From Raw Pedestrian Trajectories to Semantic Graph Structured Model-Towards an end-to-end spatiotemporal analytics framework

Lamia Karim, Azedine Boulmakoul, Karine Zeitouni

Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 184 (2021) 60–67

Procedia Computer Science

www.elsevier.com/locate/procedia

The 12th International Conference on Ambient Systems, Networks and Technologies (ANT)
March 23-26, 2021, Warsaw, Poland

From Raw Pedestrian Trajectories to Semantic Graph Structured Model–Towards an end-to-end spatiotemporal analytics framework

Lamia Karim[a,b], Azedine Boulmakoul[a,c]*, Karine Zeitouni[c]

[a]*LISA Lab. ENSA Berrechid, Hassan 1st University, Morocco*
[b]*LIM/IOS., FSTM, Hassan II University of Casablanca, Morocco*
[c]*DAVID Lab., UVSQ, Université Paris-Saclay, France*

Abstract

The conditions and context in which pedestrians are exposed to the risk of road accidents, air pollution or epidemy are important for improving safety. Spatial information on pedestrian mobility was difficult and expensive to collect so far. This is changing with the popularization of several apps either for guidance or self-monitoring. Produced Space-Time Trajectories (STT) perfectly match the concept of temporal geography, as to assess pedestrian exposure to the above risks. In this article, we propose a modeling approach of pedestrian STT. The abstract model provides a high semantic expressiveness through spatiotemporal queries and a high-level graph representation of the data. We depict a whole framework for collecting and analyzing GPS logs generated by pedestrians' devices, and discuss the system components. Among the targeted applications, a new generation of mobile applications can be developed that allow pedestrians to organize their daily activities and mobility, and minimize their exposure risks.

Keywords: Trajectory meta-model; Spatiotemporal database; Spatiotemporal data mining; Graph database model; Pedestrians risk.

* Corresponding author. Tel.: +212523315353; fax: +212523315353.
E-mail address: azedine.boulmakoul@fstm.ac.ma

1877-0509
10.1016/j.procs.2021.03.018

1. Introduction

The trajectory model and human interaction are considered very important for road safety analysis. In the ideal case, the trajectory model can be obtained by the trace of the walk of the pedestrians according to the various source-destination paths. Pedestrians' mobility, participate in the governance of travel safety and in controlling the effect of vehicle flows. Mobility trajectories are not specific to the pedestrians' movement; trajectories are ubiquitous to several application domains. Spatiotemporal navigation routes of a moving object are the most popular trajectory computation examples in the literature [3] [8] [12] [15] [20] [25]. These trajectories relate to vehicles or people. The transportation and logistics field, all modes (land, sea, rail, air) produce trajectories resulting from logistics processes. The customers' behavior in a commercial zone or the navigation of travelers in some strategic places (airports, metro or train stations, bus stations …) produce large trajectories which analysis might be insightful. A manufacturing process can also be seen as a spatiotemporal or as a semantic trajectory (i.e., events detected from sensors or generated by actuators), which exploration allows business applications analysis and inform on how a company's processes work in reality. Pedestrians trajectories generated by camera videos or other technologies are also utilized in the analysis of road safety patterns or in crowd behaviors' analytic [1] [5] [7]. Mining trajectory patterns from moving objects databases has been widely studied in various works using various approaches [9-11] [23]. These models are useful in areas such as transportation planning and traffic monitoring. Discovery of high traffic flow patterns related to urban traffic congestion mostly uses vehicle trajectories. In our previous work, trajectories were dealt with in areas such as logistics, pedestrians' safety [12] [17-18], congested networks, and HazMat monitoring [4]. The convergence of the work carried out so far has fostered the proposal of a unified generic trajectory model. In this paper, we focus on pedestrian trajectories, and propose a suitable semantic data model and depict a system architecture for collecting, storing and analyzing pedestrian trajectories. We adopt a graph representation, which provides a flexible way to augment the raw trajectory by activities and surrounding spatial objects, thus accounting for the points of interest (PoI) and the spatiotemporal events that define trajectories [3, 20]. Many interesting queries of the Graph trajectories database can be expressed. The system offers other services to measure pedestrian travel time, waiting time, crossing speed, etc.

