



**HAL**  
open science

# Towards a Granular Computing Framework for Multiple Aspect Trajectory Representation and Privacy Preservation: Research Challenges and Opportunities

Zaineb Chelly Dagdia, Vania Bogorny

► **To cite this version:**

Zaineb Chelly Dagdia, Vania Bogorny. Towards a Granular Computing Framework for Multiple Aspect Trajectory Representation and Privacy Preservation: Research Challenges and Opportunities. 17th CONFERENCE ON COMPUTER SCIENCE AND INTELLIGENCE SYSTEMS, Sep 2022, Sofia, Bulgaria. hal-03723133

**HAL Id: hal-03723133**

**<https://hal.uvsq.fr/hal-03723133>**

Submitted on 13 Jul 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Towards a Granular Computing Framework for Multiple Aspect Trajectory Representation and Privacy Preservation: Research Challenges and Opportunities

Zaineb Chelly Dagdia

Université Paris-Saclay, UVSQ, DAVID, France  
LARODEC, ISG, Université de Tunis, Tunis, Tunisia  
Email: zaineb.chelly-dagdia@uvsq.fr

Vania Bogorny

Programa de Pós-Graduação em Ciência da Computação,  
Departamento de Informática e Estatística (INE),  
Universidade Federal de Santa Catarina (UFSC),  
Florianópolis, SC, Brasil  
Email: vania.bogorny@ufsc.br

**Abstract**—In recent years, there has been a lot of research on trajectory data analysis and mining. Despite the fact that trajectories are multidimensional data characterized by the spatial, the temporal, as well as the semantic aspect, few are the studies that have taken into account all of these three dimensions. Apart from these dimensions that should be considered all together for an efficient trajectory data analysis and mining, it should be also made feasible to represent trajectories from several points of view, which is called *multiple aspect representation*. State-of-the-art works are typically restricted to a single trajectory representation, which limits the identification of a variety of key patterns. These multiple aspect trajectories are quite rich that they reveal sensitive information, making the user’s privacy vulnerable; hence, raising several challenges when it comes to privacy preservation. In this paper, we show that there is a need to consider granular computing for multiple aspect trajectory representation and privacy preservation, and present new research challenges and opportunities in this concern.

## I. INTRODUCTION AND MOTIVATION

**T**ODAY, we are living the era of movement monitoring and mining where large volumes of mobility data are being captured about our everyday lives and routines. The collected data is characterized not only by its volume and velocity but also by its heterogeneity (variety) and uncertainty, and most importantly by its sensitive aspect. The collection and enrichment of this tremendous amount of mobility data with information from several sources is made possible by the popularization and frequent use of mobile devices [1], applications, sensors, internet channels, social networks, among many other services and instruments.

For instance, via the use of social networks such as Facebook and Twitter and navigation applications such as Google Maps, detailed information about our routines can be collected including the places that we have visited, the time that we have spent at each visited place, our emotions at a specific time and place, our means of transportation, etc. These are referred to

as “*semantic trajectories*” which are characterized by their different aspects or dimensions, and their complex nature [2]. With this explosion of enriched trajectory data, emerged the problem of data privacy. The more information that is linked to mobility data, the more sensitive is the user privacy. Therefore, it became critical and essential to protect users’ privacy.

When it comes to trajectory representation, the same semantic trajectory can be represented with respect to different aspects [2]. As an example, a raw trajectory can be represented as a sequence of stops and moves, or as a sequence of transportation means, or as a sequence of weather conditions, of activities performed during the movement, and so on. This led to the “*Multiple Aspect Trajectory*” (MAT) concept.

Multiple aspect trajectory representation is a hot topic and it is very recently that the MASTER model was proposed in [3]. The MASTER model allows the representation of the trajectory with space, time and several aspects, any of which might violate the user privacy. Therefore, there is a need to consider a novel structure that permits the representation of multiple aspect trajectories while considering privacy preservation.

The aim of this paper is to highlight the perspectives of applying Granular Computing (GrC) formal settings (e.g., rough set theory, set theory, etc.) to multiple aspect trajectory representation. We present a representation based on granular computation that we named *GrC – MAT*. The paper, also, presents the important aspects of granular computation that can be considered for privacy preservation. It is to be noted that the objective of this paper is not to make a survey of the state-of-art methods that deal with MAT representation and privacy preservation, nor to discuss their limitations.

The rest of the paper is organized as follows: Section II introduces the main concepts and fundamentals of multiple aspects trajectories and granular computation. Section III highlights the representation of multiple aspect trajectory. Section IV, introduces the proposed granular computing framework, *GrC – MAT*, for multiple aspect trajectory representation

This work has received funding from the European Commission’s Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement N. 777695.

as well as covering its evolving aspect. In Section V, we discuss our vision of the future research challenges and opportunities in multiple aspect trajectory representation and privacy preservation from a granular computing perspective. Finally, Section VI concludes the paper.

## II. BACKGROUND

In this section, we introduce the main concepts and fundamentals of multiple aspects trajectories and granular computation.

### A. Trajectory data models

When an individual or an object is moving, his/her/its location data is collected over time via mobile devices or applications or any other collection instruments. This data is in the form of a sequential spatiotemporal points, called *raw trajectory*, and is denoted as  $Seq = \langle p_1, p_2, \dots, p_n \rangle$ , where each  $p_i = (x_i, y_i, t_i)$ , with  $p_i \in Seq$ ,  $x_i$  and  $y_i$  are the spatial position of the moving individual/object in space at a specific time frame  $t_i$ .

