

A New Measurement Method to Calculate Similarity of Moving Object Spatio-Temporal Trajectories by Compact Representation

Zhao XiuLi*

State Key Laboratory of Rail Traffic Control and Safety, Beijing, 100044, China;
School of Business, Shandong Polytechnic University, Jinan, 250353, China
School of Traffic and Transportation, Beijing Jiaotong University, Beijing, 100044, China

Xu WeiXiang

School of Traffic and Transportation, Beijing Jiaotong University, Beijing, 100044, China
E-mail: wxxu@bjtu.edu.cn
www.bjtu.edu.cn

Received 8 June 2011

Accepted 25 November 2011

Abstract

A new measurement method is proposed to calculate spatio-temporal trajectory similarity, which can reflect the similar degree between two moving object spatio-temporal trajectories compressed by the Maximal Bounding Boxes (MBB). Firstly, the similarity between two trajectories is replaced by the similarity of MBB sequences in respective trajectories which can dramatically decrease the storage volume of the trajectory data. Secondly, some factors affected the similar degree of MBB sequences are analyzed systematically, such as the time duration of overlap between two MBBs in different trajectories, space distance and the density of data points inside the boxes. And then, a similarity measurement formula is proposed by integrating these factors. Experiments show that the proposed measurement formula can improve the value of clustering index Dunn.

Keywords: Spatio-temporal data mining, moving object trajectories, trajectory similarity degree, similarity measurement formula

1. Introduction

With the integration of wireless communications and positioning technologies, massive data related to moving objects have been acquired and posed great challenges to the data mining community (Laube p. 2007; Roddick, J.F. *et al.* 2007). The need for spatio-temporal data mining and analysis techniques is growing. Some specific examples include managing cell phone networks, satellite tracking, dealing with the data generated by Radio Frequency Identification (RFID) Tags and designing intelligent transport system (Wuhong Wang *et al.* 2011, Xiaohua Zhao *et al.* 2011).

Mining such data can help to detect patterns for applications as diverse as intelligent traffic management, sensor networks, stock control and wild life monitoring. For example, considering the movement of users between cells of a mobile phone (or similar) network and being able to predict where large groups of users will go could make cell hand-over decisions easier or improve bandwidth management. Also, since most people own a mobile phone these days, the data could be used for fast and inexpensive population movement studies. Local governments could raise the ability to answer questions such as “how much is this park being used?”, “when and where

*Corresponding author. E-mail address: 07114212@bjtu.edu.cn

are the congested areas?” and “what are the popular routes that people take through the city?”. The latter query would help design better pedestrian and vehicle routes to take into account the main flows of people. These data have rich knowledge and great potential benefits, thus give rise to an upsurge of spatio-temporal data mining (Florian Verhein, Sanjay Chawla.2008).

Clustering moving object trajectories is an important subfield research of spatio-temporal data mining. Clustering produces rules according to the data self and makes this algorithm play a significant role in unknown areas. Clustering can give the users a basic category and a clear understanding of the massive and disorderly data as shown in Fig.1. (Miro Nanni, Dino

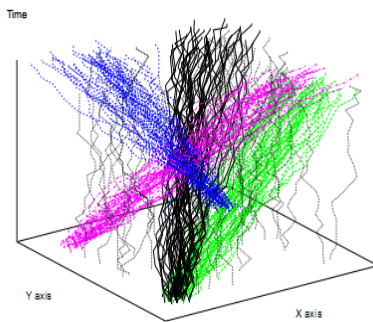


Fig.1. Illumination of trajectory clusters

Pedreschi. 2006).

Gaffney *et al.* have proposed a model-based clustering algorithm for trajectories. In this algorithm, a set of trajectories is represented using a regression mixture model. Then, unsupervised learning is carried out using the maximum likelihood principle. Specifically, the EM algorithm is used to determine the cluster memberships (Gaffney, S. *et al.*1999; Gaffney, S. *et al.*2006 ;).

