Dynamic Modeling of Trajectory Patterns using Data Mining and Reverse Engineering

Luis Otavio Alvares[1,2] **Vania Bogorny**[1] **Jose Antonio Fernandes de Macedo**[3]
Bart Moelans[1] **Stefano Spaccapietra**[3]

[1] Theoretical Computer Science Group, Hasselt University, Belgium,
Email: {vania.bogorny,bart.moelans}@uhasselt.be

[2] Instituto de Informatica – Universidade Federal do Rio Grande do Sul, Brazil,
Email: alvares@inf.ufrgs.br

[3] Ecole Polytechnique Fédérale de Lausanne, Switzerland,
Email: {jose.macedo,stefano.spaccapietra}@epfl.ch

Abstract

The constant increase of moving object data imposes the need for modeling, processing, and mining trajectories, in order to find and understand the patterns behind these data. Existing works have mainly focused on the geometric properties of trajectories, while the semantics and the background geographic information has rarely been addressed. We claim that meaningful patterns can only be extracted from trajectories if the geographic space where trajectories are located is considered. In this paper we propose a reverse engineering framework for mining and modeling semantic trajectory patterns. Since trajectory patterns are data dependent, they may not be modeled in conceptual geographic database schemas before they are known. Therefore, we apply data mining to extract general trajectory patterns, and through a new kind of relationships, we model these patterns in the geographic database schema. A case study shows the power of the framework for modeling semantic trajectory patterns in the geographic space.

Keywords: trajectory data mining, trajectory data modeling, moving object data modeling, trajectory patterns, pattern visualization

1 Introduction

Trajectories have been generally considered as the path followed by an object moving in space and time (Wolfson et al. 1998, Güting et al. 2006). Each point in this path represents one position in space and one instant in time. Typically, trajectory data are obtained from mobile devices that capture the position of an object at specific time intervals. The background geographic information on which objects are moving is not captured by these devices, but is of fundamental importance for the analysis of trajectory data in real applications. Therefore, there is an increasing necessity for a more meaningful representation of trajectories, as well as their relationships with the geographic space. An example which expresses such necessity is shown in Fig. 1. In Fig. 1 (left) we can visualize a set of trajectories, from which not

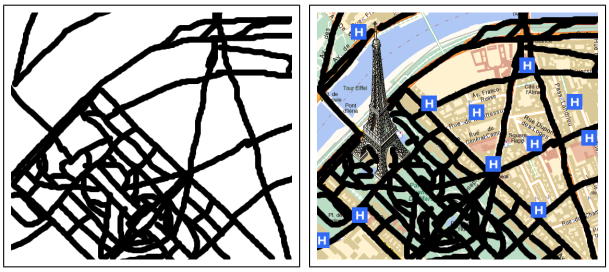

Figure 1: (left) trajectories and (right) trajectories with geographic information

much information can be extracted. In Fig. 1 (right) we have the same trajectories over the geographic space, where we can visually infer information like the geographic location (Paris) and the intersection of trajectories with the Eiffel tower and hotels.

We claim that in many application domains useful information can only be extracted from moving object data if their meaning as well as the background information is considered. Therefore, in this paper we will define trajectories from a semantic point of view (Spaccapietra et al. 2007), where trajectories are represented as a set of stops and moves.

The knowledge of moving patterns between different places in the geographic space (e.g. school to shopping, touristic place A to touristic place B) may help the user to answer queries about moving objects or movement behavior. In order to capture and model such pattern relationships, data mining techniques play an essential role.

On the one hand, data mining techniques have the objective to extract novel, useful, non-trivial, and previously unknown patterns/relationships from data (Fayyad et al. 1996). Conceptual data modeling on the other hand, has the objective to specify patterns (relationships) that are well known, and are normally specified in order to warrant the integrity of the data.

