Visualizing the Time-varying Crowd Mobility

Feiran Wu[*] Minfeng Zhu[†] Xin Zhao[‡] Qi Wang[§] Wei Chen[¶]
State Key Lab of CAD&CG, Innovation Joint Research Center for Cyber-Physical-Society System, Zhejiang University
Ross Maciejewski[||]
The School of Computing, Informatics and Decision Systems Engineering, Arizona State University

Abstract

Modeling human mobility is a critical task in fields such as urban planning, ecology, and epidemiology. Given the current use of mobile phones, there is an abundance of data that can be used to create models of high reliability. Existing techniques can reveal the macro-patterns of crowd movement, or analyze the trajectory of an individual object; however, they focus on geographical characteristics. In this paper, we employ a novel data representation, the mobility transition graph, which is generated from a citywide human mobility dataset by defining the temporal trends of crowd mobility and the interleaved transitions between different mobility patterns. We describe the design, creation and manipulation of the mobility transition graph and demonstrate the efficiency of our approach by case study.

CR Categories: I.3.8 [Computer Graphics]: Applications—Visual Analysis;

Keywords: storyline, timeline, mobility, spatio-temporal transition

1 Motivation

Understanding and exploring human mobility patterns is essential in domains such as urban planning, transportation optimization, and epidemiology [Gonzalez et al. 2008; Schneider et al. 2013]. Understanding where and how people move provides a window into how they interact with their built environment and can provide insight for planners to improve transportation routes, prepare for disasters, or various other concerns. Given the current use of mobile devices, researchers can now, more than ever, study how humans move. As such there are emerging research trends focusing on the development of data-driven pattern discovery for human mobility patterns [Barabasi 2005]. While there has been much recent work on the visual exploration of traffic patterns based on trajectory data (e.g., [Wang et al. 2013; Zeng et al. 2014]), Visualizing the mobility patterns from a citywide population is still a challenging task.

Given the ubiquity of mobile phones, measurements of human location and travel via cell signals and GPS locations are becoming readily available. Such data can be mined as a proxy to describe the daily life of citizens. Previous studies based on mobile phone data have made significant progress in extracting human behavior and mobility patterns by leveraging various approaches, including statistical physics, mobility modeling, and data mining. Though effective, these approaches focus primarily on discovering macro-scale patterns. However, studies on mid-level or micro-scale mobility patterns are missing. Such studies are of critical importance as recent work from a study of fifteen months of human mobility data for 1.5 million individuals revealed that human mobility traces are highly unique [de Montjoye et al. 2013].

In order to better explain the extracted mobility patterns, many of the previous approaches incorporate classical visualization techniques, including parallel coordinates, heat-maps, and glyphs. However, there are rarely integrated works that involve intertwined statistical methods, data mining and visualization techniques. In this paper, we contribute the design and implementation of a novel visual exploration approach for depicting the temporal evolution of human mobility patterns of 500,000 mobile phone users in a medium-sized city (3 districts, 2 county-level cities and 6 county areas, 919.7 million resident population in 2013). This dataset contains the location information of one million individuals over the course of 30 days. The location of an individual is specified at a temporal resolution of less than one minute and with a spatial resolution equal to the distribution of the mobile cell towers. This makes it feasible to characterize the mobility patterns of an individual over the 30 day period at a fine scale resolution.

The main contribution of our study is to structure and depict mobility patterns and how persons transition from one mobility pattern to another by a hybrid timeline-graph representation, called mobility transition graph (MTG), in which the vertices represent clusters of similar mobility patterns and the edges represent transitions between patterns. The advantage of mobility transition graph is its ability to simultaneously characterize the temporal evolutions and interleaved transitions of mobility patterns (i.e., nodes may represent a concept such as walking to the downtown, and then transition to a node representing riding the downtown). Besides, we incorporate the other assistant views to enable users to visually locate the patterns, analyze the mobility features and groups.

2 Data Preprocess

The dataset employed in our study is provided by a mobile phone service company. It consists of 14 billion records of 7 million phone users across 25,000 cell towers over January and February of 2014. The dataset is collected on the basis of cell towers: a record is generated whenever the following activity happens: a user enters or leaves a cell tower, a user makes a call or sends a short message, or a user stays in a cell tower for more than a given duration (e.g.,3 minutes). Each record contains multiple items: a phone ID, a cell tower ID, the activity type, a time stamp.

