A novel similarity measure for multiple aspect trajectory clustering

Iraklis Varlamis
Department of Informatics and Telematics, Harokopio University
Athens, Greece
varlamis@hua.gr

Christos Sardianos
Department of Informatics and Telematics, Harokopio University
Athens, Greece
sardianos@hua.gr

Vania Bogorny
Department of Informatics and Statistics, Universidade Federal de Santa Catarina
Florianópolis, Brazil
vania.bogorny@ufsc.br

Luis Otavio Alvares
Department of Informatics and Statistics, Universidade Federal de Santa Catarina
Florianópolis, Brazil
luis.alvares@ufsc.br

Jônata Tyska Carvalho
Department of Informatics and Statistics, Universidade Federal de Santa Catarina
Florianópolis, Brazil
jonata.tyska@ufsc.br

Chiara Renso
ISTI-CNR
Pisa, Italy
chiara.renso@isti.cnr.it

Raffaele Perego
ISTI-CNR
Pisa, Italy
raffaele.perego@isti.cnr.it

John Violos
Department of Informatics and Telematics, Harokopio University
Athens, Greece
violos@hua.gr

ABSTRACT

Multiple aspect trajectories (MATs) is an emerging concept in the domain of Geographical Information Systems, where the basic view of semantic trajectories is enhanced with the notion of multiple heterogeneous aspects, characterizing different semantic dimensions related to the pure movement data. Many applications benefit from the analysis of multiple aspects trajectories, ranging from the analysis of people trajectories and the extraction of daily habits to the monitoring of vessel trajectories and the detection of outlying behaviors. This work proposes a novel MAT similarity measure as the core component in a hierarchical clustering algorithm. Despite the many clustering methods in the literature and the recent works on MAT similarity, there are still no works that dig deeper into the MAT clustering task. The current article copes with this issue by introducing TraFoS, a new similarity measure that defines a novel method for comparing MATs. TraFos includes a multi-vector representation of MATs that improves their similarity comparison. TraFos allows us to compare MATs across each aspect and then combine similarities in a single measure. We compared TraFos with other state of the art similarity metrics in Agglomerative clustering. The experimental results show that TraFos outperforms other similarities metrics in terms of internal, external clustering metrics and training time.

CCS CONCEPTS

• **Information systems** → **Geographic information systems**; **Similarity measures**; *Clustering and classification*;

KEYWORDS

semantic trajectories; multi-aspect trajectories; trajectory similarity; trajectory clustering.

ACM Reference Format:
Iraklis Varlamis, Christos Sardianos, Vania Bogorny, Luis Otavio Alvares, Jônata Tyska Carvalho, Chiara Renso, Raffaele Perego, and John Violos. 2021. A novel similarity measure for multiple aspect trajectory clustering. In *The 36th ACM/SIGAPP Symposium on Applied Computing (SAC '21), March 22–26, 2021, Virtual Event, Republic of Korea.* ACM, New York, NY, USA, 8 pages. https://doi.org/10.1145/3412841.3441935

1 INTRODUCTION

Trajectory clustering is an important data mining method that can be useful in several application fields like object tracking in video sequences, airspace monitoring or the detection of common and outlying behaviors in vessel routes. The main purpose of trajectory clustering algorithms is to group similar trajectories or moving objects together and thus provide a better understanding of commonalities that exist between the trajectories of different objects, or of the same object at different moments. Trajectory clustering allows us to extract patterns [17] and detect common and outlying moving objects behaviors [16].

The successful application of clustering techniques strongly depends on the appropriate similarity metric. Regarding trajectory

SAC '21, March 22–26, 2021, Virtual Event, Republic of Korea

ACM ISBN 978-1-4503-8104-8/21/03…$15.00
https://doi.org/10.1145/3412841.3441935

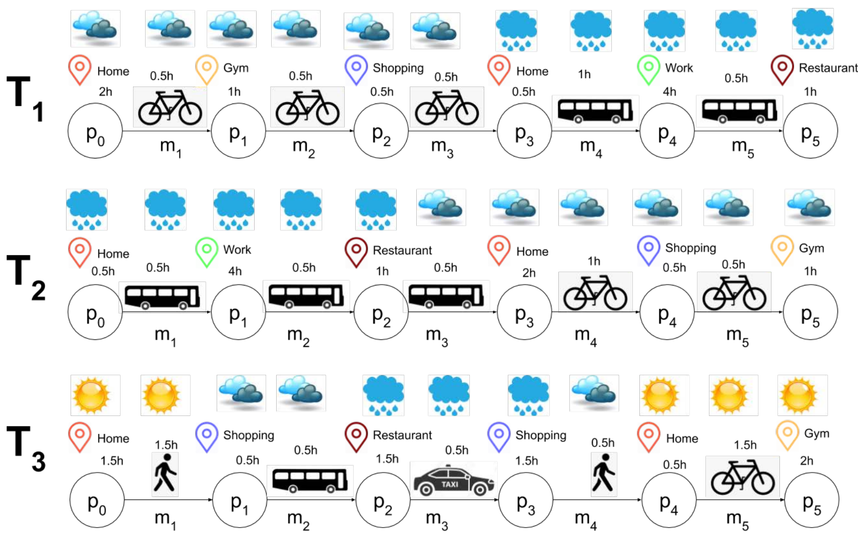

Figure 1: An example of three trajectories with their multiple aspects.

clustering, similarity criteria usually assume that trajectories are sequences of points in the 2D or 3D space with timestamp, and mainly rely on the spatial distance (e.g. Euclidean, Mahalanobis, great circle distance etc.) of the respective points. Difference in length is a specific problem when comparing two trajectories and therefore there is a need for a trajectory alignment pre-processing step, where dynamic time warping, subsequence matching, or edit distance [25] may assist.