Figure 2 shows part of a geographic database schema where the relationships between gas stations and streets, streets and counties, as well as water resources and counties are well known. This kind of relationships, when considered in association rule mining, for instance, produce rules with a 100% confidence and generate only well known patterns (Bogorny et al. 2007), such as

$$intersects\_GasStation \Rightarrow intersects\_Street$$

While many relationships between spatial feature

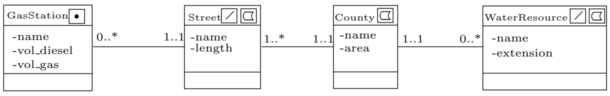

Figure 2: conceptual geographic database schema

types are well known by the geographic database designer, trajectory relationships are not. For instance, a move between gas station and water resource (e.g. lake, river) is not known a priori. Trajectory pattern relationships are hard to be visually identified by simply overlapping different layers of geographic information. In the example shown in Fig. 1 (right), it is difficult to visualize trajectory patterns. Therefore, we propose a reverse engineering approach, where we apply data mining to extract only strong and interesting patterns from trajectory data in order to provide a set of pattern relationships that are not known a priori. More specifically, we extract frequent trajectory movements by considering not only the trajectories, but also the whole geographic space in which trajectories were collected. Indeed, the discovered patterns are stored in the trajectory database and modeled in the geographic conceptual schema, in order to help the user to understand, visualize, and query trajectory pattern relationships.

1.1 Scope and Outline

The scope of this work is limited to the discovery of *binary trajectory pattern relationships* and their modeling in geographic database schemas. The focus relies on binary trajectory relationships for their representation in the schema, between two spatial feature types. In this paper we are interested in frequent movements independently of the exact time in which these movements are frequent. According to the application, time can be a filter applied in data preprocessing steps by selecting trajectories that satisfy given time intervals.

The remaining of the paper is organized as follows: Sect. 2 presents the related works and the main contributions. In Sect. 3 we introduce the basic concepts about trajectories, database schemas, and moving patterns. In Sect. 4 we present a framework to extract trajectory patterns from data as well as their modeling in geographic database schemas. In Sect. 5 we present a case study over trajectories of conference attendees to show the application and the trade off of our framework. In Sect. 6 we discuss our approach for modeling trajectory patterns and in Sect. 7 we conclude the paper and present some directions of future work.

2 Related Work and Contributions

Several data models have been proposed for efficiently querying moving objects (Wolfson et al. 1998, Güting et al. 2006, Brakatsoulas et al. 2004, Mouza & Rigaux 2005). In (Wolfson et al. 1998) the main focus relies on the geometric properties of trajectories, while both (Brakatsoulas et al. 2004) and (Mouza & Rigaux 2005) considered semantics and background geographic information.

In (Mouza & Rigaux 2005) moving patterns are extracted from data by defining the patterns a priori. For instance, find all trajectories that move from a zone A to a zone B and cross zone Z. In that work, moving patterns are the trajectories that follow a given pattern. In our work we will extract moving patterns from data which are not known a priori and which are frequent in a minimal number of trajectories.

In (Brakatsoulas et al. 2004) a semantic model for trajectories has been proposed as well as relationships of trajectories with the background geographic information or environment information. This model, however, is restricted to a specific application domain, and trajectory relationships are related to vehicles and roads.

Spaccapietra (Spaccapietra et al. 2007) proposed a general model for semantic trajectories, and introduced the concept of stops and moves, which we have adopted in this paper.

From the data mining perspective many trajectory pattern mining algorithms have been developed, such as (Tsoukatos & Gunopulos 2001, Laube et al. 2005, Verhein & Chawla 2006, Cao et al. 2006, Gudmundsson & van Kreveld 2006). Their main drawback, however, is that the semantics of the trajectory and the geographic information behind trajectories has not been considered. Some approaches find dense patterns, where moving objects are in the same region and move in the same direction (Cao et al. 2006). For semantic trajectory patterns we are interested in frequent movements between places that may be spatially sparse (e.g. from airport to touristic place).

Other approaches find long patterns (Tsoukatos & Gunopulos 2001, Gudmundsson & van Kreveld 2006), while we are interested in short patterns between two places, similarly to the approach presented by (Verhein & Chawla 2006), where trajectory association patterns are extracted between two regions.

