A New Semantic Trajectory Ontology Model for Modelling and Reasoning on Movement Data

Mohammad Akbari[*1], Ali Mousavi[†2], Seyed Ahmad Eslaminezhad [‡3], Mobin Eftekhari[§4] and Mohsen Ghorani[**5]

[1] Department of Civil Engineering University of Birjand, Birjand, Iran
[2] Department of Geomatics, University of Calgary, Calgary, Canada.
[3] Department of surveying and Geomatics Engineering, College of Engineering, University of Tehran, Tehran, Iran.
[4] Young Researchers and Elite Club, Mashhad Branch, Islamic Azad University, Mashhad, Iran.
[5] Department of surveying and Geomatics Engineering, College of Engineering, University of Tehran, Tehran, Iran.

Summary

Recently several types of research have been suggested in the last few years for modelling and analyzing raw movement data, but they mainly concentrate on studying the geometric view of raw trajectory data and do not consider the semantic facet of moving objects. Therefore, the major challenge in trajectory data modelling and reasoning is the definition of techniques to enrich raw movement data semantically. In this work, we investigate a new semantic trajectory ontology model for modelling moving objects by considering several semantic features and also reasoning on them. The model can be applied to the inference of different activity types. The reasoning mechanism is used to achieve further semantic enrichment that puts together different levels of information of the domain. The performance of the approach was tested and evaluated using a dataset, which was acquired by a user over a year in the City of Calgary.

KEYWORDS: Ontology modelling, Semantic trajectory, Ontology reasoning.

1. Introduction

In recent years, advances in positioning and tracking technologies have led to a large number of positioning data that can be represented as trajectories (Zhou et al., 2004). Some researchers have investigated some analytical techniques (Galton, 2005; Zheni et al., 2009) and computational methods (Laube, 2005; Miller, 2005) for the modelling and analyzing of movement data, but they mostly focus on studying the geometric properties of raw trajectory data, which makes the interpretation of results difficult. Mined results can be made more meaningful when the nature of the movement data is considered as context within the modelling process (Zhou et al., 2004; Gao et al., 2020). Semantic trajectory modelling is a main task of the semantic trajectory construction. It is the process of defining and analyzing data requirements to support the application of trajectories. Therefore, trajectory data need to be reconsidered not only from the geometric view but also from the meaningful semantic view as well to interpret and understand their meaning (Manaa and Akaichi, 2017). Xie et al. (2009) have proposed the influence and influence duration of POIs on trajectories. This method actually selects the POIs closest to the polyline geometry of the trajectory.This study investigated various extracted features and background information based on the model ontology to extract different activity types. For the representation and modelling of semantic trajectories, there are three different approaches; namely, data type based; design pattern based; and ontology based modelling. As a data

[*] moakbari@birjand.ac.ir
[†] mousavia@ucalgary.ca
[‡] Ahmad.eslami73@ut.ac.ir
[§] mobineftekhari@yahoo.com
[**] Mohsen.ghorani@ut.ac.ir

type based model approach, the research presented in Zheni et al. (2009) introduced an algebraic model that represents a spatiotemporal trajectory as an Abstract Data Type (ADT), encapsulating dynamic and features. It supported operations covering its spatiotemporal and semantic properties. However, a conceptual model supporting the various requirements of the applications of semantic trajectories was still needed. To fill this gap, Dodge (2001) introduced dedicated data types. They brought the minimal information common to all trajectories like the begin, end, moves, stops, as well as their sample points and interpolation functions, and encapsulated them in a generic data type. In Zhou et al. (2004), the authors introduced a design pattern based model relying on the MADS model. The model aims at the explicit representation of trajectories and their components as object types in the database schema and linking those components with application objects. Moreover, design pattern based modelling has been significantly adopted in the GeoPKDD project, where many mining techniques have used these stop-move concepts (Ong et al., 2010; Wachowicz et al., 2011). One example of a model based on ontologies is represented in Yan and Spaccapietra (2009), where the authors analyzed modelling requirements for trajectory modelling and proposed a trajectory model. In this approach, they used different ontologies to add semantics to the trajectories. In Wannous et al. (2013), a case study was presented on the use of an ontological-based approach for modelling semantic trajectories.
Choosing the right modelling approach for semantic trajectories depends on several factors. Data type modelling approaches alone are not sufficient to support the semantic trajectories application requirements. This is due to the inefficiency of using a generic data type for all application domains. Design pattern based models are even more generic than the data type based models, as they are not restricted to a specific data type. Instead, a dedicated type relevant to the application in hand can be added to the generic data types but will need the help of a database designer. Therefore, this model requires from the designer to add the semantic information specific to the application. On the other hand, ontology based models are application specific, as the ontology needs to reflect the application domain. They can represent richer semantic, and involve any kind of semantic annotations. In contrast to data type models, ontological models are naturally extensible because ontologies are designed to extend. Therefore, this paper investigates a new ontology based model for modelling and reasoning on movement data.