MATs might include trajectory related features, such as speed, direction and duration. They may also contain contextual information that is extracted from different external sources like weather or air pollution, the moving object status (e.g. heart rate), etc. During their daily routines, users commutes to work every day, visit the same gym, go to restaurants, and malls, but not always to the same ones and in the same order. Their GPS traces provide only part of the information, which mainly refers to the spatial attributes of each stop and each transition, ignoring all other semantics: the time spent at each place, the means of transportation and the duration of each move, the weather conditions that held at each segment, etc. The spatial dimension is only one of the several aspects of these MATs, which may contain more aspects that can be described by a number of different attributes (e.g. a POI is described by its category, the time that the user spends there and the weather conditions at that time). We call these cases *multiple-attribute aspects*.

In the motivating example of Figure 1, three trajectories T_1, T_2, and T_3 that begin from a user's home and have a total duration of 12 hours are depicted. It is worth noting that the spatial aspect of POIs is ignored since we assume that the semantics of stops and moves are more relevant for MATs clustering. Each trajectory has 6 stops (p_0 to p_5, with p_0 being at home) and 5 moves ($p_0 \rightarrow p_1$, ..., $p_4 \rightarrow p_5$). Each stop has 3 different aspects: i) the weather condition that can be either sunny, cloudy or rainy, ii) the stop type (or POI type) that can be one of home, gym, shopping, work or restaurant and the hours spent there that is a multiple of 0.5 hours. Similarly, the moves have three aspects: i) the means of transportation (walk, bus, taxi, bicycle), ii) the duration that is again a multiple of 0.5 hours, iii) the weather condition that, for the sake of simplicity, we assume that is the same during the whole move.

Clustering of MATs poses new challenges since aspects can comprise multiple heterogeneous attributes with dependency relations among themselves. Thus, clustering algorithms for MATs, must adopt multidimensional similarity metrics that capture such semantics [8]. Furthermore, MATs are affected by the curse of dimensionality due to their high dimensional nature. The sparsity of the attribute space, makes ineffective most of the commonly used similarity metrics. The more complex and heterogeneous the aspects are, the more flexible the similarity measure should be. A totally rigid measure that requires the perfect match on each aspect, for each point, will be too strict and would not capture the inherent heterogeneity of the human movements. On the other hand, a very flexible measure would not capture the implicit similarity among different users. It is, therefore, challenging to define a suitable trade-off between the two extremes.

Regarding the selection of the clustering algorithm, centroid-based clustering algorithms, such as k-Means, are very fast and perform well, but can only find globular shaped clusters. Spectral clustering cannot successfully cluster data that contain structures at different scales of size and density [18]. Single and average linkage agglomerative clustering algorithms are based on the pairwise comparison of MATs and does not require apriori knowledge of the number of clusters.

In this paper, we perform a study and comparison of state-of-the-art similarity measures for MATs applied to the clustering task. We also introduce a novel MAT similarity measure, called *Trajectory Forest Similarity* (TraFoS)[1], with the objective of overcoming the limitations of existing measures and providing a smooth integration with available clustering algorithms.

In summary, the contributions of this work include:

[1]The term 'Forest' comes from isolation trees and is irrelevant to the Random Forest classification algorithm.

- a comprehensive analysis of representation models for MATs, the similarity measures that can be used for their comparison, and their application in agglomerative clustering;
- a new vector representation model for semantic trajectories that captures the frequency of semantic information across the trajectory points and allows fast similarity comparisons;
- a tree-based similarity measure for trajectories (TraFoS) that employs the novel representation and a hierarchical partitioning of the initial set of trajectories. TraFoS has lower time and memory requirements than other state-of-the-art trajectory similarity measures for semantic trajectories, compares to their performance, and outperforms traditional geographical distance and edit distance measures.

In Section 2 that follows, we summarize related works in the field of similarity measures, with emphasis on the clustering of semantic trajectories. Section 3 discusses the notion of MAT and clustering. Section 4 provides the details of the MAT representation and the proposed TraFoS similarity measure. Section 5 describes the experimental evaluation of the proposed solution, its algorithmic implementation and discusses the results. Finally, Section 7, concludes the paper and highlights the next steps of our work.

2 RELATED WORK

The interest of the research community for trajectory analysis, similarity and clustering is high due to the abundance of positioning equipment and location tracing devices, enabling the track of moving objects that range from vehicles, vessels, and planes to humans and animals. However, the majority of works have mainly considered the spatio-temporal properties and derivative measures from these dimensions [3, 12, 19] with little attention to the semantic dimensions.

The first works that go beyond their geometric properties, and consider their semantic dimensions, appeared since 2007 [1, 2, 22]. In these works, semantics are assigned to special parts of a trajectory, also called sub-trajectories, such as the stops and the moves between them. More recently, we have seen the emergence of the concept of MATs [6, 15], which introduces the enrichment of each trajectory point with several layers of semantic information, called aspects. Semantic information ranges from the characteristics of a visited place (e.g. opening hours, price, rating, etc) to the relationships of a moving object with other objects [15], like the encounters between moving objects, this last one also became an important topic in the scenario of the COVID-19 pandemic.

The problem of MATs clustering has not yet been tackled in-depth in the literature. Many algorithms exist for clustering moving object trajectories however they use spatio-temporal information only [28]. It is worth noting that MATs contain rich information about the trajectory context, the semantics of movement, the surrounding elements, and many aspects related to the moving object [8]. Therefore it is important to create similarity measures that capture the heterogeneous aspects, the different nature, scale, and dimensionality of MATs various features. A recently proposed method called evolving clusters algorithm [23] is one of the few methods dealing with semantic information for trajectory clustering. It employs the spatial distance of trajectory points, graph mining algorithms (Clique and Maximal Connected Subgraph detection) and trajectories extended with a few semantic annotations. The algorithm has been evaluated in discovering unified group behaviour of moving objects (e.g. flocks or convoys), but has not been tested in detecting very similar or repeating MATs.