Since existing trajectory data mining approaches do not consider the semantics of the data, it is difficult or even impossible to understand and model trajectory patterns from the conceptual point of view. It becomes possible only when some background geographic information is considered, as in (Güting et al. 2006), where trajectories are integrated with road networks.

Recently we have proposed a novel data mining approach which uses conceptual schemas to improve spatial association rule mining (Bogorny et al. 2006, 2007). More specifically, to avoid the generation of patterns that are a priori known as non-interesting. In this paper we propose an approach that takes the opposite direction. Since we are interested in the discovery and modeling of new patterns, and which are not obvious to the database designer, we address the problem through a process of reverse engineering. In a first step we apply data mining techniques to extract patterns from trajectories considering the background geographic information. In a second step, we model the discovered patterns in geographic conceptual schemas.

Reverse engineering has been used to understand the data model in legacy systems (McKearney & Roberts 1996) and for automatic query extraction (Shoval & Shreiber 1993).

The main contributions of the work presented in this paper include: (i) the use of trajectory data from a semantic point of view, instead of a sequence of points in time; (ii) the extraction of frequent movements between *two* places, giving semantics to the discovered patterns as well as their visualization in the conceptual schema; (iii) the representation of movement patterns in geographic database schemas through a new kind of relationship, that will facilitate both querying and visualizing trajectory patterns over the geographic space.

3 Definitions and Preliminaries

In this section we present a brief overview on geographic conceptual schemas, introduce a semantic model for *trajectories* from the semantic point of view,

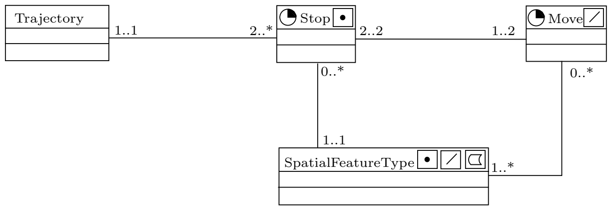

Figure 3: basic conceptual schema for trajectory

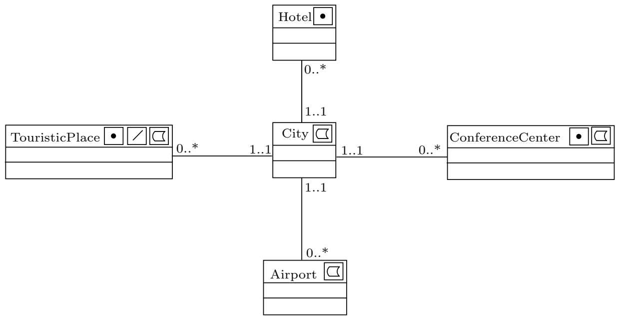

Figure 4: part of a geographic database conceptual schema

and introduce the new concept of *moving pattern* relationship.

3.1 Trajectories and Background Geographic Information

In order to extract moving patterns from trajectory data, we define a trajectory similar to the definition presented in (Spaccapietra et al. 2007).

Definition 1 (Trajectory) Semantically a trajectory T is an ordered list of stops S and moves M.

Definition 2 (Stop) A stop s is a semantically important part of a trajectory where is considered that the object has not effectively moved. A stop is represented by a spatial feature type in the geographic space and a non-empty time interval.

Definition 3 (Move) A move $\overrightarrow{s_1 s_2}$ is the part of a trajectory that has a time interval and is delimited by two consecutive stops s_1 and s_2, where consecutive stops by definition must have non-overlapping time intervals.

In our definitions, stops are interesting places specified according to the application. For instance, a traffic light may be considered a stop in a transportation management application, but probably not in a tourism application.

Definition 4 (Spatial Feature Type) A spatial feature type is a real world entity that has a location on the Earth surface (OGC 1999).

Figure 3 shows the trajectory data model that we adopt in this work. Notice that both stops and moves are related to a spatial feature type, where each stop is located in one spatial feature type and moves may have any spatial relationship (e.g. within, crosses) with one or more spatial feature types.

Spatial feature types are represented in geographic database schemas as different geographic object types. In the schema shown in Fig. 4 there are five different spatial feature types: *city, hotel, airport, conference center,* and *touristic place*, that will be used later in a case study.

