Expert Systems With Applications 92 (2018) 533–545

Contents lists available at ScienceDirect

Expert Systems With Applications

journal homepage: www.elsevier.com/locate/eswa

FrameSTEP: A framework for annotating semantic trajectories based on episodes

Tales P. Nogueira[a,*], Reinaldo B. Braga[b], Carina T. de Oliveira[b], Hervé Martin[a]

[a] Univ. Grenoble Alpes, CNRS, Grenoble INP, LIG, F-38000, Grenoble, France
[b] Laboratory of Computer Networks and Systems (LAR), Federal Institute of Education, Science and Technology of Ceará (IFCE), 62800-000, Aracati, Brazil

ARTICLE INFO

Article history:
Received 14 May 2017
Revised 1 October 2017
Accepted 2 October 2017
Available online 3 October 2017

Keywords:
Trajectory modeling
Semantic web
Trajectory annotation

ABSTRACT

We are witnessing an increasing usage of location data by a variety of applications. Consequently, information systems are required to deal with large datasets containing raw data to build high level abstractions. Semantic Web technologies offer powerful representation tools for pervasive applications. The convergence of location-based services and Semantic Web standards allows an easier interlinking and annotation of trajectories. However, due to the wide range of requirements on modeling mobile object trajectories, it is important to define a high-level data model for representing trajectory episodes and contextual elements with multiple levels of granularity and different options to represent spatial and temporal extents, as well as to express quantitative and qualitative semantic descriptions. In this article, we focus on modeling mobile object trajectories in the context of Semantic Web. First, we introduce a new version of the Semantic Trajectory Episodes (STEP) ontology to represent generic spatiotemporal episodes. Then, we present FrameSTEP as a new framework for annotating semantic trajectories based on episodes. As a result, we combine our ontology, which can represent spatiotemporal phenomena at different levels of granularity, with annotation algorithms, which allow to create instances of our model. The proposed spatial annotation algorithm explores the Linked Open Data cloud and OpenStreetMap tags to find relevant types of spatial features in order to describe the environment where the trajectory took place. Our framework can guide the development of future expert systems in trajectory analysis. It enables reasoning about knowledge gathered from large trajectory data and linked datasets in order to create several intelligent services.

1. Introduction

Location data is ubiquitous in many aspects of our digital lives. We are witnessing an increasing production of this kind of data every day by a variety of applications and devices usually equipped with location sensors. The interest in movement analysis crosses many application fields such as behavioral ecology, people's mobility using various means of transportation, sports analysis, surveillance, among others. This increasing flow of data results in a demand for information systems able to deal with a large amount of raw and noisy data in such a way that high-level abstractions are designed, analyzed and organized to offer useful functionalities.

When location data is constantly collected and associated with an instant in time, we have a trajectory. From photos to sports activities, the presence of this kind of data at different levels of granularity is constantly growing. Therefore, there is a need for novel data structures, algorithms, and techniques to deal with spatiotemporal aspects of this information.

The variety and complexity of spatiotemporal data have led to the establishment of an interdisciplinary field called Computational Movement Analysis. This field studies the development and application of computational techniques for capturing, processing, managing, structuring, and ultimately analyzing data describing movement phenomena both in geographic and abstract spaces (Laube, 2014). The main goal is to achieve a better understanding of the processes governing a movement.

Trajectory data mining became a relevant area of research in this context. It is usually subdivided into specific procedures to transform raw location data into useful knowledge. Many authors have provided definitions for elements of trajectory data mining frameworks. Laube (2014) delimited movement mining into four groups: (i) segmentation and filtering, (ii) similarity and clustering, (iii) movement patterns and (iv) exploratory analysis and visualization. Zheng (2015) described a more detailed framework containing elements such as pre-processing, indexing and retrieval,

* Corresponding author.
E-mail addresses: tales.paiva-nogueira@univ-grenoble-alpes.fr (T.P. Nogueira), reinaldo.braga@ifce.edu.br (R.B. Braga), carina.oliveira@ifce.edu.br (C.T. de Oliveira), herve.martin@univ-grenoble-alpes.fr (H. Martin).

https://doi.org/10.1016/j.eswa.2017.10.004
0957-4174/

uncertainty, pattern mining, outlier/anomaly detection, classification, as well as ways of modeling trajectories. de Vries and van Someren (2012) handled some of these problems in practice by compressing, clustering, classifying and detecting outliers of vessel trajectories. More recently, Pan, He, Wang, Xiong, and Peng (2016) proposed a multidimensional trajectory clustering algorithm to mine regular behaviors based on characteristics such as type, position, velocity, and course.

In parallel, the Semantic Web environment has experienced relevant growth in the last few years. The adoption of Semantic Web technologies can change the way that information is made available on the Internet. This shift from a Web of documents towards a Web of data allows machines to understand and reason about this data. In this scenario, smart applications can exploit the connections among dataset's entities and their interactions with other datasets.

Semantic Web technologies offer powerful representation and reasoning features for a wide range of applications. The Semantic Web standards and tools are capable of tackle issues related to data and knowledge modeling, querying, reasoning, service discovery, privacy and provenance (Ye et al., 2015). The combination of these technologies can be useful, for instance, to intelligent data analysis tasks, as demonstrated by Roda and Musulin (2014), whose work builds an ontological framework capable of representing and reasoning over temporal data.

