# Moving objects beyond raw and semantic trajectories

Maria Luisa Damiani
University of Milan, I
damiani@di.unimi.it

Ralf Hartmut Güting
FernUniversität Hagen, D
rhg@fernuni-hagen.de

Fabio Valdés
FernUniversität Hagen, D
fabio.valdes@fernuni-hagen.de

Hamza Issa
University of Milan, I
issa@di.unimi.it

## ABSTRACT

Mobile applications, for example for road traffic monitoring, mobile health and animal data ecology, call for methods enabling rich and expressive representation of moving objects. This demand motivates the increasing concern for the paradigm of semantic trajectories. In this paper, I overview related research, focusing in particular on the novel data model of *symbolic trajectories* proposed for the efficient and flexible handling of semantics-aware trajectories through a Moving Object DBMS.

## 1. INTRODUCTION

Semantic trajectories is a relatively recent paradigm developed to provide applications with knowledge about the movement of moving entities. The key idea is to supplement the raw mobility data (i.e. raw trajectories in the following) - typically sequences of GPS points - with contextual data [4]. For example, semantic trajectories can be used to describe the sequence of points of interest visited by tourists in a city, or the sequence of transportation means used by an individual to reach the working place from home. Basically a semantic trajectory consists of a raw trajectory augmented with annotations regarding the whole trajectory or parts of it. Probably because of its simplicity and naturalness, the concept of semantic trajectory has attracted the interest of numerous researchers over the last years. Current research develops along diverse streams including: ontology/conceptual modeling, mobility pattern mining for the generation of semantic annotations, semantic location privacy, and - more recently - the connection with the theories of complex networks and social analysis. The main results achieved so far are nicely summarized in the survey paper [4].

Somewhat surprisingly, one aspect that is largely ignored by the most recent literature regards the data management dimension of semantic trajectories. Put simply: how can we store and access semantic trajectories? How can we represent semantic trajectories through a rigorous data model? How can semantic trajectories interplay with raw trajectories and conventional data? These questions have been only marginally addressed. In fact no operational system enabling the management of semantic trajectories in real applications exists. We believe that this is a critical limitation especially in the light of the increasing availability of *big* raw trajectory data collected from mobile applications (e.g. LBS) that creates challenging opportunities for the application of this concept.

The research that we have undertaken in the context of the European initiative Cost Action MOVE[1] aims to fill this gap. Indeed the goal is not simply to take some existing definition of semantic trajectory and find the best way for implementing it on a DBMS, but rather to re-think of the notion of *semantically meaningful movement* while targeting the specification of a general, formal and operational framework. We imagine that in the long run this research could lead to the development of a novel class of software platforms for mobility data handling. The users of these systems will be able to organize and analyze mobility data in the same way that users now organize and analyze spatial data in a conventional GIS platform, e.g. Quantum GIS, or using one of the more recent platforms on cloud, e.g. GISCloud. While the idea in itself may sound not particularly innovative, just a restyling of GIS, we believe that these platforms, going beyond the notion of Moving Object DBMS, can greatly facilitate the development of novel and challenging applications. In what follows, the notion of semantic trajectory is presented; next the concept of symbolic trajectory is introduced along with the results achieved so far and major open issues.

## 2. SEMANTIC TRAJECTORIES

Early work on semantic trajectories was triggered by the experimental analysis of a set of raw trajectories about a group of birds [5]. By using the standard functionalities of a GIS, we found that the sequences of points, just pairs

[1] http://move-cost.info/

of timestamped coordinates, associated with birds identifiers were actually representing the migration routes from Central Europe to Africa and vice versa. Such discovery, that was somewhat unexpected, inspired the proposal of a novel model for the high level representation of movement. Since this first result, research developed along different directions, including the following:

- *Conceptual modeling.* The first conceptualization was centered on the notions of stop and move [5]. A stop represents a temporary suspension of the movement, while a move is the transfer from one stop to another stop. While this conceptualization is appropriate in many applications, there is increasing evidence that stop-and-move is just one of the possible mobility patterns. For example Yan et al. [7] present an approach to extract and represent the sequence of activities from raw trajectories. In the light of these experiences, a novel conceptual model has been recently proposed which enables the attachment of any kind of meaning (not just stop and move) to sequences of points [4].

- *Extraction of mobility patterns.* A major research direction regards the mining of mobility patterns to automatically annotate semantic trajectories. Early work by Alvares et al. [1] focuses on the identification of stops and moves. Numerous approaches can be found in literature, either explicitly related to the notion of stop-and-move or developed within different communities. A comprehensive survey can be found in [4].

- *The privacy of mobility patterns.* A different issue is to preserve the privacy of sensitive mobility patterns such as the presence in places, e.g. hospitals and religious buildings, that might reveal sensitive information about moving individuals. This problem is particularly challenging in on-line applications, e.g. LBS and geo-social networks, whereas the privacy mechanism has to rely on partial knowledge of the movement (past and current positions are known, but not future positions). The privacy of mobility patterns in an open issue [2]. An approach in this direction, focused on the protection of specific mobility pattern, i.e. sensitive places, is presented in [8].