Geographic database schemas are normally extended relational or object-oriented schemas (Shekhar & Chawla 2002). Some approaches such as (Parent et al. 2006, Borges et al. 2001, da Rocha et al. 2001) extend such models with pictograms to provide special treatment for geographic applications. In geographic conceptual data modeling, relationships among spatial and non-spatial data are represented through associations with cardinality constraints. In geographic database schemas, these associations may either represent a spatial relationship (e.g. touches, contains) or a non-spatial relationship such as association or aggregation.

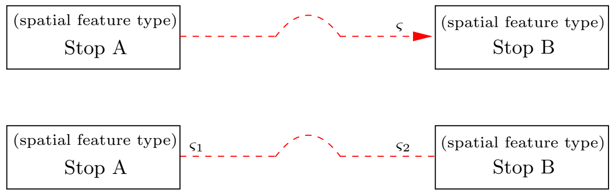

Figure 5: moving pattern structure

3.2 Moving Patterns (MP)

On the contrary to many trajectory pattern mining approaches that consider trajectory patterns as a set of ordered points that occur during the same time interval, we are interested in patterns of moves between two places. We will define a moving pattern as a frequent move between two stops in order to be able to represent such a pattern as a relationship between two spatial feature types in a geographic database schema.

Definition 5 (Support) Let $T = \{t_1, t_2, \ldots, t_n\}$ be a set of trajectories. Let $\overrightarrow{AB}$ be a move from a stop A to a consecutive stop B. We define the support of the move $\overrightarrow{AB}$, denoted by $sup(\overrightarrow{AB})$, as the fraction of trajectories t of T in which the move $\overrightarrow{AB}$ occurs. More formally,

$$sup(\overrightarrow{AB}) = \frac{\left|\{t \in T \mid \overrightarrow{AB} \in t\}\right|}{|T|}$$

where $|X|$ is the number of elements in the set X.

Definition 6 (Moving Pattern) A moving pattern is a move with support ς, where ς is higher than a given threshold, called *minsup*.

A moving pattern $\overrightarrow{AB}(\varsigma)$ is a relationship between two spatial feature types A and B which has two main properties: direction and support. The direction of a moving pattern $\overrightarrow{AB}$ is a path from A to B, in this order, and the support ς is the fraction of trajectories having this move. Considering that each stop is within a spatial feature type, a moving pattern relationship can be modeled as shown in Fig. 5.

The first pattern shown in Fig. 5 is unidirectional, from a spatial feature type A to a spatial feature type B with support ς. The second pattern is bidirectional, where the move from a spatial feature type A to a spatial feature type B has support ς_2 and the moving pattern from a spatial feature type B to a spatial feature type A has support ς_1.

4 A Framework for Trajectory Pattern Mining and Modeling

In this section we present a framework to extract frequent trajectory patterns and their modeling in

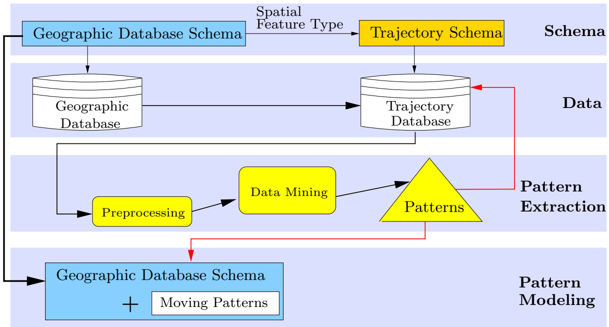

Figure 6: framework for Trajectory Pattern Extraction and Modeling

geographic database schemas. Figure 6 presents an overview of the framework, that can be analyzed in four levels: schema, data, pattern extraction, and pattern modeling.

At the schema level there is the trajectory schema, where trajectories are seen as an ordered set of stops and moves. At this level, there is also a geographic database schema that contains the background geographic information of the same region of trajectories, i.e., the spatial feature types on which objects are moving.

The data level has both trajectory data and geographic data, where stops are within spatial feature types defined in the geographic database.

