Multiple Aspect Trajectory Data Analysis: Research Challenges and Opportunities

Carlos Andres Ferrero[1,2], Luis Otavio Alvares[1], Vania Bogorny[1]

[1]Programa de Pós-Graduação em Ciência da Computação
Departamento de Informática e Estatística (INE)
Universidade Federal de Santa Catarina (UFSC), Florianópolis, SC, Brasil.

[2]Instituto Federal de Santa Catarina (IFSC), Lages, SC, Brasil.

andres.ferrero@ifsc.edu.br, vania.bogorny@ufsc.br

Abstract. *Trajectory data analysis and mining has been largely studied in the past years. Although trajectories are multidimensional data, that have space, time, and semantic information, only a few works on the literature have considered all three dimensions. Indeed, we claim in this paper that not only the three dimensions must be considered in trajectory data analysis, but that trajectories can be represented from different points of view, that we call* multiple aspect representation. *Existing works, in general, are limited to one single trajectory representation, what narrows the discovery of several types of interesting patterns. In this vision paper we show that there is a need for a change of paradigm in trajectory data analysis, and present new research challenges in movement analysis.*

1. Introduction and Motivation

We are living the era of movement tracking and mining, where huge volumes of data about our daily lives are being collected and stored in several sources and formats. Examples include our smartphones, from which Google and Apple collect all details about our daily routines, including the places we visit and the time we stay there. Facebook captures our location, stores our friendship relationships, as well as our thoughts and opinions about things and people. More recently, the Pokémon GO emerges to capture not only our movement, but photos of places we visit when capturing Pokémons, what certifies with a high accuracy where we are. In summary, when an individual is moving, the application collects his/her location over time, in the form of sequential spatio-temporal points, called *raw trajectories*, as shown in Figure 1 (left). A raw trajectory is a complex data type, which has *space* and *time* information associated with each trajectory point.

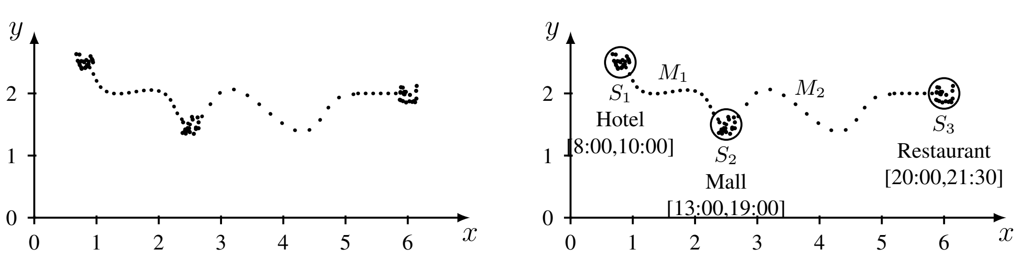

Figure 1. (left) Example of a raw trajectory T, and (right) the corresponding semantic trajectory.

Raw trajectory is the most simple trajectory representation. Since 2007, because of the explosion of social networks (e.g. Foursquare, Facebook) and navigation applications (e.g. Google Maps, Waze), raw trajectories are being enriched with more information, such as the name of the place visited by an object, called Point of Interest (POI), and the amount of time the individual stayed at each POI. With more information associated to raw trajectories, a new representation is defined, the *semantic trajectories* [Spaccapietra et al. 2008, Alvares et al. 2007, Parent et al. 2013]. In other words, the movement is a sequence of stops (visited places) and moves (spatio-temporal points between stops) [Spaccapietra et al. 2008]. An example is shown in Figure 1 (right). With the new representation of trajectories, movement becomes a more complex data type, having now more dimensions to be considered: *space, time,* and *semantics*.