2. Methodology

As depicted in **Figure 1**, this section illustrates the proposed methodology, which consists of two main parts, namely: semantic trajectory ontology modelling and prototype development.

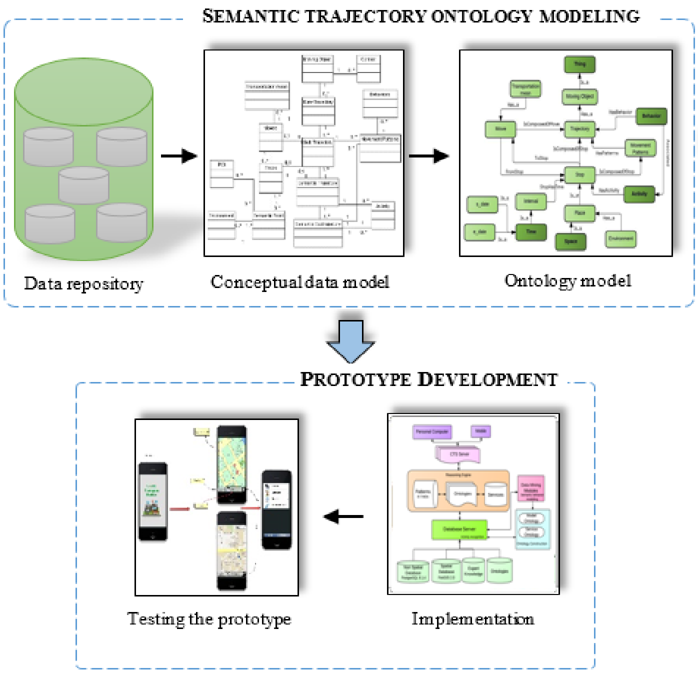

Figure 1 The proposed methodology

2.1. Semantic trajectory modelling

2.1.1. A conceptual data model for semantic trajectories

Figure 2 describes the conceptual model, which addresses the modelling requirements with the goal of analysis of semantic trajectory data. The aim of this model is to represent the concepts and relations of the movement domain where trajectory data and semantic movement patterns are to be interpreted along with the activity types. This model is an extended version (green colored box) of the conceptual framework introduced in Alvares et al. (2007), which relies on the conceptualization of stops and moves in trajectories. The conceptual model contains information related to: moving object, raw trajectory, sub-trajectory, semantic sub-trajectory, semantic trajectory, semantic place, stop, move, activity type, and behavior type.

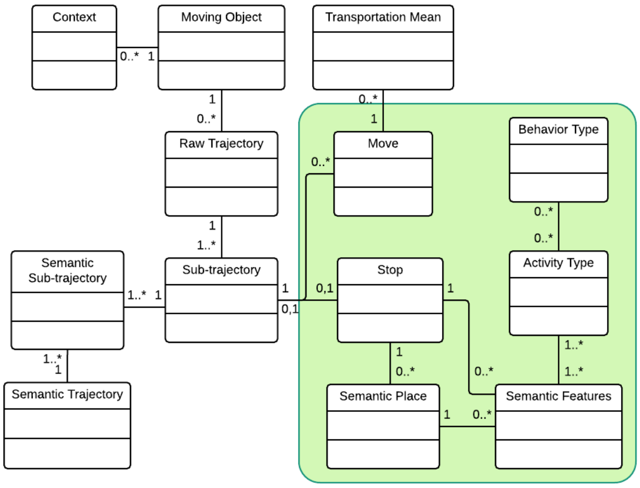

Figure 2 Extended conceptual model of semantic trajectory used in this work

A moving object generates raw trajectory, which based on different criteria can be divided into sub-trajectories. Giving meaning to the sub-trajectories, semantic trajectory is a combination of different semantic sub-trajectories. Each sub-trajectory is composed by stops and moves. Every stop is connected to an interval that represents the time of the stop. It includes begin time and end time concepts that identify when; the trajectory starts and ends. Each stop, as shown in the **Figure 2**, could have different semantic features. Semantic features are extracted from the stops and are divided into five different types such as stop frequency; average duration, stop land use type, stop POI category type, and stop begin time, which can help to infer the type of activity that occurs at each stop. An activity is what the moving objects are going to do during their movement. In other words, it is the objective of the movement, which has a start time and an end time, and it can be relative to the entire trajectory or part of the trajectory (the semantic sub-trajectory).