Distance measures for searching similar trajectories have been proposed recently. The LCSS (the Longest Common Subsequence) distance measure is proposed by (Vlachos *et al.*2002) and the distance measure EDR (Edit Distance on Real Sequence) is proposed by (Chen *et al.*2005). Both LCSS and EDR are based on the edit distance and are extended so as to be robust to noises, shifts, and different lengths that occur due to sensor failures, errors in detection techniques, and different sampling rates. EDR can

represent the gap between two similar subsequences more precisely compared with LCSS. Besides, dynamic time warping has been widely adopted as a distance measure for time series (Keogh, E. J. 2002.). These distance measures, however, are not adequate for this problem since they are originally designed to compare the whole trajectory (especially, the whole time-series sequence). In other words, the distance could be large although some portions of trajectories are very similar. Hence, it is hard to detect only similar portions of trajectories.

According to the amount of trajectory to be clustered, trajectories clustering can be classified into two different categories: single trajectory clustering and multi-trajectories clustering. The former is aimed to extract interesting places in a single trajectory (Andrey Tietbohl Palma, 2008; S. Spaccapietra *et al.*, 2007). Especially, in the article of S. Spaccapietra *et al.*, they have introduced a new model for reasoning over trajectories, which allows powerful semantic analysis, called stops and moves. A stop is a semantically important part of a trajectory that is relevant for an application, where the object has stayed for a minimal amount of time. For instance, in a tourism application, a stop could be a tour place, a hotel, an airport, etc. In a traffic management application, important places can be traffic lights, roundabouts, parking places, etc. According to the application, the minimal stop duration can vary significantly. Thus the stops and the moves fully cover the trajectory (i.e., there is no data point that belongs to neither a move nor a stop).

The latter is interested in discovering the group patterns of moving objects such that the objects in the same group are geographically close to one another for significant amounts of time (S.Y.Hwang *et al.* 2005; Aris Anagnostopoulos *et al.* 2006). Each trajectory is represented by an id T_i and a chronological sequence of MBB. This problem is how to employ a clustering methodology to extract interesting patterns from the trajectory set. It can be expressed as follows: the clustering algorithm receives a set of trajectories T_i and outputs new trajectory clusters: $\{c_1, c_2, c_3, \dots\}$, $\{c_1 = \langle T_1, T_8, T_3 \rangle, c_2 = \langle T_5, T_7, T_9 \rangle, \dots\}$, trajectory cluster centroid c_1 containing three trajectories T_1, T_8, T_3 , and trajectory cluster centroid c_2 containing three trajectories T_5, T_7, T_9 . This article belongs to the latter.

Based on the method proposed by Sigal Elnekave *et al.* (Sigal Elnekave *et al.* 2007a) the raw moving object trajectories can be compressed into some Minimum Bounding Boxes (MBB), it needs a new similarity measurement formula to calculate the similarity between two trajectories which are compressed into MBB sequences. In this article a new similarity measurement formula is proposed to calculate the similarity between two trajectories which consider the density of data point in the MBBs. Then the MBB representation trajectories are clustered by integrated the similarity of time duration, space and motion characteristics. Finally experiments show that the similarity measurement proposed in this article has a high performance.

This paper is organized as follows: in Section 2 some related work on clustering moving object trajectories are briefly introduced; in Section 3 some factors are analyzed which impact the similarity measurement of MBB representation trajectories and then integrate these factors to form a new measurement formula. In Section 4 the proposed formula are applied to cluster the moving object trajectories, and finally some conclusions are summarized in Section 5.

2. Related Work

2.1. The Compact Representation of Moving Object Trajectories

The compact representation of the moving object trajectories can be grouped as follows.