Works proposed in [22], concern human mobility and its direct influence on disease transmission. This work processes daily information on mobility and activity to measure exposure to risk factors for infection. They present some methods and an integrated ArcGIS desktop-based visual interface for the pre-processing and spatiotemporal analyses of trajectory data. They provide examples of how such processing may be used to model human space-time activities, especially with pedestrian trajectory data. In the work in [24], user-generated GPS trajectories were used to derive a location-interest graph that contains useful location-related information (location, interest, stay time, travel time, sequence classic travel). These information in a region are used to recommend a route based on the user's request. An architecture for recommending a route is provided to manage the GPS trajectories generated by the offline users and the processing of the online users' requests. The purpose of the study given in [7] is to examine the effects of EEG signals on different cognitive tasks of the pedestrians. The author developed Pedestrian 3D virtual simulator which considered in detail the decision-making cognitive tasks of pedestrians when receiving/replying SMS while walking, with simultaneous measurement of the EEG signals. Works given in [13, 14] present a model of semantic trajectories that apply in indoor space, which adapt to context-aware mobility data mining and analytics methods. This work combines aspects of semantic outdoor trajectory models, with a hierarchical symbolic representation of indoor space, which respects the IndoorGML standard of the OGC. Work in the field of databases of moving objects also addresses other problems [11]. In particular, some research focuses on the compression of trajectory-type spatiotemporal data in road networks [21]. The closest work to ours is the one proposed in [9] using graph databases for the purpose of moving objects analysis. Raw trajectories are transformed into semantic trajectories. This work also proposes the use of Neo4j graph databases to represent and store trajectory data, which can thus be analyzed using Cypher graph query languages. It shows the superiority of graph databases over the relational alternative in both expressiveness and performance of trajectory queries in many typical cases. Although a general semantic trajectory is interesting, its specialization for different sub-categories of moving objects is of a paramount importance in the application domains thereof. This was the case in maritime domain in Itani et al. [26]. Our proposal revolves around typical pedestrians' semantic trajectories, which differs from, e.g., ships or cars movement semantic.

The rest of the paper is organized as follows. Section 2 gives an overview of pedestrian trajectories meta-model. In Section 3, we present our system architecture and provide discussions followed by conclusion in Section 4.

2. Modeling pedestrian trajectories

The availability of pervasive digital devices, wireless networks and satellite positioning systems have enabled the development of Location Based Services (LBS) that provide relevant location information of mobile users. The trajectory of a moving pedestrian is defined as a sequence of spatiotemporal locations. These trajectories can also relate the semantics of locations through points of interest (semantic trajectories). It is also possible to capture the activities carried out by pedestrians during their walk (space-time trajectories). The pedestrian trajectory model is briefly discussed in this section. Only relevant elements to the present work will be considered. The class diagram given in Fig. 3, describes the leading classes of pedestrian trajectories modeling. For the sake of simplicity, we briefly express the necessary modeling elements to perceive the meta-model:

- A spatiotemporal event is made up of observables such as space, time, measure, etc.,
- An observable can be a fuzzy physical quantity defined by an algebra,
- An event is produced from sensors,
- A semantic event is an event extended with a location toponym, and it is associated with a point of interest,
- For trajectories, semantic events are categorized as begin event (departure from a place), stop event (stop during a lapse of time) and end event (end of trip),
- An event is linked (causal relation) to another event by means of the move relation,
- The perception-topology class-relation denotes the neighborhood relation of a mobile to capture nearby points of interest,
- Activity events are semantic events materializing physical activities or virtual activities,
- A trajectory is made up of an ordered set of events, it has an intention (shopping, getting to work, coming home, going to school, etc.),
- A story is a trajectory, made up of trajectories,
- An episode is a trajectory, a story is made up of episodes,
- A raw trajectory is made up of raw events (space, time, measure),
- A structured trajectory specializes a raw trajectory by creating "topological" links between raw events,
- A semantic trajectory specializes a structured trajectory, it is made up of semantic events,
- A space-time trajectory specializes a semantic trajectory, it is made up of activity-type events,
- An intelligent trajectory is intended for an intelligent actor, such as a person.