Because of the dirtiness, geographical inaccuracy and temporal sparsity of this dataset, several preprocesses are taken to the raw data:

- **Removing Ping-pong Effects** The ping-pong effect [Xiong et al. 2012] is caused by the frequent handoffs between two nearby cell towers, yielding fabricated records. We encode a

[*]e-mail:tomwfr007@gmail.com
[†]e-mail:zhuminfeng@live.cn
[‡]e-mail:zhaoxinitachi@gmail.com
[§]e-mail:wqstu1@163.com
[¶]e-mail:chenwei@cad.zju.edu.cn
[||]e-mail:rmacieje@asu.edu

SA'15 Symposium on VisHPC, November 02 – 06, 2015, Kobe, Japan.
ACM 978-1-4503-3929-2/15/11.
http://dx.doi.org/10.1145/2818517.2818540

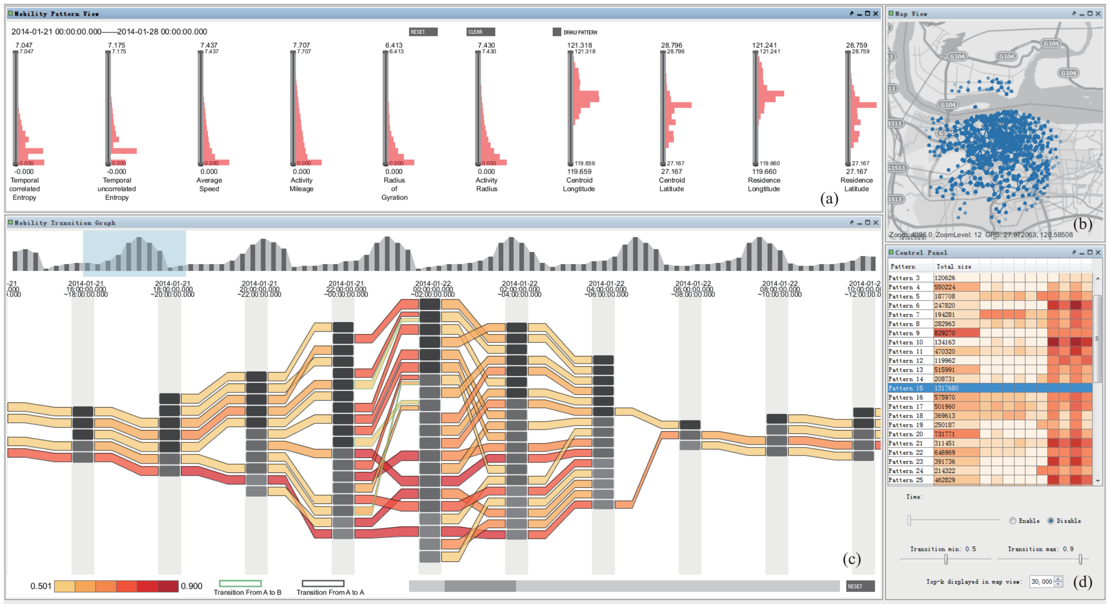

Figure 1: *The mobility patterns extracted from the mobile phone location records of 500,000 users in 7 days (Jan.21 - Jan.27, 2014). (a) The mobility patterns (clusters) are depicted with a multiple coordinate style in the mobility pattern view. (b) A map view shows the geographical scene and trajectories. In this case, the selected mobility pattern describes the inactive behavior in the downtown. (c) A mobility transition graph (MTG) is built upon the dataset. (d) Controlling widgets. When the user selects a pattern, the corresponding trajectories are shown in the map view and the statistical information are depicted in the multiple coordinate view.*

cell tower with a word, and employ an N-gram model [Gao et al. 2002] to detect and eliminate frequent changes between multiple cell towers.

- **Removing Invalid Records** The records whose the activity type or the cell tower id is unknown are removed.
- **Removing Duplicate Records** Several consecutive records of a mobile phone may be generated due to the signal intensity fluctuations or reconnections. We remove duplicated records.
- **Compensating Missed or Sparse Records** To compensate missed records of a mobile phone, we add records by interpolating the values from known records.