Several data models have been presented in the last decade to represent and augment trajectories with semantic information. These can be categorized into three main time frames. For instance, the work proposed in [4] initiated trajectory modelling via the use of sequences of events in space and over a specific time window. These are typically raw trajectories which represent the first category of trajectory data models, and are the simplest. Later on, and during the period of 2008 and 2018 which is characterized by the explosion of social networks and navigation applications, works have been around enriching trajectories with semantic information. In this concern, in [5], authors proposed the integration of geographic information while distinguishing stops and moves. The segments of a trajectory where the moving individual/object has stayed for a minimal amount of time define the stops, whereas the movement between stops define the moves. In this work, semantic trajectories can have each stop linked to Points of Interest (POI), which are usually a place name, reflecting semantic information; which are added as a third dimension to space and time. In 2014, a semantic trajectory data model called CONSTANT was proposed in [2]. The model associates the moving individual/object trajectories to a set of other information such as the visited POIs, the activities performed at a specific POI, and the transportation means. Also, within this second category, the BAQUARA framework was proposed in [6]. The semantic model enriches trajectories with ontologies and linked open data.

Very recently, in [3] and within the third category, authors introduced the MASTER model, which introduces the multiple aspect trajectory concept. Multiple aspect trajectories can be defined as the enhancement of the basic view of semantic trajectories with the notion of multiple heterogeneous aspects, characterizing different semantic dimensions related to the pure movement data. MASTER allows the augmentation of trajectories with any type of information, dubbed aspects. The model addresses the issue raised by [7], in which different

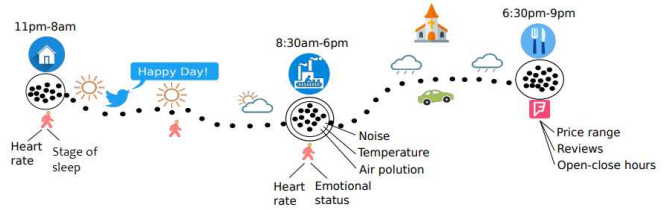


Fig. 1. Example of a multiple aspect trajectory as defined in [3].

aspects were taken into account independently. The model permits the representation of a trajectory in terms of space, time, and a variety of other aspects. Figure 1 depicts a multiple aspect trajectory with aspects that differ along the trajectory, as defined by [3].

As shown in Figure 1, the trajectory contains a wealth of information, aspects, about the moving individual, including: (1) the heart rate and the sleeping stage which are collected by a smartwatch, at home, from 11pm to 8am; (2) the humor is provided by a tweet when the individual walks to work; (3) the environmental information such as noise, temperature, and pollution are collected by sensors at a smart office; and (4) from 6:30pm to 9pm, the characteristics of the places visited by the individual, such as price, reviews, and open-close hours of a restaurant are collected as well [8].

This recent category of trajectory displays a person's very precise daily routine in depth than earlier models, while revealing lots of sensitive information. Hence, raising privacy concerns.

### B. Granular computing models

Granular Computing (GrC) [9] is an emerging information processing computing paradigm that focuses on representing and processing complex entities, known as "*granules*", which are produced from the process of data abstraction and knowledge extraction from information or data. The roots of GrC can be traced back to the works of Zadeh [10] defining a granule as a clump of points (objects) drawn together by indistinguishability or indiscernibility, similarity, proximity, or functionality.

Granules can be decomposed into a set of smaller and finer granules which are known as "*subgranules*". Levels, hierarchies, and granular structures can all be used to organize granules and subgranules. Granulation is a procedure or a process which is needed for constructing or decomposing granules. In [11], author provides a comprehensive overview of the unified principals of granular computing along with its comprehensive algorithm framework and design practices.

The origins of the granular computing ideology are to be found in the rough sets and fuzzy sets literature [12], [13], but they can also be found in some other formal settings for GrC such as interval calculus, set theory, shadowing sets, and probabilistic granules [9]. The concept of granules and granulation are defined differently in each of these settings, and the work of [14] offers a preliminary way to uncover commonality and bridge the gap across these settings.

All of these formal settings, allow, in general, the formation of granular structures. Space and temporal aspects are naturally represented via multi-level structures in which high-level granules reflect more abstract notions and low-level granules represent more particular and more specific concepts. Such granular structures are critical to our predefined goals. GrC is based on a large set of interactions and relationships as defined in [15]. These may be used to organize granules in several structures such as hierarchies, trees, and networks. A granule  $g$  is a refinement of  $G$  (or  $G$  is a coarsening of  $g$ ), denoted by  $g \preceq G$ , if all of  $g$ 's data or subgranules are contained in some subgranules of  $G$ . When just some data or subgranules of  $g$  are contained in some subgranules of  $G$ , refinement (coarsening) can be partial, and is denoted as  $g \sqsubseteq G$ .

Structuring and forming the granules in the correct and most appropriate manner is an open research question that has been inspected by numerous researchers and is rather dependent on the application's requirements. The principle of justified granularity was developed in [16], [17] as a way to evaluate the performance of informational granules. The principle of justified granularity is based on a trade-off between the coverage and specificity measures, which do not exclusively depend on the application. Formally, coverage refers to the ability to cover data, while specificity refers to the granule prototype's level of abstraction as measured by its size. The proper expression of these two measures is dependent on the nature of the set formed.

