory analysis using estimation or interpolation techniques. Zhu et al. (2021) utilized a ratio of shortest to the longest possible distance between the points. Cheng et al. (2021) used a cost for spatial components. This cost is 1 for two trajectories if they are on the same floor and is the ratio of difference of floors to total floors otherwise. These estimates and presumptions do not accurately depict the path taken by a moving object or person in a confined environment, making it impossible to determine the real spatial similarity between the two trajectories. Secondly, these studies used a single value to indicate the similarity of trajectories across multiple characteristics. However, a single value cannot accurately describe the similarity in several trajectory parameters that are used to determine the overall similarity. Consequently, the number of semantic patterns present in a database and the number of spatially similar groups of trajectories present in a given semantic pattern cannot be determined by combining two similarity values into a single value. It is crucial to keep in mind that two trajectories, in a semantic pattern, that are identical semantically may not always be similar spatially. Although, it is still expensive but, in some indoor spaces like a big conference venue or a shopping mall, it is possible to capture the movements of pedestrians. Furthermore, synthetic indoor trajectory datasets, which are proven effective in the analysis, are also available. Thus, now it is possible to incorporate the actual trajectories for the calculation of spatial similarity. Similar to outdoor trajectory studies, classical distance functions like Hausdorff and Fréchet distance can be utilized to calculate the spatial similarity between two trajectories. But, the distance functions that can be used as a similarity measure between the trajectories have different characteristics. Furthermore, due to the unique characteristics of indoor spaces and the varying degrees of movement restrictions they impose, it is still unclear which similarity measures are suitable for indoor environments.

To address the above-mentioned issues a new framework to mine the indoor trajectory patterns is presented. Given in figure 1, the proposed framework differs from the current approach, which combines the semantic and spatial characteristics to mine patterns in a single phase, by mining the semantic and spatial patterns sequentially in two different steps.

The appropriate preprocessing will be done in the first stage and the stay points will be extracted. In the second, phase, the entire trajectory database will be further separated into smaller groups in accordance with the problem statement or the necessary research questions. Then sequential patterns will be retrieved using a sequential pattern mining algorithm, and those satisfying the requirements to be referred to as semantic patterns will be chosen. The trajectories that are spatially similar within a semantic pattern will be grouped together in the following stage.

In the third step, the trajectories will be initially split out, and the trajectory segments that constitute a semantic pattern will be taken into consideration for further processing. In order to determine the spatially related groups of trajectories inside a semantic pattern, a distance function as a similarity metric will be used. The distance function will determine how similar the trajectories are to one another in an indoor space. As the use of similarity measures is subjective thus, clustering is used to assess the effectiveness of the different similarity measures. The outcomes of the clustering may

Abstracts of the International Cartographic Association, 6, 7, 2023.
31st International Cartographic Conference (ICC 2023), 13–18 August 2023, Cape Town, South Africa.
https://doi.org/10.5194/ica-abs-6-7-2023 |

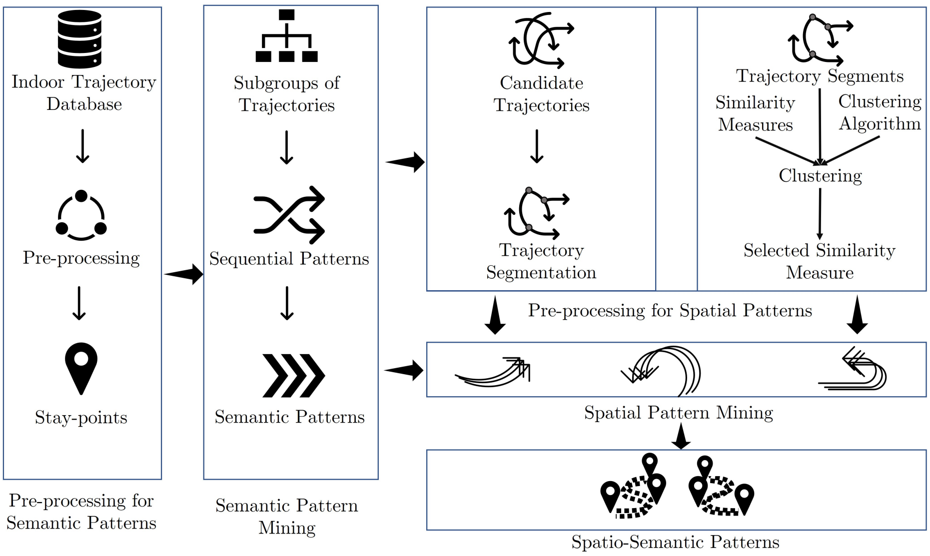

Figure 1. Indoor trajectory pattern mining framework

then be compared using evaluation metrics and the most appropriate similarity measure will then be used. Since the spatially similar clusters already originated from semantic patterns, they will be both spatially and semantically similar when mined, getting the name spatio-semantic patterns.

A synthetic trajectory dataset of an indoor conference venue is used to evaluate the proposed framework. Edit Distance with Real Penalty (ERP) performs better than other similarity measures for the given dataset or the particular problem at hand. Furthermore, one set of ground-truth data was used along with the results to predict the other i.e. type of the sub-venues of the conference. The suggested framework was successful in predicting unidentified venues, demonstrating its efficacy. A combination of the semantic and spatial components of trajectories is essential for mining indoor trajectory patterns, while the temporal aspect may provide value, and could prove to be a useful addition to the framework for indoor pattern mining in the future.

References

Cheng, D., Yue, G., Pei, T. and Wu, M., 2021. Clustering indoor positioning data using e-dbscan. *ISPRS International Journal of Geo-Information* 10(10), pp. 669.

Wang, P., Yang, J. and Zhang, J., 2022. Indoor trajectory prediction for shopping mall via sequential similarity. *Information* 13(3), pp. 158.

Zhu, J., Cheng, D., Zhang, W., Song, C., Chen, J. and Pei, T., 2021. A new approach to measuring the similarity of indoor semantic trajectories. *ISPRS International Journal of Geo-Information* 10(2), pp. 90.

https://doi.org/10.5194/ica-abs-6-7-2023 |