Fig. 1. Moving pedestrians and trajectories.

Trajectories can be represented using graph data model and loaded into a graph database such Neo4j. The graph could be generated by Arrows, a JavaScript-based graph rendering tool. Handy for presentations, Arrows tool, meshes a minimal interface with support for features like the Cypher language. When a pedestrian is moving (Fig. 1, Fig. 2), the Begin, Stop, End type events are associated with contextual information related to the pedestrian's cognitive perception [6], such as, the points of interest encountered, urban signage, advertising information, noise nuisance, etc. The collection of points of interest in the surrounding of the pedestrian's perception combined with other factors ensure the assessment of exposure to the risk of accidents, high pollution or noise hazards (Fig. 1). Urban pedestrian mobility is of interest to road safety studies. The cognitive perception of things in the proximity of pedestrians during their movements helps to under-stand risk-taking. This walkability context is collected for analytical purposes (Figures 1,2).

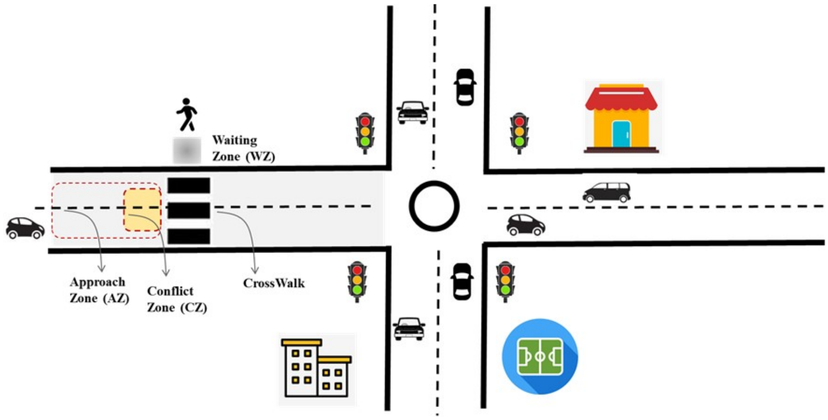

Fig.2. Pedestrian in junction context.

3. System architecture

This section describes the main nodes of our system architecture (Fig. 5), which include: spatial database (dockerized PostGIS), Outdoor Geolocation service (docker service), Graph Trajectories Factory (Neo4j docker services), and Trajectories Mining Services (Neo4j Data Science docker services). The central data exchange node consists of Apache Kafka on Kubernetes. We describe only some processing operations for the construction of the trajectories.

Data collection

Trajectory data can be collected with handheld GPS units, GPS-enabled smart phone tracking applications, as well as Assisted GPS devices (Fig. 1), in the following we describe the data collection activities:

- Trajectory data is saved in terms of time-latitude-longitude records, in GPX (GPS eXchange) file format. files are converted into commonly-used Geographic Information Systems (GIS) file format (i.e. ESRI shapefile),
- Load shapefiles from other layers (polygon, line, point) of the boundary of the study area,
- The graph generation starts with downloading the OSM data to the local machine and convert it to shapefile format, followed by geoprocessing tasks such that filtering, fixing topology, and editing data that are required before other specific processing. OpenStreetMap provides geospatial data in the .osm file format [19]. The OSM project has the necessary resources and infrastructure to insert and edit cartographic and semantic data and useful information here: traffic signal, pedestrians crossing, building, street, etc. To overcome the problems related to missing data and information, the experiment is done in an area of the OSM database with all the data available and the information required for pedestrians; adding missing items (e.g. footway) and updating existing features with required information (sidewalks, crossings, traffic signals, bus stops, steps, Traffic sign, poles, trees, entrances and footways).

Pre-processing

Pre-processing stage of Pedestrian Trajectories is described as the following:
- Visualize trajectories in the appropriate projection system. Examine raw trajectories in space and time,
- Identify errors in the data based on the shape, speed and topology of track segments.
- Select them and delete them from the original trajectories,
- Replace the points grouped in star shapes due to a weak GPS signal by the spatiotemporal centroid of these points.