Among the state-of-the-art similarity measures for trajectories, stands the Uncertain Movement Similarity (UMS) [7]. UMS has outperformed a number of previous measures developed for spatio-temporal trajectories, but it is limited to the spatial dimension. If we also consider the semantic information, then we have to mention the *Stops and Moves Similarity Measure* (SMSM) [13], which considers both spatial and semantic features, but it is limited to trajectories represented as sequences of stops and moves.

Longest Common Subsequence (LCSS) was originally introduced for sequence comparison and employed as a robust similarity measure for raw trajectories in [24]. LCSS considers all aspects equally important and relevant, and requires a strict match for all aspects of each stop or move to consider them as similar. *Edit Distance on Real sequence* (EDR) [4] is a similarity measure based on the Edit Distance, a popular metric for comparing strings. EDR measures the minimum number of inserts, deletes, and replacements of points (stops or moves in our case) needed to transform one trajectory to the other is counted. Although EDR assigns penalties (increases the distance) according to the gap (mismatch) length, it still requires an exact match in all aspects. The result is that EDR and LCSS are too restrictive for measuring similarities in multiple-aspect trajectories, such as the daily user routines.

Multidimensional similarity measure (MSM) [8] overcomes some limitations of previous measures by explicitly including the semantic dimension of trajectories in addition to space and time. MSM examines every aspect separately and supports different weights for each aspect, which assign more or less importance to each aspect based on the application needs. When comparing the stops in two trajectories, MSM seeks for every stop in the first trajectory the best match in the second. However, MSM disregards any relationships that might exist between aspects or attributes (e.g. type of stop and weather conditions), making it less robust for the case of multiple-aspect trajectories.

MUltIple-aspect TrAjectory Similarity (MUITAS) [20] overcomes the limitations of MSM by supporting composite aspects (e.g. POI type and price), which aggregate multiple attributes and provide a more comprehensive trajectory comparison. The multi-attribute aspects (also called *composite-aspects*) consider the semantic relationship between the attributes of a trajectory. In addition it supports the use of thresholds in each aspect value of the multi-attribute aspects, and the use of different weights across aspects, thus increasing the flexibility of the trajectory similarity metric.

The improved performance of MUITAS, compared to other measures, is due to the fact that trajectories are considered as sets of points, rather than sequences. However, a limitation of this measure is that the popularity of attributes shared by multiple points or the frequency of occurrence of points themselves are not explicitly considered. Moreover, the computational cost of the algorithm is very high due to the pairwise trajectory similarity measure, which is quadratic to the number of points and aspects.

The aforementioned similarity measures either assume that aspects are totally unrelated to each other, or they indiscriminately combine all attributes, even those that do not relate to others. They are also limited to the order of stops and/or moves, and seek for shorter or lengthier sub-trajectories. In addition, they are not considering the frequency of occurrence of each aspect value in the trajectory. Last but not least, their complexity is dependent on the trajectory length and the number of aspects, since they compare all points of each trajectory to all the points in other trajectories, and they do this for all aspects.

To address these limitations we introduce a novel measure that we call TraFoS, which takes into account the frequency of values in the trajectory aspects, generates a representation of all trajectories in the same vector space and considers the whole set of trajectories before comparing them in a pair-wise level. This results in a much faster and more fair calculation of trajectories similarity, which accounts for the information contained in the whole set. The measure first creates a cluster hierarchy for each aspect, using a vector representation and a vector space similarity, and then combines the hierarchies in order to provide a multiple-aspect similarity measure.

3 CLUSTERING OF MULTIPLE-ASPECT TRAJECTORIES

The clustering of multiple-aspect trajectories is an unsupervised process that may extract useful patterns or detect interesting outliers. In the same time, it is a complex task that must consider: (1) the similarity of trajectories, which includes how trajectory information is represented and compared, (2) the clustering algorithm to use. The trajectory similarity measure has to take into account many dimensions, such as i) the similarity between individual components of the trajectories (e.g. points, stops, or movements), ii) the order in which these components occur in each trajectory, and iii) their frequency of occurrence. The trajectory clustering algorithm accounts for the nature of the clustering task at hand, as well as the underlying cluster model and the properties of the produced clustering scheme.

When defining the similarity between two multiple-aspect trajectories, the semantics of each aspect and the relationships between aspects could be more important than the order of stops and moves, their spatial or temporal relations, as well as the exact match in all attributes. Also, the frequency of each stop or move, and the frequency of the respective aspect values in each trajectory is of equal importance. The user may visit places of the same type (e.g. restaurants), which may be located far away from each other, and may spend the same time (e.g. 1 hour), but in completely different moments of the day. In the example presented in Figure 1, for instance, even though trajectory T_1 has many aspect values in common with T_3, they still have many differences that make them semantically different.

One of the motivations behind developing a new measure for comparing multiple-aspect trajectories is that the current state of the art measures, such as MUITAS and MSM, do not take into account the frequency of aspect values during comparisons. They also have a high computational complexity, which limits their applicability in the comparison of long trajectories or in the clustering of large trajectory datasets.

4 TRAFOS: A NEW SIMILARITY MEASURE FOR MATS

TraFoS is based on the creation of a hierarchical tree for each aspect that is being used to measure the similarity of trajectories along that aspect. This tree results from a repetitive partitioning of the full set of trajectories using a binary partitioning algorithm. The binary partitioning algorithm splits the set of trajectories into two subsets at each step, until the resulting clusters contain only a few trajectories each. However, any other binary or multiple-way partitioning algorithm could be employed alternatively.

The hierarchical trees of TraFoS have an interesting property: very similar trajectories are in neighboring hierarchy nodes, whereas very dissimilar trajectories belong to distant nodes. As we will see in the next subsections, using this property, we define a new similarity/distance measure for multiple-aspect trajectories, which is based on the distance of trajectories in such a hierarchy.

