A Trajectory Compression Method Based on Fréchet Distance

Fan Jiang¹, Yan Zhou¹, Mengdou Qin¹ and Xu Wang¹

¹School of Resources and Environment, University of Electronic Science and Technology of China, 2006 Xiyuan Avenue, West Hi-tech Zone, Chengdu, Sichuan 611731, China.
Email: 201852070342@std.uestc.edu.cn

Abstract: To better maintain the geometric similarity between the compressed trajectory and the original trajectory, this paper proposes a trajectory compression method based on Fréchet distance. This method selects the characteristic points of the trajectory based on the Fréchet distance, and keeps the similarity of the geometric shape characteristics of the compressed trajectory and the original trajectory. Compared with the traditional DP (Douglas-Peucker) algorithm and the TD-TR (Top-Down Time-Ratio) algorithm, the method in this paper can obtain better geometric similarity between the compressed trajectory and the original trajectory. At the same time, better visualization of trajectory data is obtained.

1. Introduction

In recent years, with the development of technologies such as tracking detection and network positioning, the application of location-based services has become more and more widespread, resulting in massive trajectory data carrying important location information. How to effectively manage large-scale trajectory data has become an increasingly serious issue. Obviously, trajectory compression methods is an important way to solve the problem. Trajectory compression methods are technically divided into three categories [1]: 1) Trajectory compression based on feature point extraction, such as DP algorithm [2], which is to use the vertical Euclidean distance as an error measurement method and replace the original trajectory with several line segments to achieve the purpose of trajectory compression; TD-TR algorithm [3], it is a top-down compression algorithm based on speed and time ratio; 2) Trajectory compression based on the road network structure, Such as PRESS (Paralleled Road-Network-Based Trajectory Compression) [4], to effectively compress trajectory data under road network constraints, and proposes a novel representation for trajectories to separate the spatial representation of a trajectory from the temporal representation; Nonmaterialized algorithm [5], it uses the sequence of intersections traveling in the compressed trajectory to express the original trajectory, and the compression rate has been greatly improved; 3) Semantic-based trajectory compression: STC (Semantic Trajectory Compression) algorithm [6], this method achieves its compression rate by replacing raw, highly redundant position information from, for example, GPS sensors with a semantic representation of the trajectory consisting of a sequence of events.

Most of the existing trajectory compression methods meet the demand of data compression concerning the data compression ratio, however, there is still room for improvement in the geometry features of keeping the trajectory compression before and after. Therefore, this paper proposes a trajectory compression method based on Fréchet distance, which uses Fréchet distance to select trajectory feature points, which can not only ensure more efficient compression, but also ensure the temporal and spatial similarity of compressed trajectories as much as possible.
2. Trajectory Compression Method Based on Fréchet Distance

2.1. Definitions

Define 1: In Fréchet space \([7]\), for any two curves \(f : [0,1] \rightarrow \mathbb{R}^2\) and \(g : [0,1] \rightarrow \mathbb{R}^2\), Fréchet distance can be expressed as \([8]\):

\[
\delta(f,g) = \inf_{\alpha,\beta} \max_{t \in [0,1]} \|f(\alpha(t)) - g(\beta(t))\| \tag{1}
\]

Where \(\|\|\) represents the Euclidean distance; \(\inf\) represents the lower bound of the set; \(\alpha, \beta\) refers to the continuous non-decreasing real function of the parameter \(t\) established after the curve is parameterized, and satisfies \(\alpha(0) = \beta(0) = 0\) and \(\alpha(1) = \beta(1) = 1\).

Define 2: Two or more consecutive trajectory points on the trajectory \(Tr = (P_1, P_2, \ldots, P_n)\) form a sub-trajectory \(tr = \{P_m, P_{m+1}, \ldots, P_k\}\) (\(1 \leq m < k \leq n\)), in time sequence, where the approximate trajectory of the sub-trajectory refers to the line segment \(P_mP_k\) connecting the starting point \(P_m\) and the ending point \(P_k\) of the sub-trajectory.

