Creating Moving Objects Representations for Spatiotemporal Databases

D

José Moreira
University of Aveiro, Portugal

Paulo Dias
University of Aveiro, Portugal

Luís Paulo
University of Aveiro, Portugal

INTRODUCTION

Modern sensor and communication technologies have facilitated the generation of geo-referenced data streams and have fostered the development of a broad spectrum of novel applications requiring efficient tools to deal with spatiotemporal data. In recent years, there were many research works on the development of spatiotemporal methods for dealing with moving points, that is, moving objects whose size and shape do not need to be represented in information systems. These works cover data models and query languages to represent networked and non-networked moving points in real-time systems or in historical systems for data storage and analysis, data structures and access methods for efficient implementation of spatiotemporal queries, benchmarking, data mining and visualization.

In contrast, the developments of fully operational systems to deal with moving objects with extent, spatial (geographical) entities that may change location, size, or shape continuously over time, has received minor attention. Existing proposals on data models and query languages for the representation and management of moving objects with extent were not supported by the development of adequate data acquisition systems. Though, this is a crucial issue to carry out a consistent evaluation of the expressivity and the performance of the spatiotemporal data models, query languages and algorithms proposed in databases literature.

This article introduces the fundamental concepts of spatiotemporal data models and presents a framework for acquisition and representation of data regarding location and shape of moving objects in spatiotemporal databases. Examples of data sources include satellite images tracking the location of icebergs or ocean's oil spills, aerial images of forest fires or other images of moving areas in raster format. The emphasis is on developing methods to estimate the transformation of a moving object between two consecutive observations. The output is a representation that includes the geometry and the topology of the moving object, and the movement analytical formulas describing the transformation. The approach was evaluated using synthetic and real data sets. The results show that it is possible to obtain adequate continuous representations from a sequence of observations. They also put in evidence issues on the quality of data representations that were not mentioned in previous literature and need further investigation.

BACKGROUND

The Database Management Systems (DBMS) provide standard methods for storing, managing and querying data efficiently. They have been initially tailored for handling one-dimensional data, such as, strings, dates or numbers, but the success of the relational model and the widespread use of DBMS in many organizations have led to the development of new solutions for dealing with more complex data types. This is the case of spatial DBMS extensions that focus on the representation of geographical objects or phenomena, which can be modeled using abstractions such as points, lines or polygons. Complexity increases because spatial data is usually represented in two or higher dimensional

DOI: 10.4018/978-1-4666-5888-2.ch163

spaces and the topology of spatial entities need to be considered as well. *Oracle Spatial* and *PostGIS* are notable examples of spatial DBMS extensions currently in use.

In addition, spatial data may also be dynamic and the spatiotemporal behavior of geographical objects or entities across time is essential for several applications. This has gained particular importance with recent developments in remote sensing and communication technologies. One of the most challenging topic is the representation of continuously changing spatial data about moving objects with extent, hereafter referred to simply as moving objects.

The abstract definition of a moving object is a triple $\left(\mathcal{I},\mathcal{S},\mathcal{V}\right)$ where $\mathcal{I}\subset\mathbb{R}$ is a temporal value, $\mathcal{S}\subset\mathbb{R}^2$ is the moving object's geometry at a certain time and $\mathcal{V}:\mathbb{R}^2\times\mathbb{R}\rightarrow\mathbb{R}^2$ is a continuous function defining the transformation of $\mathcal{S}$ during $\mathcal{I}$ (Chomicki, Haesevoets, Kuijpers, & Revesz, 2003). The semantics is

$$m=\left\{\begin{matrix}\left(x,y,t\right)\in\mathbb{R}^2\times\mathbb{R}\mid\left(\exists x'\right)\left(\exists y'\right)\\ \left(\left(x',y'\right)\in\mathcal{S}\wedge t\in\mathcal{I}\wedge\left(x,y\right)=\mathcal{V}(x',y',t)\right)\end{matrix}\right\},$$

where x and y are coordinates, t is a time instant and $\mathcal{V}$ is a continuous function that describes how the object moves or changes during time. This is an abstract representation of moving objects where the spatial and the temporal values are defined as infinite set of points, which needs to be translated into discrete (finite) data models suitable for implementation in databases. The main spatiotemporal data models were based on constraints *(Grumbach, Rigaux, & Segoufin, 2001) and* Abstract Data Types (ADT) *(Güting, et al., 2000).* Subsequent works have mainly followed the ADT approach because this data model can be smoothly built into extensible DBMS such as the object-relational DBMS.

The ADT approach represents the spatiotemporal behavior of a moving object as a sequence of motion units $\mathcal{M}=\left\{\left(\mathcal{I},\mathcal{S},\mathcal{V}\right)\right\}$, such that $\mathcal{I}$ is a time interval $\left[t_b,t_e\right[$, $\mathcal{S}$ denotes the geometry of the moving object at t_b and $\mathcal{V}$ is a variability function describing the movement of each vertex in $\mathcal{S}$ during $\mathcal{I}$. $\mathcal{S}$ can be a point, a line, a polygon, or a complex geometry such a polygon with holes, a multi-line or a multi-polygon. The time intervals in $\mathcal{M}$ must be disjoint, the topology of $\mathcal{S}$ must be valid for every time instant in $\mathcal{I}$ and the geometries at the end and at beginning of two temporally connected motion units must correspond.

This data model was fully implemented on the top of Secondo which is a prototype DBMS for research and teaching *(Almeida, Güting, & Behr, 2006).* There also are partial implementations of spatiotemporal extensions based on ADT for Object-Relational DBMS currently in commercial use, namely Hermes *(Pelekis, Theodoridis, Vosinakis, & Panayiotopoulos, 2006),* STOC *(Jin & Sun, 2008; Zhao, Jin, Zhang, Wang, & Lin, 2011)* and the spatiotemporal extension proposed *by Matos, Moreira, and Carvalho (2012).* All these extensions were built into Oracle.

Figure 1 depicts two motion units describing the spatial transformation of a moving object during a certain time.

The first motion unit represents the spatial transformation of a triangle into a pentagon. As the number of vertices is different, two additional vertices were added to the initial triangle. In this case, the choice was to clone the two vertices represented as white dots. The second motion unit represents the transformation of a pentagon into a quadrilateral. In this case, there is a mapping where two vertices converge into a single one. The movement of the vertices is described by linear functions and this transformation can be represented as shown in Box 1.

The values $\mathcal{I}_i$ are timestamps; each $\mathcal{S}_i$ is a set of vertices represented as 2D coordinates; and each $\mathcal{V}_i$ is a set of 2D velocity vectors. The upper bound of the last time interval is not defined, meaning that the last observation was taken at time instant *7*. This and the value of $\mathcal{V}_3$ must be filled when a new observation arrives. For a complete example on how to store and query moving objects data in a spatiotemporal database, please refer to *Matos, Moreira, and Carvalho (2012).*

One of the main issues that arise in this context is the creation of the continuous data representations from observations, such as, a sequence of satellite or aerial images, using automatic or semi-automatic procedures. This requires the implementation of procedures for segmentation of objects from the images, which is a subject that has been widely studied in the image processing research, but it also raises novel issues