Undoubtedly, the growth of portable device's location capabilities has facilitated the development of intelligent mobile systems for the acquisition of trajectory data. The convergence of these technologies with Semantic Web standards allows the easy acquisition, interlinking, and annotation of information about trajectories.

In the context of trajectory data mining, this work positions itself in the earlier stages of the process. In other words, we propose solutions tailored for the modeling, pre-processing and annotation phases of a trajectory data mining pipeline. We define a high-level data model to represent trajectory and context with multiple levels of granularity and different options for representing spatial and temporal extents, as well as for expressing quantitative and qualitative semantic descriptions.

Ontologies provide an expressive formalism for representing data. Moreover, there is an increasing availability of datasets containing location information (Patroumpas, Giannopoulos, & Athanasiou, 2014) that can be used to enrich trajectories. Our proposal is based on the Semantic Trajectory EPisodes (STEP) ontology (Nogueira & Martin, 2015). More specifically, we evolve STEP to cover a wider range of requirements. Among the advantages of choosing a Semantic Web approach with ontologies for modeling and implementing our solution, we can highlight the possibility of integrating different data sources through federated queries, a transparent data model, the support to reasoning by inference, and the integration of data and metadata.

The contributions of this paper are complemented by the proposal of a framework based on the STEP ontology which serves as an interface between applications and our data model. Additionally, we present an example of an annotation algorithm that uses our framework and Semantic Web resources to enrich movement data. Also, we introduce a high-level data model to represent trajectory episodes and contextual elements with multiple levels of granularity and different options of representation of spatiotemporal extents as well as the ability to express quantitative and qualitative semantic descriptions.

The remainder of this paper is organized as follows: Section 2 gives an overview of trajectory ontologies and conceptual models. Section 3 presents the improvements on top of the STEP ontology and its main differences compared to relevant work. Section 4 presents the main components of FrameSTEP and Section 5 shows examples of its usage. Finally, Section 6 contains our conclusions about this work.

2. Trajectory models

The term Semantic Trajectory has been used for referring to models that can be used to enrich trajectories beyond latitude, longitude and timestamp information. Semantic Trajectories do not necessarily refer to the Semantic Web. However, there is an intersection in some cases where Semantic Web tools are used to model trajectories.

One of the most influential works in trajectory modeling was presented by Spaccapietra et al. (2008). In this work, the authors elicit some requirements for trajectory modeling. First, three types of trajectories are identified: (i) metaphorical trajectory, consisting in changes of a time varying attribute of an entity (e.g., the career trajectory of a person or the price for a stock); (ii) naive trajectory, consisting in geographic trajectories that are not defined in terms of spatial coordinates, and (iii) spatiotemporal trajectory, consisting in changes of position in a 2D or 3D space of an object. Then, two facets of trajectories are defined: the *geometric facet* – the spatiotemporal recording of a travelling point's position – and the *semantic facet* – the information that conveys the application-oriented meaning of the trajectory and its related characteristics.

Spaccapietra et al. (2008) also give a generic characterization of trajectories semantics. The main conceptualization of their work consists in defining trajectories as a sequence of stops and moves. The authors summarize that a conceptual trajectory model should cover the characterization of trajectories and their components with attributes, semantic constraints, topological constraints, and links to application objects.

2.1. Semantic Trajectory ontologies

The Stop/Move model has then inspired works that translated similar ideas into ontologies. Baglioni, Macedo, Renso, and Wachowicz (2008), for instance, proposed an ontological approach for semantic modeling and reasoning on trajectories. They highlighted the need for transforming raw traces into high-level representations, i.e. semantic trajectories. The authors have formulated OWL-DL (Web Ontology Language-Description Logics) axioms for identifying malicious behavior in a recreational game activity based on the duration and location of stops.

The same strategy was followed by Baglioni, Macedo, Renso, Trasarti, and Wachowicz (2009) to specify touristic activities. However, they aimed to detect frequent patterns. Similarly, Renso, Baglioni, Macedo, Trasarti, and Wachowicz (2013) proposed a slightly modified version of this ontology as a core model containing the entities *SyntacticTrajectory, Stop, Move, Time, Place*, and *Pattern* to represent human behavior. Behaviors in these cases are predefined sequences of stops manually annotated and encoded as ontological axioms, e.g. Home to Work behavior.

Yan, Macedo, Parent, and Spaccapietra (2008) present a framework for structuring, modeling, and querying trajectory data. They present an ontological infrastructure with three modules: geometric, geographic, and application domain ontologies. The Geometric Trajectory Ontology is application-independent and encloses four sub-ontologies: *Spatial, Temporal, Spatiotemporal*, and *Trajectory* ontologies. Spatial and temporal concepts are reused from the Modeling of Application Data with Spatio-temporal features (MADS) specifications (Parent, Spaccapietra, & Zimányi, 2006), which define well-known simple concepts of these two domains (*Point, Line, Interval, Instant*) and are equivalent to other spatial and temporal ontologies that we use in our work. The *Spatiotemporal* ontology defines pairwise combinations of spatial and temporal concepts (e.g.,