Earth Observation and Geomatics Engineering 5(2) (2021) 132-144

EARTH OBSERVATION AND GEOMATICS ENGINEERING
website: https://eoge.ut.ac.ir

The credibility evaluation of the trajectory clustering results using a user-defined similarity

Ali Moayedi[1], Rahim Ali Abbaspour[2*], Alireza Chehreghan[3], Pooya Mojtabaee [4]

[1,2]School of Surveying and Geospatial Engineering, College of Engineering, University of Tehran, Tehran, Iran
[3]Mining Engineering Faculty, Sahand University of Technology, Tabriz, Iran
[4]Faculty of Geomatics, K. N. Toosi University of Technology, Tehran, Iran.

Article history:
Received: 12 April 2020, Received in revised form: 26 August 2021, Accepted: 1 September 2021

ABSTRACT

Evaluation of the cluster analysis results in spatio-temporal trajectories is more sophisticated than the same procedure in other data. Measuring clusters' compactness and separateness requires defining an appropriate similarity function. Similarity definitions in trajectories are diverse and application-based. Positional similarity and the similarities of speed and direction, as elemental features of moving objects, are fundamental concepts in the trajectory similarity definition. In this paper, we present a new framework for evaluating trajectory clustering results based on the expert's opinion on the definition of similarity. Specifically, the meaning of similarity is defined by the experts using the AHP method and based on the application context. Moreover, we propose a new index, which is utilized in estimating the optimal cluster number. Based on the obtained results, taking the application and the data structure into consideration is very influential in the evaluation process. To verify that they are not random, the one-way ANOVA test is carried out at the confidence interval of 95% to provide the significance test of the results.

KEYWORDS

Spatio-temporal
Similarity functions
Trajectory clustering.
Evaluation

1. Introduction

The increasing availability of positioning devices equipped by GPS receivers, such as smartphones, which can be used to track moving objects, results in a significant rise in capturing and storing spatio-temporal data. The moving objects under study can be taxis (Kan et al., 2019; Yu et al., 2019; Zhao & Stefanakis, 2018), animals (Ardakani et al., 2019; De Cáceres et al., 2019), or aircraft (Hurter et al., 2014; Olive & Morio, 2019). Therefore, it is important to pay attention to effective methods for extracting useful and relevant knowledge from this large volume of data.

Cluster analysis is an efficient technique commonly used to explore and extract interesting patterns in huge datasets. Since trajectories are traces of moving objects in space and time, we can gain knowledge of their typical behavior if we cluster them properly, which benefits modelling, simulation, and prediction of movement. Trajectory clustering aims to identify clusters in which the trajectories have the highest level of similarity while having the most substantial difference from those in other clusters. In order to detect similar trajectories and evaluate the similarity between two trajectories, a similarity function, i.e., the inverse of distance, is required. In contrast to regular point data, similarity measurement between trajectories is challenging and calls for an exact definition of the similarity concept (Magdy et al., 2017). Trajectories hold both spatial and temporal characteristics, which must be considered in the similarity measurement process. The definition of similarity can vary depending on the application domain and the user's needs (Dodge et al., 2008). In other words, the similarity is defined in a way that fulfils the user's perspective on the meaning of similarity.

Extensive reviews of indices, including internal and external indices, can be found in the previous works. (Deborah et al., 2010) is among studies on the evaluation of

* Corresponding author
E-mail address: abaspour@ut.ac.ir (Rahim Ali Abbaspour)
DOI: 10.22059/EOGE.2022.327815.1100

trajectory clustering results that review research on the validation of clustering results. Besides, in (Arbelaitz et al., 2013), 30 different indices were examined in varied environments and characteristics. In (Mao et al., 2017), a new distance named SDTW and external validity indices were utilized for the evaluation purpose. In (Niu et al., 2019), a new method is presented for trajectory clustering of road networks in which trajectories are modelled using the dual graph. They used internal indices such as Davies-Bouldin Index (DB), Akaike Information Criterion (AIC), and Bayesian Information Criterion (BIC) to evaluate the clustering results.

Over the past few years, some studies devoted considerable efforts to the clustering evaluation issue, e.g. (Zhu & Ma, 2018) in which a new variance based clustering validity index (VCVI) was proposed. Moreover, in (Cheng et al., 2018), a new local core-based index was suggested, which assesses the clustering results for data of the arbitrary shape. Lee et al. (Lee et al., 2018) presented an index grounded on a support vector data description that computes the compactness of clusters in the kernel space and is independent of the cluster's shape and noise. Some research has also been conducted on the evaluation of time-series clustering. In (Košmelj & Batagelj, 1990), a method for time-series clustering evaluation was suggested, and a method was presented to estimate the optimal number of clusters (Baragona, 2001). An Overview of the most important literature is presented in Table 1.

Table 1. Overview of the reviewed sources

Authors	Employed Method(s)	Limitations
Mao et al., 2017	external validity indices	Not applicable for real life applications
Niu et al., 2019	DB, AIC, and BIC	Not considering the specific characteristics of trajectory data and cannot be used to for clusters with arbitrary shapes.
Lee et al., 2018	A support vector data description (SVDD)-based index	Not considering the specific characteristics of trajectory data
Zhu & Ma, 2018	VCVI	
Košmelj & Batagelj, 1990	A cross-sectional approach	

The conventional techniques of clustering evaluation are unable to consider the unique characteristics of spatio-temporal data. These traditional methods simply use a distance function to measure the similarity between the point data. Spatio-temporal data, however, are multi-dimensional, complicated, and difficult to be visually displayed. Unlike time-series data that have a single, homogeneous space, this type of data owns a multiple heterogeneous space and might contain local shifts, noise, and outliers; hence, the analysis methods of time-series data should be cautiously used for spatio-temporal data. As far as we know, an appropriate framework for evaluating the trajectory clustering results seems to be missing.

As we discussed above and regarding the importance of considering the user's need to evaluate the clustering results, this study proposes a new framework for trajectory clustering evaluation based on the definition of similarity by taking users' judgment on the meaning of similarity into consideration. Moreover, an index for trajectory clustering validation is presented, which is novel with respect to considering the special characteristics of trajectory data. Our study examines the proposed framework on massive volume datasets of taxis, vessels, animals, aircraft, and synthetic data to investigate the impact of data characteristics on the final results.

2. Basic Concepts

2.1. Trajectory

A trajectory is defined as a path of a moving object in spatial and temporal dimensions. The trajectory Tr can be formulated as Equation (1). Each sampling point on the path represents the captured position (x, y) at a specific time instance (t).

$$Tr = \{(x_1, y_1, t_1), (x_2, y_2, t_2), \ldots, (x_n, y_n, t_n)\} \qquad (1)$$

2.2. Clustering and Similarity Functions

Numerous clustering algorithms have been employed for trajectory clustering purposes. Among them, K-means and Agglomerative Hierarchical Clustering (AHC) are the most well-known and widely used ones. K-means is used to partition a dataset into K groups automatically, and it starts by selecting K initial clusters and then iteratively refining them to reach final cluster centers. On the other hand, AHC, the most common type of hierarchical clustering, groups objects into clusters based on their similarities. Each data point is initially treated as a separate cluster, and in the next step, combined with other clusters in each iteration until K clusters are formed.