One is mapping time to locations in which a trajectory is decomposed into a set of linear functions, one for each disjoint time interval (S.Y.Hwang *et al.*, 2005). The derivative of each linear function yields the direction and the speed in the associated time interval. For example, a trajectory of a moving object on a 2-D space may consist of the following two linear pieces:

$$[(x=t-3) \wedge (y=t+3) \wedge (0 \leq t < 2)] \cup [(x=-1) \wedge (y=t+3) \wedge (2 \leq t < 6)]$$

Another method is to use linear interpolation (Aris Anagnostopoulos *et al.*, 2006) which takes the sampled points as the endpoints of the line segment of the polylines and the movement of an object is represented by an entire polyline in 3-D space. A trajectory is a

sequence: $\langle (x_1, y_1, t_1), (x_2, y_2, t_2), \dots, (x_k, y_k, t_k) \rangle$, in which t_1, t_2, \dots, t_k is a time sequence and (x_i, y_i) is locations. Objects are assumed to move straight between the observed points at a uniform speed as shown in Fig.2.

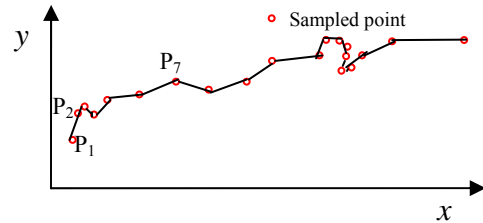


Fig.2. Illumination of a polyline trajectory

The other methods are constructing Minimal Bounding Rectangle (MBR) or Minimal Bounding Box (MBB) (Sigal Elnekave *et al.* 2007a; Sigal Elnekave *et al.* 2007b) to compress the trajectories into simplified form as shown in Fig.3.

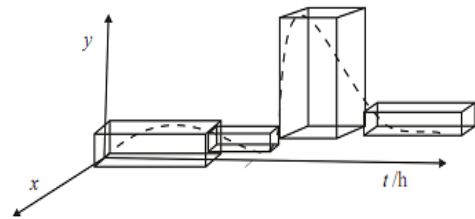


Fig.3. Illumination of a trajectory represented by MBBs

According to prior defined rule the spatio-temporal data points in trajectories are assigned to different MBBs and the MBB sequence is used to approximate trajectories. We adopt this method to compress the trajectories and form the MBB based compact moving object spatio-temporal trajectories.

2.2. The Similarity Measurement of the MBB Based Compact Moving Object Spatio-Temporal Trajectories

Since the compact representation of the moving object spatio-temporal trajectories transforms the raw trajectories into a sequence of MBBs, the similarity measurement between two trajectories accordingly replaced by calculating the similarity between two corresponding MBB sequences (Sigal Elnekave *et al.*, 2007a). The first method is proposed in article (Aris Anagnostopoulos *et al.*, 2006) which calculates the distance between the rectangles at every time point,

and the distances between all MBRs is the sum of distances between them at every time instance. The second method is proposed by Sigal *et al* (Sigal Elnekave *et al.*, 2007b) in which integrate the distances in x axis, y axis and time axis between two MBB representation trajectories. This method has a serious fault that the similarity equals naught when the MBBs don't overlap on the x axis and y axis even they have overlap on time axis. Later Sigal *et al.* proposes another similarity measurement (Sigal Elnekave *et al.*, 2007a). They add the factor of the amount of data points in the MBBs. However, after thorough analysis, it can find that there still exist inconsistencies in similarity measurement. Intuitively the distance between two MBBs belonged to different trajectories reduces with time overlap duration and increases with the space faraway. The time distance and space distance is an inverse relationship, so multiplying them to calculate the distance turns out inconsistency.

To solve above problems, we analyze these factors which impact the similarity measurement of MBB representation trajectories and then integrate these factors to form a new measurement method.

3. A New Similarity Measurement between Two MBB Representation Trajectories

Similarity measurement is the basis of clustering. If there is no reasonable and scientific measurement method, it is impossible to obtain good clustering results. Based on above analysis, it is necessary to design a new measurement method to calculate the similarity between two MBB representation trajectories.

3.1. The Factor Analysis of the Similarity Measurement between Two MBB Representation Trajectories

In this article we use the generalized distance to measure the similarity between the two MBBs. In order to describe and express conveniently, without loss of generality, we choose two-dimensional space plus time dimension as the research domain.

