

A Method for extracting multi-scale stay feature of trajectory based on OPTICS¹

Liu Minshi^{a, b}, Long Yi^{a, *}, Zhang Ling^a, He guifang^b

^a School of Geography Science, Key Laboratory of Virtual Geographic Environment of Ministry of Education, Nanjing Normal University, Nanjing, China, liuminshi1983@126.com, Long Yi, longyi@njnu.edu.cn, Zhang Ling, zhangling@njnu.edu.cn.

^b School of Geographic Information and Tourism, Chuzhou University, Chuzhou, China

* Corresponding author: longyi@njnu.edu.cn, Tel.: +86-138-5182-3107

Keywords: stay feature, multi-scale, OPTICS method, moving trajectory

Abstract:

With the popularity of mobile positioning devices, mobile trajectory data is rapidly increasing. The trajectory stay feature extraction is to find the staying part of the moving object from the moving trajectory, which is the basis of the semantic matching of the staying segment. The trajectory stay segment extraction can be used to analyze tourist visit recommendation and mode, special landmark extraction such as gas station, destination prediction, rental car passenger advice, and even traffic guidance.

Most existing researches extract the single-scale stay feature of the trajectory through the spatiotemporal characteristics of the trajectory or the geographical environment. However, the stay feature of the trajectory is multi-scale. For example, at a city scale, the stay of a tourist trajectory is composed of several scenic spots; however, at the scenic scale, a scenic stop may contain a number of subtle attraction. Therefore, this paper proposes to extract trajectory multi-scale stay feature based on OPTICS method. Since the trajectory data is different from the conventional discrete point set, but an ordered point string, the OPTICS method needs to be modified to accommodate the clustering of the trajectory point string. Method improvements include the following four aspects:

(1) The correction of the two-point distance definition of the trajectory. Since the trajectory points are ordered, the distance between the two points of the trajectory is not the linear distance between the two points, but the sum of the lengths of the trajectory polylines between the two points.

(2) The count calculation of points in the ϵ neighborhood of the center points needs to be improved. When the time interval of the trajectory is equal, the calculation method is consistent with OPTICS; and when the time interval of the trajectory sampling is not equal, the number of points represented by one trajectory points should be related to its time interval. When the time interval is large, the point should represent more points. Therefore, a time coefficient is defined, which is defined as the ratio of the track point time interval to the basic time interval. The total number of the points is the sum of time coefficient m of each points. By improving as above, the algorithm can be adapted to trajectory data sampled at unequal time intervals. In addition, when searching for the field of points, it is not necessary to search the entire set of trajectory points, but only need to start with the center point, and search forward and backward until the distance between the two points is greater than the threshold ϵ .

(3) Improvement of clustering order method: When generating the clustering order, the OPTICS needs to center the current point in order, and correct the reachable distance value of each point in the ϵ neighborhood of the point, then the point at the minimum distance is selected as the subsequent point. For the trajectory data, the subsequent point on the trajectory is obviously the point which the reachable distance is the shortest in the ϵ neighborhood of the current point, and so the clustering order is consistent with the trajectory points sequence. Therefore, the improved algorithm only needs to calculate the reachable distance of a subsequent point of the current point, thereby quickly generating a clustering sequence.

(4) The scale of the staying feature is automatically controlled by reachable distance. Although the OPTICS is capable of generating multi-level clustering results, it does not measure the hierarchy. In this paper, the scale of each stay feature is measured by the reachable distance range of a cluster, and the relationship between the reachable distance and the stay feature scale is established, so that the multi-scale stay feature can be obtained automatically.

This paper compares this method with other methods using a variety of trajectory data. The results show that the method can extract the track stay feature faster and more accurately, and can automatically control the stay feature extraction level.

¹ The research introduced in this paper is supported by the National Natural Science Foundation of China (grant No.: 41601499)