More recently, in 2014, Bogorny [Bogorny et al. 2014] proposed a new trajectory representation model, where the same trajectory can be represented according to several aspects. In other words, the same trajectory can have *multiple aspect representations*. For instance, a raw trajectory can be represented as a sequence of stops and moves, a sequence of transportation means used during the movement, the sequence of weather conditions, the sequence of activities performed during the movement, and so on. In another recent work [Noël et al. 2015], the authors propose a semantic trajectory data model composed of multiple aspects, which are all different points of view from which a trajectory can be observed. They apply the model for life trajectories considering several high level aspects as residential, professional, and familial, where each aspect contains specific information. Although the works of [Bogorny et al. 2014] and [Noël et al. 2015] address the need for multiple aspects, they are limited to a model for multiple aspect representation, and do not present the challenges related to multiple aspect trajectory data analysis and mining.

To the best of our knowledge, there is no work in the literature on trajectory data mining and similarity analysis that considers different representations for a single trajectory. So far, existing works consider either raw trajectories or semantic trajectories in the form of stops and moves. Most of the existing works do not even consider the three dimensions of space, time and semantics. In similarity analysis, for instance, only the work of [Furtado et al. 2016] considers all three dimensions, while works as [Vlachos et al. 2002, Chen et al. 2005, Ying et al. 2010, Pelekis et al. 2012, Liu and Schneider 2012] consider only two dimensions as space and time, or space and semantics. In trajectory clustering, for instance, only the single raw trajectory representation has been considered [Lee et al. 2007, Abul et al. 2010, Hung et al. 2015] and a few works are starting semantic trajectory mining as [Pelekis et al. 2011, Lv et al. 2013, Xiao et al. 2014, Ying et al. 2014, Wu et al. 2015, Cai et al. 2016].

By considering only one aspect in movement analysis and mining, several types of interesting patterns cannot be discovered. For instance, how do individuals move when it is raining, by car, bike, or public transportation? Do groups of friends visit specific places only by bus and with good weather? How do weather conditions affect traffic jams? Which is the transportation pattern at a beach town on a rainy weekend and a sunny weekend? In this paper we analyze some existing works and show their limitations related to multiple aspects trajectory representation, and that there is a need for a change of paradigm on trajectory analysis and mining, which is not limited to only a few trajectory attributes.

The rest of this paper is organized as follows: Section 2 presents a multiple aspect trajectory data analysis, that introduces the similarity analysis on multiple aspects. It also presents a comparative study of similarity measures proposed in the literature and their limitations, and the need of new proposals to multiple aspect trajectory data mining. In Section 3 we present a discussion of our vision about the future in trajectory data analysis methods.

2. Multiple Aspects Trajectory Data Analysis

Let us consider three trajectories, P, Q, and R, shown in Figure 2. These trajectories can be represented as, for instance, four different aspects: as raw trajectories, in Figure 2(a); as stops and moves (Figure 2(b)), where the labeled parts are the stops; as transportation means (Figure 2(c)); and according to weather conditions (Figure 2(d)).