Firstly, the time factor is considered. Due to time is without the movement of non-reversibility (or one-way) and the nature order. Therefore, only the time overlap with two MBBs in the similarity measure will be meaningful. The longer time duration overlaps, the greater the similarity is. The length of time overlap can

be obtained by $t_n - t_m$, where t_m is the time when the two MBBs start to overlap and t_n is the time when the two MBBs end the overlap, as shown in Fig. 4A.

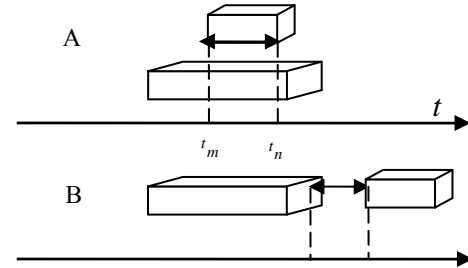


Fig. 4. A The time overlap between two MBBs, B The minimal space distance between two MBBs

Secondly, the space factor is considered. Only the distance between compressed trajectories is not greater than the distance between the original trajectories, this distance can be used as meaningful similarity measurement. The minimal space distances between two MBBs meet this constraint (Aris Anagnostopoulos *et al.*, 2006) as shown in Fig.4B. We take the sum of the minimal distances in x axis and y axis between two MBBs as the distance of space distance. The smaller this distance is, the greater the similarity is.

Finally, the density of data points in the MBBs is considered. As illustrated in Fig. 3, each MBB contains partial raw trajectory. The more number of data point in a MBB is, the slower the speed is, vice versa. Due to the difficulty to integrate movement direction, we take the density of MBB as the speed factor. In Sigal Elnekave *et al.*, 2007a the authors directly take the number of data points as the speed factor. We consider their proposal is not so appropriate because that each MBB has different size and the bigger size MBB will has more number of data points. In order to avoid this influence we take the density of data points in the MBBs as the speed factor to remove the bias of size.

3.2. Parameters Setting for Similarity Measurement between MBB Representation Trajectories

Based on above analysis, parameters setting for similarity measurement between two MBB representation trajectories include three parts: time overlap duration t_{dur} , space distance $D(MBB_i, MBB_j)$ and

density of data points in MBBs $\rho(MBB_i)$. These parameters are calculated as follows.

$$t_{dur} = t_n - t_m \quad (1)$$

Where t_{dur} denotes the time interval of two MBBs overlap on time axis; t_m is the start timestamp of the time overlap of MBBs; t_n is the end timestamp of the time overlap of MBBs.

$$D(MBB_i, MBB_j) = D_x(MBB_i, MBB_j) + D_y(MBB_i, MBB_j) \quad (2)$$

Where $D(MBB_i, MBB_j)$ denotes the space distance between two MBBs; $D_x(MBB_i, MBB_j)$ denotes the space distance between two MBBs on x axis; $D_y(MBB_i, MBB_j)$ denotes the space distance between two MBBs on y axis.

$$D_x(MBB_i, MBB_j) = \max(0, (\max(MBB_{ix\min}, MBB_{jx\min}) - \min(MBB_{ix\max}, MBB_{jx\max}))) \quad (3)$$

$$D_y(MBB_i, MBB_j) = \max(0, (\max(MBB_{iy\min}, MBB_{jy\min}) - \min(MBB_{iy\max}, MBB_{jy\max}))) \quad (4)$$

In equations (3) (4) $MBB_{ix\min}$, $MBB_{ix\max}$, $MBB_{iy\min}$, $MBB_{iy\max}$ separately denotes the minimum and maximum on x axis and y axis of MBB_i

$$D\rho(MBB_i, MBB_j) = |\rho(MBB_i) - \rho(MBB_j)| \quad (5)$$

$$\rho(MBB_i) = num(MBB_i) / vol(MBB_i) \quad (6)$$

$$vol(MBB_i) = (MBB_{ix\max} - MBB_{ix\min}) \times (MBB_{iy\max} - MBB_{iy\min}) \times (MBB_{it\max} - MBB_{it\min}) \quad (7)$$

Where $D\rho(MBB_i, MBB_j)$ denotes the difference of density between two MBBs; $\rho(MBB_i)$ denotes the density of MBB_i ; $num(MBB_i)$ denotes the number of data points in MBB_i ; $vol(MBB_i)$ denotes the volume of MBB_i .

