TrajAnalyst: Matching Data to Trajectory Analysis Modules via a Conceptual Framework

Song Gao[1], Jiue-An Yang[1,2], Krzysztof Janowicz[1], Yingjie Hu[1], Bo Yan[1]

[1]STKO Lab, Department of Geography, University of California, Santa Barbara, CA, USA
Email: {sgao, jiueanyang, jano, yingjiehu, boyan}@geog.ucsb.edu
[2]Department of Geography, San Diego State University, CA, USA

1. Introduction

With the fast growth of location awareness devices and sensors (e.g., GPS, mobile phones and RFID tags) as well as location-based social networks (e.g., Foursquare, Jiepang), large-scale movement data (or trajectories) have become available to GIScience community and related domains. Researchers have developed amount of theories, models and techniques for movement pattern analysis (MPA) and applied in many interdisciplinary/multidisciplinary studies, including but not limited to investigating human mobility (González et al. 2008, Song et al. 2010, Kang et al. 2012), travel behaviour and traffic flow estimation (Gao et al. 2013, Liu et al. 2013), animal interactions (Miller 2012), and hurricane trajectory similarity (Dodge et al. 2012). The movement trajectories can be generalized as a series of spatiotemporal points with diverse contexts. As trajectory data often share common structures and semantics, such as ***location fix***, ***order***, ***timestamp, segment, direction, speed and so on*** (Hu et al. 2013), a number of generic analysis methods reoccur in the trajectory data processing across different domains. However, these generic analysis modules and methods are still isolated from GIS integration perspective, although there is a large group of researchers using R package (Kranstauber and Safi 2014) or ArcMET[1] for movement patterns in animal ecology. Previous trajectory studies primarily focused on space-time integration and visualization but little on generic functionality. We still lack tools supporting trajectory analyses (Goodchild 2013).

Thus, the motivation of this research is to take a first-step towards creating an integrated trajectory analysis framework and tools which can be plugged in a widely used GIS platform to better support the processing and spatial analysis of trajectory data. **A theoretical question is how to match trajectory data to corresponding functionality. We present a conceptual framework for linking the data structure to the analyses and supporting better human-computer interaction**.

2. Trajectory Analysis Framework (TrajAnalyst)

2.1 Key Parameters for Trajectory Analysis Functionality

As discussed in the work by Dodge et al. (2008), the development of effective trajectory data mining algorithms and visualization methods need accurate formalization of the trajectory properties. A trajectory usually comprises a series of location points ordered by timestamp or rank, associated with various parameters and contextual attributes in different domains. Based on a literature survey (Table 1), it is clearly found that most trajectory analysis modules rely on three core parameters: *spatial location (L), timestamp (T), and sequential order (O).* We also can calculate the derived parameters such as *distance, time duration, and speed* from the primary parameters: spatial location and timestamp. Note that not all of these parameters are

[1] http://www.movementecology.net/

always necessary for MPA functionalities. It would be valuable to help trajectory data analysts to find what modules and functions are available after setting the parameters. A research on matching the trajectory data to certain types of analytical functions and thus triggering specified analysis modules in GIS environment needs to be investigated.

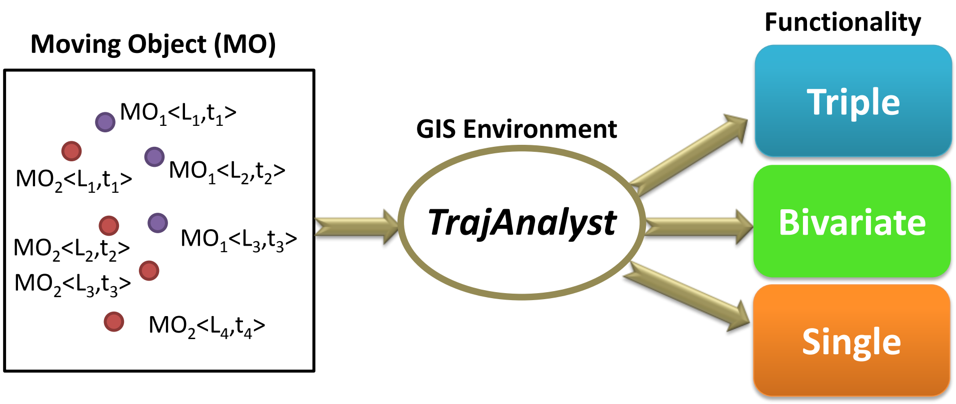
Table 1 Selected trajectory analysis modules and required parameters. (Note that not all of the trajectory analysis modules are listed here because of the space limits.)