Similarity must be calculated before the clustering of trajectories. In this regard, various functions have been proposed, each with its unique features. The most cited ones, which are well-established and highly popular for trajectory clustering, are Euclidean (Priestley, 1980), DTW (Myers et al., 1980), Hausdorff (Huttenlocher et al., 1993), EDR (Chen et al., 2005), and ERP (Chen & Ng, 2004). We employed these similarity functions in our work, although their complete descriptions and equations are missed for the sake of space.

2.3. The AHP

The AHP is a multi-criteria decision-making technique that simplifies complex decisions through a series of pairwise comparisons between criteria. Those criteria are determined by experts who are trying to achieve a specific goal. In our case, with the goal of defining the meaning of similarity for trajectory clustering, the term "experts" refers to a group of people who set the desired application of the clustering and aim to extract relevant information regarding this application. For the sake of space and length, this document does not include a thorough description of AHP, and the readers are referred to (Teknomo, 2006) for more details.

3. The Proposed Method

3.1. Motivations

Internal validity indices assess the quality of clustering by means of properties intrinsic to the data without relying on any external information or ground truth. Compactness and separation are the two main criteria used in internal indices of quality assessment. From the spatial point of view, well-defined clusters have higher within-cluster similarity and lower between cluster similarity. However, the way the concept of similarity is defined remains a vague question that should be addressed. Definition of similarity must be set by experts based on the purpose of the clustering.

Figure 1 demonstrates how considering the meaning of similarity is crucial in the clustering process. If the position is the most important parameter in defining similarity, we will have two clusters (C_1 = [Tr_1, Tr_2, Tr_3], C_2 = [Tr_4, Tr_5]). On the other hand, if the similarity of trajectories implies specifically the similarity of their directions, we will have three clusters (C_1 = [Tr_1], C_2 = [Tr_2, Tr_4]), C_3 = [Tr_3, Tr_5]). Consequently, we decided to propose a framework to evaluate the credibility of clustering results by employing experts' opinions on the meaning of the similarity. In this regard, a credible clustering yields suitable results for a certain application context and satisfies the user's perspective on the clustering purpose.

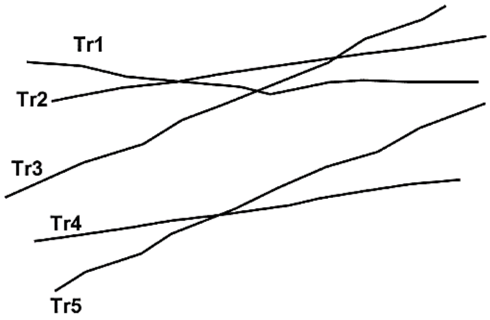

Figure 1. Influence of the meaning of similarity in the clustering process

3.2. The Employed Criteria

The movement has temporal, spatial, and spatio-temporal dimensions. Due to the existing time parameter in the definition of trajectory, enriching trajectory with movement characteristics such as speed and direction is essential to obtain a more meaningful similarity (Dodge et al., 2009). Adding such parameters, which are critical to the definition of movement, modifies the meaning of similarity.

In this study, in order to define similarity, we employ three main components of movement, including position, speed, and direction. This categorization enables the user to consider different aspects of similarity separately or simultaneously. In numerous applications, positional information is of higher importance than movement parameters such as speed and direction. Moreover, in some applications, clustering trajectories based on the spatio-temporal similarity are quite restricting and might lead to either extraction of irrelevant patterns or missing some interesting ones. On the other hand, in many other applications, such as traffic monitoring, leaving parameters like speed and direction leads to unsatisfactory outcomes. Besides, in some contexts, other parameters might be important. For instance, flight directors might be interested in ROT (rate of turn). However, in the following, we explored the most common criteria for computing and defining the similarity in more detail, which are crucial for all applications.

3.2.1. The Positional Similarity

In many applications, the similarity of trajectories implies specifically the closeness of their physical location, and the information such as the speed and direction are of less importance. The Euclidean distance between trajectories is a simple yet efficient method for positional similarity measurement regardless of the temporal dimension. As a result, the Positional Distance (PD), the inverse of positional similarity, between trajectories Tr_1 and Tr_2 is calculated by Equation (2).

$$PD(Tr_1, Tr_2) = \frac{1}{m}\sum_{i=1}^{m} \text{Euclidean}\,(Tr_i^1, Tr_i^2) \tag{2}$$

where Tr_i^1 denotes the ith point on the trajectory Tr_1. Two trajectories must be of the same lengths; otherwise, both are resampled at the same space interval by the interpolation method, and m is their length after this interpolation.

3.2.2. The Speed Similarity

Two trajectories are called similar regarding their speed when they move at the same speed, even if they are not close or in the same direction. Along with the positional similarity, speed similarity between two trajectories is important in some applications like the detection of different transit modes. In clustering pedestrians, in order to distinguish those who run from those who walk, considering the speed similarity is more pivotal than the positional and direction similarities. In this similarity, first, the trajectory Tr is represented as a vector in an F-dimensional space:

$$V_{Tr} = \{v_1, v_2, \ldots, v_i, \ldots, v_{n-1}\} \subseteq R^F \qquad (3)$$

where n is the number of points in the trajectory and v_i refers to the movement speed between two consecutive points i and i+1 and is obtained by the equation below:

$$v_i = \frac{\sqrt{\left(x_{Tr}(i+1) - x_{Tr}(i)\right)^2 + \left(y_{Tr}(i+1) - y_{Tr}(i)\right)^2}}{t_{Tr}(i+1) - t_{Tr}(i)} \qquad (4)$$

Thus, the Speed Distance (SD), the inverse of the speed similarity, between two trajectories Tr_1 and Tr_2 can be computed by Equation (5).

$$SD = \frac{1}{m}\sum_{i=1}^{m}\left|V_i^{Tr_1} - V_i^{Tr_2}\right| \qquad (5)$$

In this equation, m is the number of extracted speeds between two consecutive sampling points after equalization of the lengths of two trajectories.

3.2.3. The Direction Similarity

In some applications, e.g., animals' path clustering, in order to extract food search patterns, utilizing the positional similarity is not sufficient, but assigning trajectories with a similar direction to the same cluster is a more suitable approach. Direction illustrates how moving objects turn and twist through their paths in a spatial reference system. A trajectory can be shown as a sequence of directions between each two consecutive sampling points (Equation (6)). Expressing trajectories in this way made it independent from the starting point, i.e., from the trajectories' spatial closeness.

$$DT = [(dir_1), \ldots, (dir_i), \ldots, (dir_{n-1})] \qquad (6)$$

In this equation, n refers to the number of points in the trajectory, and dir_i can be obtained by exploiting Equation (7). In this equation, we employ the notion of azimuth, which is a simple way to demonstrate the orientation of a line and is measured clockwise from North. In case it is located in the first quadrant, dir is shown in Figure 2

