Automatic Construction and Multi-level Visualization of Semantic Trajectories (Demo Paper)

Zhixian Yan
EPFL, Switzerland
zhixian.yan@epfl.ch

Lazar Spremic
EPFL, Switzerland
lazar.spremic@epfl.ch

Dipanjan Chakraborty
IBM Research, India
cdipanjan@in.ibm.com

Christine Parent
University of Lausanne
christine.parent@unil.ch

Stefano Spaccapietra
EPFL, Switzerland
stefano.spaccapietra@epfl.ch

Karl Aberer
EPFL, Switzerland
karl.aberer@epfl.ch

ABSTRACT

With the prevalence of GPS-embedded mobile devices, enormous amounts of mobility data are being collected in the form of *trajectory* - a stream of (x,y,t) points. Such trajectories are of *heterogeneous* entities - vehicles, people, animals, parcels etc. Most applications primarily analyze raw trajectory data and extract *geometric* patterns. Real-life applications however, need a far more comprehensive, *semantic* representation of trajectories. This paper demonstrates the automatic construction and visualization capabilities of *SeMiTri* - a system we built that exploits 3rd party information sources containing geographic information, to semantically enrich trajectories. The *construction* stack encapsulates several spatio-temporal data integration and mining techniques to automatically compute and annotate *all meaningful parts* of heterogeneous trajectories. The *visualization* interface exhibits different levels of data abstraction, from low-level raw trajectories (i.e. the initial GPS trace) to high-level semantic trajectories (i.e. the sequence of interesting places where moving objects have passed and/or stayed).

Categories and Subject Descriptors

H.2.8 [**Database Management**]: Database Applications—*data mining, spatial databases and GIS, visualization, Web*

General Terms

Algorithms, Experiments, Demonstration

Keywords

Trajectory Construction, Trajectory Visualization

1. INTRODUCTION

Recent years have seen a tremendous surge in applications and services with location feeds. This is possible due to the large-scale embedding of GPS equipped mobile devices. Trajectories today have become ubiquitous data sources for analysis along several dimensions. Examples are behavioral analysis of wild-life, tracking, traffic analysis, analysis of urban landuse through people trajectories etc.

Several data mining techniques have been applied towards abstraction and discovery of interesting mobility patterns over such real-life GPS data, such as *clustering*, *classification*, *outlier detection*, *finding convoys* and *sequential rule-driven mining*. [1] provides an excellent review of these studies. Interesting prototypes such as *MoveMine* [2], *GeoLife* [5] have emerged from these works. These studies however, focus on raw trajectories (spatio-temporal records only), ignoring the context of background geographic information that carries significant semantic knowledge about movement. As a result, without the consideration of semantics that are available from background geographic data, it is still very difficult to have a complete interpretation of movement behaviors based on the extracted patterns from the raw mobility traces - only the geometric perspective.

Recently, *semantic trajectory analysis* receives significant attention for representing trajectories from such *semantic* viewpoint. E.g., [3] proposes a conceptual view on trajectories, modeling trajectory as a sequence of *stops* and *moves*. [4] designs a platform to progressively compute such *spatio-semantic trajectories*. The GeoPKDD (*Geographic Privacy-aware Knowledge Discovery and Delivery*) and MODAP (*Mobility, Data Mining, and Privacy*) projects emphasize the need to address semantic behaviors of moving objects.

This paper demonstrates a system - *SeMiTri* - that constructs and visualizes trajectories at different levels of data abstraction and semantics. Such semantics can be extracted from 3rd party data sources that contain geographic information. Our objective is to *automatically* construct *meaningful parts* of *heterogeneous* trajectories with additional semantics from relevant geographic sources. This is achieved by a set of *computation* and *annotation* algorithms that construct trajectories at different levels of abstraction. In addition, SeMiTri provides a Web interface for users to easily query and visualize multi-level trajectories.

2. SYSTEM ARCHITECTURE

Fig. 1 presents the detailed systematic architecture of the *SeMiTri* system. It follows a layered structure that progressively abstracts higher-level semantic trajectory concepts from lower-level raw GPS feeds.

ACM GIS '10, November 2-5, 2010. San Jose, CA, USA

- *Trajectory Computation Layer* - Performs preliminary works such as trajectory data cleaning (e.g. smoothing data and filtering outliers), dividing a trajectory into several episodes, and computing structural-level trajectories (e.g. a sequence of stops/moves [3]).
- *Semantic Annotation Layer* - Annotates trajectories with data of *semantic places* available from heterogeneous 3rd party geographic sources (e.g. OpenStreetMap). We design suitable annotation algorithms (e.g. *spatial join* for region, *map matching* for line and *hidden Markov model* for point), addressing challenges involved in modeling trajectories of different nature.
- *Semantic Trajectory Analytics Layer* - Computes additional statistical information (e.g. distribution values such as *mean, variance, max, min*) for trajectories.
- *Web Interface* - Presents users with a visual and integrative way to query and retrieve the enriched semantic trajectories at different abstracted levels.