Analysis Module/Function	**Parameters**	**Pattern**	**Reference**
Radius of gyration	< L >	mobility	González et al. 2008
Entropy	< L, T>	mobility	Song et al. 2010
Lock-step measures	< L, T, O>	similarity	Ranacher and Tzavella 2014
Fréchet distance	< L, T, O>	similarity	
Common paths	< L, T, O>	similarity	
Hausdorff distance	< L, O>	similarity	
Coefficient of sociality	< L, T>	dynamic interaction	Miller 2012
Path correlation coefficient	< L, O>	similarity	Long and Nelson 2013
Flock and convoy identification	< L >	individual and group dynamics	
Spatio-temporal edit distance	< L, T, O>	similarity	Yuan and Raubal 2014

2.2 A Conceptual Design of TrajAnalyst

We design a conceptual framework to match the fitting analytic modules for specified trajectory data inputs. In this framework (see Figure 1), the core component *TrajAnalyst* has the capability to guide the user to specify whether the input trajectory data contains one field or more fields of the primary parameters <L, T, O> for any moving object (e.g., vehicle, pedestrian, animal). Then, it can return corresponding MPA functionality candidates that fit the data input. For instance, if the data input only has the location fields (latitude/longitude) or the user doesn't specify other fields, only *single* variable <L> supporting modules such as *radius of gyration, flock and convoy identification etc.* should be populated in the analysis toolbox. In addition to the location, if the data contains another field of timestamp <L, T> or sequential order <L, O>, both *single* variable and *bivariate* modules including *entropy*, *hausdorff distance etc.* could be added in the triggered function list. Similarly, if the user has specified all three primary fields <L, T, O>, several trajectory segment/path analysis modules which need spatio-temporal footprints and orders can be added. Also, the mechanism supports triggering other modules which need both primary and derived parameters (e.g., *speed, time cost).*

The goal of this conceptual framework is to guide the technical implementations of various MPA functions and modules across domains, and to support better human-computer interaction, i.e., let the user know what analysis functions could be executed when a user load a trajectory dataset in GIS environment.



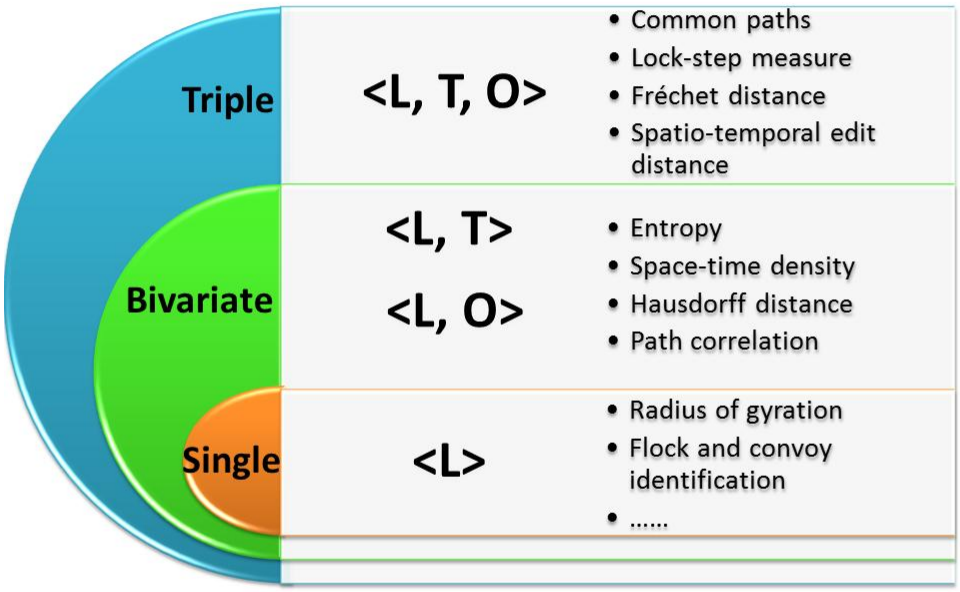


Figure 1. A conceptual framework of matching movement data to trajectory analysis modules and functionalities.

3. Prototype Development

We have been developing a *Trajectory Analyst* Toolbox as a plugin in desktop-ArcGIS to support trajectory-based geoprocessing using Python scripts. The benefits of adding the *TrajAnalyst* module on the mainstream GIS platform have three folds. First, it helps handle large-scale trajectory data with the efficient loading and geovisualization capabilities. Second, it is valuable to make use of existing spatial analysis functions (e.g., clip, spatial join) for data filtering and preprocessing. The flexible built-in architecture also supports researchers to customize their own functions and modules based on domain knowledge, especially for MPA. Last but not least, users can share and reuse the *TrajAnalyst* modules through Web services using ArcPy.

Figure 2 is a screenshot about the GIS plugin. After loading in a trajectory data (e.g., taxi GPS tracks), the user can specify the parameter fields for analysis in the graphical user interface (GUI). Then, the *TrajAnalyst-control* will evaluate the inputting parameters and return a key-pair (0:None, 1:<L>, 2:<L, T>, 3:<L, O>, 4:<L, T, O>) for triggering the corresponding group of MPA functions and modules.