3 The Mobility Transition Graph

Before designing a visual representation for studying citywide mobility patterns, we firstly represent the mobility of a mobile phone user with an array of features that is derived from the movement and geographical context. Considering the commonly used and the representative features proposed in recent years, 8 feature descriptors (10 dimensions) are employed as follows:

- **Temporal-uncorrelated Entropy** An entropy describes the degree of predictability. Many definitions concerning entropy have been proposed [Song et al. 2010] to measure the predictability in human mobility. We adopt the temporal-uncorrelated entropy because of its capability of distinguishing different movements. Usually, high entropy suggests the movements with dramatic changes. The temporal-uncorrelated entropy $S^{unc} = -\Sigma_{j=1}^{N} p(j) log_2 p(j)$ describes the heterogeneity of visitation patterns, where $p(j)$ denotes the historical probability of the visit from the mobile phone user to the jth cell tower.
- **Temporal-correlated Entropy** We additionally employ a temporal-correlated entropy S^{tc}, which is similar to S^{unc} despite that $p(j)$ is temporal-correlated. If the ith mobile phone stays a long time in the jth cell tower, $p(j)$ will be high.
- **Average Speed** is defined as the average speed of valid movement records. We set the speed limit to be 200km per hour, and clean the records whose speed is higher than the limit.
- **Activity Mileage** is defined as the distance the mobile phone moves.
- **Centroid Location** is the geographic centroid (the longitude, the latitude) of the sequence of records in a frame.
- **Radius of Gyration** is the typical distance traveled by a mobile phone user around the centroid of mass of the trajectory [Gonzalez et al. 2008]. Which is a synthetic and easy to compute parameter leveraged to describe the movement span of the user.
- **Residence Location** is the estimated location of the home of a mobile phone user. It is defined as the most frequently appeared location (the longitude, the latitude) from 00:00 a.m. to 6:00 a.m.
- **Activity Radius** is the average distance a mobile phone user travels from his residence location. It measures the activity scope in t_i.

Second, the human mobility in a period of time can be regarded as a composition of multiple segments, and each segment encodes a specific event or activity in a time interval. The transitions among consecutive segments may indicate temporal variations, periodicity or abnormality of the mobility. Third, the mobility in a segment is reasonable if and only if the result has a statistical significance. Consequently, we divide the record sequence of each mobile phone user into segments with a fixed time interval, then extract the features of each segment and cluster all segments of all mobile phone users by K-means. Each cluster encodes the average mobility of a group of mobile phone users can be regarded as a mobility pattern. By assuming that the mobility pattern transition can be characterized as the transition between the consecutive segments, we incorporate the k-th ($k = 1$) order Markovian assumption and calculate the transition likelihood between the mobility patterns in two consecutive periods of time by the method in [Song et al. 2009].

3.1 Visualization

Visualizing the constructed mobility transition graph is straightforward. Essentially, a mobility transition graph (MTG) is a hybrid timeline-graph representation that characterizes the interconnected mobility transition in a sequential way. To emphasize on the temporal transitions, we pack the mobility patterns vertically as a stacked bar chart sorted by the value of a descriptor or the size (Figure 2 (a)). Each pattern is encoded as a grey rectangle, whose darkness encodes the count of segments belong to the mobility pattern. The user may filter and display the necessary patterns (e.g., according to the likelihood of the transitions which connects to these patterns).

Then we sequentially place the stacked bars from left to right to represent the flow of mobility over time. The transitions between two consecutive time periods are encoded by ribbons that connect pairs of mobility pattern. The ribbons are colored with respect to the transition likelihood. To highlight the transition in the identical pattern, the edge of the ribbon is colored in black while the edges in green represents the transitions between different patterns. Normally, the width of a ribbon is set to be identical. In cases that multiple transitions start or end in a node, the width of each ribbon is set to be proportional to the transition likelihood of the corresponding transition (see the region in red rectangle of Figure 2 (a)).

Further, an enhanced bar chart is employed to summarize the mobility transition graph as an overview. Each bar in dark grey corresponds to a time period in the MTG, and its height encodes the number of the displayed patterns in the period. Each bar in light grey encodes the transitions between consecutive time periods. A semitransparent rectangle mask in blue indicates the interval of the overview that represents the shown portion of the underlying MTG (Figure 2 (c)).

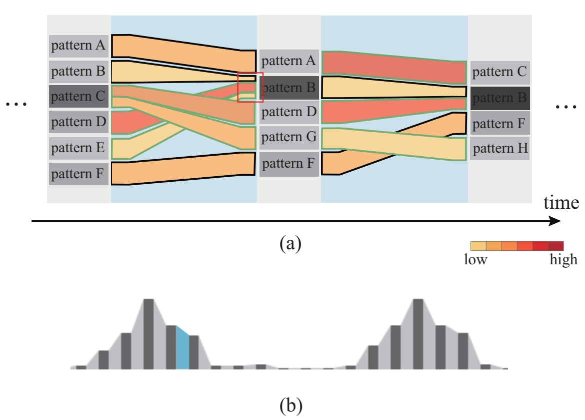

Figure 2: *(a) The stacked flow view of a MTG. The color bar at the right is used to color the links based on the transition likelihood; (b) A bar chart overviews the MTG.*

3.2 The Interface

Figure 1 illustrates the interface of our system. The map view and the mobility pattern view are designed to display the corresponding information of a selected mobility pattern. While the MTG view concentrates to demonstrate the transitions between the mobility patterns and supports the user to explore with interactive operations.