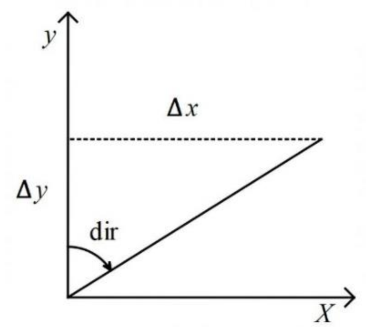

Figure 2. The visual demonstration of dir in the first quadrant

$$dir_{Tr,i} = \begin{cases} \arctan\left(\frac{x_{Tr}(i+1) - x_{Tr}(i)}{y_{Tr}(i+1) - y_{Tr}(i)}\right); \; x_{Tr}(i+1) \geq x_{Tr}(i) \text{ and } y_{Tr}(i+1) \geq y_{Tr}(i) \\ \pi - \arctan\left(\frac{x_{Tr}(i+1) - x_{Tr}(i)}{y_{Tr}(i+1) - y_{Tr}(i)}\right); \; x_{Tr}(i+1) \geq x_{Tr}(i) \text{ and } y_{Tr}(i) \geq y_{Tr}(i+1) \\ \pi + \arctan\left(\frac{x_{Tr}(i+1) - x_{Tr}(i)}{y_{Tr}(i+1) - y_{Tr}(i)}\right); \; x_{Tr}(i) \geq x_{Tr}(i+1) \text{ and } y_{Tr}(i) \geq y_{Tr}(i+1) \\ 2\pi - \arctan\left(\frac{x_{Tr}(i+1) - x_{Tr}(i)}{y_{Tr}(i+1) - y_{Tr}(i)}\right); \; x_{Tr}(i) \geq x_{Tr}(i+1) \text{ and } y_{Tr}(i+1) \geq y_{Tr}(i) \end{cases} \qquad (7)$$

To achieve a more precise direction similarity, trajectories must be smoothed employing the moving average filter. We denote the Direction Distance, the inverse of the direction similarity, between two smoothed trajectories STr_1 and STr_2 as $DD(STr_1, STr_2)$, which can be calculated as:

$$DD(STr_1, STr_2) = \frac{1}{m}\sum_{i=1}^{m}\left|dir_i^{STr_1} - dir_i^{STr_2}\right| \qquad (8)$$

where $dir_i^{STr_1}$ is the movement direction between two consecutive points i and i+1 in the smoothed trajectory STr_1, and m is the number of components of the movement direction in DT.

3.3. The MDBT index

Firstly, experts' opinions on the relative pairwise comparison of positional, speed, and direction similarities are taken and then aggregated by the AHP so that the weights for each of the aforementioned similarities are adjusted. After calculating these coefficients and normalizing PD, VD, and DD distances between 0 and 1, the similarity between two trajectories is defined by Equation (9).

$$TD = W_p \times PD + W_v \times VD + W_d \times DD + W_o \times OD \quad (9)$$

In this equation, the W_p, W_v, W_d coefficients are constant values related to the weights of positional, direction, and speed criteria, respectively. The term PD refers to the positional distance, VD shows the speed distance, and DD indicates the direction distance of trajectories obtained by Equations (2), (5), and (8), respectively. Moreover, experts are allowed to consider other important criteria, such as rate of turn and acceleration, based on their desired application using W_α and OD, in which OD is Other Distance functions and W_o is its weight.

After defining the similarity, Modified Davies-Boulding for Trajectories (MDBT) is established, inspired by the idea of the classical Davies-Boulding index (Davies & Bouldin, 1979), and is obtained as follows:

$$MDBT = \frac{1}{k}\sum_{i=1}^{k} \max_{j=1,\dots,k,i\neq j}\left\{\frac{diam(c_i)+diam(c_j)}{TD(z_i,z_j)}\right\} \quad (10)$$

where k is the number of clusters, c_i refers to the ith cluster, and z_i is the representative trajectory of the cluster ci. Representative trajectories are trajectories that have the least TD distance with from other group members. If n_i exhibits the size of the ith cluster, the diam(.) is computed using Equation (11). Figure 3 briefly describes the proposed framework.

$$diam(c_i) = \frac{1}{n_i}\sum_{x\in c_i} TD(x, z_i) \quad (11)$$

This index facilitates the comparison between different clustering methods by offering an intuitive numerical criterion that can also be implemented for all similarity functions and clustering algorithms. It is defined as the ratio of the within-cluster scatter to the between-cluster separation; consequently, the lower the index value, the better the clustering result. Since credibility measures how well clustering results match the user's need, this index is an efficient one that assesses the credibility of the clustering result. That is because it is established based on the definition of the similarity between trajectories according to the user's perspective.

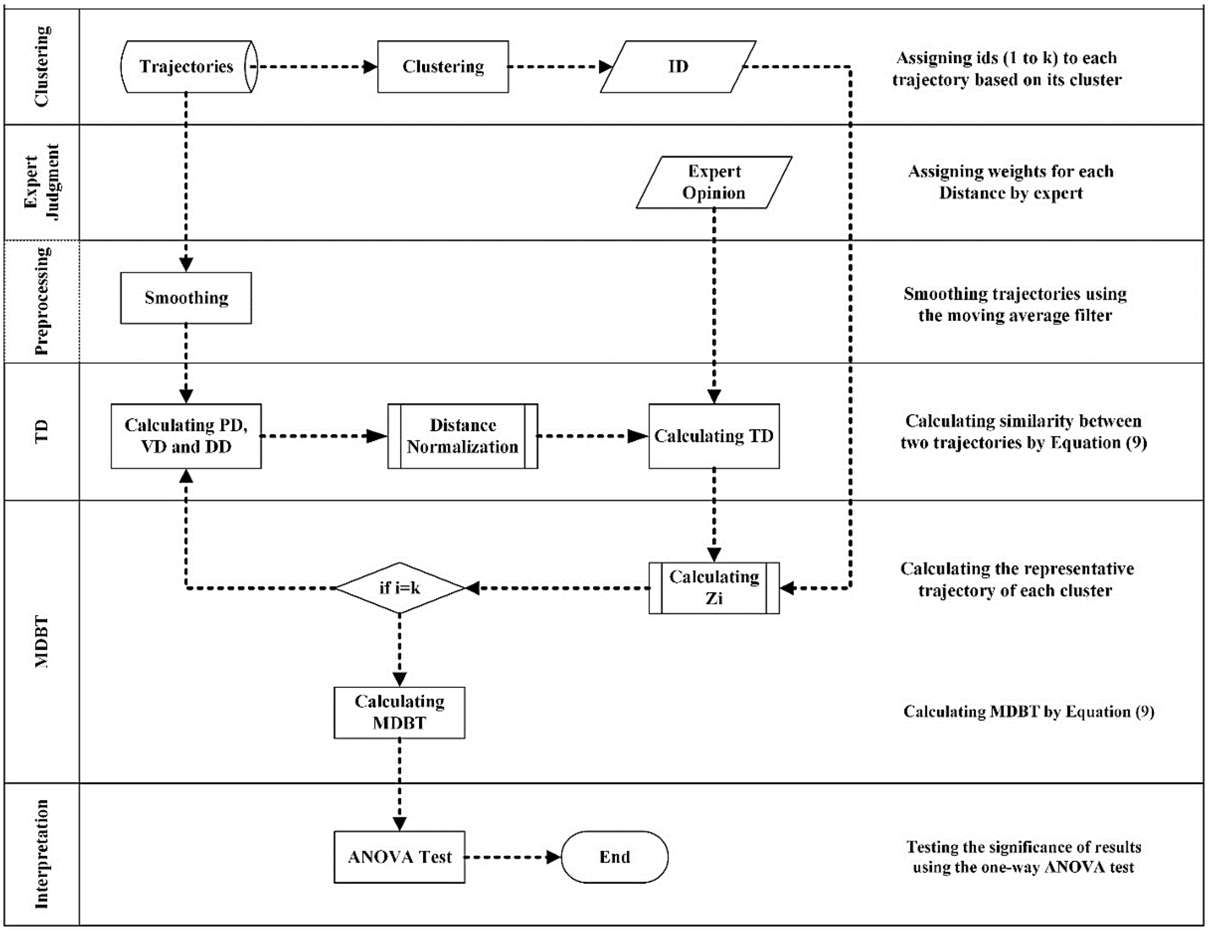

Figure 3. The general overview of the calculation of the MDBT index.