At the pattern extraction level data are preprocessed and transformed to the input format required by the data mining algorithm. In the data mining step the user may define mining parameters such as minimum support, in order to extract only patterns that pass these constraints. Trajectory pattern relationships are identified in this step, which will be explained in more detail in Sect. 4.1.

Since the new knowledge is identified after the mining process, the user may choose the new relationships to be represented in the geographic database schema. For instance, only patterns with support higher than 50%. Details of this step are presented in Sect. 4.2.

4.1 From Data to Patterns

Different kinds of patterns can be discovered considering stops and moves in trajectories. For instance, (i) the most frequent stops during a certain period of time, (ii) frequent stops that have a duration higher than a given threshold, (iii) frequent moves at a certain time interval, (iv) most frequent moves inside a certain region, (v) frequent moves that intersect a given spatial feature type, and so on.

In this paper we have the objective to find one specific kind of patterns from trajectory data: *frequent moves between two stops*. For this purpose we take as an input of our pattern extraction algorithm the dataset of moves, where each move is represented as a pair of ordered stops among which an object has moved during a given time interval. For all moves, we count in how many trajectories each move appears. If a move is a moving pattern, i.e., has support higher than a given *minsup* threshold, then this move is considered frequent and is stored in a set of moving patterns in the trajectory database, with its respective support. Since all patterns have been generated and stored, the user may choose the most interesting moving patterns to be modeled in the geographic database schema.

The patterns extracted by our method are at a high level of abstraction (e.g. Hotel to TouristicPlace). This is one of the main strengths of our data mining method. Because of this abstraction we discover moves that are frequent in sparse geographic locations, while most trajectory data mining algorithms extract trajectory patterns in dense locations.

Figure 7: moving pattern relationship from Hotel to TouristicPlace

Table 1: moves sample dataset

Mid	Tid	stop 1	stop 2	time interval
1	1	Airport	Hotel	10:05–11:10
2	1	Hotel	TouristPlace	14:08–14:30
3	1	TouristPlace	Hotel	18:03–19:05
1	2	Hotel	ConferenceCenter	07:50–08:32
2	2	ConferenceCenter	TouristPlace	11:04–11:15
3	2	TouristPlace	Hotel	12:03–12:15
4	2	Hotel	TouristPlace	14:08–15:07
5	2	TouristPlace	Hotel	18:58–20:03
1	3	Hotel	ConferenceCenter	08:10–08:35
2	3	ConferenceCenter	Hotel	16:00–16:38
3	3	Hotel	TouristPlace	17:36–18:08
4	3	TouristPlace	Hotel	18:40–19:15
1	4	Hotel	ConferenceCenter	08:06–08:20
2	4	ConferenceCenter	Airport	17:03–18:20
1	5	Airport	ConferenceCenter	09:25–10:18
2	5	ConferenceCenter	Hotel	18:27–19:05
1	6	Hotel	TouristPlace	08:05–08:55
2	6	TouristPlace	TouristPlace	12:08–12:29
3	6	TouristPlace	Hotel	17:05–18:07
1	7	Hotel	ConferenceCenter	08:10–08:19
2	7	ConferenceCenter	TouristPlace	15:55–16:25
3	7	TouristPlace	Hotel	17:28–18:03
4	7	Hotel	Airport	21:03–22:05
1	8	Airport	Hotel	09:58–10:35
2	8	Hotel	TouristPlace	11:07–11:20
3	8	TouristPlace	TouristPlace	14:05–14:38
4	8	TouristPlace	Hotel	17:06–18:12
1	9	Hotel	ConferenceCenter	07:55–08:33
2	9	ConferenceCenter	Airport	18:25–19:38
1	10	Airport	ConferenceCenter	08:32–09:02
2	10	ConferenceCenter	TouristPlace	11:03–16:27
3	10	TouristPlace	Airport	17:28–18:15

4.2 From Patterns to Relationships: The Reverse Engineering Approach

Patterns extracted from trajectories express frequent moves over several trajectories. For instance, a pattern such as $\overrightarrow{Hotel\ TouristicPlace}$ (50%) characterizes a move from a Hotel to a Touristic Place, in this sequence, for 50% of the trajectories in the database. The representation of this pattern in the geographic database schema is shown in Fig. 7.