3.3. The Formula of Similarity Measurement for MBB Representation Trajectories

The sampled points are used to express the raw trajectory. Formula (8) illustrates a trajectory T with n sampled points.

$$T = [(x_1, y_1, t_1), (x_2, y_2, t_2), \dots, (x_n, y_n, t_n)] \quad (8)$$

The method proposed in (Sigal Elnekave *et al.*, 2007a) is used in this article to compress raw trajectories. Fig. 3 shows the raw trajectory T is partitioned into four MBBs. Formula (9) shows the four MBB expressed compact trajectory T.

$$T = [MBB_1, MBB_2, \dots, MBB_k], k \ll n \quad (9)$$

Each MBB can be express by the coordinate of endpoints as shown formula (10).

$$\begin{aligned} MBB_i &= \langle (i.x_{\min}, i.x_{\max}), (i.y_{\min}, i.y_{\max}), (i.t_{\min}, i.t_{\max}) \rangle \\ i.x_{\min} &= \min(\forall data \in i, data.x_{\min}) \\ i.x_{\max} &= \max(\forall data \in i, data.x_{\max}) \\ i.y_{\min} &= \min(\forall data \in i, data.y_{\min}) \\ i.y_{\max} &= \max(\forall data \in i, data.y_{\max}) \\ i.t_{\min} &= \min(\forall data \in i, data.t_{\min}) \\ i.t_{\max} &= \max(\forall data \in i, data.t_{\max}) \end{aligned} \quad (10)$$

Where MBB_i denotes the i^{th} MBB ; $i.x_{\min}$ denotes the minimum value of all data points in i^{th} MBB on x axis; $i.x_{\max}$ denotes the maximum value of all data points in i^{th} MBB on x axis; $\forall data$ denotes all data points in i^{th} MBB .

With above notations and the analysis in section 3.1, we can draw the similarity measurement formula between $T_i = [MBB_{i1}, MBB_{i2}, \dots, MBB_{ik}, \dots]$ and $T_j = [MBB_{j1}, MBB_{j2}, \dots, MBB_{jg}, \dots]$ as shown formula (11).

$$\begin{aligned} sim(T_i, T_j) &= \sum_{t=t_0}^{t_{end}} \frac{t_{dur} \times D\rho(MBB_{ik}, MBB_{jg})}{D(MBB_{ik}, MBB_{jg})}, \\ &\text{iff } MBB_{ik} \text{ overlap with } MBB_{jg} \end{aligned} \quad (11)$$

In order to discern “data-amount-based” distance, we named $sim(T_i, T_j)$ in formula (11) “data-density-based” distance.

3.4. How to Gain the “Data-Density-Based” Distance in Similarity Measurement

In this section we use a toy example to show how effective the similarity measurement formula (11) will work out.

Fig. 5 shows two trajectories with same time span.

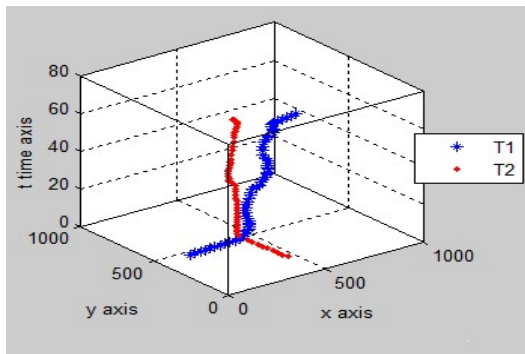


Fig.5. Two spatio-temporal trajectories

We firstly compress trajectories into MBB sequences by using the method in (Sigal Elnekave *et al.*, 2007b). The formula $bound = (\max(D) - \min(D)) \times b$ ($b=0.2$) provides the calculation of the bounding value for each dimension. Thus these two trajectories can be express into six MBBs separately. Fig.6 illustrates this progress.