Space-Time-Path Trajectory factory

Trajectories factory requires the building layer, the urban road network layer and the points of interest layer. First, the spatial objects of the stop type are developed corresponding to the stop for a lapse of time of a pedestrian. A PoI is considered as a stop when a moving object remains in it for a length of time above some threshold, in which points of a trajectory that are located inside the PoI are transformed to the spatial object that represents this stop. Afterwards, the other geographic objects are produced as well as the relations linking them. Using Neo4j with the spatial API facilitates the construction of trajectories and the search for neighborhood objects. We use Neo4j library that facilitate spatial operations. In particular, it allows adding spatial indexes and performing spatial operations such as searching for data in specified regions or at a specified distance from a point of interest. It also provides services for exporting data to geospatial data servers.

Trajectories Analytics

The need for a trajectory analysis service is eagerly awaited by decision-makers. Various possibilities are available. They are based on graph mining algorithms and spatial data mining in general. Below we describe some ways just to draw the scheduled choices:

- Trajectory Prediction–TP [8]: Given the (incomplete) trajectory of a moving object, consisting of its time-stamped locations recorded at past i time instances, and a set C of constraints, predict the anticipated trajectory at next (i+1) time instances of the object, which is consistent with C,
- Future Location Prediction–FLP [8]. Given the (incomplete) trajectory of a moving object, consisting of its time-stamped locations recorded at past i time instances, and an integer value j ≥1, predict the anticipated trajectory at next j time instances of the object,
- Trajectory Clustering [2] [6], in order to discover meaningful trajectories,
- Trajectory Exploration: geo-visual analytics for exploring movement data that links three-dimensional visual representations, a virtual globe and trajectory data mining algorithms,
- Several platforms are also available for the analytics of trajectories modeled by graph structures. So, two very interesting libraries, SNAP† and Neo4j Graph Data Science‡, are useful for graphs analysis. Stanford Network Analysis Platform (SNAP) is a general-purpose network analysis and graph mining library. The Neo4j Graph Data Science Library that enable reasoning about network structure: Community Detection, Centrality, Similarity, Heuristic Link Prediction, Pathfinding & Search.

