1 Introduction

The growing use of GPS receivers and WIFI embedded mobile devices equipped with hardware for storing data enables us to collect a very large amount of data, that has to be analyzed in order to extract any relevant information. The complexity of the extracted data makes it a difficult challenge. Trajectory clustering is an appropriate way of analyzing trajectory data, and has been applied to pattern recognition, data analysis, machine learning, etc. Additionally, trajectory clustering is used to gather temporal spatial information in the trajectory data and is widespread used in many application areas, such as motion prediction Chen et al. (2010) and traffic monitoring Atev et al. (2006)

Trajectory data is recorded in different formats depending on the type of device, object motion, or even purpose. In certain specific circumstances, other object-related properties such as direction, velocity or geographical information are added [3, 4]. This kind of multidimensional data is prevalent in many fields and applications, for example, to understand migration patterns by studying trajectories of animals, predict meteorology with hurricane data, improve athlete's performance, etc. Given different types of analysis tasks and moving object data applications, calculating the distance between moving object trajectories is a common technique for most tasks and applications. Therefore, distances are a fundamental component of those tasks and applications of trajectory analysis, allowing us to determine effectively how close two trajectories are. Unlike other simple data types, however, such as ordinal variables or geometric points where the distance description is straightforward, the distance between the trajectories must be carefully defined to represent the true underlying distance. It is because trajectories are basically high-dimensional data attached to both spatial and temporal attributes which need to be considered for distance measurements. As such the literature contains dozens of distance measurements for trajectory data. For example, distance measurements measure the sequence-only distance between trajectories, such as Euclidean distance and Dynamic Time Wrapping Distance (DTW); there are trajectory distance measurements measure both spatial and temporal dimensions of two trajectories as well. In order to extract useful patterns from high-volume trajectory data, different methods, such as clustering and classification, are usually used. Clustering is an unsupervised learning method that combines data in groups (clusters) based on distance [5, 6]. The

aim of trajectory clustering is to categorize trajectory datasets in cluster groups based on their movement characteristics. The trajectories existing in each cluster have similar characteristics of movement or behavior within the same cluster and are different from the trajectories in other clusters [7, 8, 9].

In general, two main approaches can be used for clustering complex data such as trajectories. First, identify trajectory-specific clustering algorithms and second, use generic clustering algorithms that use trajectory-specific distance functions (DFs). Measurement of similarity is the central focus of the clustering problem; thus, similarity (inverse distance) should be calculated prior to grouping. Distance definition in spatial trajectories is much more complicated than point data. Trajectories are sequences of points in several dimensions that are not of the same length. Thus, in order to compare two trajectories, a comprehensive approach is needed to fully determine their distance. Depending on the analysis purpose and also the data type, different DFs are presented. The concept of similarity is domain specific, so different DFs is defined in order to address different aspects of similarity such as spatial, spatio-temporal, and temporal. Spatial similarity is based on spatial parameters like movement path and its shape whereas spatio-temporal similarity is based on movement characteristics like speed, and temporal similarity is based on time intervals and movement duration. For instance, in order to extract the movement patterns in trajectories like detecting the transportation mode, besides the trajectory's geometry, their movement parameters should also be considered. Among all defined Distance Functions so far, Euclidean, Fréchet, Hausdorff, DTW, LCSS, EDR, and ERP distances are the basic functions in similarity measurements from which so many other functions are generated.

1.1 Predecessor

The template was developed sources provided at http://www.uef.fi/en/cs/latex-pohja

2 Making a Bibliography

References

Atev, S., Masoud, O., & Papanikolopoulos, N. (2006). Learning traffic patterns at intersections by spectral clustering of motion trajectories. In *2006 ieee/rsj international conference on intelligent robots and systems* (pp. 4851–4856).

Chen, Z., Shen, H. T., Zhou, X., Zheng, Y., & Xie, X. (2010). Searching trajectories by locations: an efficiency study. In *Proceedings of the 2010 acm sigmod international conference on management of data* (pp. 255–266).