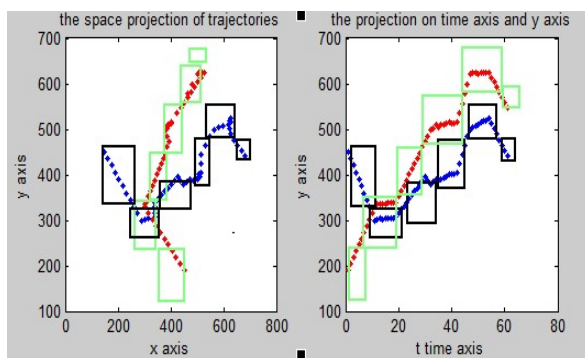


Fig.6. The projections of two trajectories.

According to (11), we can obtain the similarity value formula includes three factors: time similarity, space similarity and speed similarity, the movement characters of trajectories can be presented clearly. In the following section we will use the effective index Dunn to show how well it performs.

4. Experiments and Result Analysis

4.1. The Datasets

In order to compare with “data-amount-based” distance (Sigal Elnekave *et al.*, 2007a), we use two datasets,

both synthetic and real. The synthetic dataset is produced by the CENTER data generator. The software CENTER can be freely downloaded from <http://www-kdd.isti.cnr.it/>. We ran 70 simulations each representing a moving object. The first 40 runs simulated daily movements (trajectories) of a mobile object during 25 days, and the other 30 runs simulated daily movements (trajectories) of a mobile object during 45 days. These trajectories belonged to seven movement pattern reached at least three identical locations at identical times. The location of each object was sampled at least 40 times during each day.

The real world dataset we use consists of satellite tracking of a number of Barren-ground Caribou living in the Northwest Territories of Canada (Space for species. <http://www.spaceforspecies.ca/>). In total, 19 Caribou are tracked for various lengths of time during the period from 18th September 2001 to 17th January 2006. The Caribou were fitted with satellite transmitters in three groups, and the length of time these functioned varied widely between one month and two and a half years—the average being about 14months. Therefore, the number of animals being tracked at any one time was between zero and seven. On average, roughly three animals were being tracked at most times. Unfortunately, this is a small number of objects—and this is typical of such datasets. The techniques we describe in this paper are designed and suited for many more objects, as the goal of our work is to find patterns that groups of objects follow, not individuals.

4.2. The Clustering Algorithms

Based on the spatio-temporal trajectory similarity measurement formula proposed in this paper, we use the k-means and hierarchical clustering algorithms to group the trajectories. The procedure has two steps: first one is to compress the trajectories into MBB sequences and the second one is to cluster the MBB sequences into similar trajectory groups. In the first step we use the method proposed in (Sigal Elnekave *et al.* 2007b) shown as the Fig.7.