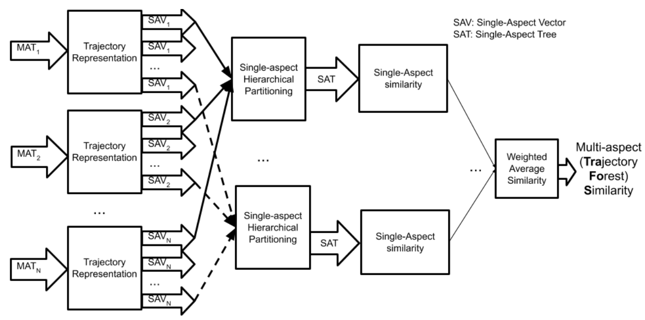

Figure 2: The steps of TraFoS similarity for MATs.

4.1 Trajectory representation

A fundamental step of the TraFoS clustering method is to transform multiple aspect trajectories into frequency vectors representing the number of occurrences of each value of a given aspect. In this representation, trajectories are therefore bag-of-points counting the number of points that share that aspect value. Trajectory points may correspond to stops or moves.

Let us better introduce this idea by using the example of Figure 1. Trajectory T_1, regarding the weather aspect on points, contains 3 cloudy and 3 rainy stops (and no sunny stops). Therefore, the representation in the 3-Dimensional space $< cloudy, rainy, sunny >$ of the weather values for stops will be $< 3, 3, 0 >$. The respective vectors for T_2 and T_3 will be $< 3, 3, 0 >$ and $< 1, 2, 3 >$. Similarly, the vectors for weather conditions during moves are $< 3, 2, 0 >$, $< 3, 2, 0 >$ and $< 2, 1, 2 >$ for T_1, T_2 and T_3 respectively. Although nominal attributes are used in all examples of this study for simplicity, the method can also handle numeric attributes with continuous values. Apart from the k-means clustering that is actually used, any other unsupervised discretization technique [5] (e.g. multiple correlation clustering) can be employed.

Without loss of generality, when the compared trajectories differ substantially in length, we can normalize the absolute frequency described above by dividing it by the trajectory length. Afterwards, we can easily define the single-aspect similarity of two trajectories, e.g. using cosine similarity or any other similarity measure defined in a vector space. Similarly, Manhattan, Euclidean or any other Minkowski distance can also be used.

A vector-based representation measures the similarity of trajectories, ignoring the order of points that contribute to each value. It is important to clarify here that since trajectories are represented using frequency (absolute or relative) vectors, which are built on the values of an aspect, the vector dimensions are all the possible values that the aspect can take.

4.2 Single-aspect tree-based trajectory similarity

Given a set of multiple-aspect trajectories, their single-aspect vector (SAV) representations, and their pair-wise vector similarity, the next step for TraFoS is to construct one hierarchical partitioning tree for each aspect. In a single-aspect tree (SAT), trajectories that are very similar in this aspect are expected to be in the same partition for many levels, whereas very dissimilar trajectories will be separated in the top levels of the respective tree.

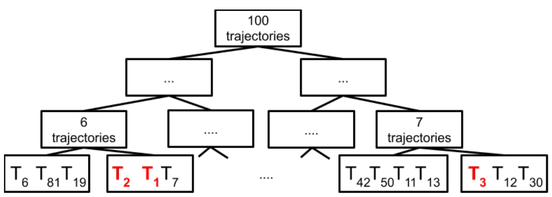

Figure 3: A sample single-aspect tree created using hierarchical partitioning.

For example, if we partition a set of 100 trajectories using the weather aspect the result will be similar to the one depicted in Figure 3. In this hierarchical partitioning process we assume that all trajectories belong to the same node (i.e. cluster or partition) in the root of the tree, and, as we move towards the leaf level, fewer and fewer trajectories will be on each node due to the recursive binary partitioning of the trajectory sets. All trajectories will be assigned to some leaf, where each leaf represents a small group of very similar trajectories. Trajectories from neighboring leaf clusters will be more similar than trajectories from distant leaf clusters.

The node partitioning process may continue until each leaf contains only one trajectory. In order to avoid this extreme partitioning of the dataset, we can set an early stop criterion, or apply post-pruning to the resulting tree. The partitioning of nodes can be done using any function that splits the initial set of trajectories in two subsets.

The single-aspect tree created for each aspect, using the process explained above, is the basis for the single-aspect similarity measure of TraFoS, which has an interesting property, compared to flat partitioning schemes. With flat partitioning, in general, we can easily infer that trajectories that end-up in the same partition are similar. However, it is not straightforward to define the similarity of trajectories lying in different partitions. With hierarchical partitioning schemes, we can use path distance measures to compare trajectories that fall in different, neighboring, or distant leaves. For this purpose, we employ the Wu and Palmer similarity measure for hierarchies [26], which is based on the depth of two items (that are trajectories in our case) in the tree and the depth of their least common ancestor (LCA), which is, in our case, the deepest partition in the tree that contains them both.

Equation 1, defines the Wu and Palmer similarity for two trajectories T_i and T_j, LCA(T_i,T_j) is the deepest cluster that contains both trajectories, L(T_x) is the leaf cluster that contains trajectory T_x and $d(c)$ function returns the depth of a cluster c.

$$Sim_{wup}(T_i, T_j) = \frac{2 * d(LCA(T_i, T_j))}{d(L(T_i)) + d(L(T_j)) + 2 * d(LCA(T_i, T_j))} \quad (1)$$

The creation of the single-aspect tree and the definition of the tree-based similarity for trajectories (based on the Wu & Palmer similarity measure) provides a relaxed comparison for trajectories, which allows "birds of a feather to flock together" and in the same time allows to define multi-aspect trajectory similarity as it is explained in the following section.