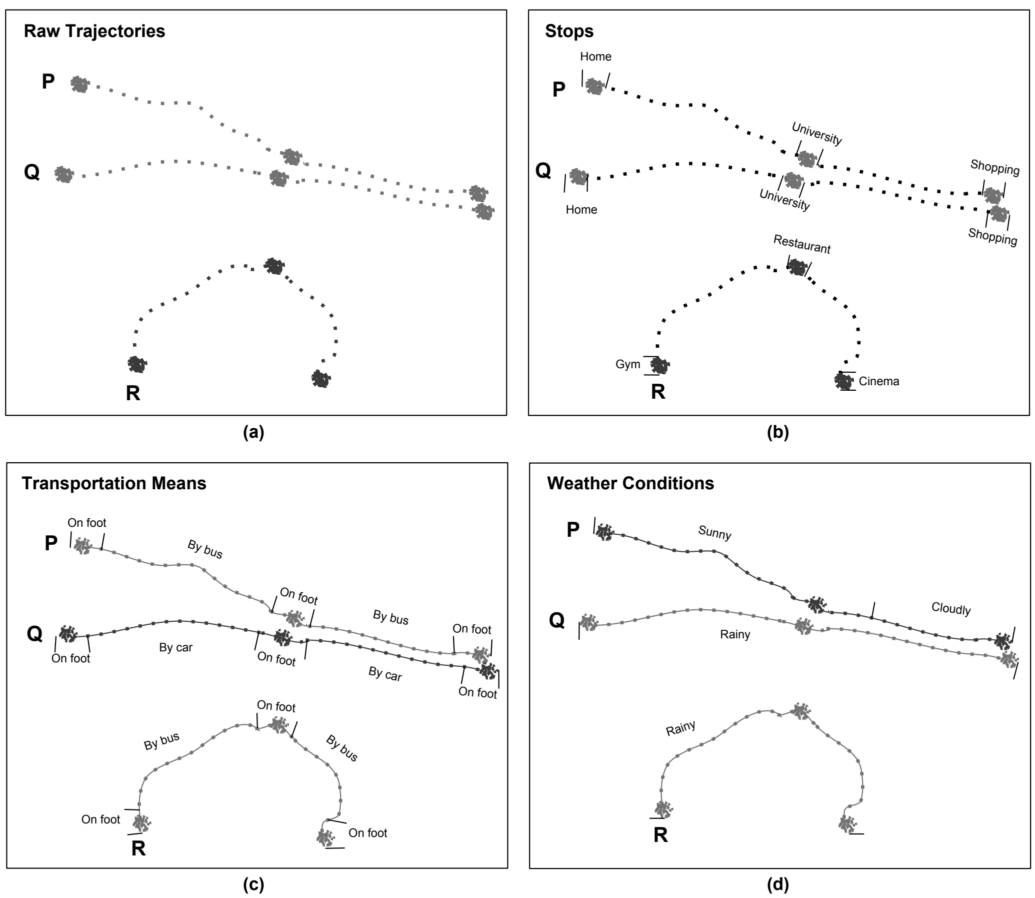

Figure 2. Multiple trajectory representation.

By considering every aspect separately, and considering only the *space* and *semantic* dimensions, excluding time for simplification, trajectories would be analyzed as follows:

Raw Trajectories: from this aspect (Figure 2(a)), trajectories P and Q are spatially closer than P and R or Q and R. For any spatial point of P, the closest spatial point is always from Q, not from R, and vice versa. So P and Q are the most similar.

Stops: from this aspect (Figure 2(b)), trajectories P and Q visit the same POI types $Home$, $University$, and $Shopping$, and in the same order. On the other hand, trajectory R visits different POI types (Gym, $Restaurant$, and $Cinema$). So from the Stops and moves aspect trajectories P and Q are the most similar.

Transportation Means: from this aspect, P moves $On\ foot$ and $By\ bus$ while Q moves $On\ foot$ and $By\ car$. Trajectory R uses exactly the same transportation means of P $On\ foot$ and $By\ bus$, and in the same sequence. So in this aspect, trajectories P and R are the most similar.

Weather Conditions: from this aspect (Figure 2(d)), P and Q occur at different weather conditions: P under $Sunny$ and $Cloudy$, and Q under $Rainy$. On the other hand, trajectories Q and R occur under the same weather condition, $Rainy$. This is possible because P and Q occur at different days. Then, Q and R are the most similar.

Figure 3 summarizes the similarity analysis for every aspect for the trajectories P, Q, and R. For instance, P and Q have higher similarity for the aspects ***Raw Trajectory*** and ***Stops*** than for the other two aspects. Trajectories P and R show higher score for the aspect ***Transportation Means***, while Q and R are more similar on ***Weather Conditions***.