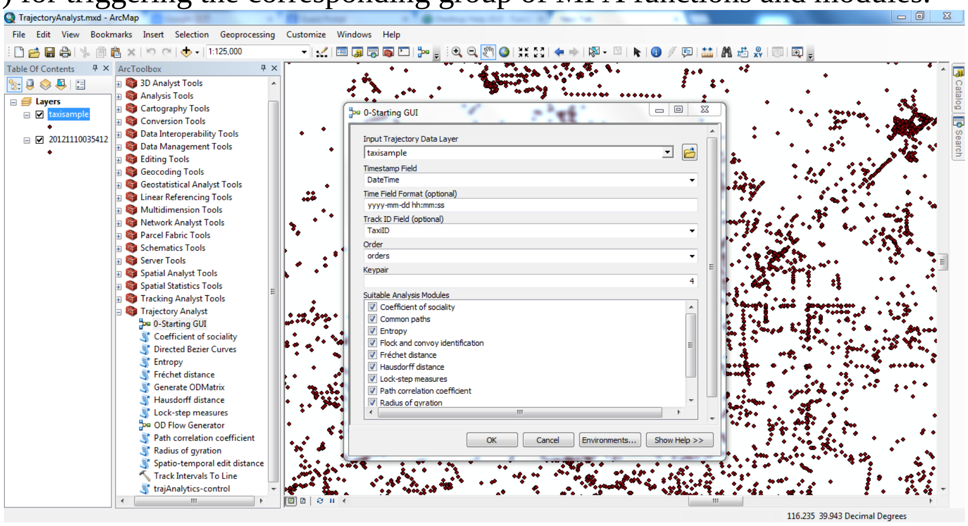


Figure 2. A screenshot of a TrajAnalyst toolbox developed in ArcGIS.

4. Conclusion

In this research, we present a conceptual framework of matching data to generic trajectory analysis modules based on the combination of MPA primary parameters <L, T, O>. A prototype was developed as a plugin for a mainstream GIS platform. This analytical framework is supposed to be applied not only in GIScience but also in more broad disciplines such as social science and health studies. In future work, additional modules and functions from multi-domains (e.g., computational geometry) for generic trajectory data analysis could be integrated by applying this conceptual framework. In addition, given the complexity of selecting suitable spatial analysis tasks, the semantics of trajectory data analysis in different contexts need to be addressed as well, such as individual-based or group-based tasks.

References

Dodge S, Laube P, and Weibel R, 2012, Movement similarity assessment using symbolic representation of trajectories. *International Journal of Geographical Information Science*, 26(9), 1563-1588.

Dodge S, Weibel R, and Lautenschütz AK, 2008, Towards a taxonomy of movement patterns. *Information Visualization*, 7(3-4), 240-252.

Gao S, Wang Y, Gao Y, and Liu Y, 2013, Understanding urban traffic-flow characteristics: a rethinking of betweenness centrality. *Environment and Planning B: Planning and Design*, 40(1), 135-153.

González MC, Hidalgo CA, and Barabási, AL, 2008, Understanding individual human mobility patterns. *Nature*, 453(7196), 779-782.

Goodchild MF, 2013, Prospects for a Space–Time GIS, *Annals of the Association of American Geographers*, 103(5), 1072-1077.

Hu Y, Janowicz K, Carral D, Scheider S, Kuhn W, Berg-Cross G, Hitzler P, Dean M, and Kolas D, 2013, A Geo-Ontology Design Pattern for Semantic Trajectories. In *Spatial Information Theory* (pp. 438-456). Springer International Publishing.

Kang C, Ma X, Tong D, and Liu Y. 2012, Intra-urban human mobility patterns: An urban morphology perspective. *Physica A: Statistical Mechanics and its Applications*, 391(4), 1702-1717.

Kranstauber B, and Safi K. 2014, (in press), *Analysis and mapping of animal movement in R*. CRC Chapmann and Hall.

Liu Y, Kang C, Gao S, Xiao Y, and Tian Y, 2012, Understanding intra-urban trip patterns from taxi trajectory data. *Journal of geographical systems*, 14(4), 463-483.

Long JA, and Nelson, TA, 2013, A review of quantitative methods for movement data. *International Journal of Geographical Information Science*, 27(2), 292-318.

Miller, JA, 2012, Using spatially explicit simulated data to analyze animal interactions: a case study with brown hyenas in northern Botswana. *Transactions in GIS*, 16(3), 271-291.

Ranacher P, and Tzavella K, 2014, How to compare movement? A review of physical movement similarity measures in geographic information science and beyond. *Cartography and Geographic Information Science*, 41(3), 286-307.

Song C, Qu, Z Blumm N, and Barabási, AL, 2010, Limits of predictability in human mobility. *Science*, 327(5968), 1018-1021.

Yuan Y, and Raubal M, 2014, Measuring similarity of mobile phone user trajectories–a Spatio-temporal Edit Distance method. *International Journal of Geographical Information Science*, 28(3), 496-520.