- **The Map View** employs OpenStreetMap as the map, and shows the geographic-related information in the map (Figure 1 (b)). The trajectories of mobility phone users are shown in the map to help study the mobility descriptors and mobility patterns. To avoid heavy visual clutter, trajectories of top-k mobile phone records that are the closest to the cluster centroid are shown. k is a user-adjustable parameter. The stroke weight of trajectory encodes the distance to the cluster centroid.

- **The Mobility Pattern View** displays the detailed information of 10 dimensions of 8 mobility descriptors by means of a multiple coordinates plot (Figure 1 (a)). The histogram on each coordinate axis characterizes the statistical distribution of each dimension of the selected mobility pattern. A pattern list panel (Figure 1 (d)) employs a matrix view to show the dimensions of each mobility pattern. The color of each matrix cell encodes the numeric value of the corresponding dimension. The user can either select a row in the list panel or choose the ribbons in the mobility pattern view to specify and study a mobility pattern. The user can also filter each dimension by adjusting the range slider on the associated axis.

- **The MTG View** The MTG view visualizes the constructed mobility transition graph. The user can navigate, reorder, filter and manipulate with the patterns, links and axes of the representation (Figure 1 (c)).

4 Implementation

The backend takes the responsibility of the data preprocessing and MTG model constructing. We store and calculate the data with 12 computing nodes. Each node drives 8 core and 20GB memory. We employ the Apache Spark as the data processing engine which takes about 1h to fulfill all preprocess steps. The model constructing is faster which spends about 12m clustering the mobility patterns and 2m calculating the transition between them. The frontend runs on a PC with 3.4GHz dual core, 16GB memory, which is implemented in JAVA and uses Processing for rendering. Because the construction of MTGs is pre-computed, exploring and analyzing the mobility patterns and their transitions can be performed in real-time.

5 Case Study

5.1 Case 1: The General Exploration of The Mobility Patterns

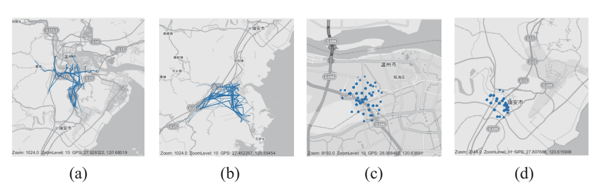

Figure 3: *(a)(b): The patterns with high speed and entropy are distributed along the highway and railway; (c)(d) The patterns with low speed and entropy may distribute in the residential zone.*

The first case study is designed to study the general citywide mobility patterns. We select the records collected from Feb. 4th 2014 to Feb. 10th 2014. By the mobility pattern view, we may investigate the distribution of each mobility description and find the radius of

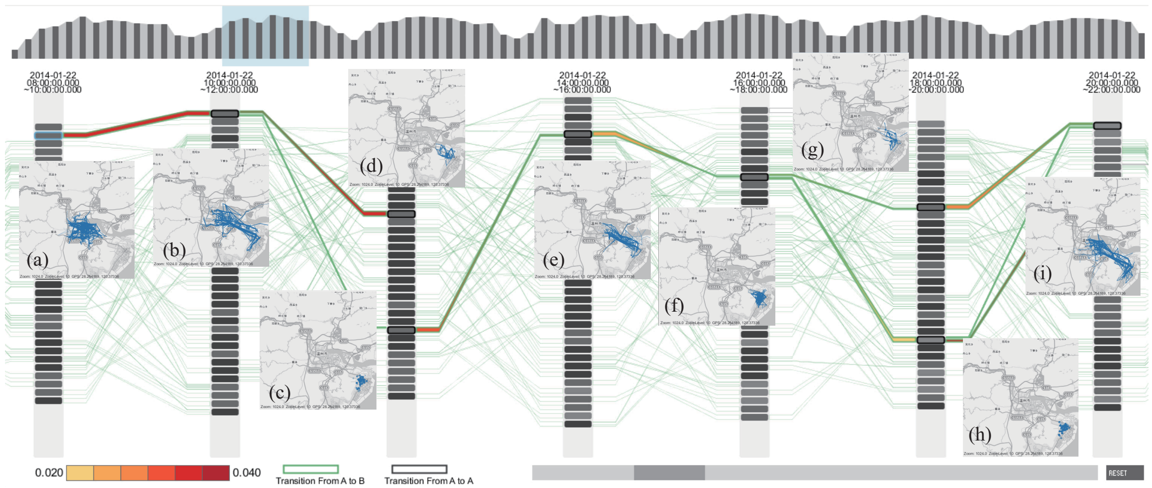

Figure 4: *The mobility transition graph of case 2. The mobility pattern transition are highlighted and can be traced from (a) to (i).*