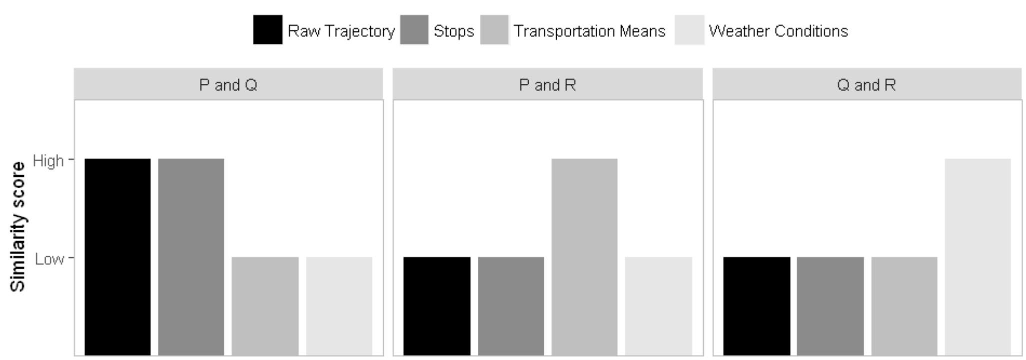

Figure 3. Similarity scores between pairs of trajectories, considering the aspects: raw data, stops, transportation means, and weather conditions.

In the following sections we analyze the multiple aspect trajectory data representation from two perspectives: similarity analysis and trajectory data mining.

2.1. Multiple Aspect Similarity Analysis

In this section we compare several trajectory similarity measures, applied to the trajectories of Figure 2, which mainly consider two aspects: raw trajectories and stops. Table 1 shows the results. This table shows the aspect considered by each approach and which dimensions are taken into account, where Ti represents time, Sp represents space and Se represents semantics.

The measures DTW [Vlachos et al. 2002], LCSS [Vlachos et al. 2002], and EDR [Chen et al. 2005] consider only raw trajectories, specifically the spatial dimension. DTW returns the distance between points, so the closest trajectories are P and Q. LCSS

Table 1. Similarity Values of existing approaches.

		Raw Data		**Stops**			**Transp. Means**			**Weather Cond.**			
#	**Measure**	Ti	Sp	Ti	Sp	Se	Ti	Sp	Se	Ti	Sp	Se	The Most Similar Trajectories
1	DTW Distance [Vlachos et al. 2002]		✓										P and Q
2	LCSS Ratio [Vlachos et al. 2002]		✓										P and Q
3	EDR Ratio [Chen et al. 2005]		✓										P and Q
4	MSTP [Ying et al. 2010]					✓							P and Q
5	[Liu and Schneider 2012]				✓	✓							P and Q
6	[Lv et al. 2013]			✓		✓							P and Q
7	MTM [Xiao et al. 2014]			✓		✓							P and Q
8	MSM [Furtado et al. 2016]			✓	✓	✓							P and Q
	Multiaspect Similarity Measure	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	
	Raw Data	✓	✓										P and Q
	Stops			✓	✓	✓							P and Q
	Transp. Means						✓	✓	✓				P and R
	Weather Cond.									✓	✓	✓	Q and R

also considers trajectories P and Q as the most similar, because this measure uses the sequence of nearest spatial points between two trajectories to calculate their similarity. EDR also considers trajectories P and Q as the most similar, because the cost to transform P into Q is lower than the cost to transform P in R or Q in R. The similarity measures proposed by Ying [Ying et al. 2010], Liu [Liu and Schneider 2012], Lv [Lv et al. 2013], Xiao [Xiao et al. 2014], and the MSM [Furtado et al. 2016] were developed for the stops and moves representation, so they all consider the semantic dimension. Therefore, all these measures return a high similarity for P and Q, and low similarity for P and R, and Q and R.

As can be seen in the last rows of Table 1, the multiple aspect similarity measures should first consider all three dimensions: space, time, and semantics. By considering all dimensions and more aspects, not only trajectories P and Q would be similar, but also P and R and Q and R, depending on the aspect(s) considered. From a similarity analysis point of view, we have two problems: first, the trajectories have several *aspects* to be considered, such as raw data, stops, transportation means, activities, weather conditions, and others; second, existing measures consider only one representation, either raw trajectories or stops and moves. In the following section we show how the multiple aspect trajectory data analysis can lead to new types of trajectory patterns that cannot be detected so far by existing data mining methods.

2.2. Multiple Aspect Trajectory Data Mining

In this section we show two simple examples of interesting patterns that can only be discovered when considering trajectories over multiple aspects, not simply stops and moves or raw trajectories.