The expressive power of a geographic conceptual data model with moving pattern relationships provides a high abstraction level of trajectory patterns that cannot be identified through data visualization techniques, as shown in the example in Fig. 1 (right). For instance, a general moving pattern from Hotel to Touristic Place can only be identified by considering stops at a general granularity level, and not the specific instances (e.g. Ibis_Hotel, HollidayInn_Hotel, Eiffel_Tower, Louvre_Museum). This is possible because stops and moves in our trajectory model are related to *spatial feature types*. Because of data generalization we are able to match data mining patterns with conceptual modeling.

5 Case Study

In this section we present a case study over trajectories of conference attendees. There is a database of attendees's trajectories where one person may have many trajectories. The moves in these trajectories are shown in Table 1, where each row is a move of one trajectory. In this dataset we observe people arriving, going to the conference center, visiting touristic places, moving in different directions.

Table 2: moves and respective support

Move			Sup
Airport	,	Hotel	2/10
Hotel	,	TouristicPlace	5/10
TouristicPace	,	Hotel	6/10
Hotel	,	ConferenceCenter	5/10
ConferenceCenter	,	Hotel	2/10
Airport	,	ConferenceCenter	2/10
TouristicPlace	,	TouristicPlace	3/10
ConferenceCenter	,	Airport	2/10
ConferenceCenter	,	TouristicPlace	4/10
TouristicPlace	,	Airport	2/10
Hotel	,	Airport	1/10

Table 3: moving pattern with minimum support 50%

Move			Sup
Hotel	,	TouristicPlace	5/10
Hotel	,	ConferenceCenter	5/10
TouristicPlace	,	Hotel	6/10

Besides the trajectory data of visitors that will attend the conference, there is also a geographic database of the city where the conference takes place. This database has many spatial feature types including hotels, conference center, touristic points, and airport, which might be interesting to discover trajectory patterns. Part of the conceptual schema of this database is shown in Fig. 4. In this figure, well known relationships specified by the database designer are the associations of city with hotels, airports, convention center, and touristic places. No more relationships are known. Without mining the trajectory database we are not able to visualize frequent places visited by people attending the conference.

In our dataset shown in Table 1, stops are the spatial feature types which are represented in the geographic database of the city where the conference takes place. In (Alvares et al. 2007) we have presented an algorithm to integrate trajectories and the background geographic information, generation stops and moves.

In our case study we are only interested in frequent moves, and not in specific time intervals in which the conference attenders move into the city.

As can be observed in Table 1, the data are represented at general granularity level, which we call *feature type granularity*. In order to extract moving patterns, data need to be represented at a feature type granularity level (e.g. Hotel, TouristicPlace). At a feature instance granularity (e.g. Ibis_Hotel, Louvre_Museum), moves may not be frequent enough to find strong and general patterns. Nevertheless, it would not be possible to represent the discovered patterns in the conceptual model.

The frequent moves between two stops considering the dataset shown in Table 1, are shown in Table 2, with their respective support.

Considering the dataset shown in Table 1 and $minsup = 50\%$, we obtain the moving patterns shown in Table 3. The first pattern shows that 50% of the trajectories of people attending the conference have in their trajectory a move that goes from a hotel to a touristic place. The second pattern shows that 50% of all trajectories have at least one move in their trajectory that goes from a hotel to the conference center. In the last pattern we observe that 60% of the trajectory database has a move going from a touristic place to a hotel.

