A SURVEY ON VISUAL ANALYTICS IN URBAN MOBILITY

by

QIAOMU SHEN

Department of Computer Science and Engineering
The Hong Kong University of Science and Technology
Supervised by Prof. Huamin Qu

October 2017, Hong Kong

TABLE OF CONTENTS

A SURVEY ON VISUAL ANALYTICS IN URBAN MOBILITY

by

QIAOMU SHEN

Department of Computer Science and Engineering

The Hong Kong University of Science and Technology

ABSTRACT

Urban mobility indicates the human movement patterns in a city. With the rapid urbanization, mobility gradually raises more and more attention on a variety of researches areas. The studying of mobility is beneficial to both individual residents and city management, and it has a long history. Previous researches on the mobility are greatly limited to data viable. Thanks to the sensing technologies, more and more types of data can be collected to describe the human movement, providing new chances to study the mobility from different perspectives.

Currently, many automatic methods have been proposed on the mobility and a lot of meaningful results has been found. However, the mobility patterns differ from time, space and situations, which greatly needs the involvement of the domain experts. Visual analytics bridge the gap between the techniques and domain knowledge.

In this survey, we first summarize the data modeling methods for visual analytics from a new perspective as well as the corresponding research challenges. Then we will focus on the general techniques used in different stages of analysis. After that, we classify the previous work based on common application problem and discuss how the problem is solved through the specific modeling and techniques. At the end, we conclude with some future direction of mobility research.

CHAPTER 1

INTRODUCTION

In this chapter, we first briefly introduce the background of the research on human mobility as well as the corresponding visualization. Then we summarize the data source that are used to describe the human movement. After that, we summarize the general tasks of the human mobility analysis. At last, an overview of the whole survey is given.

1.1 Background

Movement is the fundamental characteristics of all the species, especially for the human-being, the movement range greatly represents the evolution history as well as the development of society and technology, while the movement patterns can also reflect the activities in relative small scale. Especially in the recent 150 years, with the rapid development of modern traffic, the movement pattern tends to be more diversified and complicated. With the advanced sensing and location-based technology, more and more movement data are captured to describe humans mobility. Nowadays, the movement data could precisely describe human activities in the urban, and raise an increasing attention of researchers in different research domains like urbanology, meteorology, sociology and economics.

However, the analysis of the movement data is difficult, in addition to the increasing data size, the movement data has spatial-temporal features which is independent to other attributes. And the movement patterns also differ from time and region. Besides, some reasoning tasks from the data are also require domain knowledge, which needs the involvement of domain expert.

Visualization provides the methods that leverage the distinct capabilities of machine and human for the exploration tasks. Hundreds of years ago, the visualization to human movement has been designed that enable human to understand the movement patterns.

A classic case of the mobility visualization is Napoleons campaign to Russia(Figure 1.1) which is depicted by French civil engineer Charles Joseph Minard. This well designed figure not only the present the spatial-temporal features of the marching route, but also the moving directions, number of troops as well as the temperature.

With the rapid development of computer techniques, the more advanced visual analytics are designed based on powerful computing capability as well as the efficient algorithms.

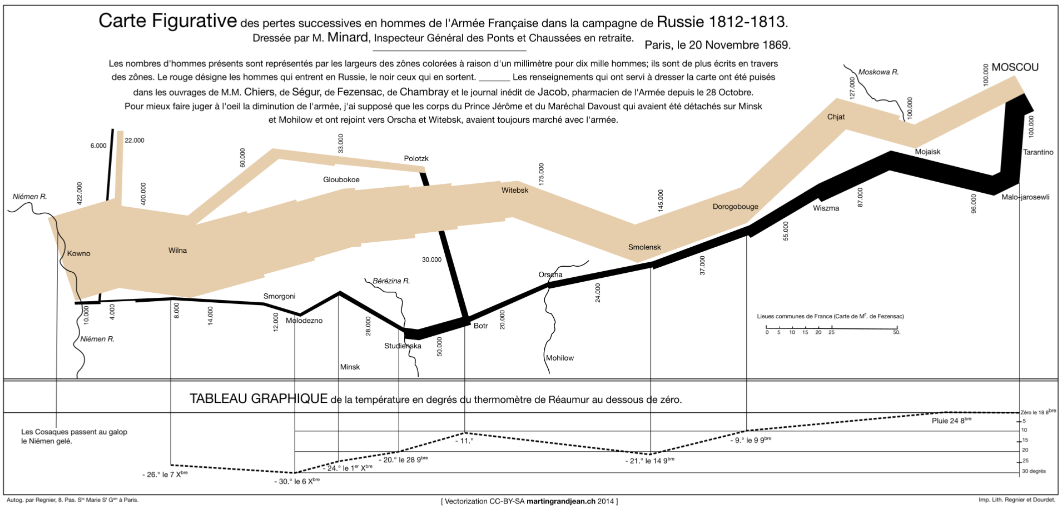

Figure 1.1: Charles Minard's map of Napoleons campaign to Russia of 1812. The band illustrates the marching route from Kaunas to Moscow. The position on the figure present the relative position of geo-map. The band width illustrates the troop number and the color indicates the directions of departure and withdraw.