For crisp sets, a measure of coverage may be  $Cov(P) = \frac{1}{N} \text{card}\{X_k | x_k \in P\}$ , but for fuzzy sets, the sum of the degree of memberships of the elements can be used  $Cov(P) = \frac{1}{N} \sum_{k=1}^N \mu_P(X_k)$ .  $Cov(P)$  should, ideally, be equal to 1, indicating that the prototype covers all data. The intervals must be as narrow (specific) as feasible to achieve specificity. An interval's specificity can be assessed in a variety of ways. A specificity measure must meet two criteria: it must achieve a maximum value for a single element, and the larger the interval, the lower the specificity measure. The two concepts of coverage and specificity are at odds. To visualize their relationship, consider arranging them in the form of a coverage-specificity plot, which may also be parameterized, and calculating the area under the curve to derive a global measure of quality.

Another criterion for granule design is the principal of uncertainty level preservation as defined in [18], [19], which is primarily concerned with assessing the quality of the granulation itself.

This principle considers the quantification of uncertainty as an invariant property to be retained during the granulation process by considering information granulation as a mapping between some input and output. The discrepancy between the input and output entropy is considered as an error that must be decreased in order to achieve optimal information granulation [20].

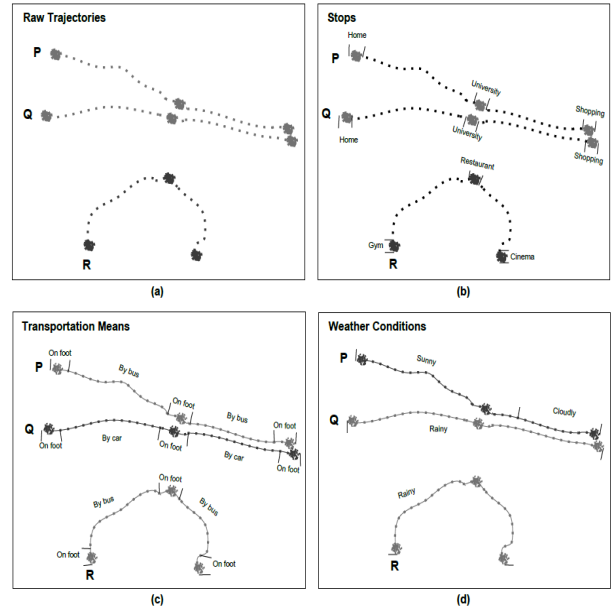


Fig. 2. Multiple trajectory representation [7]

### III. MULTIPLE ASPECT TRAJECTORY REPRESENTATION

In this Section, we replicate the example given in [7], illustrated in Figure 2, to show the challenge in representing multiple aspect trajectories.

Figure 2 shows three trajectories,  $P$ ,  $Q$ , and  $R$ , which are represented using four different aspects: (1) as raw trajectories (Figure 2(a)); (2) as stops and moves (Figure 2(b)), where the labeled parts are the stops; (3) as transportation means (Figure 2(c)); and (4) according to weather conditions (Figure 2(d)).

If every aspect is considered separately, and by considering only the space and semantic dimensions, excluding time for simplification, the three trajectories would be analyzed as follows:

- If *Aspect = raw trajectory* then  $P$  and  $Q$  are spatially closer than  $P$  and  $R$  or  $Q$  and  $R$ .
- If *Aspect = stops* then  $P$  and  $Q$  are the most similar.
- If *Aspect = transportation means* then  $P$  and  $R$  are the most similar.
- If *Aspect = weather condition* then  $Q$  and  $R$  are the most similar.

Two issues can be reported from the standpoint of similarity analysis: (1) the trajectories must consider multiple aspects, such as raw data, stops, transportation means, activities, weather conditions, and others; (2) existing similarity measures only consider a single representation (aspect). Now from a mining view point, the following important issue can be reported: multiple aspect trajectory data analysis can lead to new types of trajectory patterns that cannot be detected so far by existing data mining methods. More details about the illustration of these issues via examples can be found in [7].

Now in terms of representation, one would wonder why not simply combine all relevant data into a single trajectory

representation. When evaluating the stops representation, for example, one may claim that it is straightforward to enrich the trajectory with all relevant information such as weather, transportation means, activities, and so on. The issue is not as straightforward as it appears.

Let us assume the moving individual is walking and the weather changes from sunny to rainy during one stop. In the activity aspect, the moving individual changed his/her activity from meeting to coaching when the weather is rainy. All of these changes happen at a single stop (POI); labelled office.

It would be very hard to correctly split and annotate the stop into two weather conditions, each one having a different start and end time, splitting it in different activities [7]. The same stop would have multiple semantic labels for weather, transportation modes, stop names, activity names, and so on, as well as multiple time intervals associated with each semantic label, such as the start and end times of a stop, the duration of a transportation mode, the distance traveled by one transportation mode, and possibly different space information [7].

As it can be noted, the issues reported from the standpoint of similarity analysis and from a mining view point, are connected to the challenges linked to the multiple aspect trajectory representation. This leads to the ultimate challenge of how to efficiently represent trajectories with all this information while offering an appropriate way to conduct similarity analysis and mining tasks.

#### IV. A GRANULAR COMPUTING FRAMEWORK FOR MULTIPLE ASPECT TRAJECTORY REPRESENTATION

The concept of “multiple aspect” trajectories can be modeled using the hierarchical structure of granular computing. Data granules are formal entities that aid in the organization of data and relationship knowledge. The use of data granules to represent the different aspects of MAT and the characteristics of each aspect can help to detect hidden patterns, as well as increase the logicity, systematicity, and efficiency of decision-making [21], [22]

In this section, we present a tentative representation of multiple aspect trajectory using granular computation, called  $GrC - MAT$ . In this representation, we consider that every aspect (or dimension) can be seen as a granule, and each granule can be, in its turn, represented using sub-granules. All of these granules are connected between each other to ensure an overall representation of the multiple aspect trajectory.