From a semantic point of view, two trajectories are similar in an aspect when they express highly similar frequencies of this aspect. Since the focus is on frequencies, the chronological order of values in the trajectory is inherently ignored. However, the proposed method allows the definition of aspects that examine short sequence of values (e.g. pairs or triples) in analogy to word n-grams in text sequences, which can capture short-term ordering on aspects of interest. Similarly, it can support composite aspects that examine the duration of stops or moves of each type and allow the comparison of temporarily annotated trajectories [9] as explained in the following subsection.

4.3 Multiple-aspect trajectory similarity

The set of all single-aspect trees created using the above process constitutes the multiple-aspect trajectory forest, which can provide a multi-aspect trajectory similarity measure. The similarity of the two trajectories can be the (weighted) sum of single-aspect similarities as we can see in Figure 2. Trajectories that are very similar, will expose high similarity in multiple aspects and thus they will be in the same of neighboring leaves in multiple trees, which means that their (weighted) average similarity will be high.

Figure 4 presents an illustrative multiple-aspect trajectory forest, constructed on two basic aspects (i.e. stop duration, POI type) and a composite aspect (i.e. stop duration and POI type), of a dataset with 100 trajectories. The forest comprises one tree for each stop aspect, and an extra tree for the composite aspect. As depicted in the figure, the set of 100 trajectories is likely to be partitioned differently by each tree and two trajectories may be in neighboring or in the same leaf in one tree, which denotes high similarity in one aspect and in distant leaves in another tree, denoting high dissimilarity.

The use of binary clustering trees allows us to quickly separate similar trajectories from dissimilar ones and provides a fast, but comprehensive, representation of the value space of each aspect. It also allows bi-aspect or multiple-aspect trees to be created, using composite aspects and values as input. For example, trajectories of the form (morning, sunny), (noon, cloudy), (afternoon, sunny), (night, cloudy), ..., (night, rainy) can simply be represented as vectors in a high-dimensional space.

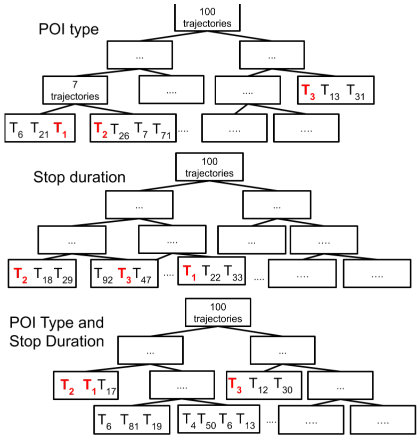

Figure 4: A sample set of single- and multiple-aspect trees that constitute the multiple-aspect forest.

5 EXPERIMENTAL EVALUATION

The experimental evaluation assesses the performance of different similarity measures in the multiple-aspect trajectories clustering of several users in a way that the trajectories of the same user cluster together. All experiments were run on a DELL Inspiron laptop, with an Intel Core i7 processor and 16 GB RAM, running Windows.

Dataset: We use a real-world dataset containing trajectories of Foursquare users enriched with semantic aspects. The dataset, also used in [20], is based on the dataset presented in [27], which was enriched with the addition of semantic aspects related to venue information and weather conditions. The venue information includes the spatial position, rating, and price tier, collected using the Foursquare API[2].

The dataset was pre-processed to remove noisy, duplicate and incomplete information records. The trajectories were split to create weekly trajectories for each user and labeled with the user id in order to provide a weak supervised ground-truth for clustering evaluation. This ground-truth assumes high self-similarity among the weekly trajectories of each user and dissimilarity to the trajectories of other users. Trajectories with less than 10 check-ins and users with less than 10 trajectories were removed from the dataset. The final dataset contains a total of 66,962 check-ins distributed in 3,079 weekly trajectories of 193 different users, with an average length of approximately 22 check-ins per trajectory and an average of approximately 16 trajectories per user [20].

For each check-in in the user trajectory, we keep the following features, which are later used to compose the trajectory aspects: i) latitude (numeric), ii) longitude (numeric), iii) time of the day (numeric), iv) day of the week (nominal), v) check-in type (nominal), vi) check-in category (nominal, broader than type), vii) price level (ordinal), viii) rating level (ordinal), ix) weather (nominal). The time of the day has been converted to a nominal features by using a discretization to 24 bins of equal frequency and another one using 24 equi-width bins (1 hour each).

[2]https://developer.foursquare.com/. The weather information has been collected for each check-in, via the Weather Wunderground API[3]

Clustering algorithm: We employed agglomerative clustering using single and average linkage. Both of these approaches are available through the Scikit-learn machine learning library[4]. To provide a fair comparison for all similarity measures, we always employed the default hyper parameters for the clustering algorithms and set the number of clusters to be equal to the number of different users (i.e. 193). We run every clustering algorithm 10 times and report on the mean internal and external evaluation metrics.

Evaluation metrics: Clustering evaluation is a well-known issue in the literature. This is due to the fact that clustering is an unsupervised method and we don't usually have a ground truth to compare with. However, in the case of trajectory similarity, we can assume that trajectories of the same user are likely to belong to the same cluster, as indicated by [10] and already used in other state-of-the-art works like [14]. Therefore, the external evaluation of the clustering method is based on this ground truth. For the internal clustering validity metrics, we assume that the best clusters are those that are well separated and compact, as described in [21].