---

† http://snap.stanford.edu/

‡ https://neo4j.com/graph-data-science-library/

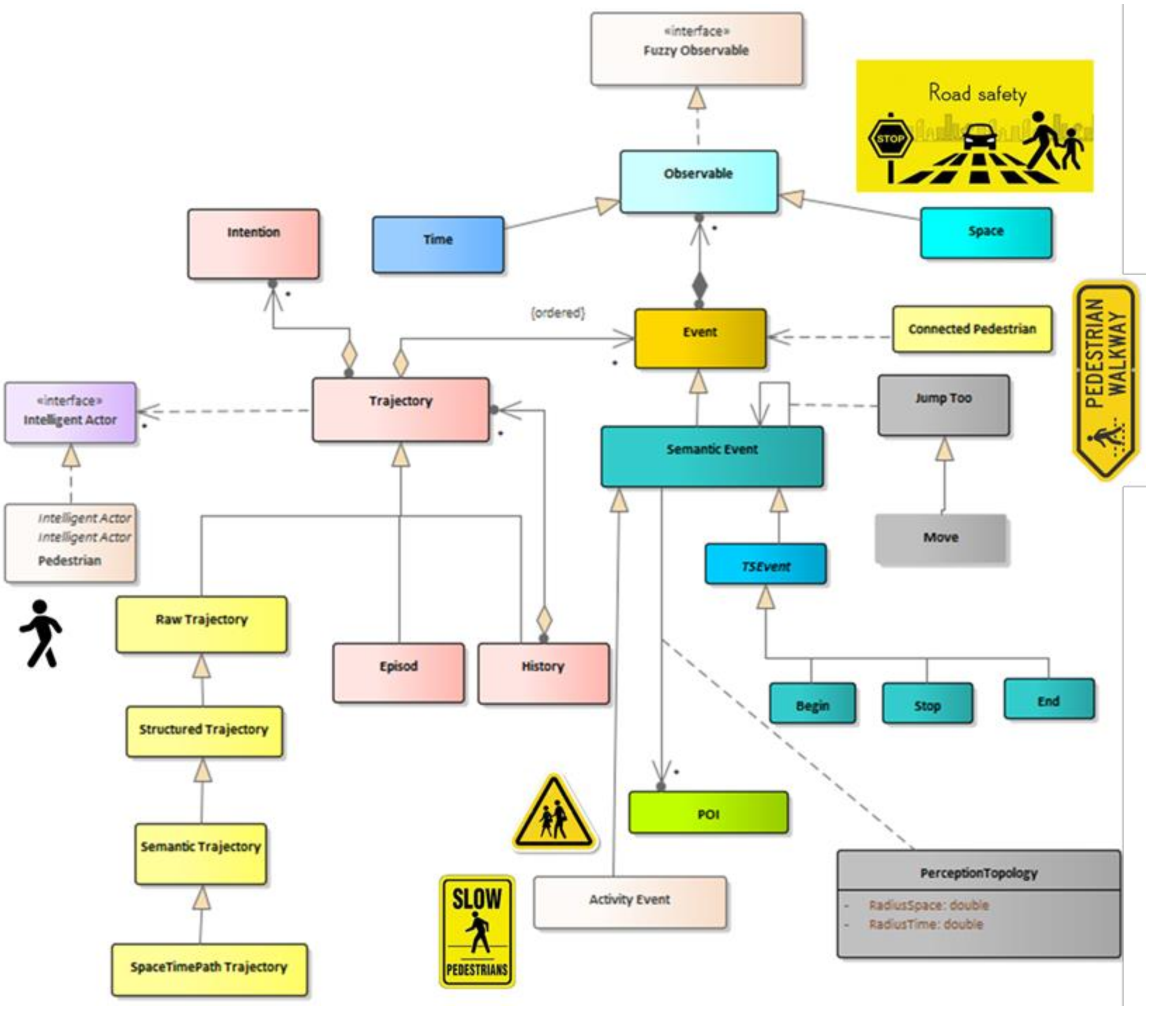

Fig. 3. Simplified pedestrians' trajectories meta-model.

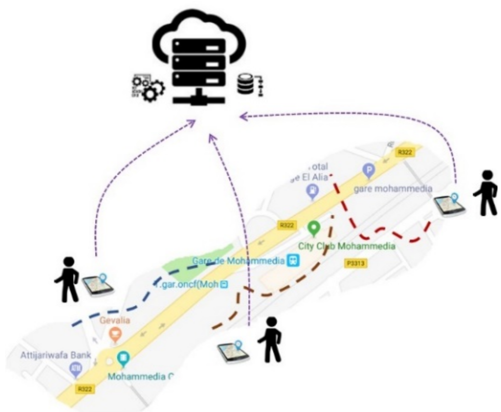

Fig. 4. Cloud system deployment.