##### A. Definitions

Introducing GrC into multiple aspect trajectory generates the following related concepts:

**Definition 1 (MAT Data Granule):** Multiple aspect trajectory data granule, denoted as  $Mat - Gr$ , refers to a mobility data chunk defined via a set of mobility data elements drawn together by space, time, proximity, or indistinguishability [23].

**Definition 2 (MAT Data Granulation):** Multiple aspect trajectory data granulation refers to the process that partitions the mobility rich data into fine grained and semantically clear

mobility data granules with respect to a set of criteria which are associated with spatial and temporal scale features together with other considered aspects.

**Definition 3 (MAT Data Granular Layer):** Multiple aspect trajectory data granular layer, denoted as  $Mat - Lay$ , is composed of a set of mobility data granules with respect to a certain set of granulation criteria.

**Definition 4 (MAT Data Granular Structure):** Multiple aspect trajectory data granular structure, denoted as  $Mat - GrS$ , refers to the relational structure generated by the links between several mobility data granules corresponding to various granulation criteria.

**Definition 5 (MAT Granular Computing):** Multiple aspect trajectory granular computing refers to the process that uses mobility data granules to describe, analyze, and solve multiple aspect trajectory data mining problems from different scales and perspectives.

##### B. A granular representation of multiple aspect trajectories: a tentative representation

Multiple aspect trajectory data granule is a complete entity with multiple dimensions, such as space, time, weather, transportation means, emotion status, and activity. As all of these dimensions are related to the spatial and temporal scales, then the granular representation of multiple aspect trajectories should consider the spatiotemporal aspect. This can be guaranteed by integrating the time and space dimensions, as attributes of data granules, into a single systematic model.

Figure 3 represents the multiple aspect trajectory data granule structure. Every considered data granule  $Mat - Gr$  is represented via a layer  $Mat - Lay_l$ , where  $l \in \{1, \dots, L\}$ , and  $L$  is the maximum number of layers for the different considered dimensions. With respect to this structure, every  $Mat - Gr$  can be defined in different scales; denoted as  $Mat - Gr_l$ . The data granules in each upper scale  $Mat - Gr_l$  are transformed into those in the lower scale  $Mat - Gr_{l+1}$  using a granulation criteria (rule)  $GCr_i$ , where  $i \in \{1, \dots, L - 1\}$ . The data granules decrease (increase) as the scale decreases (increases). A correlation, representing the connection between the different aspects of the multiple aspect trajectory, is reflected by the different edges between the various  $Mat - Gr_l$ . All  $Mat - Gr_l$  are connected between each other as well as between their corresponding granules. For a simplified view, in Figure 3, only few edge are represented.

**Example 4.1 (Example of a granular representation of a multiple aspect trajectory):** Let us suppose that during one stop the object is moving on foot and the weather condition changes from Sunny to Rainy. Figure 4 represents a simplified view of a type-2 granular structure of such representation. The edges represent the different correlations between the different granules; i.e., temporal-spatial, temporal-activity, space-weather, ..., and space-temporal-weather-activity.

With respect to this granular structure, we can define the following roots:

- At POI, Weather = "Sunny" during "Start-time" and "End-time", and Activity = "Foot"