Figure 4(a) shows several raw trajectories, where the moving objects are traveling from region A to region B. By applying a clustering technique over these raw trajectories we can obtain three clusters, as shown in Figure 4(b), since most works as [Lee et al. 2007, Lee et al. 2008, de Vries and van Someren 2010, Liu et al. 2010, Liu et al. 2013, Chen et al. 2014] consider proximity in *space*. If we consider the *stops* representation, as all trajectories have the same two stops, first A and after B, probably they will be in the same cluster, as shown in Figure 4(c).

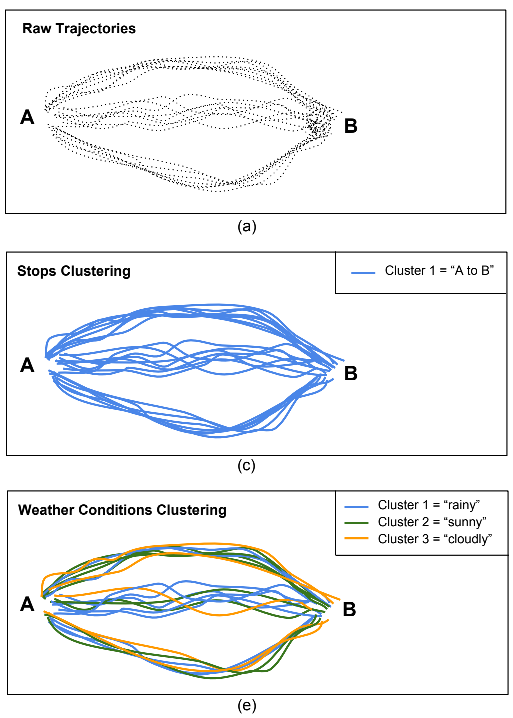

Figure 4. (a) raw trajectories between A and B, (b) clusters of raw trajectories, (c) clustering under stops representation, (d) clustering by transportation means, and (e) clustering by weather conditions.

Now let us consider the aspects that have not been considered so far, as transportation means. With this representation we can obtain two clusters, as shown in Figure 4(d), one cluster at the center that represents objects moving *by car*, and the second cluster (in blue) objects moving *by bus*, indicating that there are two different bus lines to move from region A to region B. Another possible clustering is the one shown in Figure 4(e), where the trajectories were grouped by weather conditions. Notice that there are three clusters, corresponding to *rainy*, *sunny*, and *cloudy*. Existing clustering methods would not be able to detect these clusters, since most of them consider the spatial distance as similarity measure.

Considering the examples in Figure 4, by mixing the different representations, different clusters could be generated. Besides, we could also include many other variables in the clustering analysis, as the duration time and the average speed of each transportation means, or the duration and the length of each weather condition, and so on.

For other data mining techniques, as classification, several aspects besides raw trajectories and stops are very important. And more interesting is the combination of different aspects in the same mining step. For instance, if we want to distinguish two classes of transportation means: *on foot* and *not on foot*, a classification algorithm could find patterns as:

- $if\ (temperature\ >\ 35\ degrees)\ then\ class = \text{"}not\ on\ foot\text{"}$
- $if\ (weather = \text{"}rainy\text{"}\ and\ length > 1000m)\ then\ class = \text{"}not\ on\ foot\text{"}$
- $if\ (weather = \text{"}sunny\text{"}\ and\ temperature < 30\ degrees\ and$ $length < 5000m)\ then\ class = \text{"}on\ foot\text{"}$

The weather conditions are also very important in the analysis of traffic jams since in most cities they are more frequent and stronger when it is raining. To the best of our knowledge, existing trajectory similarity and mining methods are limited to a few trajectory attributes, and one single aspect. None of the existing works have considered multiple aspects. In the following section we present a discussion about the challenges behind considering multiple aspects together for similarity analysis and trajectory data mining.

