

Supporting Visual Exploration of Massive Movement Data

Natalia Andrienko

Gennady Andrienko

Fraunhofer Institute IAIS (Intelligent Analysis and Information Systems)
Schloss Birlinghoven; 53754 Sankt Augustin, Germany, +49 2241 142486
{natalia.andrienko, gennady.andrienko}@iais.fraunhofer.de

ABSTRACT

To make sense from large amounts of movement data (sequences of positions of moving objects), a human analyst needs interactive visual displays enhanced with database operations and methods of computational analysis. We present a toolkit for analysis of movement data that enables a synergistic use of the three types of techniques.

Categories and Subject Descriptors

H.1.2 [User/Machine Systems]: Human information processing – Visual Analytics; I.6.9 [Visualization]: information visualization.

Keywords

Movement data, trajectory, movement patterns, movement behavior, visual analytics, exploratory data analysis, visualization, interactive displays, cluster analysis, aggregation.

1. INTRODUCTION

Thanks to the recent advent of inexpensive positioning technologies, data about movement of various mobile objects are collected in rapidly growing amounts. Potentially, these data are a source of valuable knowledge about behavioral and mobility patterns. To gain an understanding of these patterns, an analyst needs a visual representation of the data, which is the most effective way to support human perception, cognition, and reasoning. However, purely visual methods of analysis (e.g. Hägerstrand 1970), even being enhanced with interactive techniques (Andrienko et al. 2000, Kraak 2003, Kapler and Wright 2005), are not scalable to large datasets. Such methods need to be combined with database operations and computational analysis techniques helping to handle large amounts of data. Some approaches have been suggested recently. Forer and Huisman (2000) and Dykes and Mountain (2003) summarize movement data into surfaces, but this is not suitable for analyzing routes. Buliung and Kanaroglou (2004) envelop bunches of trajectories and compute the central tendency, which works well for similar and close trajectories. Laube et al. (2000) combine visualization with data mining methods oriented to specific types of patterns.

In (Andrienko et al. 2007) we have presented a framework and a

toolkit for analysis of movement data based on a synergy of visualization, database operations and computations. Here, we focus on the visual and interactive components of the toolkit.

2. MAKING SENSE FROM POSITION SEQUENCES

Movement data acquired by position tracking usually lack any semantics. The records basically consist of time stamps and coordinates. In particular, there are no explicitly defined trips with specified origins and destinations and no semantically identifiable places. To understand the data, an analyst should be able to link them to his/her prior knowledge and interpretable information from other sources. Visualization is essential for this purpose.

2.1 Finding significant places

One important task in analysis of movement data is to extract and interpret the places of stops. Our toolkit supports this task in the following way. First, the positions of stops with user-specified minimum duration are extracted from the database. Second, a spatial clustering tool is applied to find groups of spatially close positions, which indicate repeatedly visited places. Third, the results are shown on a map where the positions are marked by colored point symbols (each cluster receives a unique color). The map provides the spatial context and thereby helps the analyst to interpret the places. Additional help may come from temporal histograms showing the distribution of the stops within temporal cycles (daily, weekly, etc.). Thus, the histograms in Figure 1 show the frequencies of stops of a personal car for minimum 3 hours by days of the week (A) and by hours of the day (B). The colored bar segments represent the results of the spatial clustering of the stops. It is vividly seen that the stops of the “blue” cluster occur only on the working days and mostly in the morning times. The stops of the “red” cluster occur all days and mostly in the evenings. A plausible conclusion is that the “blue” cluster of positions is situated near the place where the person works and the “red” cluster is near person’s home.

2.2 Extracting trips and exploring the routes

The sequence of position records representing the movement of an entity needs to be partitioned into subsequences corresponding to trips. The notion of trip may be application- and goal-dependent. Our toolkit allows the users to divide data in several ways: by stops, by spatial gaps, by temporal cycles, and by places of interest. The division is done by means of database operations. After that, repeated trips and typical routes may be detected by means of clustering, which groups together trips having something in common, depending on the distance function chosen (e.g. closeness of the origins and destinations, similarity of the routes). The analyst may select one or a few clusters and refine

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

Conference'04, Month 1–2, 2004, City, State, Country.
Copyright 2004 ACM 1-58113-000-0/00/0004...\$5.00.

them by re-applying the clustering tool with a different distance function or different parameter settings. To represent massive movements on a map display, we apply computational summarization of moves. The algorithm is described in (Andrienko et al. 2007). The results look as shown in Figure 2.

2.3 Exploring movement dynamics

3D views where two dimensions represent space and one represents time (Hägerstrand 1970) are good for exploring the speed of movement and its variation over time. This approach can be used even for multiple trajectories if they do not intersect (in particular, if they follow the same route). One of the distance functions in our toolkit groups trajectories by similarity of the routes and similar dynamics of the movement. Selected clusters can be explored and compared in a 3D view as shown in Figure 3, which is quite legible despite the number of trajectories displayed.

3. CONCLUSION

Interactive visual displays play the key role in supporting sense-making from movement data but are insufficient when the data are large. Our framework combines visualization with database operations and computations. The generic database techniques enable handling large datasets and are used for basic data processing and extraction of relevant objects and features. The computational techniques, which are specially devised for movement data, aggregate and summarize these objects and features and thereby enable the visualization of large amounts of information. The visualization enables human cognition and reasoning, which, in turn, direct and control the further analysis by means of the database, computational, and visual techniques.