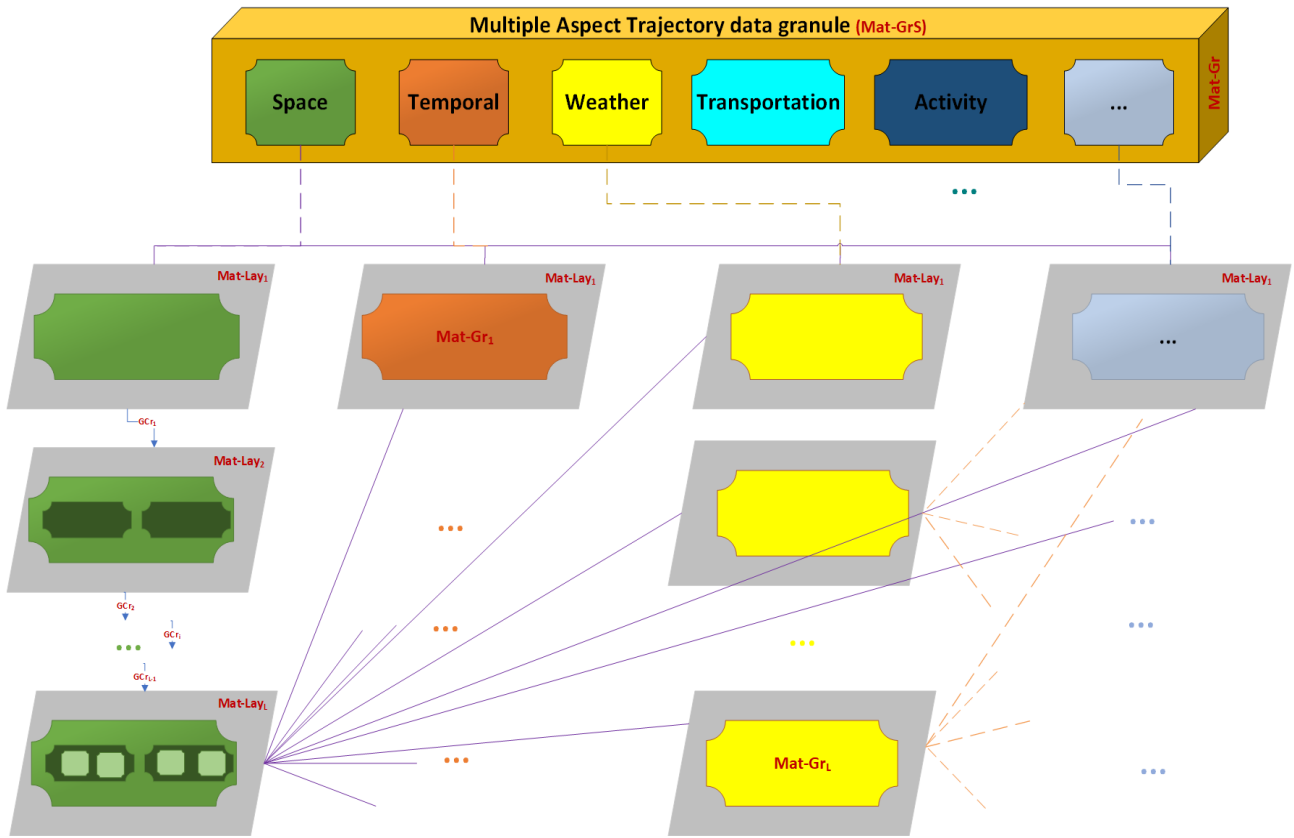


Fig. 3. A granular representation of multiple aspect trajectories (inspired by [24])

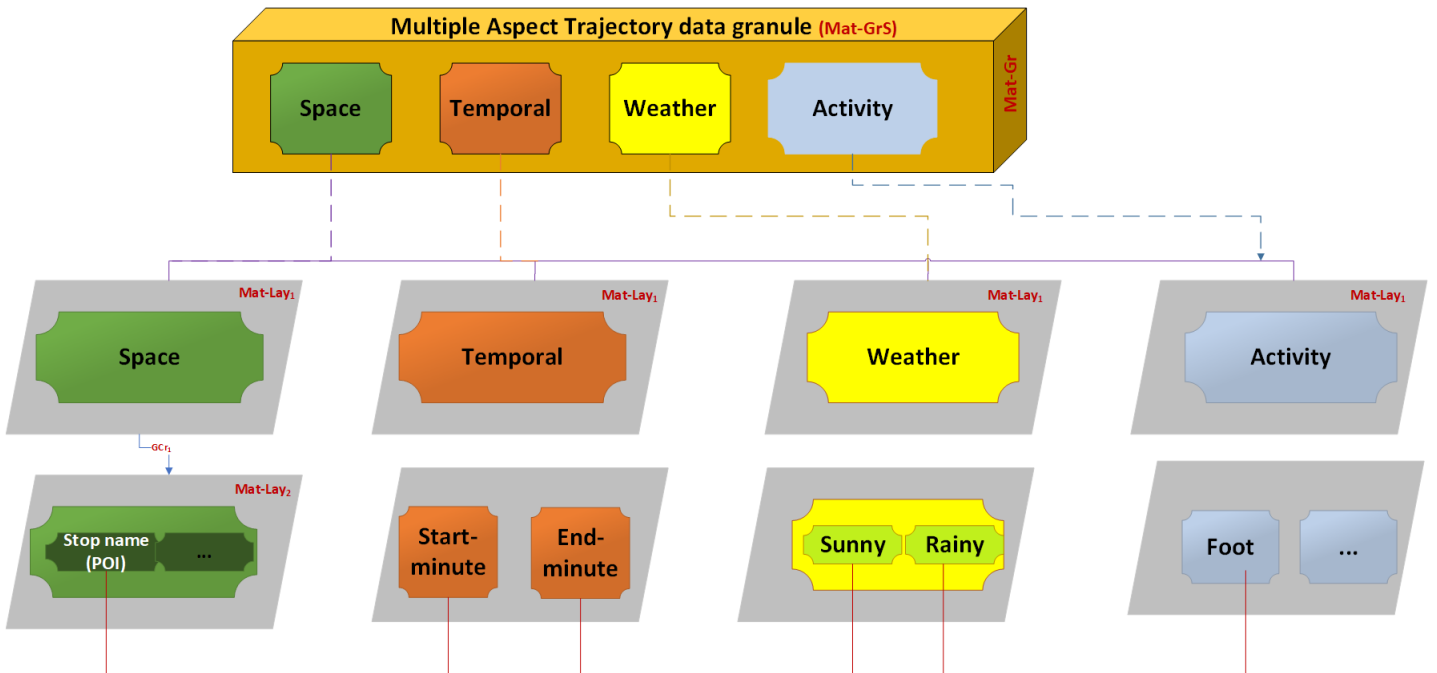


Fig. 4. Example of a granular representation of a multiple aspect trajectory

- At POI, Weather = "Rainy" during "Start-time" and "End-time", and Activity = "Foot"

With respect to the need of representing different aspects at a time, sub-granular structures can be extracted from Figure 3. This is to achieve landscape law mining at a specific scale for different granules.