3. Research Challenges and Opportunities

In this section we present some major challenges on multiple trajectory representation analysis and how they lead to new research opportunities. Tables 2, 3 and 4 give an overview of the complexity related to multiple aspect trajectory data analysis. Every table shows just a few examples of different features that can be extracted from trajectories over one single aspect. For instance, in Table 2 (the stops representation), information as stop duration, traveling time between stops, stop name, route followed between stops, etc, can be extracted. The complexity relies on the amount of information that can be obtained over each dimension: space, time, semantics, and the combination of dimensions, as for instance, the name and location of the stops with duration above 1 hour. This example refers to one aspect and three dimensions.

Now consider the combination of multidimensional features of three aspects, such as (i) the *average speed* of the moving object when traveling *by car* when it is *raining*; (ii) average traveled distance by car under rain; (iii) average traveled time by bus under rain; (iv) total traveled distance on foot in a sunny day.

As mentioned previously, existing works in the literature do not support multiple aspect trajectory data analysis. But one may ask why not simply include all aspects information into a unique trajectory representation? For instance, considering the *stops representation* one could argue that it is simple to load (enrich) the trajectory with all aspect information such as weather conditions, transportation means, activities, etc. The problem is not so simple. Let us suppose that during one stop the object is moving on foot

Table 2. Examples of information to be extracted for the aspect of stops.

Information	Dimensions
Stop duration	Time
Traveled time between the stops	Time
Traveled distance between the stops	Space
Geographical position of the stops	Space
Visited POI types	Semantics
Amount of time at each POI type	Time and Semantics
Name and location of the stops with duration greater than 1 hour	Time, Space and Semantics

Table 3. Examples of information to be extracted for the aspect of transportation means.

Information	Dimensions
Duration of each transportation means	Time
Traveled distance of each transportation means	Space
Type of transportation means	Semantics
Total duration of each transportation means	Time and Semantics
Distance traveled on foot	Space and Semantics
Average speed by car	Time, Space and Semantics

Table 4. Examples of information to be extracted for the aspect of weather conditions.

Information	Dimensions
Travel duration under rain	Time and semantics
Traveled distance under rain	Space and Semantics
Average speed under sun	Time, Space, and Semantics

and the weather condition changes from $Sunny$ to $Rainy$. And in another aspect as *Activity* the object changed its activity from $drinking\ coffee$ to $teaching$ when the weather is $Rainy$. All these things happening at one single stop with label *University*. It would be very hard to correctly split and annotate the stop into two weather conditions, each one having a different start and end time, splitting it in different activities. The same stop would have several semantic labels for weather, transportation means, stop name, activity name, etc, and several time intervals associated to each semantic label, as for instance, the start time and end time of a stop, duration of a transportation means, distance traveled by one transportation means, etc., and probably different space information as well.

There are six main challenges in multiple aspect trajectory data analysis:

Multiple Aspect Representation. The first point is how to compute different aspect information since it involves heterogeneous data sources. For instance, climate conditions can be taken from the web, stops labeling from Open Street Maps, transportation means from raw trajectories or manual annotation, activity inference from social networks, etc. While some aspects may be easy to represent and to obtain as weather conditions, others are more complex, as transportation means and activity inference, which by their own are still challenging open research fields. Each aspect has space, time, and semantic information as well as their combinations. The second point is how to represent trajectories with all this information.

Feature Extraction. New efficient algorithms must be developed for trajectory segmentation and feature extraction, considering each aspect and their combinations. The algorithms should handle different aspects and different segmentation forms, avoiding trajectory feature recomputation, mainly when dealing with large trajectory datasets.

Data Storage. New efficient algorithms are needed to store, besides raw trajectories, the information of different aspects and the extracted features. New indexes and data structures have to be proposed for efficiently storing and querying these complex data.

Similarity Analysis and Data Mining. Similarity measures and data mining algorithms do not yet consider all three dimensions (space, time and semantics) or are still limited to raw trajectories or stops and moves representation. There is a need for similarity measures that are not limited to a predefined set of variables, and which do not only give a similarity score to express how similar two trajectories are. Note that a low global similarity can hide a strong similarity for a specific aspect. New data mining algorithms are needed to find more complex patterns than a group of objects moving together in space and time or that visit similar places. Algorithms should infer if the objects moving together have a relationship, how much moving objects are aware of each other in the group, and how the movement of a single individual influences the group.