Table 1: Agglomerative Clustering using Single Linkage

Sim. Meas.	Homogen.	Complet.	V-measure	Mut. Info	Rand	F.M.	Silhouette	C.H.	D.B.
Mean TraFoS	0.08	0.72	0.14	0.14	0.00	0.08	-0.94	0.99	1.01
Max TraFoS	0.08	0.72	0.14	0.14	0.00	0.08	-0.94	0.99	1.01
W/Thr TraFoS	0.08	0.72	0.14	0.14	0.00	0.08	-0.94	0.99	1.01
MUITAS	0.06	0.58	0.11	0.11	0.00	0.07	-0.19	0.73	1.18
MSM	0.06	0.58	0.11	0.11	0.00	0.07	-0.13	0.70	1.22
EDR	0.08	0.72	0.14	0.14	0.00	0.08	-0.94	0.99	1.00
LCSS	0.08	0.72	0.14	0.14	0.00	0.08	-0.94	0.96	1.02

Table 2: Agglomerative Clustering using Average Linkage

Sim. Meas.	Homogen.	Complet.	V-measure	Mut. Info	Rand	F.M.	Silhouette	C.H.	D.B.
Mean TraFoS	0.24	0.32	0.27	0.27	0.00	0.01	-0.95	0.96	3.66
Max TraFoS	0.24	0.32	0.27	0.27	0.00	0.01	-0.95	0.96	3.67
W/Thr TraFoS	0.24	0.32	0.27	0.27	0.00	0.01	-0.95	0.96	3.67
MUITAS	0.08	0.56	0.13	0.13	0.00	0.07	-0.13	1.21	2.21
MSM	0.07	0.58	0.12	0.12	0.00	0.07	-0.07	1.63	1.83
EDR	0.17	0.28	0.21	0.21	0.00	0.02	-0.94	0.99	2.86
LCSS	0.19	0.30	0.23	0.23	0.00	0.02	-0.95	0.96	3.19

The external (supervised) and internal (unsupervised) clustering validity metrics [11] employed, comprise: i) homogeneity score (external), ii) completeness score (external), iii) v-measure score (external), iv) adjusted mutual info score (external), v) adjusted rand score (external), vi) Fowlkes Mallows score (external), vii) silhouette score (internal), viii) Calinski Harabaz score (internal), ix) Davies-Bouldin Index (separation) (internal). In Tables 1 and 2 we use light gray shading for the columns that depict external metrics and dark gray shading for those reporting the internal metrics.

6 RESULTS AND DISCUSSION

TraFoS creates the trajectory forest, by constructing each binary partitioning tree using the Spectral Clustering algorithm (with the number of cluster = 2) and nominal or ordinal features and their combinations only. Any other partitioning clustering algorithm (e.g. k-means with k=2) could be used. The latitude and longitude of check-ins were not used.

For the 3079 trajectories in the dataset, we construct 14 trees in total, 8 for the different aspects, each one examined separately,

[4]https://scikit-learn.org/stable/

and another 6 for binary or ternary aspect combinations of the most strongly correlated aspect combinations: (i) check-in general category, price and rating level, ii) weather, day of the week and check-in general category). This correlation was intuitively chosen based on the nature of the task (e.g. the cost of a service is usually associated by users with the service quality). Of course, any other data-driven inference technique (e.g. correlation analysis) can be employed.

Each tree has on average 1740 nodes and 870 leaves and a maximum tree height between 14 and 18. The average construction time for each partitioning tree, including the time to calculate all the pairwise (Wu and Palmer, [26]) similarities for all trajectories is 181 seconds. All experiments were run on a PC equipped with an Intel i5 CPU and 8GB of Memory, running MS Windows. Using the binary partitioning forest, we evaluate three different TraFoS variations: i) the first considers the average Wu and Palmer [26] similarity across all trees, ii) the second takes the maximum similarity across trees and iii) the third sets a threshold on the average similarity across all trees and filters out (i.e. sets to zero) the scores that are lower than 0.6[5]).

The time for pre-computing the pairwise similarity matrix for all trajectories for the first TraFoS variation was less than 10 minutes (590 seconds), 90 seconds for the second (i.e. max similarity across all trees) and less than a second for the third (which simply thresholds the first similarity matrix).

All the experiments were based on computing the pairwise similarity matrix for all trajectories in the dataset. It is worth noting the different execution time needed to compute the similarity matrix for the above similarity measures. The execution time for EDR and LCSS was the fastest among all, since it took less than 20 minutes to compute all the pairwise similarities. The complexity of MSM and MUITAS similarity measures is high due to the all-to-all comparison of all the stops, which for long trajectories may be extremely demanding in resources. Running both MUITAS and MSM to compute the pairwise similarity matrix took more than 4 hours. In the case of TraFos, the construction of the 14 binary partition trees took 42.5 minutes to complete. Adding the 10 minutes needed to compute the similarity matrix from the trees, sums to less than one hour, which is 4 times faster than MUITAS and MSM.

When the grouping of trajectories by user is employed as our ground truth for (external) clustering evaluation, MUITAS and MSM have their worst performance (e.g. the lowest V-measure scores). This can be explained by the very low homogeneity score, which means that a few large clusters dominate the clustering scheme.

The performance of TraFoS using Hierarchical Agglomerative Clustering (HAC) with single linkage between clusters (i.e. the similarity between two clusters is the maximum pairwise similarity between individual trajectories, one from each cluster) shows very high completeness scores, compared to MUITAS or MSM, and similar scores to EDR and LCSS. High completeness means that the trajectories of the same user are clustered together. However, the same cluster may contain trajectories from multiple users, which results in a low homogeneity score. The results of TraFoS are also good in terms of homogeneity and worse than other similarity measures in terms of completeness, when full linkage (i.e. average pairwise similarity) is used, with MUITAS and MSM beign the most stable in performance and TraFoS showing a better balance between homogeneity and completeness. Low homogeneity scores happen with all similarity measures, which is an indication of the limit of the hierarchical clustering algorithms, which creates a few large clusters and many smaller ones.