*Example 4.2 (Example of reasoning):* The "Weather" granule can be considered as a "localized or a particular view or aspect" of the MAT presented in Figure 4. "Sunny" and "Rainy" are examples of granules of the sub-layer ( $Mat - Lay_2$ ) of the "Weather" granular layer ( $Mat - Lay_1$ ). Data granules in  $Mat - Lay_2$ , for instance, can be viewed as an "information table", with a binary values representation, and different granules can be further generated from it. For example, if we consider "Rough Set Theory" as a GrC theory, we can generate three disjoint regions (granules) by applying the "lower approximation" and "upper approximation"; namely, the "positive region" which comprises those data objects certainly related with the decision class, the "negative region" which comprises those objects certainly not related with the decision class, and the "boundary region" which comprises those objects possibly related with the decision class. The connections and relationships between the granules may be interpreted by the rough set "dependency degree". All granules of the different granular layers provide a collective description of the MAT.

### C. An evolving granular representation

So far we have created a hierarchical granular structure resembling the hierarchies existing between the different aspects that can be considered in multiple aspect trajectories. To accommodate other aspects that can be further considered in multiple aspect trajectories, either in  $Mat - Gr$  or/and in  $Mat - Lay_1$ , we can leverage on the concept of evolvable granules [25], [26].

In [26], the behavior of evolving granules is reported via Figure 5, and explained as follows, when considering only time and space: the time window shows different objects registered by sensors in three time slices,  $t$ ,  $2t$ , and  $3t$ . The top of the figure shows three granular structures, where granules denote the position of the objects in the space dimension. During  $time = t$ , singleton granular structures are created comprising  $\{a\}$ ,  $\{b\}$  and  $\{c\}$ . During  $time = 2t$ , two additional objects,  $\{d\}$  and  $\{e\}$ , are recognized and merged with the existing granules to form higher level granules  $\{a, d\}$ ,  $\{b, e\}$ , and  $\{c\}$ . This process continues to iterate in the next time windows, and, as output, granular structures evolve with granules that can be either merged, split, removed or new granules formed. In [26], a complete formalism to deal with splitting and merging criteria in the case of granules created with Fuzzy c-means can be found.

To accommodate other aspects in  $Mat - Gr$  or/and  $Mat - Lay_1$ , we should reason on the evolution of granular structures. To do so, the transitions between  $Mat - Lay_1$  and  $Mat - Lay_{1+1}$  should be considered with respect to the applied granulation criteria  $GCr_i$ . This can be achieved depending on

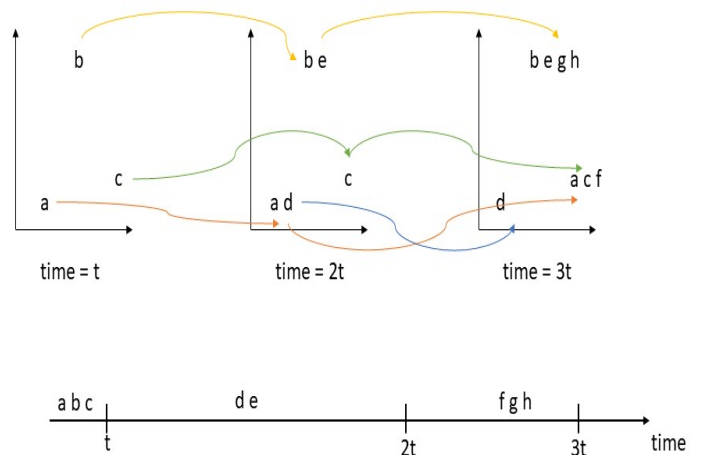


Fig. 5. Behavior of evolving granules by granulating time and space [20]

previous knowledge about the rules ( $GCr_i$ ) that govern the evolution of phenomena under observations and/or the actions that can enable situations transitions.

## V. RESEARCH CHALLENGES AND OPPORTUNITIES

In [7], authors presented a set of challenges and opportunities tied to multiple aspect trajectory data analysis. These include multiple aspect representation, feature extraction, data storage, similarity analysis and data mining, visualization, and privacy protection. The later challenge was further detailed in [8] where authors presented a survey of the state-of-the-art trajectory anonymization methods for privacy preservation, together with a description of a set of challenges for anonymizing multiple aspect trajectories. Therefore, in this section, we mainly focus on presenting the main challenges tied to multiple aspect trajectory representation and privacy preservation from a granular computing perspective.

Representing multiple aspect trajectories reveals many challenges when it comes to defining the concrete structure and connections, between the different considered aspects. Among these challenges, we mention the uncertainty and the imprecision that can be found in mobility data. Let us consider a basic scenario in which people wander around a city and reveal their locations twice an hour. To avoid stalking, the disclosed location is chosen at random from within a one-kilometer radius circle that contains the user's location. Not being conscious of uncertainty may lead to erroneous deductions. For example, we could mistakenly believe that a group of individuals have met or that someone has stayed at a privacy-sensitive location. If we take uncertainty in account, such erroneous conclusions can be avoided. For instance, if an individual was more than one kilometer away from the location of an accident, we may safely assume that that individual was not involved in that accident [27].

By considering a granular computing representation, our  $GrC - MAT$  model can efficiently handle such type of uncertainty in mobility data. Settings related to GrC implementation, including fuzzy sets, rough sets, and inclusion degree theory,

contain all the results of uncertain reasoning (e.g., inclusion degree [28] based on conditional probability as is a form of uncertain reasoning).

A resulting challenge from the above is how to formally define the granular computing model based on the formal setting used to implement GrC.

Now from a building point of view, the definition of the granules, the construction of the granule layers in an optimized way using the granulation criteria – which also needs to be defined–, the definition of the edges between the different granules as well as the strength (degree) of these edges, present all major challenges.