4. Implementation and Results

4.1. *Datasets*

Because small datasets could not reveal the actual clustering performance, five high-volume datasets are employed in this research. Due to the distinct navigation environments and geographic contexts, different types of trajectories differ in dynamic behavior, e.g., speeds, directions, sampling rates, and lengths. Therefore, trajectories of animals, ships, taxis, and aircraft together with synthetic trajectories are explored with each having unique characteristics. The data are introduced in Table 2 and displayed in Figure 4.

Table 2. Data Introduction

Title	Number of Trajectories	Number of Points on Each Trajectory	Data Type	Reference
Antarctic	75	64	Animal	(Tarroux et al., 2016)
AIS	1200	1450	Ship	https://sites.google.com/site/moving objectsatsea/data-challenge
CROSS	1900	13	synthetic	(Morris, 2011)
MIT	3500	115	Taxi	(Wang et al., 2011)
Flight tracks	4700	148	Aircraft	http://www.flyquietsfo.com

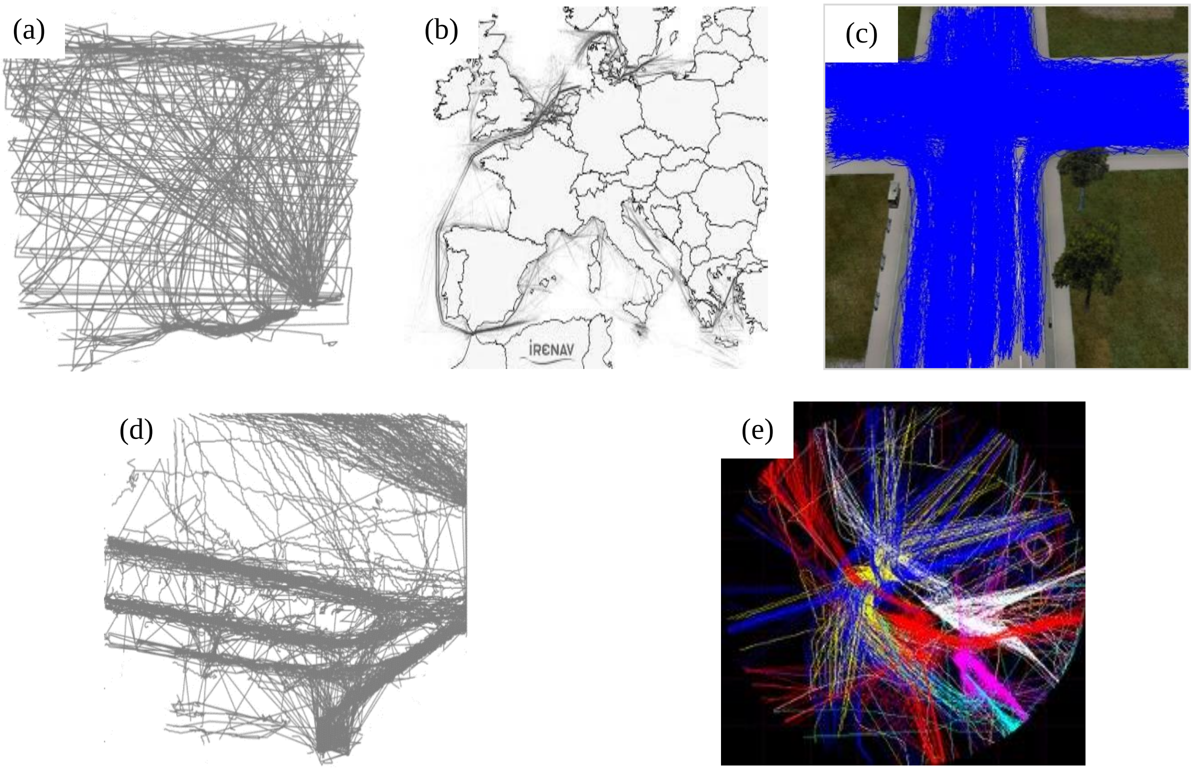

Figure 4. Visualization of Dataset: (a) Antarctic, (b), AIS (c) CROSS, (d) MIT, and (e) Flight.

4.2. Scenario Definition

To enlighten the influence of the dataset on the performance of clustering algorithms, two major potential scenarios in trajectory clustering are assumed: first, with the main focus on the positional similarity of trajectories, and second, with the main focus on the speed and direction. It should be noted that we define Scenario A and Scenario B just for testing results, and the values in the pairwise comparison are hypostatical.

Scenario A: Clustering the movement paths of workers in a workshop with the main focus on the closeness of the movement paths from the spatial aspect. The assumption is that the hypothetical experts believe that the importance of positional similarity is five times more than the speed similarity and three times more than the direction similarity, while the importance of direction similarity is two times higher than the speed similarity.

Scenario B: Clustering birds' flying paths with the focus on speed and direction similarities. Here the assumption is that the hypothetical experts think that speed and direction similarities are of equal importance, and their significances are five times more than the positional similarity.

The pairwise comparison matrices for scenarios A and B are shown in Tables 3 and 4, respectively. Both of the tables were compatible according to the ratio $\frac{CI}{RI}$ <0.1. These tables represent weights for each criterion of similarity.

Table 3. The pairwise comparison of criteria for Scenario A.

	PD	SD	DD	W
PD	1	5	3	0.65
SD	0.2	1	0.5	0.12
DD	0.33	2	1	0.23
Sum	1.53	8	4.5	1

Table 4. The pairwise comparison of criteria for Scenario B

	PD	SD	DD	W
PD	1	0.33	0.33	0.14
SD	3	1	1	0.43
DD	3	1	1	0.43
Sum	7	2.33	2.33	1

Furthermore, in the following, to examine the effect of varying definitions of similarity concept on the evaluation of clustering results, seven major scenarios with their respective coefficients are introduced in Table 5. It should be noted these scenarios are hypostatical and we define them to analyze the performance of our framework. As an example, for scenario 6, in which the key parameters for similarity are speed and direction, WS and WD are adjusted to 0.5, whereas 0 is assigned to WP.