As far as it concerns the internal clustering validity metrics (i.e. silhouette score, Calinski-Harabaz score, and Davies-Bouldin index), there are several interesting facts that arise from the cluster analysis. First, the silhouette coefficient that evaluates cluster compactness versus separation is negative in all cases, and it is close to zero only for MUITAS and MSM when hierarchical clustering (with full linkage) is employed. This is mainly explained by the nature of full linkage clustering, which aims for the maximum pairwise similarity of the trajectories of each cluster when agglomerating clusters, and thus leads to clusters of the highest possible compactness and lowest possible separation.

TraFoS exploits a binary partition strategy for building hierarchies of trajectory clusters, one hierarchy per aspect, and then employs the tree-based similarities across all aspect hierarchies in a combined multiple-aspect similarity score. Results show that TraFoS: (1) outperforms the traditional geographical distance and edit distance measures (LCSS, EDR); (2) performs much faster than MUITAS, which has a higher execution cost; (3) outperforms MUITAS using the Hierarchical Agglomerative Clustering algorithm (HAC).

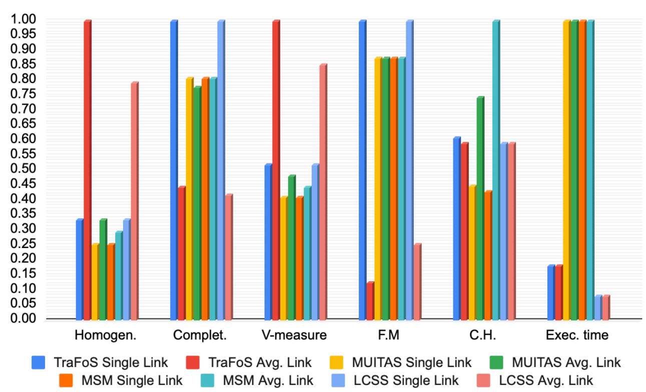

Figure 5: Summarizing evaluation results and execution time.

Figure 5 illustrates the main results of the cluster similarity measures and the execution time. The maximum absolute scaler has been applied in order to highlight the relative differences of the experimental results. We can see that TraFoS surpasses the other similarity measures in internal clustering metrics. MSM with agglomerative average linkage clustering has the best results in the external evaluation metric but it takes approximately ten times longer for execution than TraFoS. We conclude that TraFoS is the best option based on the internal evaluation metrics and it provides results with a very good tradeoff between external evaluation

[5]The value has been chosen empirically after experimentation.

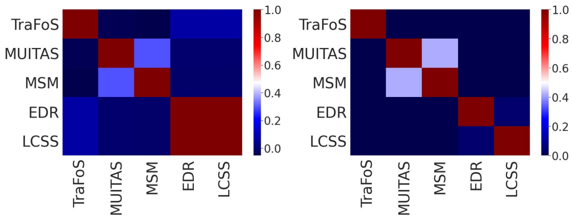

Figure 6: Pairwise agreement scores (Adjusted Rand Index) of the clustering schemes (Single linkage-left, Average linkage-right) using the different similarity metrics.

metrics and execution time. Finally, Figure 6 depicts the pairwise agreement for the two HAC versions. The metrics demonstrate low agreement, with the exception of MSM-MUITAS, and LCSS-EDR which have commonalities in the definition of similarities.

7 CONCLUSIONS

The main contribution of this work is the TraFoS method for computing multi-aspect trajectory similarity, which is combined with agglomerative clustering algorithm for MATs clustering. The main advantage of the proposed method is that it supports a more relaxed MAT comparison, either at a single-aspect and at multiple-aspect level. It is also more flexible with respect to the order (which is ignored) and the frequency of occurrence (which is accounted) of matching stops and moves in the compared trajectories. So the MAT comparison is based on a trajectory representation with frequency vectors, and a tree-based similarity measure computed over a set of binary trees (or isolation trees), each one build for each aspect.

The evaluation of TraFoS in a trajectory clustering task is performed against traditional trajectory similarity measures like LCSS and EDR, and state-of-the-art similarity measures for multi-aspect trajectories like MUITAS and MSM. Results show that TraFoS outperforms the other similarity measures in terms of internal, external evaluation metrics and training time.

Since the potential of the clustering forest is not fully exploited in this work, it is among our next steps to examine supervised techniques that will allow our algorithm to learn on which aspects and aspect combinations to rely. Future work include the inference of the weight using a supervised learning task. In the assumption that the cluster should correspond to users, we can label the trajectories with the user label in the training set and learn the weights.

ACKNOWLEDGEMENTS

This work has been developed in the frame of the MASTER project, which has received funding from the European Union' s Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 777695.

REFERENCES

[1] Luis Otavio Alvares, Vania Bogorny, Bart Kuijpers, Jose Antonio Fernandes de Macedo, Bart Moelans, and Alejandro Vaisman. 2007. A model for enriching trajectories with semantic geographical information. In *Proceedings of the 15th International Symposium on Advances in GIS*. ACM, 22.

[2] Luis Otavio Alvares, Vania Bogorny, Bart Kuijpers, Bart Moelans, Jose Antonio Fern, ED Macedo, and Andrey Tietbohl Palma. 2007. Towards semantic trajectory knowledge discovery. *Data Mining and Knowledge Discovery* 12 (2007).

[3] Jinyang Chen, Rangding Wang, Liangxu Liu, and Jiatao Song. 2011. Clustering of trajectories based on Hausdorff distance. In *2011 International Conference on Electronics, Communications and Control (ICECC)*. IEEE, 1940–1944.

[4] Lei Chen, M Tamer Özsu, and Vincent Oria. 2005. Robust and fast similarity search for moving object trajectories. In *Proceedings of the 2005 ACM SIGMOD international conference on Management of data*. ACM, 491–502.

[5] James Dougherty, Ron Kohavi, and Mehran Sahami. 1995. Supervised and unsupervised discretization of continuous features. In *Machine Learning proceedings 1995*. Elsevier, 194–202.