In such a dataset one could expect a moving pattern from the conference center to a touristic place. However, for this specific dataset and minimum support, this move in not frequent. Conference attenders

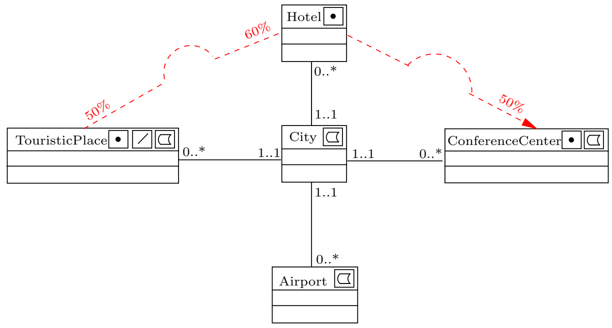

Figure 8: moving pattern relationships in geographic database schemas

in this dataset have a moving pattern that goes from hotel to touristic place instead.

The extracted moving patterns shown in Table 3 are modeled in the geographic conceptual database schema shown in Fig. 8. The representation of the patterns in the geographic database schema provides to the user the relationships of trajectory data with the background geographic information. Indeed, only strong relationships which pass the support constraint and were selected by the user as most relevant are represented.

A moving pattern relationship represented in the schema through a bidirectional association between *TouristicPlace* and *Hotel* with its exact support could not be modeled by the database designer without knowing the data.

6 Discussion

The use of data mining techniques to extract patterns from trajectory data has some general advantages. The storage of these patterns will help the user to efficiently write queries over moving object patterns. The representation of a moving pattern in the geographic database schema besides providing support for modeling trajectory pattern relationships, it helps the user to understand moving behavior in the context of the background geographic information. Trajectory pattern modeling can be semantically more expressive and powerful than visualization toolkits.

If we compare Fig. 1 (left) and the conceptual model with trajectory patterns shown in Fig. 8, it is easy to understand the importance of mining trajectories considering background geographic information and its semantics. From the visualization perspective we can observe that a conceptual schema is more powerful to express trajectory patterns at a high level of abstraction.

From the conceptual modeling point of view, the semantics of a moving pattern relationship is more expressive than a relationship with cardinality $0..n$ or $1..n$ used to represent static relationships. For instance, in relationships with cardinalities $0..n$ the probability for the relationship hold varies from zero to all instances of the related object type. For a relationship with a cardinality $1..n$ we know that the relationship always holds, but we do not how many times. A moving pattern relationship, however, although it is dynamic and may change when the database is updated, it specifies the exact percentage of trajectories that have a move between two specific stops. In Fig. 8, for instance, we know that exactly 50% of all trajectories in the database have a move from hotel to touristic place, even if we do not know how many hotels are in the database.

7 Conclusions

In this paper we presented a framework for mining and modeling moving patterns from a semantic point of view. The use of data mining methods to extract trajectory patterns has some general benefits: helps the database designer to model moving patterns that are not known a priori; provides a visual representation of the discovered patterns, what helps to understand moving behavior in a certain geographic space; and patterns can be stored and reused for future mining and querying of trajectory data.

Our framework is dynamic because new patterns can be extracted from trajectories when the database is updated. Such characteristic facilitates the analysis of the evolution of trajectory patterns and moving behavior over long time periods.

The discovery of semantic trajectory pattern relationships would neither be possible by simply mining raw moving object data nor by visualizing different layers of geographic information.

Future works include the evaluation of moving patterns for specific time intervals as well as the extraction of frequent relationships with different spatial feature types. For instance, during the movement of an object from a stop A to a stop B, this object may have many spatial relationships with different spatial feature types that might be interesting to extract and represent in the geographic conceptual schema (e.g. crosses river).

Acknowledgments

This research has been funded by the Brazilian agency CAPES, the European Union (FP6-IST-FET programme, Project n. FP6-14915, GeoPKDD: Geographic Privacy-Aware Knowledge Discovery and Delivery), and by the Research Foundation Flanders (FWO-Vlaanderen), Research Project G.0344.05.

References

Alvares, L. O., Bogorny, V., Kuijpers, B., de Macedo, J. A. F., Moelans, B. & Vaisman, A. (2007), A model for enriching trajectories with semantic geographical information, *in* 'ACM-GIS', ACM Press, New York, NY, USA.

Bogorny, V., Camargo, S., Engel, P. & Alvares, L. O. (2006), Mining frequent geographic patterns with knowledge constraints, *in* 'ACM-GIS', ACM Press, New York, NY, USA, pp. 139–146.