The reported work has been partly funded by EU in the project GeoPKDD - Geographic Privacy-aware Knowledge Discovery and Delivery (IST-6FP-014915; <http://www.geopkdd.eu>).

4. REFERENCES

- [1] Andrienko, G., Andrienko, N., Wrobel, S. 2007. Visual Analytics Tools for Analysis of Movement Data. ACM SIGKDD Explorations, 9 (2), (in press)
- [2] Andrienko, N., Andrienko, G., Gatalsky, P. 2000. Supporting Visual Exploration of Object Movement. In Proc. Working Conf. Advanced Visual Interfaces AVI 2000 (Palermo, Italy, May 2000), ACM Press, 217-220, 315
- [3] Buliung, R.N., Kanaroglou, P.S. 2004. An Exploratory Data Analysis (ESDA) toolkit for the analysis of activity/travel data. Proceedings of ICCSA 2004, LNCS 3044, Springer, Berlin, 1016-1025
- [4] Dykes, J. A., Mountain, D. M. 2003. Seeking structure in records of spatio-temporal behaviour: visualization issues, efforts and applications, Computational Statistics and Data Analysis, 43, 581-603
- [5] Forer, P., Huisman, O. 2000. Space, Time and Sequencing: Substitution at the Physical/Virtual Interface. In Information, Place and Cyberspace: Issues in Accessibility (Eds: Janelle, D.G., Hodge, D.C.), Springer, Berlin, 73-90
- [6] Hägerstrand, T. 1970. What about people in regional science? Papers of the Regional Science Association, 24, 7-21
- [7] Kapler, T., Wright, W. 2005. GeoTime information visualization, Information Visualization, 4(2), 136-146

- [8] Kraak, M.-J. 2003. The space-time cube revisited from a geovisualization perspective. In Proc. 21st Int. Cartographic Conf. (Durban, South Africa, Aug. 2003), 1988-1995
- [9] Laube, P., Imfeld, S., Weibel, R. 2005. Discovering relative motion patterns in groups of moving point objects. Int. J. Geographical Information Science, 19(6), 639-668

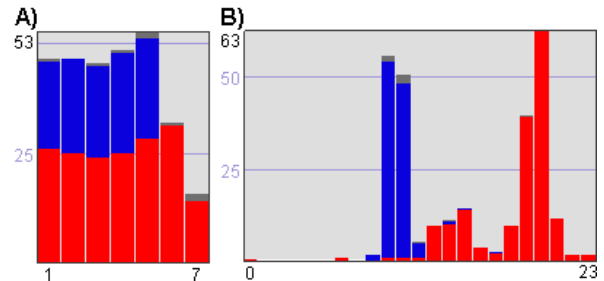


Figure 1. Temporal histograms show results of clustering of stop positions.

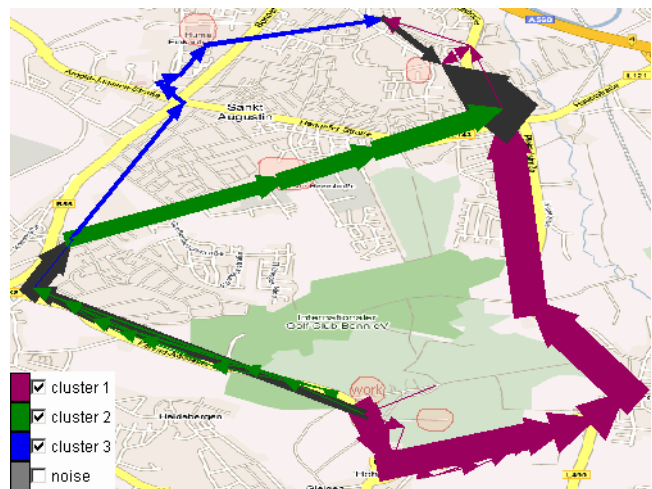


Figure 2. Three clusters of trips are represented in a summarized form.

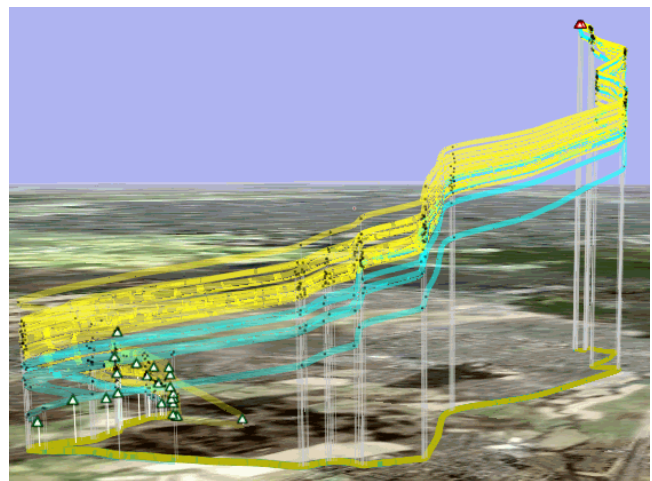


Figure 3. Two clusters of trips follow the same route but differ in the dynamics.