Activity types are highly influenced by users' location. For instance, if a person is close to a university, the most probable activity types would be studying, teaching, or working. Therefore, there is an association between places and activity types and according to the conceptual model; an activity is typically performed in a place. Activity types are also correlated to time and, in particular, to the time of day and the duration that a user spends at each place. Different activity types might have different timetables and durations. For instance, if the place is a restaurant, different time periods may be interpreted as different activity types. For instance, the period of 15 to 30 minutes would be interpreted as a delivery, since it is not enough time to stay in and eat. If the time period is between 30 minutes and 3 hours, it would be interpreted as dining and if the time period is between 3 hours and 8 hours, it would be interpreted as working. Also, stop frequency and average duration are important features to find out the type of activity. As a general example, if the stop frequency is more than five days a week and the average duration is more than eight hours, it could infer that the place would be either where the person works or lives. Therefore, this research hypothesizes a functional relationship (1) for Activity Types (AT) based on different features as shown below.

$$AT=f(P, L, S_f, T_b, S_d) \qquad (1)$$

Where P is the POI type that is around the stop, L is the land use type where the stop has occurred, S_f is the frequency of the stop in a week, T_b is the begin time of the stop in the place, and S_d is the average duration of the stop in a week. For instance, for a specific stop, if the land use type is residential, the POI type is null, the begin time is evening, stop frequency per week is more than six, and the average duration is more than ten hours per week then the moving object is 'spending time at home', i.e., AT= Return Home (Actually return home not imply that the moving object was not starting from at home). At this step some rules can be defined on the captured data in order to extract different activity types.

2.1.2. Semantic trajectory ontology model

Semantic Trajectory Ontology Model (STOM) is built based on the proposed conceptual data model. **Figure 3** shows a very partial version of this ontology with only the most important concepts and relationships. Shapes represent the main concepts whereas arrows represent relationships between two concepts.

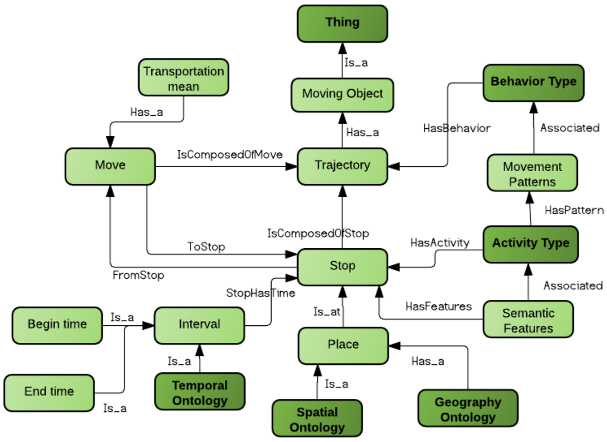

Figure 3 Semantic trajectory ontology model

The STOM consists of different ontologies such as spatial ontology, temporal ontology, geography ontology, and the theme ontology. The spatial ontology holds generic concepts for the description of the geometry component of a trajectory. The Temporal ontology is another source of information, which integrates time concepts and rules for modelling semantic trajectories. OwlTime ontology (Mousavi and Hunter, 2012), which is being developed by the World Wide Web Consortium (W3C) is chosen. The geography ontology describes the places where people move through and includes a variety of land use types, road networks, and POIs layer. The POI represents the specific categories such as shopping center, park, etc. The road network represents the interconnections of different road types designed within urban areas. The land use represents different regions and their utilization such as agricultural, residential, recreational or other purposes. Therefore, this ontology is used to aid potential interpretation of each stop, i.e., why the moving object stopped. The theme ontology model gathers a wide range of application dependent concepts. The understanding of trajectories profoundly depends on their relationships to application objects not just the moving object itself. As it can be seen in **Figure 4**, the model describes the concept of activity type that is of interest within a particular application context, which considers users' semantic features. It includes stop, trajectory, semantic features, movement patterns, activity type, and behavior type concepts. Therefore, integrating these ontologies together provides the semantic description of application-relevant trajectories with their domain specific semantic meaning. These ontologies are integrated into a unique ontology by setting up rules between them.