[6] Carlos Andres Ferrero, Luis Otávio Alvares, and Vania Bogorny. 2016. Multiple aspect trajectory data analysis: research challenges and opportunities.. In *GeoInfo*. 56–67.

[7] Andre Salvaro Furtado, Luis Otavio Campos Alvares, Nikos Pelekis, Yannis Theodoridis, and Vania Bogorny. 2018. Unveiling movement uncertainty for robust trajectory similarity analysis. *International Journal of Geographical Information Science* 32, 1 (2018), 140–168.

[8] Andre Salvaro Furtado, Despina Kopanaki, Luis Otavio Alvares, and Vania Bogorny. 2016. Multidimensional similarity measuring for semantic trajectories. *Transactions in GIS* 20, 2 (2016), 280–298.

[9] F Giannotti, M Nanni, F Pinelli, and D Pedreschi. 2006. *T-patterns: temporally annotated sequential patterns over trajectories.* Technical Report. ISTI-CNR.

[10] Marta C Gonzalez, Cesar A Hidalgo, and Albert-Laszlo Barabasi. 2008. Understanding individual human mobility patterns. *nature* 453, 7196 (2008), 779.

[11] Maria Halkidi, Yannis Batistakis, and Michalis Vazirgiannis. 2001. On clustering validation techniques. *Journal of IIS* 17, 2-3 (2001), 107–145.

[12] Jae-Gil Lee, Jiawei Han, and Kyu-Young Whang. 2007. Trajectory clustering: a partition-and-group framework. In *Proceedings of the 2007 ACM SIGMOD international conference on Management of data*. ACM, 593–604.

[13] Andre L Lehmann, Luis Otavio Alvares, and Vania Bogorny. 2019. SMSM: a similarity measure for trajectory stops and moves. *International Journal of Geographical Information Science* (2019), 1–26.

[14] Lucas May Petry, Camila Leite Da Silva, Andrea Esuli, Chiara Renso, and Vania Bogorny. 2020. MARC: a robust method for multiple-aspect trajectory classification via space, time, and semantic embeddings. *International Journal of Geographical Information Science* (2020), 1–23.

[15] Ronaldo dos Santos Mello, Vania Bogorny, Luis Otavio Alvares, Luiz Henrique Zambom Santana, Carlos Andres Ferrero, Angelo Augusto Frozza, Geomar Andre Schreiner, and Chiara Renso. 2019. MASTER: A multiple aspect view on trajectories. *Transactions in GIS* 23, 4 (2019), 805–822.

[16] Fanrong Meng, Guan Yuan, Shaoqian Lv, Zhixiao Wang, and Shixiong Xia. 2019. An overview on trajectory outlier detection. *AI Review* 52, 4 (2019), 2437–2456.

[17] Brendan Morris and Mohan Trivedi. 2009. Learning trajectory patterns by clustering: Experimental studies and comparative evaluation. In *2009 IEEE Conference on Computer Vision and Pattern Recognition*. IEEE, 312–319.

[18] Boaz Nadler and Meirav Galun. 2007. Fundamental limitations of spectral clustering. In *Advances in neural information processing systems*. MIT Press, 1017–1024.

[19] Mirco Nanni and Dino Pedreschi. 2006. Time-focused clustering of trajectories of moving objects. *Journal of Intelligent Information Systems* 27, 3 (2006), 267–289.

[20] Lucas May Petry, Carlos Andres Ferrero, Luis Otavio Alvares, Chiara Renso, and Vania Bogorny. 2019. Towards semantic-aware multiple-aspect trajectory similarity measuring. *Transactions in GIS* 23, 5 (2019), 960–975.

[21] Eréndira Rendón, Itzel Abundez, Alejandra Arizmendi, and Elvia M Quiroz. 2011. Internal versus external cluster validation indexes. *International Journal of computers and communications* 5, 1 (2011), 27–34.

[22] Stefano Spaccapietra, Christine Parent, Maria Luisa Damiani, Jose Antonio de Macedo, Fabio Porto, and Christelle Vangenot. 2008. A conceptual view on trajectories. *Data & knowledge engineering* 65, 1 (2008), 126–146.

[23] George S. Theodoropoulos, Andreas Tritsarolis, and Yannis Theodoridis. 2020. EvolvingClusters: Online Discovery of Group Patterns in Enriched Maritime Data. In *Multiple-Aspect Analysis of Semantic Trajectories*, K. Tserpes, C. Renso, and S. Matwin (Eds.). Springer International Publishing, Cham, 50–65.

[24] Michail Vlachos, George Kollios, and Dimitrios Gunopulos. 2002. Discovering similar multidimensional trajectories. In *Proceedings 18th international conference on data engineering*. IEEE, 673–684.

[25] Haozhou Wang, Han Su, Kai Zheng, Shazia Sadiq, and Xiaofang Zhou. 2013. An effectiveness study on trajectory similarity measures. In *Proceedings of the 24th Australasian Database Conference*, Vol. 137. Australian CS, Inc., 13–22.

[26] Zhibiao Wu and Martha Palmer. 1994. Verbs semantics and lexical selection. In *Proceedings of the 32nd annual meeting on Association for Computational Linguistics*. Association for Computational Linguistics, 133–138.

[27] Dingqi Yang, Daqing Zhang, Vincent W Zheng, and Zhiyong Yu. 2014. Modeling user activity preference by leveraging user spatial temporal characteristics in LBSNs. *IEEE Transactions on Systems, Man, and Cybernetics: Systems* 45, 1 (2014), 129–142.

[28] Guan Yuan, Penghui Sun, Jie Zhao, Daxing Li, and Canwei Wang. 2017. A review of moving object trajectory clustering algorithms. *Artificial Intelligence Review* 47, 1 (2017), 123–144.