Table 5. Different similarity concepts based on the scenario

Scenarios	key parameters	W_P	W_S	W_D
Scenario 1	only location	1	0	0
Scenario 2	only speed	0	1	0
Scenario 3	only direction	0	0	1
Scenario 4	location and speed	0.5	0.5	0
Scenario 5	location and direction	0.5	0	0.5
Scenario 6	speed and direction	0	0.5	0.5
Scenario 7	location, speed, and direction	0.33	0.33	0.33

4.3. Evaluating the Clustering Results Using the Proposed Framework

4.3.1. Clustering

In the first step, raw data (from Table 2) were prepared and pre processed. The outliers are removed by applying the 3-Sigma rule while the noise is eliminated from the data using a moving average filter (Figure 5). Then, datasets were clustered using the K-means and AHC algorithms and five highly ranked similarity functions, namely Euclidean, DTW, Hausdorff, EDR, and ERP. Datasets are clustered in different numbers of clusters (4, 7, 10, 13, and 16) in order to remove the effect of cluster numbers on the evaluation process.

4.3.2. Evaluation

This section describes an investigation of the proposed framework and demonstrates how it can be effective and beneficial in two separate cases. The evaluation results of the first case, in which the aim is to clarify the influence of the dataset on choosing the appropriate clustering algorithms, are illustrated in Tables 6 and 7. It is clear from these tables that changing the datasets will lead to a change in the optimal similarity function.

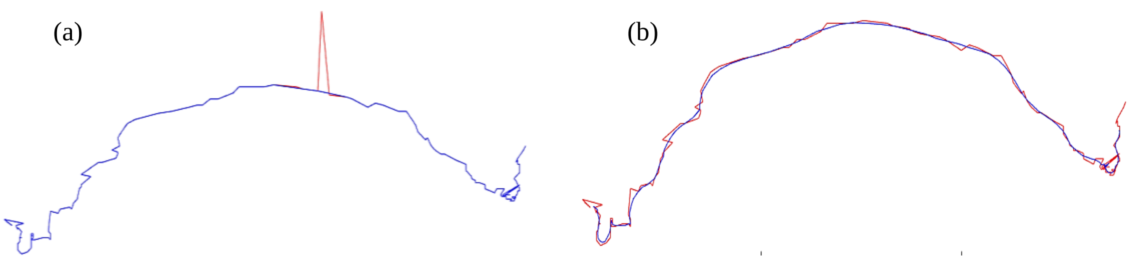

Figure 5. (a) The initial trajectory (red) and the trajectory after the outlier removal (blue) and (b) the initial trajectory (red) and the trajectory after the noise removal (blue).

Table 6. The average values of the MDBT index for Scenario A.

	Antarctic		**AIS**		**CROSS**		**MIT**		**Flight**	
	K-means	AHC	K-means	AHC	K-means	AHC	K-means	AHC	K-means	AHC
Euclidean	33.785	1.006	0.674	0.485	0.985	2.404	1.201	1.325	1.471	1.461
DTW	1.354	0.913	0.774	0.624	1.018	2.633	1.255	1.102	1.595	1.566
Hausdorff	2.974	1.791	1.487	0.638	2.5014	3.061	1.972	1.859	2.536	1.556
EDR	1.442	1.165	0.990	1.035	1.8784	2.259	1.861	1.197	2.264	1.349
ERP	1.425	1.338	1.325	1.308	1.344	2.756	1.519	1.483	1.805	1.789

Table 7. The average values of the MDBT index for Scenario B.

	Antarctic		**AIS**		**CROSS**		**MIT**		**Flight**	
	K-means	AHC	K-means	AHC	K-means	AHC	K-means	AHC	K-means	AHC
Euclidean	32.857	1.569	1.674	1.554	1.561	1.651	1.497	1.625	٢,٢30	2.291
DTW	1.719	1.548	1.767	1.516	1.532	1.906	1.527	1.411	2.294	2.439
Hausdorff	2.619	1.617	2.742	1.950	2.196	2.200	1.975	2.219	2.791	2.106
EDR	1.753	1.565	1.761	1.272	1.995	1.448	1.715	1.244	2.640	1.350
ERP	1.772	1.729	2.640	1.802	1.628	2.099	1.510	1.582	2.253	2.507

For instance, in Table 6, in the case of AHC clustering, for the Antarctic, AIS, CROSS, MIT, and Flight datasets, the best results are achieved from DTW, Euclidean, EDR, DTW, and EDR, respectively. However, in the case of K-means clustering, the proper similarity function is less influenced by the dataset. According to this table, for the Antarctic dataset, DTW, and for the other datasets, Euclidean distance has accomplished the best results.

When we change the scenario from A to B, although the superior distance functions for Antarctic and Flight datasets remain unchanged, the best distance functions for AIS, CROSS, and MIT datasets are altered. The other interesting idea here is that K-means demonstrates its best performance when it uses Euclidean as a distance function, and AHC gains its superior results if EDR is employed.

In the second case, which intends to investigate the role of the application domain in the performance of clustering algorithms, all datasets are clustered using K-means and AHC algorithms and with the consideration of scenarios defined in Table 5. The corresponding evaluation results are illustrated in Table 8.

The results of Table 8 are summarized in Figure 7. The horizontal axis shows scenarios 1 to 7, and the vertical axis demonstrates the number of scenarios in which a similarity function had been superior. For example, in scenario 1, the Euclidean function had performed better than others, and among all datasets and two clustering algorithms was six times the superior one. According to this figure, for scenarios two to seven, the optimum similarity functions are as follows: EDR, EDR, Euclidean, Euclidean, EDR, and DTW, respectively. These results highlight the impact of different definitions of similarity concept on the selection of the suitable similarity function.

4.4. Statistical Analyses

Although the acquired results in Tables 6, 7 and 8 indicate varied efficiency of similarity functions in different datasets and applications, it is of key importance to determine the significance of superiority of a specific distance function compared to others and to verify that they are not random. To address this issue and to provide the significance test of results, the one-way ANOVA test is carried out at a confidence interval of 95%.

The one-way ANOVA compares the means between the groups and determines whether any of those means are statistically significantly different from each other. Its null hypothesis is as follows:

$$H_0 = \mu_1 = \mu_2 = \cdots = \mu_K \qquad (12)$$

Table 8. Comparison of the clustering results using the MDBT index for Scenarios 1-7.