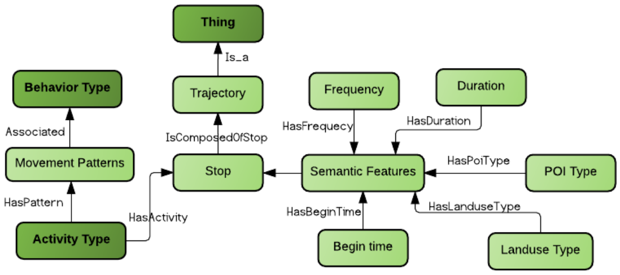

Figure 4 Theme ontology model describing activity type

The activity types are defined using axioms based on different semantic features included in the ontology model to express relations between the ontologies (see **Table 1**). For instance, in rule number one, if the land use type is residential, the POI type is null, the begin time is evening, the stop frequency per week is more than 5 and the average duration is more than 10 hours per week then the moving object is 'spending time at home', i.e., AT= Return home.
Another example is rule number 2: if the moving object stops once within a residential land use type per week, with an average stop duration of more than 30 min and the begin time is evening or night, then the moving object is 'visiting a friend', i.e., AT= Socializing. Therefore, applying the axioms outlined above, different activity types can be inferred.

Table 1 Axioms associated to activity types

Land use Type	POI Category Type	Features			Activity Type
		T_b	S_f	S_d	
Residential	-	Evening or night	≥ 5	≥ 9 hours	Return home
Residential	-	Evening or night	≥ 1	≥ 30 minutes	Socializing
Commercial	Shopping	Evening or night	≥ 1	≥ 30 minutes	Shopping
Commercial	Daily Shopping	Evening or night	≥ 1	≥ 30 minutes	Daily Shopping
Any Type	Any Type	Morning	≥ 5	≥ 8 hours	Work Full-Time

The added value of having such an ontology based approach, allows one to define axioms in terms of high-level semantic concepts, abstracting away from the geometry coordinates of the geographical features. Indeed, in this approach, each stop is treated as a semantic concept instead of using spatial coordinates.

2.2. Prototype development

The Prototype development consists of three steps. The first step is the data preparation, where the GPS data is cleaned and the daily and weekly basis trajectories are identified. The second step is the semantic enrichment process. The final step is the ontology-based activity model. The information retrieved from the previous steps is used to populate STOM for reasoning the activity type. As shown in **Figure 5**, the overall system has several essential components such as a database server, an ontology constructor, a mining module, a reasoning engine, and a user interface. The database server receives the current position data and stores it. In the ontology constructor, the ontology model, STOM is built. The extracted semantic features from the prototype development step are used to populate STOM and then the reasoning engine is executed in order to classify the activity types.

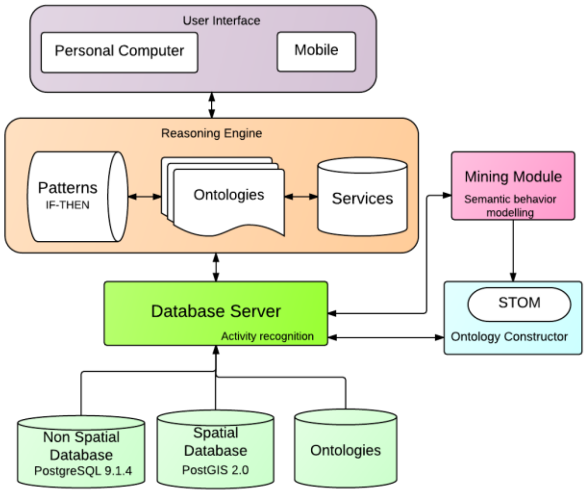

Figure 5 Main components of the system prototype

3. Experiment and results

The performance of the ontology model was evaluated using a dataset, which was captured by a user in the city of Calgary for a year in 2018. It has a total of 862,046 GPS records. The attributes collected include: user id, date, speed, heading, mode, and location of the user. The land use data includes different types such as commercial, urban development, residential, institutional, industrial, parks, major infrastructure, and transportation. The POIs were downloaded from the OSM. Data preparation procedure was applied to the collected data. First, trajectories, which were not in the monitored area, were removed. Next, the dataset was cleaned from the inconsistencies such as empty values, duplicates and outliers. As a result of the stop detection, 1,237 sub-trajectories with 832 moves and 801 stops over the dataset were produced. The algorithm in (Marketos et al., 2008) was used to annotate stops with the land use and POI types.