Visualization. Tools to visualize trajectories from different aspects, and their information, is crucial. It is important to note that each aspect has, in general, time, space and semantics dimensions, and the relationship between these dimensions should be treated by visualization approaches. In addition, there is a need of visualization techniques to show the patterns found in data mining tasks, to make them easier to evaluate and validate.

Privacy Protection. Using more information about the moving objects is a crucial problem related to the privacy preserving of users and protecting their sensitive information. Multiple aspects reveal more details about users, so privacy preserving data mining methods become more challenging.

We believe that new methods and tools are needed to simultaneously process multiple aspect information. This will lead to a new era in movement data processing and to the discovery of more complex and interesting patterns which have not been addressed so far. We believe that new similarity measures will need to output not only a single number that represents the similarity degree of a pair of objects considering two or three features, but in which aspect the movement of individuals is more or less similar. These measures will allow answering questions as: (i) In which aspect two trajectories T1 and T2 are more similar? (ii) In which aspect two trajectories are less similar? (iii) In which aspect two trajectories T1 and T2 have a similarity degree higher than δ? (iv) Which trajectories are more similar in a given aspect α? (v) Which trajectories are more similar considering all aspects?

In summary, we strongly believe that multiple aspect representation is a big issue in future trajectory data analysis and a challenge for researchers to develop new concepts and methods in this promising area.

Acknowledgment

This work was supported by CNPq, UFSC and IFSC-Lages.

References

Abul, O., Bonchi, F., and Nanni, M. (2010). Anonymization of moving objects databases by clustering and perturbation. *Information Systems*, 35(8):884–910.

Alvares, L. O., Bogorny, V., Kuijpers, B., de Macedo, J. A. F., Moelans, B., and Vaisman, A. (2007). A model for enriching trajectories with semantic geographical information. In *Proceedings of the 15th ACM SIGSPATIAL*, pages 22:1–22:8, New York, NY, USA. ACM.

Bogorny, V., Renso, C., de Aquino, A. R., de Lucca Siqueira, F., and Alvares, L. O. (2014). CONSTAnT - A Conceptual Data Model for Semantic Trajectories of Moving Objects. *Transactions in GIS*, 18(1):66–88.

Cai, G., Lee, K., and Lee, I. (2016). Discovering common semantic trajectories from geo-tagged social media. In *Proceedings of the 29th International Conference on Industrial Engineering and Other Applications of Applied Intelligent Systems: Trends in Applied Knowledge-Based Systems and Data Science*, pages 320–332, Cham. Springer International Publishing.

Chen, L., Özsu, M. T., and Oria, V. (2005). Robust and fast similarity search for moving object trajectories. In *Proceedings of the 2005 ACM SIGMOD international conference on Management of data (SIGMOD '05)*, pages 491–502, New York, NY, USA. ACM.

Chen, Y., Shen, H., and Tian, H. (2014). Clustering subtrajectories of moving objects based on a distance metric with multi-dimensional weights. In *Proceedings of the Sixth International Symposium on Parallel Architectures, Algorithms and Programming*, pages 203–208.

de Vries, G. and van Someren, M. (2010). *Clustering Vessel Trajectories with Alignment Kernels under Trajectory Compression*, pages 296–311. Springer Berlin Heidelberg, Berlin, Heidelberg.

Furtado, A. S., Kopanaki, D., Alvares, L. O., and Bogorny, V. (2016). Multidimensional similarity measuring for semantic trajectories. *Transactions in GIS*, 20:280–298.

Hung, C.-C., Peng, W.-C., and Lee, W.-C. (2015). Clustering and aggregating clues of trajectories for mining trajectory patterns and routes. *The VLDB Journal*, 24(2):169–192.