Define 3: The total description length \(L_{total}\) of the trajectory refers to the time-space relationship between the current sub-trajectory \(tr = \{P_m, P_{m+1}, \ldots, P_k\}\) and the approximate trajectory \(L(simtr)\) obtained from its mapping, and is composed of the description length \(L(simtr)\) of the approximate trajectory and the space-time error \(L(tr|simtr)\) between them. The expression is as follows:

\[
L_{total}(P_m, P_k) = L(simtr) + L(tr|simtr) \tag{2}
\]

The description length \(L(simtr)\) of the approximate trajectory is expressed as follows:

\[
L(simtr) = \log_2(d(p_m, p_k)) \tag{3}
\]

\[
d(p_m, p_k) = \sqrt{\left(x_m - x_k\right)^2 + \left(y_m - y_k\right)^2} \tag{4}
\]

The space-time error between the sub-trajectory and the approximate trajectory is expressed as follows:

\[
L(tr|simtr) = \log_2(\delta_{\delta}(tr, simtr)) \tag{5}
\]

Define 4: The trajectory description length reference value \(L(tr)\) refers to the total trajectory description length value of the sub-trajectory (except the starting point) without other characteristic points, the expression is as follows:

\[
L(tr) = \log_2\left(\sum d(P_i, P_{i+1})\right), \quad i = m, \ldots, k-1 \tag{6}
\]

2.2. Proposed Method

Our proposed method uses the trajectory description length to define an effective trajectory data feature point extraction criterion, and applies the Fréchet distance as a distance measurement to the feature point extraction criterion. Compared with the Euclidean distance between line segments, this method can better constrain the similarity between the original trajectory and the compressed trajectory, and achieve the purpose of efficient compression of large-scale trajectory data. As shown in Figure 1, the judgment criteria of the trajectory feature points in the trajectory compression method based on the Fréchet distance are described as follows:

1. Initialize the sub-trajectory to be compressed. Transfer the newly generated trajectory points to the sub-trajectory to be compressed in sequence, among them, for each complete trajectory, the starting point is a significant trajectory point and must be retained.

2. Judge the similarity between the current sub-trajectory and the curve of the corresponding approximate trajectory. Set the corresponding threshold \(\lambda\), calculate the ratio \(\varepsilon\) of the reference value
$L(tr)$ of the trajectory description length of the sub-trajectory to the description length $L(simtr)$ of the approximate trajectory. If $\varepsilon > \lambda$, execute criterion (3) to further judge the characteristic points; otherwise, pass in the next trajectory point in time sequence and continue to execute step (2).

(3) Judgment and extraction of current sub-trajectory feature points. Compare the relationship between the total trajectory description length $L_{\text{cost}}$ and the trajectory description length reference value $L(tr)$ on the current sub-trajectory. When $L_{\text{cost}} > L(tr)$ is satisfied, that is, $L(simtr) + L(tr|simtr) > L(tr)$, mark a trajectory point before the end point of the current sub-trajectory as a feature point; otherwise, transfer the next trajectory point in time sequence, and continue to perform step (2).

(4) It’s not until the current trajectory point becoming the termination point of the original trajectory that the algorithm stops, among them, the termination point is an important trajectory point and must be retained. Meanwhile, the trajectory feature point is extracted and the compressed trajectory is output.

![Figure 1. Trajectory compression algorithm](image)

3. Experiment

3.1. Datasets

The experimental data in this article comes from the real GPS trajectory data of taxis in Chengdu of China in one day. There are 13493 taxi trajectories in Chengdu in one day. More than 91% of the trajectory data set is obtained by high-density sampling with a time interval of 1~5 seconds or a distance of 5~10 meters. The information contained in each trajectory is shown in Table 1:
3.2. Evaluation of Results
Select the two sets of trajectory data shown in Table 2 to perform trajectory compression experiments,

| Data  | Original trajectory points | Trajectory points after compression | Compression ratio |
|-------|----------------------------|------------------------------------|-------------------|
| Data1 | 1302                       | 170                                | 87%               |
| Data2 | 1522                       | 200                                | 85.2%             |

and the results are shown in Figure 2. The red points represent the original trajectory points, and the blue points represent the compressed trajectory points. Both experimental results show that the shape of the compressed trajectory is approximately the same as the original one. That is to say, it can retain the shape characteristics of the original trajectory.

To better evaluate the accuracy of the trajectory compression algorithm, the proposed algorithm is compared with DP algorithm and TD-TR algorithm. For the same trajectory data, under the same compression rate, the Average Euclidean Distance Error (AEDE) [9], Average Synchronous Euclidean Distance Error (ASE) [10], and Average Velocity Error (AVE) [11] and Average Direction Error (ADE) [12] are used to measure the accuracy of trajectory compression.