		Antarctic		**AIS**		**CROSS**		**MIT**		**Flight**	
		K-means	AHC	K-means	AHC	K-means	AHC	K-means	AHC	K-means	AHC
Euclidean	1	40.847	0.717	0.496	0.267	0.859	2.757	1.238	1.256	1.380	1.451
	2	6.662	5.668	58.278	39.028	3.389	96.19	3.455	3.637	3.356	3.757
	3	2.590	1.972	4.240	3.850	1.6734	1.322	1.598	1.679	4.752	5.127
	4	39.623	0.894	0.863	0.409	1.3936	3.121	1.698	1.767	1.754	1.875
	5	23.735	1.310	0.874	0.722	1.0712	1.806	1.194	1.247	1.760	1.963
	6	2.796	1.978	4.571	4.133	1.6735	1.511	1.758	1.932	2.847	3.197
	7	23.056	1.303	1.024	1.043	1.338	1.942	1.357	1.431	1.793	1.931
DTW	1	1.398	0.898	0.655	0.396	0.845	3.331	1.373	0.996	1.570	1.411
	2	6.305	4.492	33.843	29.246	3.452	136.775	3.217	2.826	3.339	2.990
	3	2.389	2.048	3.940	2.655	1.511	1.172	1.495	1.409	4.689	4.157
	4	1.449	1.069	0.801	0.706	1.404	3.670	1.732	1.449	1.788	1.821
	5	1.494	1.080	1.043	0.823	1.123	2.038	1.232	1.050	۲,۸۶۴	1.756
	6	2.484	2.308	4.039	3.681	1.725	1.499	1.742	1.573	2.839	2.799
	7	1.547	1.200	1.178	0.856	1.338	2.173	1.358	0.974	1.898	2.044
Hausdorff	1	3.137	2.532	1.174	0.275	2.544	3.894	2.803	2.420	2.571	1.754
	2	7.974	5.042	80.405	56.267	4.805	62.921	3.212	2.872	3.253	2.732
	3	3.094	2.250	7.163	12.504	3.581	1.631	1.934	1.969	6.635	4.615
	4	3.320	1.984	1.702	0.290	2.578	4.113	2.443	2.369	2.318	1.821
	5	3.302	1.776	1.762	0.759	2.452	2.532	1.727	1.743	3.165	1.715
	6	3.343	2.300	6.445	11.844	2.644	1.755	2.042	2.332	3.235	2.714
	7	2.777	1.741	1.922	0.896	2.276	2.729	1.976	2.092	2.608	1.859
EDR	1	1.377	1.267	1.024	1.012	2.592	3.376	2.501	1.846	2.476	2.151
	2	4.984	2.916	26.278	14.897	3.603	19.175	3.430	2.064	3.076	1.975
	3	2.465	1.984	4.166	1.678	5.215	0.864	1.606	1.188	4.905	2.496
	4	1.454	1.166	1.261	1.134	2.049	3.327	2.538	1.734	2.177	1.098
	5	1.558	1.247	1.277	1.109	2.125	1.633	1.578	1.168	۲,۷۲۱	1.600
	6	2.492	1.826	4.350	1.739	2.238	1.131	1.771	1.397	3.024	1.741
	7	1.534	1.278	1.204	1.127	1.820	1.796	1.735	1.247	2.311	1.255
ERP	1	1.586	1.433	0.621	1.159	1.655	3.430	2.107	2.198	1.763	1.856
	2	5.323	6.644	51.956	44.851	3.520	136.938	3.185	2.940	2.988	3.954
	3	2.248	2.351	5.893	3.279	2.362	1.426	1.374	1.567	4.181	4.196
	4	1.628	1.514	1.314	0.847	1.969	3.596	2.265	2.006	1.751	2.197
	5	1.435	1.442	1.453	0.925	1.459	2.158	1.291	1.282	2.067	2.202
	6	2.363	2.423	6.546	3.130	1.793	1.553	1.594	1.740	2.573	2.959
	7	1.532	1.440	1.562	3.341	1.498	2.185	1.483	1.440	2.047	2.153

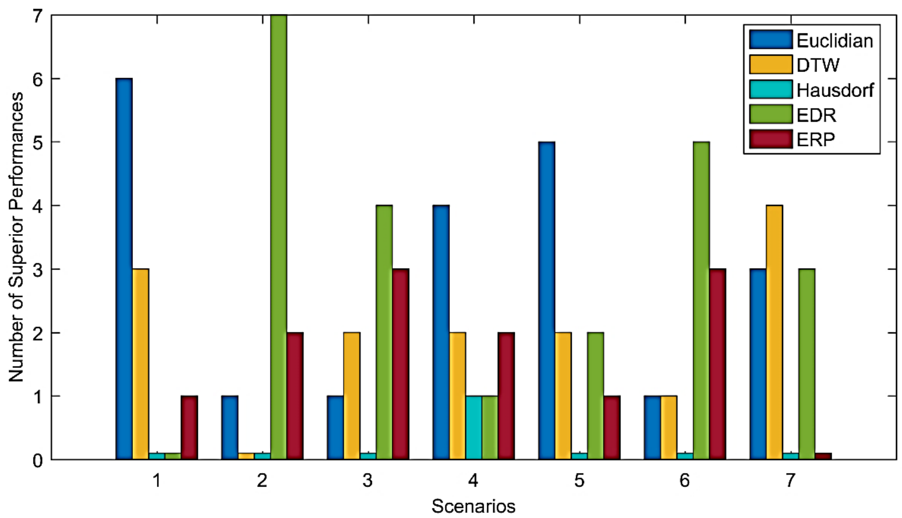

Figure 7. A summary of Table 8 on comparing the performance of different similarity functions

The results of the statistical test for Tables 6 and 7 can be seen in Table 9. According to these results, in scenario B, the superiority of DTW and EDR in AHC clustering of Antarctic and AIS datasets, respectively, are not statistically significant, implying that there are no significant differences in the MDBT index of superior distances than that of other distances. Nevertheless, the results of this table emphasize the statistically significant difference between other similarity functions in Tables 6 and 7, which highlight the substantial effect on the outcomes in case of inappropriate selection of the similarity function. For the sake of space limitation, the ANOVA test results for Table 8 are not indicated here.

Table 9. Analysis of the one-way ANOVA for the results of Tables 6 and 7

Data	**Scenario A**		**Scenario B**	
	K-means	AHC	K-means	AHC
Antarctic	0.000	0.000	0.000	0.707
AIS	0.007	0.002	0.006	0.368
CROSS	0.008	0.000	0.006	0.001
MIT	0.000	0.000	0.000	0.000
Flight	0.000	0.006	0.000	0.005

4.5. Estimating the Number of Clusters

This section aims to study the dependence of the optimal number of clusters on the intended application context. In some clustering algorithms, such as K-means, it is required that the number of clusters (K) is known in advance. The chosen value for K can vary depending on the intended data and application. In case the user is formerly ignorant about the value of K, the number of clusters must be determined in a way that suits the desired application.

Clearly, the quality of the results is highly affected by the proper choice of K. One of the solutions for estimating the optimal number of clusters is calculating a validity index for different numbers of clusters and comparing the results. The user can simply use the visual display of index values to select the appropriate value for K.

Take the case of scenarios (A) and (B), the optimal number of clusters is estimated using the MDBT index in Figures 8 and 9. In this regard, AIS2009 is clustered (using K-means and DTW) with different numbers of K, and then, the MDBT index value is calculated for each k. As it can be seen from Figure 8, if the experts have the same intention of clustering as the scenario (A), K=8 and K=18 are more appropriate than their surrounding ks.