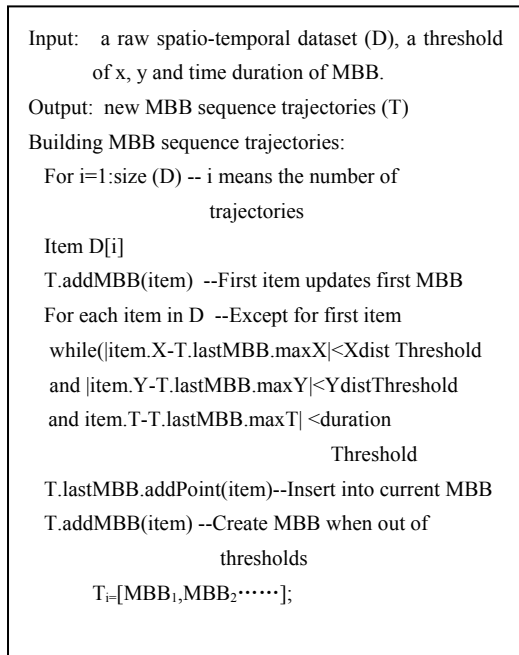


Fig.7. diagram of trajectory clustering

We set the threshold in each dimension according to the formulae: $bound = (\max(D) - \min(D)) \times b$, $b=0.01$. In the second step we use the k-means algorithm to cluster the synthetic data and use the hierarchical clustering algorithm to the Caribou data. The whole procedure is shown as the Fig. 8.

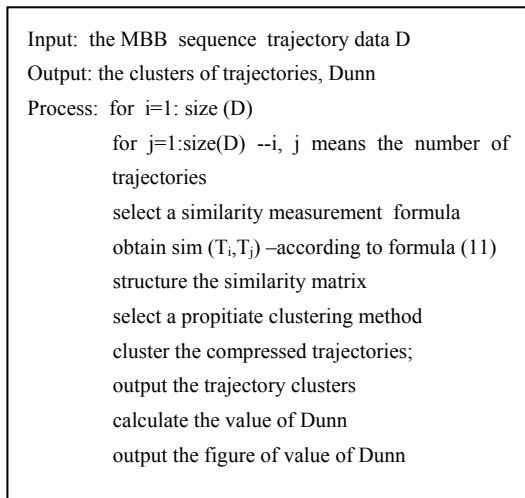


Fig.8. diagram of trajectory clustering

4.3. The Result

We adopt the clustering Dunn index to compare the validity of different similarity measurement. The Dunn index measures the worst-case compactness and separation of a clustering, with higher values being better. $Dunn = D_{min} \div D_{max}$, D_{min} is the minimum distance between any two objects in different clusters (separation) and D_{max} is the maximum distance between any two items in the same cluster (homogeneity). In this experiment we use the method in (Sigal Elnekave *et al.* 2007a) to compress trajectories and then use K-Means algorithm to cluster the MBB compressed trajectories of the CENTER data generator. The experimental result is shown in Fig. 9.

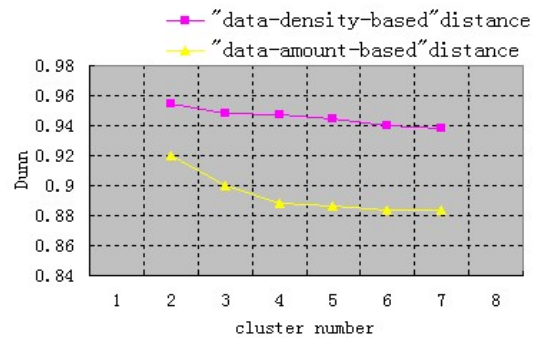


Fig.9. the Dunn value calculated by clustering with data-density-based and data-mount-based similarity of Center data

Then we run the hierarchical clustering algorithm on Barren-ground Caribou data and obtain the Dunn values calculated by clustering with data-density-based similarity and data-mount-based similarity, the result is shown in Fig. 10.



Fig.10. the Dunn value calculated by clustering with data-density-based and data-mount-based distance of Caribou data.

5. Conclusion

Considering three influence factors of time, space and motion characteristic, a new similarity measure formula is proposed to calculate the similarity between the *MBB* representation trajectories. Experiments show that this formula can improve the accuracy of trajectory clustering. Moving object trajectory clustering is a new field of study, the compression representation of the trajectory, similarity measurement, clustering methods, are worthy of further research.

Acknowledgments

The authors gratefully acknowledge the editor and anonymous reviewers for their valuable comments and constructive suggestions. This research was supported by Beijing Municipal Science & Technology Commission key project (Z090506006309011), National Science and Technology Support Program topics (2009BAG12A10), the State Key Laboratory of Rail Traffic Control and Safety (RCS2009ZT007).