Now when it comes to the privacy preservation challenge, the hierarchical or tree or network structure adopted to build the *GrC – MAT* model can be used to hide sensitive/private information. This can be achieved by several means such as moving from a granular layer to another to avoid specificity and ensure generality and vice-versa, by ignoring/not reporting some edges (correlations) connecting the different granules, or by switching a granule by another. Also, the fact of using the fundamentals of the GrC theories to report some granules instead of the actual granules that need to be reported can deal with the privacy issues. For instance, when referring to rough set theory, we can report the “*boundary region*”, as a granule, comprising those objects *possibly* related with the decision class, instead of reporting the “*positive region*” which comprises those data objects *certainly* related with the decision class or instead of the “*negative region*” which comprises those objects *certainly* not related with the decision class. The challenge, in this concern, is to formally study the feasibility of these assumptions.

The GrC structure can still lead to a privacy leak as the deductive route from non-sensitive to sensitive features can be mined; which is another challenge to be considered. The sensitive features can be defined as the set of features that should be hidden using algorithms in such a way that they cannot be reasonably approximated using the set of non-sensitive features. This is quite a challenge since the non-sensitive features may contain concepts on which sensitive features inferentially depend; these are known as “*quasi-identifiers*”. To cut the deductive route to any potential privacy leak, these quasi-identifiers must be appropriately hidden. However, they must be uncovered first. In this concern, rough set theory as a GrC example, seemed to be one among the right tools to be used to undermine the deductive route from non-sensitive to sensitive features by modelling and analyzing the dependency “non-sensitive  $\rightarrow$  sensitive” [29]. Rough set theory provides numerical quantities that measure the degree of dependency of attributes. This study was presented in [30], where authors have learned the quasi-identifiers, computed a granulation of the information system that maximizes the distribution of sensitive features in each granule, and masked the deductive route from non-sensitive to sensitive attributes. This rises the challenge of how to undermine this deductive route within a multiple aspect scenario. It would be also interesting to investigate this challenge when coupling it with

the evolving structure of the GrC model.

Another challenge is how to define similarity in multiple aspect trajectories based on granular computation. As previously mentioned, in the era of big data, large amounts of semantically rich mobility data became available. This called for the need for new trajectory similarity metrics in the context of multiple-aspect trajectories. Existing techniques have several constraints concerning the links between characteristics and their semantics. These techniques are too rigid, demanding a match on all features, or too lenient, treating all features as unrelated. Granular computing formal settings can be used to represent the appropriate dependencies between the features and hence new similarity measures dedicated to multiple aspect trajectories can be defined.

Granular computing methods offer plenty of opportunities to handle the challenges tied to multiple aspect trajectory representation and privacy preservation. This will present a considerable addendum to the mobility data management, analytic and privacy communities and fields.

## VI. CONCLUSION

In this paper, we presented a granular computing view for multiple aspect trajectory representation. Such representation will allow a more natural, a flexible, and a complete representation of multiple aspect trajectories. Additionally, by using the fundamentals of some granular formal settings such as fuzzy sets and rough sets, there are several opportunities and possibilities to deal with the privacy aspect and hence, ensure multiple aspect trajectory privacy preservation.

From a practical and applied perspectives, the potential advantages of this *GrC – MAT* approach can be reflected by the following aspects: (i) a good semantic scalability ensured by data granularization based on the extendable granulation criteria which cover not only the spatiotemporal scales but also the other aspects of mobility data with clear semantics; (ii) the adaptation to vertical dynamic changes in the data. This implies when introducing modifications to granular layers. This dynamic adaptation is ensured by the independency of the granulation criteria of each layer, which implies that the only required change will be the update of the granulation criteria between the affected adjacent granular layers; hence, not affecting other granular layers; and last but not least (iii) the adaptation to horizontal dynamic changes in the data. The granular structure of the *GrC – MAT* model allows dynamic incremental data expansion in a specific granular layer without affecting other adjacent layers. As previously mentioned, granular computing will present a considerable addendum to the mobility data management, mobility data mining, analytic and privacy communities and fields.

As future works, we will focus on the discussed challenges of the multiple aspect trajectory representation and privacy preservation from a granular computing perspective, to find efficient approaches that handle all these aspects.



## ACKNOWLEDGMENT

This work has received funding from the European Commission's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement N. 777695.