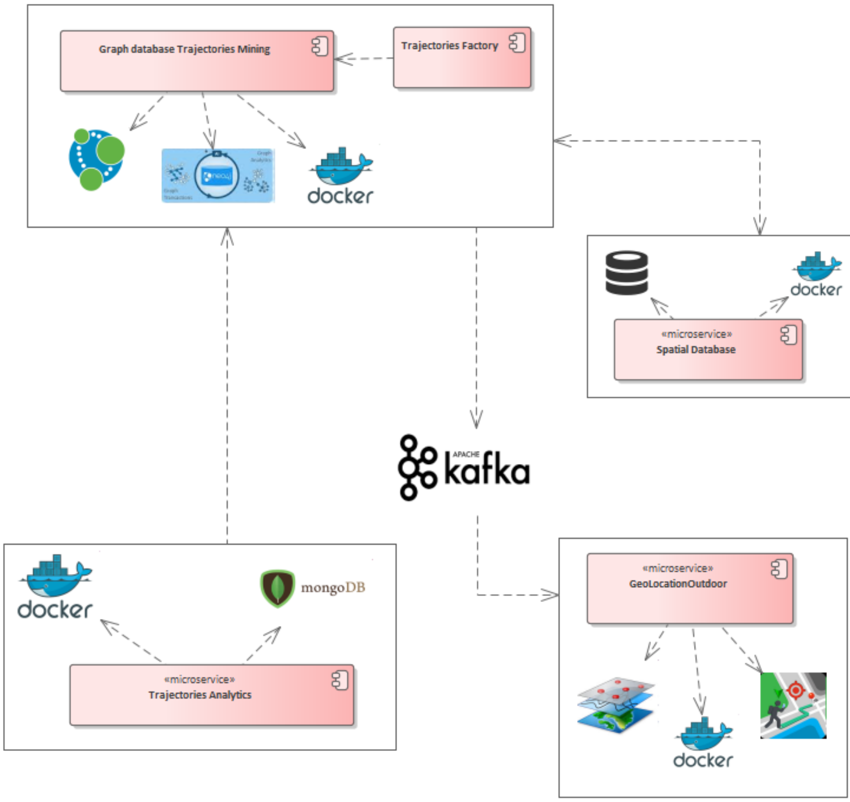

Fig. 5. Components' architecture.

4. Conclusion

Nowadays, pedestrian mobility data is not as widely available. However, this situation is changing rapidly as GPS-enabled mobile devices become more and more common for pedestrian applications and provide data collection. GPS data collected by pedestrians is a big challenge during preprocessing as errors are massive due to buildings proximity and frequent indoor stops. The objective being to characterize the pedestrian risks according to traveled paths. The expected results will make it possible to work on road signs, urban architecture and information on road users. In this work we approach the modeling of pedestrians' ST trajectories. The modeled Space-Time trajectories will be specified and stored in a Neo4j graph. Therefore, spatiotemporal analytics will be expressed using Cypher query language. Details of specifications and implementations will be presented in our future work. Trajectory analysis and exploration algorithms will be applied to study the vulnerabilities and risks incurred by pedestrians, in order to support an action plan to improve safety.

Acknowledgment: This work was partially funded by Ministry of Equipment, Transport, Logistics and Water–Kingdom of Morocco, The National Road Safety Agency (NARSA) and National Center for Scientific and Technical Research (CNRST). Road Safety Research Program# An intelligent reactive abductive system and intuitionist fuzzy logical reasoning for dangerousness of driver-pedestrians interactions analysis: Development of new pedestrians' exposure to risk of road accident measures.

References

[1] Bodeker, M. (2018) "Walking and Walkability in Pre-Set and Self-Defined Neighborhoods: A Mental Mapping Study in Older Adults", Int. J. Environ. Res. Public Health, 15(7), 1363; https://doi.org/10.3390/ijerph15071363.