References

1. Laube, P. 2005. Spatio-Temporal Data Mining - Coping with the Increasing Availability of Motion Data in Geography. Proceedings of the Seventeenth Annual Colloquium of the Spatial Information Research Centre. P.A. Whigham (Ed), Dunedin, New Zealand. pp. 189-193.
2. Roddick, J.F., Hornsby, K. And Spiliopoulou, M., 2007, An updated bibliography of temporal, spatial, and spatio-temporal data mining research. 2007 of Lecture Notes in Artificial Intelligence. pp. 147-163.
3. Wuhong Wang, Hongwei Guo, Katsushi Ikeuchi, Heiner Bubb, 2011. Numerical simulation and analysis procedure for digital driving dependability in intelligent transport system, KSCE Journal of Civil Engineering, 2011 15(5),891-898
4. Xiaohua Zhao, Xingjian Zhang, Jian Rong, Jianming Ma. 2011. Identifying Method of Drunk Driving Based on Driving Behavior. International Journal of Computational Intelligence Systems.2011.4(3).pp.336-369
5. Florian Verhein, Sanjay Chawla.2008. Mining spatio-temporal patterns in object mobility databases, Data mining and knowledge discovery 2008(16):pp.5-38.
6. Miro Nanni, Dino Pedreschi. 2006. Time-focused clustering of trajectories of moving objects. Journal Intelligent Information Systems,2006 (27) pp.267-289
7. Gaffney, S. and Smyth, P. 1999, "Trajectory Clustering with Mixtures of Regression Models," In Proc. 5th ACM SIGKDD Int'l Conf. on Knowledge Discovery and Data Mining, San Diego, California, pp. 63-72
8. Gaffney, S., Robertson, A., Smyth, P.2006, Camargo, S.,and Ghil, M., Probabilistic Clustering of Extratropical Cyclones Using Regression Mixture Models, Technical Report UCI-ICS 06-02, University of California
9. Vlachos, M., Gunopulos, D., 2002.and Kollios, G., Discovering Similar Multidimensional Trajectories. In Proc. 18th Int'l Conf. on Data Engineering, San Jose, California, pp. 673-684.
10. Chen, L., Özsu, M. T., and Oria, V., 2005. Robust and Fast Similarity Search for Moving Object Trajectories. In Proc. 2005 ACM SIGMOD Int'l Conf. on Management of Data, Baltimore, Maryland, pp. 491-502.
11. Keogh, E. J. 2002. Exact Indexing of Dynamic Time Warping. In Proc. 28th Int'l Conf. on Very Large Data Bases, Hong Kong, China, pp. 406-417.
12. Andrey Tietbohl Palma, Vania Bogorny, *et al.*2008. A clustering-based approach for discovering interesting places in trajectories. Proceedings of the 2008 ACM symposium on Applied computing, pp. 863-868
13. S. Spaccapietra, C. Parent, M.-L. Damiani, J. A. F. de Macedo, F. Porto, and C. Vangenot.2007. A conceptual view on trajectories. Technical report, Ecole Polytechnique Federal de Lausanne, April 2007.
14. S.Y.Hwang, Y.H.Liu, J.K.Chiu, F.P.Lim .2005. Mining Mobile Group Patterns: A Trajectory-based Approach. Lecture Notes in Artificial Intelligence, PAKDD 2005, (3518) pp.713-718
15. Aris Anagnostopoulos, Michial Vlachos *et al.* 2006. Global Distance-Based Segmentation of Trajectories. KDD'06, Philadelphia, Pennsylvania, USA
16. Sigal Elnekave, Mark Last, O. Maimon. 2007a. Incremental Clustering of Mobile Objects, STD07, IEEE
17. Sigal Elnekave, Mark Last, Oded Maimon,2007b. A compact representation of spatio-temporal data. Proceedings of the Seventh IEEE International Conference on Data Mining Workshops, pp.601-606