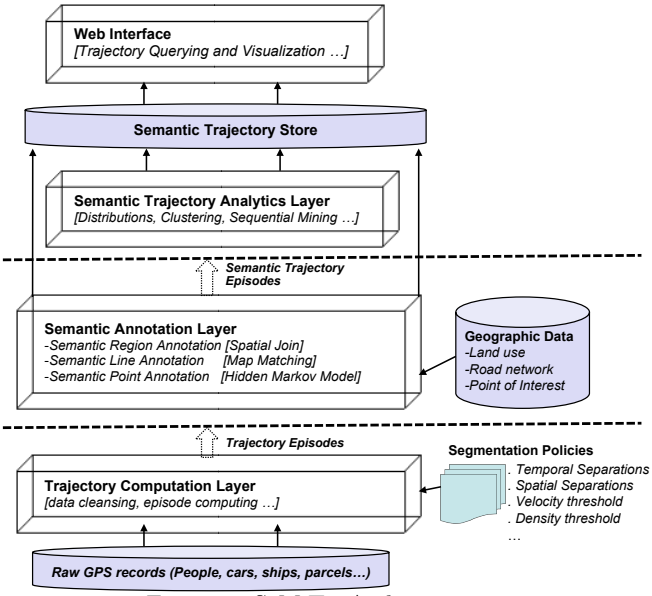

Figure 1: SeMiTri Architecture

In SeMiTri, trajectory computation and annotation algorithms are implemented in Java, and the Web interface is deployed on Apache Tomcat. Users access the system via a Web browser with the Google Earth Plugin.

3. VISUALIZATION CAPABILITIES

We have experimented with a large data store of heterogeneous trajectories (e.g. vehicle trajectories of 17243 taxies/cars and people trajectories of 6 mobile phone users) and constructed semantic abstractions of these trajectories using SeMiTri. The demonstration presents the query and visualization interface through a Web browser, and showcases the following capabilities of SeMiTri:

- *Spatio-Semantic Trajectories* - Demonstrates multiple levels of trajectory data abstraction, showing *raw GPS tracks*, *raw trajectories* (exploiting space/time gaps), *structured trajectories* (e.g. stops/moves), and *semantic trajectories* (e.g. home-office-supermarket-home).
- *Semantic Places* - Demonstrates diverse geographic resources that are associated to trajectories - *Landuse*, *Road network*, and *Point of interest* (POI) data.
- *User Interactions* - Provides a friendly Web interface to query and visualize trajectories (e.g. daily tracks) at different abstraction levels.
- *Analytics Results* - Highlights *statistical analytics* results of semantic trajectories, e.g. the average speed when user is moving, Landuse distribution where user has stopped, most frequent transportation modes, etc.

Fig. 2 shows SeMiTri's Web interface with a schematic example of a daily trajectory (at 30/3/2010) for a given user (id 205). The interface displays four parts: (a) the query input, (b) the visualization with Google Earth plugin, (c) check boxes allowing user to incrementally visualize different parts/levels of trajectories and semantic places they traverse, (d) aggregated statistics associated with the trajectories from different perspectives.

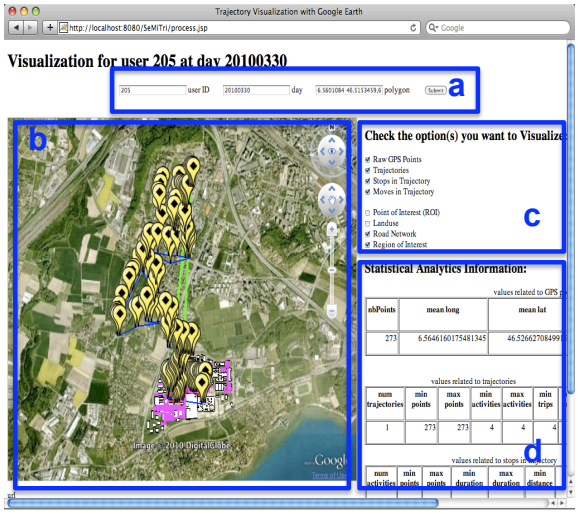

Figure 2: Web interface of SeMiTri

4. CONCLUSION

The paper demonstrated SeMiTri - a system for automatic construction and multi-level visualization of semantic trajectories from raw mobility traces. Our experiments with different vehicle and people trajectories suggested that users and applications derive significant benefits from such trajectory construction, querying and visualization. SeMiTri thus contributes towards better understanding of mobility behaviors of moving objects.

5. REFERENCES

[1] J. Han, J.-G. Lee, H. Gonzalez, and X. Li. Mining Massive RFID, Trajectory, and Traffic Data Sets. In *KDD Tutorial*, 2008.

[2] Z. Li, M. Ji, J.-G. Lee, L.-A. Tang, Y. Yu, J. Han, and R. Kays. MoveMine: Mining Moving Object Databases. In *SIGMOD*, pages 1203–1206, 2010.

[3] S. Spaccapietra, C. Parent, M. L. Damiani, J. A. de Macedo, F. Porto, and C. Vangenot. A Conceptual View on Trajectories. *Data and Knowledge Engineering*, 65:126–146, 2008.

[4] Z. Yan, C. Parent, S. Spaccapietra, and D. Chakraborty. A Hybrid Model and Computing Platform for Spatio-Semantic Trajectories. In *ESWC*, pages 60–75, 2010.

[5] Y. Zheng, X. Xie, and W.-Y. Ma. GeoLife: A Collaborative Social Networking Service among User, Location and Trajectory. *IEEE Data Eng. Bull.*, 33(2):32–39, 2010.