Figure 6 shows all the stops annotated with different land use types. The commercial type was annotated in 42 different places, whereas the institutional type was annotated in three different places.

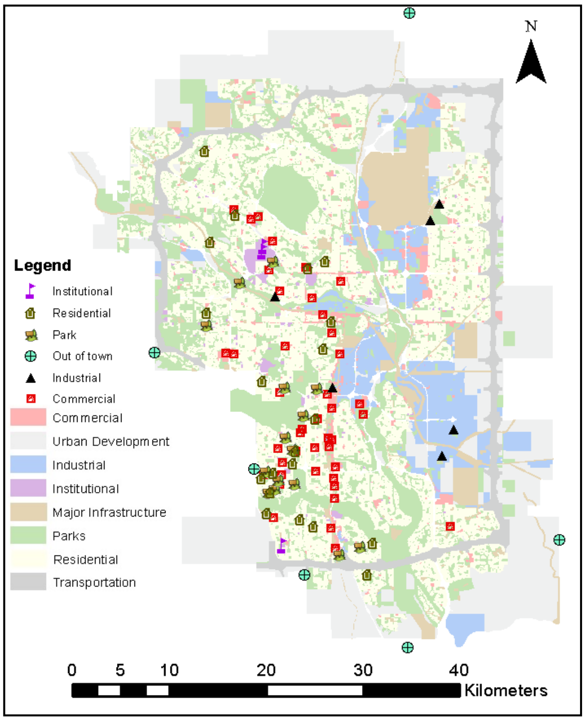

Figure 6 All detected stops in different land use types

In the POI category type annotation, as shown in **Figure 7**, most of the stops belonged to the shopping (30.1%), business services (18.9%) and food (17.7%) categories. The STOM was populated using the extracted information in the previous steps.

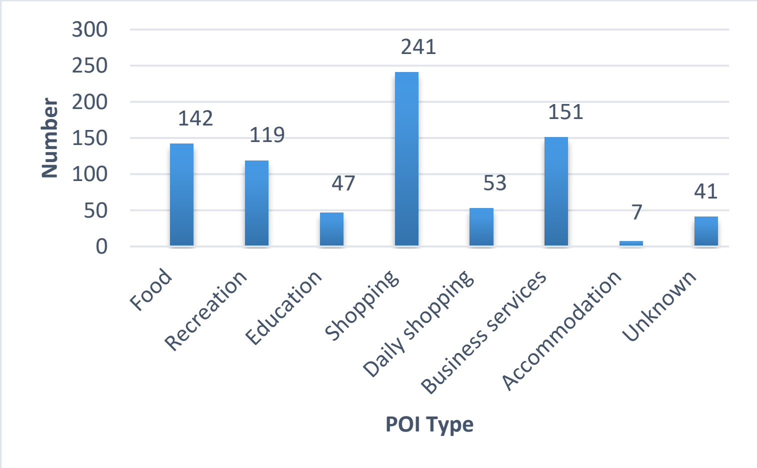

Figure 7 Number of POI types assigned to the stop trajectories

Some extracted semantic features are shown in **Table 2**. For instance, the user had stopped six times a week in a residential land use type named "Residential1" in the evenings with the average of 614 minutes per week.

Table 2 Some semantic features in the stop ontology

Stops	Frequency	Begin time	Average (min)
Commercial1	1	Evening	221
Commercial29	2	Night	14
Institutional1	5	Morning	441
Residential1	6	Evening	614
Residential14	1	Evening	22

4. Results and evaluation

The reasoning step was executed by the reasoner using the axioms that had been defined for some activity type. **Table 3** shows some inferred activity types.