gyration follows the power law which is consistent with the conclusion in [Gonzalez et al. 2008]. By selecting the patterns, the map view shows their geographic representation. The patterns with high speed and entropy are distributed along the highway and railway (Figure 3 (a)(b)) while the low speed and entropy may distribute in the residential zone (Figure 3 (c)(d)). Meanwhile, the MTG and its overview exhibit a periodicity over time (Figure 1 (c)). That is, there is routinely a peak from 0:00 a.m. to 2:00 a.m. of the transition likelihood. The transitions among these patterns are quite stable from 00:00 am to 04:00 am. In the map view, these mobility patterns are slow and inactive. We deduce that these transition patterns describe the people sleep in this time period.

5.2 Case 2: Visualizing The Interregional Movement

In this case, we show a movement pattern between two main districts selected by our system in Figure 4. Figure 4 (a) describes the movement in the downtown from 8:00 a.m. to 10:00 a.m. Then in the next 2 hour, this mobility pattern transfers to the pattern in Figure 4 (b), which describes the movement from the downtown to the development zone of this city. Next, this pattern splits into two patterns depicted in Figure 4 (c) and Figure 4 (d), respectively. From 2:00 p.m. to 4:00 p.m. the pattern describes the movement between the downtown and the development zone appears again and connects the previous pattern in Figure 4 (c). In this way, the user may trace the transition of the mobility pattern and observe the relevant information on the map view.

6 Conclusion

We present a novel visual representation that characterizes the statistical transitions of mobility patterns of crowd in daily life. The implemented visualization scheme not only allows for intuitive understanding of mobility patterns, but also provides a mechanism for studying the transition modes in a situation-aware fashion. We exemplify our approach with case studies on a real dataset and demonstrate the efficiency of our approach.

Acknowledgements

Supported by National 973 Program of China (2015CB352503), the Fundamental Research Funds for the Central Universities.

References

BARABASI, A.-L. 2005. The origin of bursts and heavy tails in human dynamics. *Nature 435*, 7039, 207–211.

DE MONTJOYE, Y.-A., HIDALGO, C. A., VERLEYSEN, M., AND BLONDEL, V. D. 2013. Unique in the crowd: The privacy bounds of human mobility. *Scientific reports 3*.

GAO, J., GOODMAN, J., CAO, G., AND LI, H. 2002. Exploring asymmetric clustering for statistical language modeling. In *Proceedings of the Association for Computational Linguistics*, 183–190.

GONZALEZ, M. C., HIDALGO, C. A., AND BARABASI, A.-L. 2008. Understanding individual human mobility patterns. *Nature 453*, 7196, 779–782.

SCHNEIDER, C. M., BELIK, V., COURONNÉ, T., SMOREDA, Z., AND GONZÁLEZ, M. C. 2013. Unravelling daily human mobility motifs. *Journal of The Royal Society Interface 10*, 84, 2013–2046.

SONG, L., KOLAR, M., AND XING, E. P. 2009. Time-varying dynamic bayesian networks. In *Advances in Neural Information Processing Systems*, 1732–1740.

SONG, C., QU, Z., BLUMM, N., AND BARABÁSI, A.-L. 2010. Limits of predictability in human mobility. *Science 327*, 5968, 1018–1021.

WANG, Z., LU, M., YUAN, X., ZHANG, J., AND V. D. WETERING, H. 2013. Visual traffic jam analysis based on trajectory data. *IEEE Transactions on Visualization and Computer Graphics 19*, 12, 2159–2168.

XIONG, H., ZHANG, D., AND GAUTHIER, V. 2012. Predicting mobile phone user locations by exploiting collective behavioral patterns. In *IEEE Conference on Ubiquitous Intelligence and Computing*, 164–171.

ZENG, W., FU, C.-W., ARISONA, S. M., ERATH, A., AND QU, H. 2014. Visualizing mobility of public transportation system. *IEEE Transactions on Visualization and Computer Graphics 20*, 12.