[2] Boulmakoul, A., Fazekas, Z., Karim, L., Gáspár, P., Cherradi, G. (2020) "Fuzzy similarities for road environment-type detection by a connected vehicle from traffic sign probabilistic data", Procedia Computer Science, Volume 170, 2020, Pages 59-66, ISSN 1877-0509.
[3] Boulmakoul, A., Karim, L., El Bouziri, A., Lbath, A. (2015) "A System Architecture for Heterogeneous Moving-Object Trajectory Metamodel Using Generic Sensors: Tracking Airport Security Case Study". IEEE Systems Journal 9(1): 283-291.
[4] Boulmakoul, A., Karim, L., Lbath, A. (2016) "T-Warehousing for hazardous materials transportation". Ingénierie des Systèmes d Inf. 21(1):39-59.
[5] Cerin, E.; Nathan, A.; van Cauwenberg, J.; Barnett, D.W.; Barnett, A. (2017) "The neighbourhood physical environment and active travel in older adults: A systematic review and meta-analysis". Int. J. Behav. Nutr. Phys. Act. 14.
[6] Coulton, C.J.; Jennings, M.Z.; Chan, T. (2013) "How big is my neighborhood? Individual and contextual effects on perceptions of neighborhood scale". Am. J. Community Psychol, 51, 140–150.
[7] Erkan, I. Cognitive Analysis of Pedestrians (2017) "Walking While Using a Mobile Phone", Journal of Cognitive Science 18(3):301-319, DOI: 10.17791/jcs.2017.18.3.301.
[8] Georgiou H. et al. (2018) "Moving Objects Analytics: Survey on Future Location &Trajectory Prediction Methods", Technical Report, arXiv:1807.04639.
[9] Gómez, L.I., Kuijpers, B., Vaisman, A. (2019) "A. Analytical queries on semantic trajectories using graph databases", Transactions in GIS, Volume 23, Issue 5, John Wiley & Sons Ltd.
[10] Güting, R. H., Schneider, M. (2005) "Moving objects databases". San Francisco, CA: Morgan Kaufmann.
[11] Güting RH, de Almeida VT, Ding Z. (2006) "Modeling and querying moving objects in networks". VLDB J 15(2):165–190 (2006).
[12] Karim, L., Boulmakoul, A., Mandar, M. Lbath, A. Nahri, M. (2020) "A new pedestrians' intuitionistic fuzzy risk exposure indicator and big data trajectories analytics on Spark-Hadoop ecosystem", Procedia Computer Science, Volume 170, 2020, Pages 137-144, ISSN 1877-0509.
[13] Kontarinis, A., Zeitouni, K., Marinica, C., Vodislav, D., Kotzinos, D. (2021) "Towards a Semantic Indoor Trajectory Model: Application to Museum Visits", to appear in GeoInformatica
[14] Kontarinis, A., Zeitouni, K., Marinica, C., Vodislav, D., Kotzinos, D. (2019) "Towards a Semantic Indoor Trajectory Model", in the Workshop Proceedingsof the EDBT/ICDT 2019 Joint Conference (March 26, 2019, Lisbon, Portugal) on CEUR-WS.org. https://hal.archives-ouvertes.fr/hal-02314572/document.
[15] Lee, W. & Krumm, J. (2011) "Trajectory preprocessing. In: Computing with Spatial Trajectories"., Zheng, Y., Zhou, X., eds., Springer, Bucher, 3-34.
[16] Maguerra, S., Boulmakoul, A., Karim, L., Badir, H. (2019) "A Distributed Execution Pipeline for Clustering Trajectories Based on a Fuzzy Similarity Relation". Algorithms 12(2): 29.
[17] Mandar, M., Boulmakoul, A., Lbath, A. (2017) "Pedestrian fuzzy risk exposure indicator", Transportation Research Procedia, Volume 22, 2017, Pages 124-133, ISSN 2352-1465.
[18] Mandar, M., Karim, L., Boulmakoul, A., Lbath, A. (2017) "A. Triangular intuitionistic fuzzy number theory for driver pedestrians' interactions and risk exposure modeling", Procedia Computer Science, Volume 109, 2017, Pages 148-155, ISSN 1877-0509.
[19] Open Geospatial Consortium, Inc. (2012) "OGC KML documentation". http://www.opengeospatial.org/standards/kml/.
[20] Parent, C., Spaccapietra, S., Renso, C. Andrienko, G., Andrienko, N., Bogorny, V., Yan, Z. (2013) "Semantic trajectories modeling and analysis". ACM Computing Surveys, 45(4), 42:1–42:32.
[21] Popa, I.S., Zeitouni, K., Oria, V., Kharrat, (2015) "A. Spatiotemporal compression of trajectories in road networks". GeoInformatica 19(1):117-145.
[22] Qi, F. and Du, F. (2013) "Trajectory Data Analyses for Pedestrian Space-time Activity Study", J Vis Exp. (72), 50130.
[23] Vecchio, P., Secundo, G., Maruccia, Y., Passiante, G. (2019 "A system dynamic approach for the smart mobility of people: Implications in the age of big data", Technological Forecasting & Social Change 149 -119771.
[24] Yoon H., Zheng Y., Xie X., Woo W. (2010) "Smart Itinerary Recommendation Based on User-Generated GPS Trajectories". In: Yu Z., Liscano R., Chen G., Zhang D., Zhou X. (eds) (2010) "Ubiquitous Intelligence and Computing". UIC 2010. Lecture Notes in Computer Science, vol 6406. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-16355-5_5.
[25] Zhong, R.Y., Huang, G.Q., Shulin, L., Dai, Q.Y., Xu, C., Zhang, T. (2015) "A big data approach for logistics trajectory discovery from RFID-enabled production data". Int. J. Prod. Econ. 165, 260–272.
[26] Itani, A., Ray, C., El Falou, A. and Issa, J., (2019) "Mining Ship Motions and Patterns of Life for the EU Common Information Sharing Environment (CISE)". In MTS/IEEE Oceans 2019, Marseille, France, 7 p.