As shown in Figure 3(a), the AEDE of the proposed algorithm and the DP algorithm is basically the same to a certain extent. However, compared with TD-TR, the AEDE of the proposed algorithm is significantly smaller. As shown in Figure 3(b), compared with DP algorithm, the ASE of the proposed algorithm is smaller. Moreover, the ASE of the proposed algorithm is similar to that of the TD-TR algorithm. As shown in Figure 3(c), the AVE of the proposed algorithm is always lower than that of DP algorithm, and compared with TD-TR algorithm, the proposed algorithm has better performance in maintaining speed characteristics. As shown in Figure 3(d), compared with DP algorithm and TD-TR algorithm, the performance of the proposed algorithm is more stable on ADE, which means that the
Figure 3. Comparison of the algorithm, DP algorithm and TD-TR algorithm compression metrics.

The compressed trajectory of the proposed algorithm can more effectively retain the important directional characteristics of the original trajectory. Based on the comparison results of the above four compression metrics, compared with DP algorithm and TD-TR algorithm, the proposed algorithm not only maintains the shape characteristics of the trajectory well, but also has better performance in retaining the important velocity and direction characteristics of the trajectory data.
Take the original trajectory shown in Figure 4 as an example, and use the algorithm in this paper, the classic Douglas-Peucker algorithm and the TD-TR algorithm to conduct trajectory compression comparison experiments. Figure 4(b), Figure 4(c), and Figure 4(d) respectively show the geometric characteristics of the trajectory points before and after compression of the three algorithms when the compression ratio is set to 80%. Comparing the details of the compressed trajectory, it can be seen that the trajectories compressed by the algorithm in this paper, the DP algorithm and the TD-TR algorithm retain the basic shape characteristics of the original trajectory. From the details circled in orange in Figure 4(c) and Figure 4(d), we can see that, compared with DP algorithm and TD-TR algorithm, the geometric feature points of the proposed algorithm are more consistent with the shape features of the original trajectory.

4. Conclusions
Based on Fréchet distance to measure the geometric similarity of tracks before and after compression, a trajectory compression algorithm based on Fréchet distance is proposed in this paper. The experimental results under the real data set show that this method ensures the shape similarity between the compression trajectory and the original trajectory as far as possible, and has a better trajectory visualization effect compared with the traditional compression method under the condition of satisfying the compression ratio.
5. Acknowledgments
This work was supported by National Natural Science Foundation of China (No. 41871321, 41471332 and 41871314), the National Key Research and Development Program of China (No. 2018YFB0505501 and 2016YFB0502303).

6. References
[1] Jiang J W and Wang X L. Overview of trajectory compression[J]. Journal of East China Normal University (Natural Science Edition), 2015(5): 61-76
[2] Douglas H and Peucker T K. Algorithms for the Reduction of the Number of Points Required to Represent a Digitized Lineor Its Caricature[J]. Cartographica, 1973, 10(2):112–122
[3] Meratnia N and Rolf A. Spatiotemporal Compression Techniques for Moving Point Objects[J]. Advances in Database Technology-EDBT 2004 Volume 2992 of the series Lecture Notes in Computer Science Springer Berlin Heidelberg, 2004, 2992:765-782
[4] Song R, Sun W and Zheng B. PRESS: A Novel Framework of Trajectory Compression in Road Networks [J]. Proceedings of the Vldb Endowment, 2014, 7(9):661-672
[5] Cao H and Wolfson O. Nonmaterialized Motion Information in Transport Networks[C]. Database Theory-ICDT 2005, International Conference, Edinburgh, UK, January 5-7,2005, Proceedings. 2005:173-88
[6] Schmid F, Richter K F and Laube P. Semantic Trajectory Compression[J]. Symposium on Large Spatial Databases, 2009:411-416
[7] Brakatsoulas S, Pfoser D and Salas R. On map-matching vehicle tracking data[C]. Proceedings of the 31st VLDB Conference. 2005: 853-864
[8] Alt H and Godau M. Measuring the Resemblance of Polygonal Curves[C]. Proceedings of the eighth annual symposium on Computational geometry. DBLP, 1992
[9] Meng Q B. Research on Contour Maintaining Oriented Trajectory Compression Algorithm [D]. Dalian University of Technology. 2017
[10] Wu Q H. Online GPS trajectory compression algorithm based on relatively synchronized Euclidean distance filtering[J]. Computer Applications and Software, 2018(3):282-288
[11] Jiang J W. Compression technology and system implementation of trajectory [D]. East China Normal University, 2016
[12] Yao Y. Research on trajectory compression algorithm for moving objects [D]. Lanzhou Jiaotong University, 2017