Table 3 Some of the inferred activity types

Land use Type	POI Category Type	Features			Activity Type
		T_b	S_f	S_d	
Residential	-	Evening	6 days per week	10.2 hours	Return Home
Residential	-	Evening	1 day per week	45 min	Visiting
Commercial	Shopping	Afternoon	2 days per week	41 min	Shopping
Institutional	-	Morning	5 days per week	8.2 hours	Go to work

A web interface application was developed to visualize daily semantic trajectories and collect users' feedback to validate the proposed methodology. Therefore, the inferred activity types using the proposed method are compared to the collected feedback of the users. The accuracy as seen in (2) is the number of correctly inferred activity types over the number of total inferred activity types from the dataset.

$$Accuracy = \frac{No.of\ correctly\ \inf erred\ activity\ type}{No.of\ total\ \inf erred\ activity\ type} \tag{2}$$

The experimental outcome and the evaluation results are depicted in **Table 4**. It shows the accuracies per activities i.e. the percentage of activities correctly identified w.r.t. the number of declared activities (of the same type). For example, good results for activities of type "business services" (the method recognized 97.3 % of them) were obtained, while the method was unable to identify "daily shopping" (the method recognized 35.9 % of them). It was observed that these results were related to the availability of the POIs around the stops.

Table 4 Accuracy of extracted activities using user's feedback

Activity Type	Accuracy (%)
Eating	88.4
Recreational	86.1
Education	69.5
Shopping	91.3
Daily shopping	35.9
Business services	97.3
Go to work	93.8
Trip	88.6
Socializing	90.6
Return home	89.1

Figure 8 illustrates a web interface application to visualize daily semantic trajectories and collecting users' feedback to validate the proposed methodology for prototype development in this research work. At the bottom of the box, the user is asked to verify if the inferred activity type is correct or not.

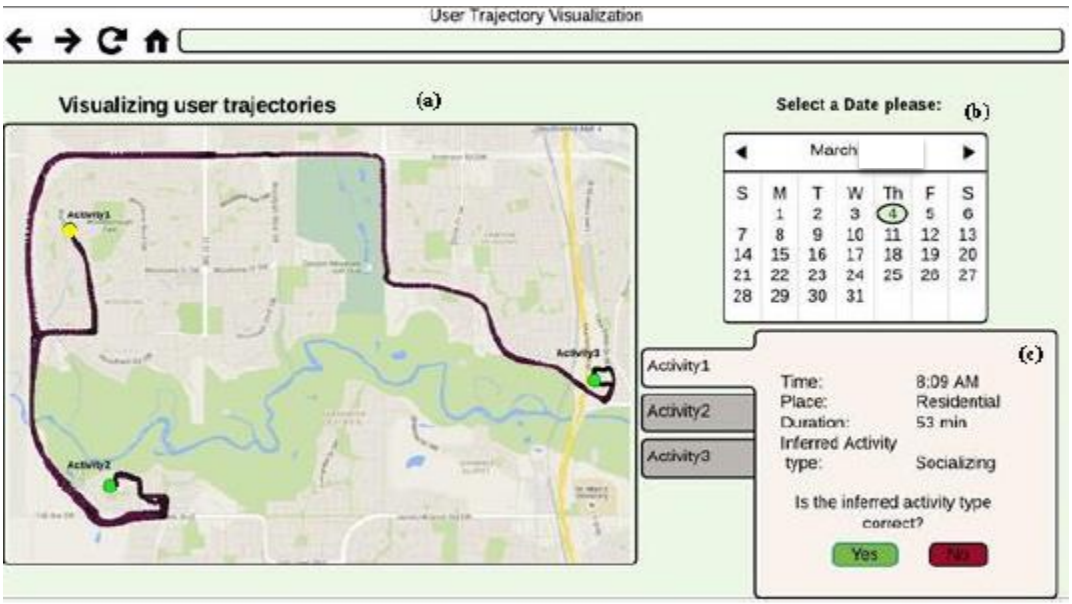

Figure 8 User interface to visualize user's trajectories in order to get his/her feedback.

5. Conclusions

This research proposed a new ontology based approach for modelling and reasoning on movement data. In comparison with other approaches, several semantic features were added to the proposed ontology model to enrich the relationship between objects in the model. These semantic features were duration, frequency, land use type, POI type, and start time. Finally, to evaluate the model, this study investigated various extracted features and background information based on the model ontology to extract some activity types from a dataset, which was collected for a year by a user. The proposed ontology model was used to define different axioms using common sense rules to infer the activity types. The results were promising and inspired us to extend the number of axioms in order to infer more types of activity. As a limitation, this research currently does not focus on multimodal transportation data and it just concentrates on users who drive their own cars. The results of this research can be identified using a rule-based approach without the use of ontology, which is one of the suggestions of future research.