# 3. SYMBOLIC TRAJECTORIES

Semantic trajectories are often considered the result of an analytical process conducted on raw trajectories. We believe that the notion of semantic trajectory is valuable on its own, independently of how these trajectories are generated. For example, annotations can be deliberately attached by individuals (e.g. user can specify the transportation means) or even the annotation can be automatically attached by the location tracking system (e.g. locations in indoor settings have natural semantics, such as room 1 and building A). Moreover, even in those cases in which semantic trajectories are obtained from an analytical process, the problem remains of how to encode them in a machine readable form. This is the focus of our current research that we briefly present in what follows.

## 3.1 The data model

We have defined a simple generic data model able to capture different types of semantics called symbolic trajectory [6]. In essence the idea is to represent semantic information in terms of names or labels. For example an activity (running, walking ) and points of interest (Colosseum, Louvre) can be straightforwardly described by labels while sensor readings, e.g. temperature, need first to be turned into qualitative values, e.g. high, low temperature. Formally, a symbolic trajectory is an ordered sequence of pairs

$$(i_1\ l_1), ..(i_n\ l_n)$$

called *units* when each unit $u_j = (i_j\ l_j)$ consists of a time interval $i_j$ and a label $l_j$. The label $l_j$ describes the movement in the time interval $i_j$. Symbolic trajectories are provided as abstract data types and integrated into the ADT model defined in [3]. For example a symbolic trajectory describing places and the transportation means used to reach those places, can be as follows:

```
(2013-01-17-9:02:30 2013-01-17-9:05:51) "home")
(2013-01-17-9:05:51 2013-01-17-9:08:44) "bus")
(2013-01-17-9:08:44 2013-01-17-9:50:02) "train")
(2013-01-17-9:50:02 2013-01-17-17:50:02) "work")
....
```

The core technical contribution is a novel language for pattern matching and rewriting on symbolic trajectories. The pattern language enables the extraction of subsequences from symbolic trajectories. Patterns are defined as regular expressions that can be matched by single units or sequences of units. For example, the query: *Which are the trajectories in which the individuals take more than 1 hour to move from home to work?* can be solved specifying the following pattern:

```
*(_ home ) Z* (_ work)*// getDuration(Z.time)> 3600
```

where:

- $Z$ is a variable denoting a sequence of units, the symbol $*$ denotes a sequence of zero or more units,

- $(\_\ home)Z * (\_\ work)$ is the pattern

- $getDuration(X.time) > 3600$ is the condition that must be met by the matching sequences, in this case the duration in seconds of the transfer from home to work.

An important feature of the language is that it is embedded into an existing Moving Object DBMS (i.e. Secondo). The pattern language at work is illustrated in a video[2].

# 4. CONCLUDING REMARKS

Capturing and representing the meaning of movement is a challenging issue that calls for novel solutions. We are working on the definition of the symbolic trajectory data model for the representation of time-varying textual descriptions. A number of issues are still open. For example, a major issue is integrating - whenever it is meaningful - the symbolic dimension with the geometric dimension of the movement. Another major issue regards the usability of the system that is fundamental for an effective deployment of symbolic trajectories in real applications.

[2]http://molle.fernuni-hagen.de/DfnA/SymbolicTrajectories.mp4

## 5. REFERENCES


[1] L. Alvares, V. Bogorny, B. Kuijpers, B. de Macedo, J.and Moelans, and A. Vaisman. A model for enriching trajectories with semantic geographical information. In *Proc. ACM GIS*, GIS '07, 2007.

[2] M.L. Damiani. *European Data Protection : Coming of Age ?*, chapter Privacy enhancing techniques for the protection of mobility patterns in LBS: research issues and trends. Springer, 2013.

[3] R. H. Güting, M. Böhlen, M. Erwig, C. Jensen, N. Lorentzos, M. Schneider, and M.Vazirgiannis. A foundation for representing and querying moving objects. *ACM Trans. Database Syst.*, 25(1):1–42, 2000.

[4] C. Parent, S. Spaccapietra, C. Renso, G. Andrienko, N. Andrienko, V. Bogorny, M.L. Damiani, A. Gkoulalas-Divanis, J. Macedo, N. Pelekis, Y. Theodoridis, and Z. Yan. Semantic trajectories modeling and analysis. *ACM Comput. Surv.*, 45(4):42:1–42:32, Aug. 2013.

[5] S. Spaccapietra, C. Parent, M.L. Damiani, J. de Macedo, F. Porto, and C. Vangenot. A conceptual view on trajectories. *Data Knowl. Eng.*, 65(1):126–146, 2008.

[6] F. Valdés, M.L. Damiani, and R. Güting. Symbolic trajectories in secondo: Pattern matching and rewriting. In *DASFAA*, 2013.

[7] Z. Yan, D. Chakraborty, C. Parent, S. Spaccapietra, and K. Aberer. Semitri: a framework for semantic annotation of heterogeneous trajectories. In *Proc. EDBT*, 2011.

[8] E. Yigitoglu, M.L. Damiani, O. Abul, and C. Silvestri. Privacy-preserving sharing of sensitive semantic locations under road-network constraints. In *IEEE MDM*, 2012.