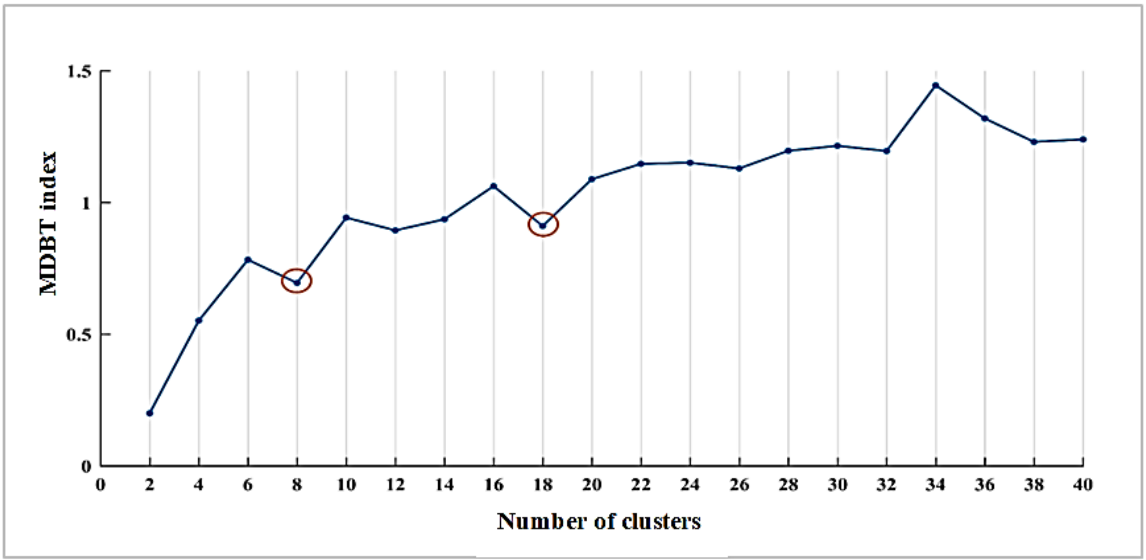

Figure 8. Determination of the numbers of clusters based on the MDBT index for Scenario A.

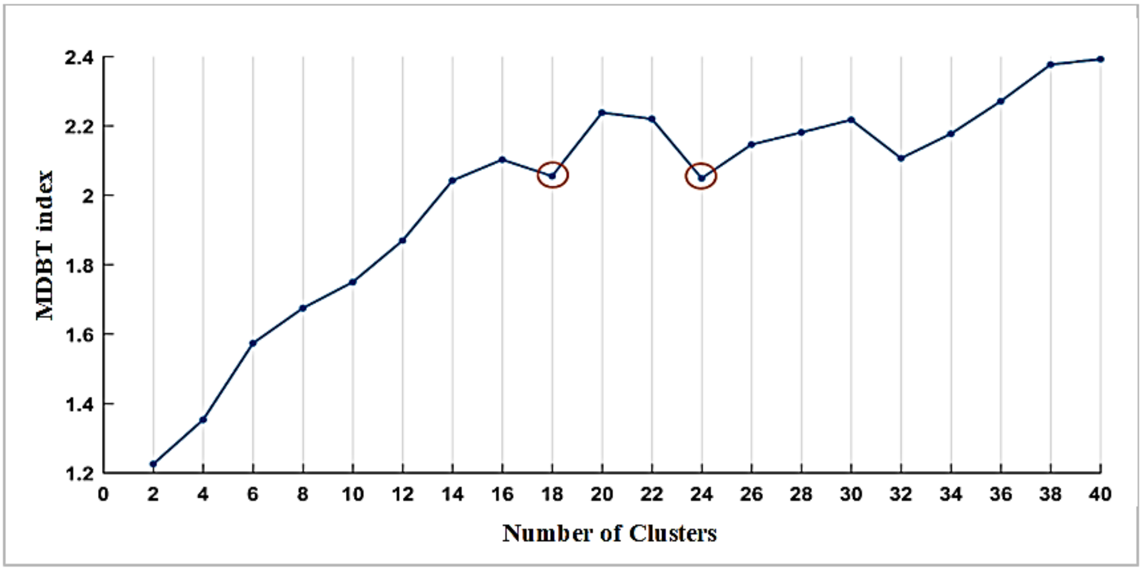

Figure 9. Determination of the numbers of clusters based on the MDBT index for Scenario B.

4.6. Discussion

The results obtained in this study confirm that employing the evaluation framework that considers the application domain for assessing the trajectory clustering results is crucial. In other words, relying on similarity functions that defined the meaning of similarity based on their assumption without considering the user's perspective on similarity will bring about a misleading evaluation.

Moreover, according to the results in Tables 6 and 7, the performance of clustering algorithms is highly related to the data type. For instance, in scenario A on the Antarctic data, DTW with the MDBT index of 1.354 surpasses the other similarity functions when using K-means clustering. On the other hand, by shifting the data to MIT, the Euclidean distance with the MDBT index of 1.201 exhibits superiority over the rest. According to our results, some datasets behave similarly, and their superior similarity function and clustering algorithm are almost identical. Among our datasets, Flight and MIT act in the same way, suggesting they might have similar movement characteristics.

Overall, in scenario A, with the primary emphasis on the positional similarity, the performance of the Euclidean distance is higher for most of the datasets, indicating that if clustering aims at partitioning data based on their spatial closeness, the Euclidean distance yields the best result. On the contrary, in scenario B, in which the focus is not on positional similarity, the superior similarity function varies by different datasets. The other interesting fact here is that AHC exhibits its best performance when we use EDR distance.

A salient point in these results is the superiority of EDR in application contexts with an exclusive focus on direction similarity and speed similarity. Moreover, when users intend to have clusters in which trajectories are similar with regard to both speed and direction similarity, and regardless of their positional similarity, EDR distance surpasses all similarity functions. Nevertheless, when positional similarity is of high importance, like in scenarios 1, 4, and 5, Euclidean distance is the best option. DTW was the superior similarity function when all aforementioned similarities have equal importance.

5. Conclusion

The diversity of similarity functions and clustering algorithms, with each claiming dominance over the other, led this study to launch an inquiry into the evaluation of these similarity functions and algorithms. This research established a new framework for evaluating the performance of trajectory clustering algorithms which examined if outcomes correctly meet users' needs and their application contexts. Due to the complexity in the definition of the similarity concept in trajectory data, evaluation is a challenging undertaking. The meaning of similarity can be highly diverse depending on the users' perspective and the intended application, and thus assuming it as a constant concept will be fallacious. However, if we want to clarify the results of this research in a simple way and for practical applications, employing Euclidean distance is suggested whenever positional similarity is of high importance. On the other hand, in applications where direction and speed similarities are more important, using EDR is recommended.

In this paper, we presented a new framework for evaluating trajectory clustering results by introducing the MDBT index, which uses the expert's opinion in defining the concept of similarity between members of clusters. This new similarity employs three main types of similarity, including positional, speed, and direction similarities as its criteria whose weights are adjusted by the expert's viewpoint and the AHP technique. Another contribution of this study is proposing a method for estimating the appropriate cluster number concerning the clustering application. This can be really helpful in real-world applications, in which determining the number of clusters has been always challenging. Moreover, in this research, the impact of data diversity on clustering performance is examined by using different datasets with huge volumes.

It should be noted that, along with determining the optimal cluster number, the MDBT index can also be used to detect anomalies, compress, and segment trajectories based on the user's perspective. In addition to the employed criteria of similarity, for future research, it is suggested that other similarities, such as semantic, rate of turn, and acceleration be considered in the evaluation of trajectory clustering results. Moreover, using fuzzy AHP might lead to more accurate results.

High computational time was one of the most important limitations of this research. Although the sample size of the data used in this study was higher than many similar studies, access to more powerful and faster processing systems will lead to more reliable results. Another major limitation of this research is the precise definition of an expert. In this study, we tied clustering to expert opinion. However, determining the exact meaning of this phrase in different applications can be challenging.