References

Alvares, L. O., Bogorny, V., Kuijpers, B., de Macedo, J. A. F., Moelans, B. and Vaisman, A. (2007) A model for enriching trajectories with semantic geographical information. In *Proceedings of the 15th annual ACM international symposium on Advances in geographic information systems*, p. 22.

Dodge, S., (2001) *Exploring Movement Using Similarity Analysis.* PhD thesis, Universitaet Zuerich.

Galton, A. (2005) Dynamic collectives and their collective dynamics. Spat. Inf. Theory, pp. 300–315.

Gao, Y., Huang, L., Feng, J., & Wang, X. (2020). Semantic trajectory segmentation based on change-point detection and ontology. International Journal of Geographical Information Science, 34(12), 2361-2394.

Laube, P. (2005) *Analysing Point Motion Spatio-Temporal Data Mining of Geospatial Lifelines*. PhD thesis, Universitz of Zurich.

Manaa, M., & Akaichi, J. (2017). Ontology-based modeling and querying of trajectory data. *Data & Knowledge Engineering*, 111, 58-72.

Marketos, G., Frentzos, E., Ntoutsi, I., Pelekis, N., Raffaetà, A. and Theodoridis, Y. (2008) Building real-world trajectory warehouses. In *Proceedings of the seventh ACM international workshop on data engineering for wireless and mobile access*, pp. 8–15.

Miller, H. J., (2005) A measurement theory for time geography. Geogr. Anal., vol. 37, no. 1, pp. 17–45.

Mousavi, A. and Hunter, A. J. S. (2012) Pointers Extraction of Trajectory Data for Semantic Knowledge Discovery. *presented at the GIScience 2012 - Seventh International Conference on Geographic Information Science*, Columbus, Ohio.

Ong, R., Wachowicz, M., Nanni, M. and Renso, C. (2010) From Pattern Discovery to Pattern Interpretation in Movement Data. In *Proceedings of the 2010 IEEE International Conference on Data Mining Workshops*, pp. 527–534.

Spaccapietra, S., Parent, C., Damiani, M. L., de Macedo, J. A., Porto, F. and Vangenot, C. (2008) A conceptual view on trajectories. Data & Knowl. Eng., vol. 65, no. 1, pp. 126–146.

Wachowicz, M., Ong, R., Renso, C. and Nanni, M., (2011) Finding moving flock patterns among pedestrians through collective coherence. Int. J. of Geogr. Inf. Sci., vol. 25, no. 11, pp. 1849–1864.

Wannous, R., Malki, J., Bouju, A. and Vincent, C. (2013) Time integration in semantic trajectories using an ontological modelling approach. In *New Trends in Databases and Information Systems*, Springer, pp. 187–198.

Yan, Z. and Spaccapietra, S. (2009) Towards Semantic Trajectory Data Analysis: A Conceptual and Computational Approach. In *VLDB PhD Workshop*.

Xie, K., Deng, K., & Zhou, X. (2009) From trajectories to activities. a spatio-temporal join approach. Proceedings of the International Workshop on Location Based Social Networks. ACM. pp. 25–32.

Zheni, D., Frihida, A., Ghezala, H. B. and Claramunt, C. (2009) A semantic approach for the modeling of trajectories in space and time. In *Advances in Conceptual Modeling-Challenging Perspectives*, Springer, pp. 347–356.
Zhou, C., Frankowski, D., Ludford, P., Shekhar, S. and Terveen, L. (2004) Discovering personal gazetteers: an interactive clustering approach. In *Proceedings of the 12th annual ACM international workshop on Geographic information systems*, pp. 266–273.

Biographies

Mohammad Akbari, PhD in GIS, Assist. Proffessor, Dept. of Civil Engineering, University of Birjand, Birjand, Iran, Location Based Services, Spatila Data mining, context awareness.

Ali Mousavi, PhD in GIS, Graduated at University of Calgary. Dept. of Geomatics and Geospatial Information, University of Calgary. Location Based Services, Ontology, Spatial Data mining.

Seyed Ahmad Eslaminezhad, Department of surveying and Geomatics Engineering, College of Engineering, University of Tehran, Tehran, Iran, Location Based Services, Data mining.

Mobin Eftekhari, Young Researchers and Elite Club, Mashhad Branch, Islamic Azad University, Mashhad, Iran, Data mining, Spatio-temporal analysis.

Mohsen Ghorani Department of surveying and Geomatics Engineering, College of Engineering, University of Tehran, Tehran, Iran, Spatial Analysis, Data mining,