Lee, J.-G., Han, J., Li, X., and Gonzalez, H. (2008). Traclass: Trajectory classification using hierarchical region-based and trajectory-based clustering. *VLDB*, 1(1):1081–1094.

Lee, J.-G., Han, J., and Whang, K.-Y. (2007). Trajectory clustering: A partition-and-group framework. In *Proceedings of the 2007 ACM SIGMOD International Conference on Management of Data*, SIGMOD '07, pages 593–604, New York, NY, USA. ACM.

Liu, H. and Schneider, M. (2012). Similarity measurement of moving object trajectories. In *Proceedings of the 3rd ACM SIGSPATIAL International Workshop on GeoStreaming (IWGS)*, pages 19–22, New York, NY, USA. ACM.

Liu, S., Liu, Y., Ni, L. M., Fan, J., and Li, M. (2010). Towards mobility-based clustering. In *Proceedings of the 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD)*, pages 919–928, New York, NY, USA. ACM.

Liu, S., Wang, S., Jayarajah, K., Misra, A., and Krishnan, R. (2013). Todmis: Mining communities from trajectories. In *Proceedings of the 22Nd ACM International Conference on Information & Knowledge Management (CIKM)*, pages 2109–2118, New York, NY, USA. ACM.

Lv, M., Chen, L., and Chen, G. (2013). Mining user similarity based on routine activities. *Information Sciences*, 236:17–32.

Noël, D., Villanova-Oliver, M., Gensel, J., and Le Quéau, P. (2015). Modeling semantic trajectories including multiple viewpoints and explanatory factors: Application to life trajectories. In *Proceedings of the 1st International ACM SIGSPATIAL Workshop on Smart Cities and Urban Analytics*, UrbanGIS'15, pages 107–113, New York, NY, USA. ACM.

Parent, C., Spaccapietra, S., Renso, C., Andrienko, G. L., Andrienko, N. V., Bogorny, V., Damiani, M. L., Gkoulalas-Divanis, A., de Macêdo, J. A. F., Pelekis, N., and Yan, Z. (2013). Semantic trajectories modeling and analysis. *ACM Computing Surveys (CSUR)*, 45(4):42.

Pelekis, N., Andrienko, G., Andrienko, N., Kopanakis, I., Marketos, G., and Theodoridis, Y. (2012). Visually exploring movement data via similarity-based analysis. *Journal of Intelligent Information Systems*, 38(2):343–391.

Pelekis, N., Kopanakis, I., Kotsifakos, E. E., Frentzos, E., and Theodoridis, Y. (2011). Clustering uncertain trajectories. *Knowledge Information Systems*, 28(1):117–147.

Spaccapietra, S., Parent, C., Damiani, M. L., de Macedo, J. A., Porto, F., and Vangenot, C. (2008). A conceptual view on trajectories. *Data and Knowledge Engineering*, 65(1):126–146.

Vlachos, M., Kollios, G., and Gunopulos, D. (2002). Discovering similar multidimensional trajectories. In *Proceedings of the 18th International Conference on Data Engineering*, pages 673–684, San Jose, CA, USA. IEEE.

Wu, X., Zhu, Y., Xiong, S., Peng, Y., and Peng, Z. (2015). A new similarity measure between semantic trajectories based on road networks. In *Proceedings of the 17th Asia-Pacific Web Conference*, pages 522–535, Guangzhou, China. Springer International Publishing.

Xiao, X., Zheng, Y., Luo, Q., and Xie, X. (2014). Inferring social ties between users with human location history. *Journal of Ambient Intelligence and Humanized Computing*, 5(1):3–19.

Ying, J. J.-C., Lee, W.-C., and Tseng, V. S. (2014). Mining geographic-temporal-semantic patterns in trajectories for location prediction. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 5(1):2:1–2:33.

Ying, J. J.-C., Lu, E. H.-C., Lee, W.-C., Weng, T.-C., and Tseng, V. S. (2010). Mining user similarity from semantic trajectories. In *Proceedings of the 2nd ACM SIGSPATIAL*, pages 19–26, New York, NY, USA.