Overall, based on the acquired results of this research, it is highly recommended that the application context be taken into attention during the clustering evaluation process.

References

Arbelaitz, O., Gurrutxaga, I., Muguerza, J., Pérez, J. M., & Perona, I. (2013). An extensive comparative study of cluster validity indices. Pattern recognition, 46(1), 243-256.

Ardakani, I., Hashimoto, K., & Yoda, K. (2019). Context-based semantical vector representations for animal trajectories. Advanced Robotics, 33(3-4), 118-133.

Baragona, R. (2001). A simulation study on clustering time series with metaheuristic methods. Quaderni di Statistica, 3, 1 26.

Besse, P. C., Guillouet, B., Loubes, J.-M. & Royer, F. (2016). Review and perspective for distance-based clustering of vehicle trajectories. IEEE Transactions on Intelligent Transportation Systems, 17(11), 3306 3317.

Chen, L., & Ng, R. (2004). On the marriage of lp-norms and edit distance. In Proceedings of the Thirtieth international conference on Very large data bases-Volume 30 (pp. 792-803).

Chen, L., Özsu, M. T., & 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 (pp. 491 502).

Cheng, D., Zhu, Q., Huang, J., Wu, Q. & Yang, L. (2018). A Novel Cluster Validity Index Based on Local Cores. IEEE transactions on neural networks and learning systems, 30(4), 985 999.

Davies, D. L. & Bouldin, D. W. (1979). A cluster separation measure. IEEE transactions on pattern analysis and machine intelligence, (2), 224-227.

De Cáceres, M., Coll, L., Legendre, P., Allen, R. B., Wiser, S. K., Fortin, M. J., Condit, R. & Hubbell, S. (2019). Trajectory analysis in community ecology. Ecological Monographs, 89(2), e01350.

Deborah, L. J., Baskaran, R. and Kannan, A. (2010). A survey on internal validity measure for cluster validation. International Journal of Computer Science & Engineering Survey, 1(2), 85-102.

Dodge, S., Weibel, R. & Forootan, E. (2009). Revealing the physics of movement: Comparing the similarity of movement characteristics of different types of moving objects. Computers, Environment and Urban Systems, 33(6), 419-434.

Dodge, S., Weibel, R. & Lautenschütz, A.-K. (2008). Towards a taxonomy of movement patterns. Information visualization, 7(3 4), 240 252.

Hurter, C., Alligier, R., Gianazza, D., Puechmorel, S., Andrienko, G. & Andrienko, N. (2014). Wind parameters extraction from aircraft trajectories. Computers, Environment and Urban Systems, 47, 28-43.

Huttenlocher, D. P., Klanderman, G. A. & Rucklidge, W. J. (1993). Comparing images using the Hausdorff distance. IEEE Transactions on pattern analysis and machine intelligence, 15(9), 850 863.

Kan, Z., Tang, L., Kwan, M.-P., Ren, C., Liu, D. & Li, Q. (2019). Traffic congestion analysis at the turn level using Taxis' GPS trajectory data. Computers, Environment and Urban Systems, 74, 229-243.

Košmelj, K. & Batagelj, V. (1990). Cross-sectional approach for clustering time varying data. Journal of Classification, 7(1), 99-109.

Lee, S.-H., Jeong, Y. S., Kim, J.-Y. & Jeong, M. K. (2018). A new clustering validity index for arbitrary shape of clusters. Pattern Recognition Letters, 112, 263-269.

Magdy, N., Sakr, M. A. & El-Bahnasy, K. (2017). A generic trajectory similarity operator in moving object databases. Egyptian Informatics Journal, 18(1), 29 37.

Mao, Y., Zhong, H., Xiao, X. & Li, X. (2017). A segment-based trajectory similarity measure in the urban transportation systems. Sensors, 17(3), 524.

Moayedi, A., Abbaspour, R. A. & Chehreghan, A. (2019). An evaluation of the efficiency of similarity functions in density-based clustering of spatial trajectories. Annals of GIS, 25(4), 313-327.

Montazeri-Gh, M. & Fotouhi, A. (2011). Traffic condition recognition using the k-means clustering method. Scientia Iranica, 18(4), 930-937.

Morris, B. T. & Trivedi, M. M. (2011). Trajectory learning for activity understanding: Unsupervised, multilevel, and long-term adaptive approach. IEEE transactions on pattern analysis and machine intelligence, 33(11), 2287-2301.

Myers, C., Rabiner, L. & Rosenberg, A. (1980). Performance tradeoffs in dynamic time warping algorithms for isolated word recognition. IEEE Transactions on Acoustics, Speech, and Signal Processing, 28(6), 623-635.

Niu, X., Chen, T., Wu, C. Q., Niu, J., & Li, Y. (2019). Label-based trajectory clustering in complex road networks. IEEE Transactions on Intelligent Transportation Systems, 21(10), 4098-4110.

Olive, X. & Morio, J. (2019). Trajectory clustering of air traffic flows around airports. Aerospace Science and Technology, 84, 776-781.

Ossama, O., Mokhtar, H. M. & El-Sharkawi, M. E. (2011). An extended k-means technique for clustering moving objects. Egyptian Informatics Journal, 12(1), 45-51.

Pelekis, N., Kopanakis, I., Kotsifakos, E. E., Frentzos, E. & Theodoridis, Y. (2011). Clustering uncertain trajectories. Knowledge and information systems, 28(1), 117-147.

Priestley, M. (1980). State- dependent models: a general approach to non- linear time series analysis. Journal of Time Series Analysis, 1(1), 47-71.

Tarroux, A., Weimerskirch, H., Wang, S. H., Bromwich, D. H., Cherel, Y., Kato, A., ... & Descamps, S. (2016). Flexible flight response to challenging wind conditions in a commuting Antarctic seabird: do you catch the drift?. Animal Behaviour, 113, 99-112.

Teknomo, K. (2006). Analytic hierarchy process (AHP) tutorial. Revoledu. com, 6(4), 1-20.

Wang, X., Ma, K. T., Ng, G. W., & Grimson, W. E. L. (2011). Trajectory analysis and semantic region modeling using nonparametric hierarchical bayesian models. International journal of computer vision, 95(3), 287-312.

Xu, H., Zhou, Y., Lin, W., & Zha, H. (2015). Unsupervised trajectory clustering via adaptive multi-kernel-based shrinkage. In Proceedings of the IEEE international conference on computer vision (pp. 4328-4336).

Yao, D., Zhang, C., Zhu, Z., Hu, Q., Wang, Z., Huang, J. & Bi, J. (2018). Learning deep representation for trajectory clustering. Expert Systems, 35(2), e12252.

Yu, D., Wang, J. & Wang, R. (2019). Exploring traffic condition based on massive taxi trajectories. International Journal of High Performance Computing and Networking, 14(1), 30-41.

Zhao, D. & Stefanakis, E. (2018). Mining massive taxi trajectories for rapid fastest path planning in dynamic multi-level landmark network. Computers, Environment and Urban Systems, 72, 221-231.

Zhu, E. & Ma, R. (2018). An effective partitional clustering algorithm based on new clustering validity index. Applied Soft Computing, 71, 608-621.