Bogorny, V., Kuijpers, B. & Alvares, L. O. (2007), 'Reducing uninteresting spatial association rules in geographic databases using background knowledge: a summary of results', *International Journal of Geographical Information Science* .

Borges, K. A. V., Davis, C. A. & Laender, A. H. F. (2001), 'Omt-g: An object-oriented data model for geographic applications', *GeoInformatica* **5**(3), 221–260.

Brakatsoulas, S., Pfoser, D. & Tryfona, N. (2004), Modeling, storing, and mining moving object databases, *in* 'IDEAS '04: Proceedings of the International Database Engineering and Applications Symposium (IDEAS'04)', IEEE Computer Society, Washington, DC, USA, pp. 68–77.

Cao, H., Mamoulis, N. & Cheung, D. W. (2006), Discovery of collocation episodes in spatiotemporal data, *in* 'ICDM', IEEE Computer Society, pp. 823–827.

da Rocha, L. V., Edelweiss, N. & Iochpe, C. (2001), Geoframe-t: A temporal conceptual framework for data modeling, *in* W. G. Aref, ed., 'ACM-GIS', ACM, pp. 124–129.

Fayyad, U. M., Piatetsky-Shapiro, G. & Smyth, P. (1996), 'From data mining to knowledge discovery in databases', *AI Magazine* **17**(3), 37–54.

Gudmundsson, J. & van Kreveld, M. (2006), Computing longest duration flocks in trajectory data, *in* 'GIS'06: Proceedings of the 14th annual ACM International Symposium on Advances in Geographic Information Systems', ACM Press, New York, NY, USA, pp. 35–42.

Güting, R. H., de Almeida, V. T. & Ding, Z. (2006), 'Modeling and querying moving objects in networks.', *VLDB J.* **15**(2), 165–190.

Laube, P., Imfeld, S. & Weibel, R. (2005), 'Discovering relative motion patterns in groups of moving point objects', *International Journal of Geographical Information Science* **19**(6), 639–668.

McKearney, S. & Roberts, H. (1996), Reverse engineering databases for knowledge discovery, *in* 'KDD', pp. 375–378.

Mouza, C. & Rigaux, P. (2005), 'Mobility patterns', *GeoInformatica* **9**, 297–319(23).

OGC (1999), 'Topic 5, opengis abstract specification - features (version 4) (1999)', Available at: http://www.OpenGIS.org/techno/specs.htm. Accessed in August (2005).

Parent, C., Spaccapietra, S. & Zimanyi, E. (2006), *Conceptual Modeling for Traditional and Spatio-Temporal Applications – The MADS Approach*, Springer.

Shekhar, S. & Chawla, S. (2002), *Spatial Databases: A Tour*, Prentice Hall.

Shoval, P. & Shreiber, N. (1993), 'Database reverse engineering: From the relational to the binary relationship model', *Data Knowl. Eng.* **10**, 293–315.

Spaccapietra, S., Parent, C., Damiani, M.-L., de Macedo, J. A. F., Porto, F. & Vangenot, C. (2007), A conceptual view on trajectories, Technical report, Ecole Polytechnique Federal de Lausanne.

Tsoukatos, I. & Gunopulos, D. (2001), Efficient mining of spatiotemporal patterns, *in* C. S. Jensen, M. Schneider, B. Seeger & V. J. Tsotras, eds, 'SSTD', Vol. 2121 of *Lecture Notes in Computer Science*, Springer, pp. 425–442.

Verhein, F. & Chawla, S. (2006), Mining spatio-temporal association rules, sources, sinks, stationary regions and thoroughfares in object mobility databases, *in* M.-L. Lee, K.-L. Tan & V. Wuwongse, eds, 'DASFAA', Vol. 3882 of *Lecture Notes in Computer Science*, Springer, pp. 187–201.

Wolfson, O., Xu, B., Chamberlain, S. & Jiang, L. (1998), Moving objects databases: Issues and solutions, *in* M. Rafanelli & M. Jarke, eds, 'SSDBM', IEEE Computer Society, pp. 111–122.