## REFERENCES

- [1] P. Kucharski, D. Sielski, T. Jaworski, A. Romanowski, and J. Kucharski, "Use of fuzzy cognitive maps for enhanced interaction with multiple mobile devices." in *FedCSIS (Communication Papers)*, 2018, pp. 217–222.
- [2] V. Bogorny, C. Renso, A. R. de Aquino, F. de Lucca Siqueira, and L. O. Alvares, "Constant—a conceptual data model for semantic trajectories of moving objects," *Transactions in GIS*, vol. 18, no. 1, pp. 66–88, 2014.
- [3] R. d. S. Mello, V. Bogorny, L. O. Alvares, L. H. Z. Santana, C. A. Ferrero, A. A. Frozza, G. A. Schreiner, and C. Renso, "Master: A multiple aspect view on trajectories," *Transactions in GIS*, vol. 23, no. 4, pp. 805–822, 2019.
- [4] K. S. Hornsby and S. Cole, "Modeling moving geospatial objects from an event-based perspective," *Transactions in GIS*, vol. 11, no. 4, pp. 555–573, 2007.
- [5] S. Spaccapietra, C. Parent, M. L. Damiani, J. A. de Macedo, F. Porto, and C. Vangenot, "A conceptual view on trajectories," *Data & knowledge engineering*, vol. 65, no. 1, pp. 126–146, 2008.
- [6] R. Fileto, C. May, C. Renso, N. Pelekis, D. Klein, and Y. Theodoridis, "The baquara2 knowledge-based framework for semantic enrichment and analysis of movement data," *Data & Knowledge Engineering*, vol. 98, pp. 104–122, 2015.
- [7] C. A. Ferrero, L. O. Alvares, and V. Bogorny, "Multiple aspect trajectory data analysis: research challenges and opportunities." in *GeoInfo*, 2016, pp. 56–67.
- [8] T. T. Portela, F. Vicenzi, and V. Bogorny, "Trajectory data privacy: Research challenges and opportunities." in *GEOINFO*, 2019, pp. 99–110.
- [9] W. Pedrycz, A. Skowron, and V. Kreinovich, *Handbook of granular computing*. John Wiley & Sons, 2008.
- [10] L. A. Zadeh, "Toward a theory of fuzzy information granulation and its centrality in human reasoning and fuzzy logic," *Fuzzy sets and systems*, vol. 90, no. 2, pp. 111–127, 1997.
- [11] W. Pedrycz, *Granular computing: analysis and design of intelligent systems*. CRC press, 2018.
- [12] G. Wang, Q. Liu, Y. Yao, and A. Skowron, "Rough sets, fuzzy sets, data mining, and granular computing," *RSFDGrC, Springer 2003, Chongqing, China*, pp. 41–48, 2003.
- [13] Z. Pawlak, "Granularity of knowledge, indiscernibility and rough sets," in *1998 IEEE International Conference on Fuzzy Systems Proceedings. IEEE World Congress on Computational Intelligence (Cat. No. 98CH36228)*, vol. 1. IEEE, 1998, pp. 106–110.
- [14] D. Dubois and H. Prade, "Bridging gaps between several forms of granular computing," *Granular Computing*, vol. 1, no. 2, pp. 115–126, 2016.
- [15] Y. Yao, "A triarchic theory of granular computing," *Granular Computing*, vol. 1, no. 2, pp. 145–157, 2016.
- [16] W. Pedrycz and W. Homenda, "Building the fundamentals of granular computing: A principle of justifiable granularity," *Applied Soft Computing*, vol. 13, no. 10, pp. 4209–4218, 2013.
- [17] W. Pedrycz, A. Gacek, and X. Wang, "Clustering in augmented space of granular constraints: a study in knowledge-based clustering," *Pattern Recognition Letters*, vol. 67, pp. 122–129, 2015.
- [18] L. Livi and A. Sadeghian, "Data granulation by the principles of uncertainty," *Pattern Recognition Letters*, vol. 67, pp. 113–121, 2015.
- [19] —, "Granular computing, computational intelligence, and the analysis of non-geometric input spaces," *Granular Computing*, vol. 1, no. 1, pp. 13–20, 2016.
- [20] G. D'Aniello, A. Gaeta, V. Loia, and F. Orciuoli, "A granular computing framework for approximate reasoning in situation awareness," *Granular Computing*, vol. 2, no. 3, pp. 141–158, 2017.
- [21] K. Zhou, Z. Tian, and Y. Yang, "Periodic pattern detection algorithms for personal trajectory data based on spatiotemporal multi-granularity," *IEEE Access*, vol. 7, pp. 99 683–99 693, 2019.
- [22] F. J. Cabrerizo, J. A. Morente-Molinera, S. Alonso, W. Pedrycz, and E. Herrera-Viedma, "Improving consensus in group decision making with intuitionistic reciprocal preference relations: A granular computing approach," in *2018 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*. IEEE, 2018, pp. 1471–1476.
- [23] L. A. Zadeh, "Some reflections on soft computing, granular computing and their roles in the conception, design and utilization of information/intelligent systems," *Soft computing*, vol. 2, no. 1, pp. 23–25, 1998.
- [24] C. Yunxian, L. Renjie, Z. Shuliang, and G. Fenghua, "Correction: Measuring multi-spatiotemporal scale tourist destination popularity based on text granular computing," *PloS one*, vol. 15, no. 5, p. e0233068, 2020.
- [25] M. Antonelli, P. Ducange, B. Lazzarini, and F. Marcelloni, "Multi-objective evolutionary design of granular rule-based classifiers," *Granular Computing*, vol. 1, no. 1, pp. 37–58, 2016.
- [26] W. Pedrycz, "Evolvable fuzzy systems: some insights and challenges," *Evolving Systems*, vol. 1, no. 2, pp. 73–82, 2010.
- [27] C. Silvestri and A. Vaisman, *Mobility and Uncertainty*. Cambridge University Press, 2013, p. 83–102.
- [28] W. Xu, J. Mi, and W. Wu, "Granular computing methods and applications based on inclusion degree," *Beijing: Science Press*, 2017.
- [29] M. Bishop, J. Cummins, S. Peisert, A. Singh, B. Bhumiratana, D. Agarwal, D. Frincke, and M. Hogarth, "Relationships and data sanitization: A study in scarlet," in *Proceedings of the 2010 New Security Paradigms Workshop*, 2010, pp. 151–164.
- [30] C. Wafo Soh, L. Njilla, K. Kwiat, and C. Kamhoua, "Learning quasi-identifiers for privacy-preserving exchanges: A rough set theory approach," *Granular Computing*, vol. 5, no. 1, pp. 71–84, 2020.