Instead the effectiveness of visual design, the modern visual analytics also focus on the how to bridge the gap between users and computers. As for the exploration itself, the interactive analysis on complicated tasks should be considered.

1.2 Data classification

Nowadays, with the development of sensing technology, more and more types of data reflecting the human movement can be collected. The variety type of data enables researchers to analyze the human mobility from different perspective and granularity. Here we summarize the data into the following categories according to the characteristic, these characteristic highly related to how the data are collected:

Data collected by professional location device: Including the vehicle trajectory(Liu et al., 2017), aircraft trajectory(Lampe & Hauser, 2011) and the vessel trajectory(Willems et al., 2009). This kind of movement data is commonly collected by professional location devices which can provide continuous and fine-grained position records. For example, most taxis are equipped with GPS device, a set of discrete location points are generated and recorded. Due to the fine spatial continuity, this type of data can meet the study of different granularity.

Data collected by the device with fixed position: Some devices cannot generate position information by themselves, however, when pass through specific regions or devices, the locations can be recorded, this type of data provides a chance to analyze the human mobility among the fixed regions or devices. For example, the using of public bicycle in London(Wood et al., 2011), when users hire or release bicycle, the records will be generated by docking stations. This is similar to the metro ticketing(Itoh et al., 2016) data and the traffic monitoring data(Guo et al., 2011). Another special case is the mobile phone using data. The location of mobile phone can be recorded through the mobile stations. Even through the triangle localization, the relative accurate location can be calculated, however, due to some practical issue, the movement of mobile phone is always analyzed among the stations(Wu et al., 2016). In summary, this type of data always analyzed in a way of origin-destination which will be discussed in detail in the Chapter 4.

Data collected by software on mobile device: More and more people today rely on social media to post their status or connect friends. When messages are post on the social media platform, the position information is also located. The social media provides a more sparse way to record the movement of users, and in many cases, only one message is posted in a relative short time for one specific user, which can only show the presence of people instead of the movement. Even though the sparse spatial based record is difficult to reflect the fine-grained movement, it can describe the general spatial distribution of people(Chae et al., 2014), and further extract the semantic information of spaces(Andrienko et al., 2013b) or movement(Itoh et al., 2016). This type of data can be analyzed in sparse way like origin-destination or spatial points.

Other data: In addition to the data mentioned previous, there are more data can be used, including the questionaire(Zhao et al., 2008), migration and refugee(Wood et al., 2010) records.

1.3 Tasks of mobility visualization

Better enriched: According to Andrienko(Andrienko et al., 2013a), though there is a variety of data can capture the movement of human, all of them share the following elements: movers, spaces, spatial events and time. Based on the previous study of movement, Andrienko further summarize the tasks as follows:

Mover-oriented perspective: This subbranch of tasks mainly focuses on the movers, including the attributes of movers as well as the spatial-temporal context of the movers.

Event-oriented perspective: This subbranch of tasks mainly focuses on the events, including the attributes of events as well as the spatial-temporal context of the events.

Place-oriented perspective: This subbranch of tasks mainly focuses on the spaces, including the movement related thematic attributes of spaces.

Time-oriented perspective: This subbranch of tasks mainly focuses on the time, including the movement related thematic attributes of specific time point or time range.

1.4 Overview

The whole survey is organized as follows:

Chapter 2 introduces the existed taxonomies which have been widely accepted. Generally, these taxonomies are designed based the movement patterns as well as the applications. Then we propose our own taxonomy based on data modeling.

Chapter 3 focus on the mobility analysis based on spatial points. We shall introduce the techniques used in the spatial points modeling. Then, we classify the research in this branch into two groups according to the entity to analyze: mover and event.

Chapter 4 focus on the analysis based on origin-destination(OD) modeling. We first introduce the data processing and interaction techniques under this modeling, then we further divide this branch into two subbranches: single origin/destination, and multiple origin/destination. In each subbranch, the visualization techniques and appropriate application is discussed.

Chapter 5 focus on the analysis based on trajectory. Same to OD, we will introduce the general data processing techniques and interactions, then classify it into massive trajectory and individual trajectory and discuss the application and visualization.

Chapter 6 makes a conclusion and discusses the future research directions.

CHAPTER 2

TAXONOMY

In this chapter, we first briefly introduce two existed taxonomies about movement and mobility analysis respectively given by Dodge et al.(Dodge et al., 2008) and Andrienko et al.(Andrienko et al., 2017). And then we propose our own taxonomy which is an integration of data modeling, techniques and applications.

2.1 Taxonomy by Dodge et al

Dodge et al.(Dodge et al., 2008) proposed a taxonomy on visualization of movement based on many previous surveys. The taxonomy present a hierarchical structure of the patterns that can be extracted from a variety of movement data, these patterns can be further considered into the mining algorithms design and visualization design, as Figure 2.1 shows:

Generic patterns and behavioral patterns: As the two sub-branches of the root, the generic patterns and behavioral patterns differ in the specificity of the movement. The generic patterns are the low level building blocks, these patterns exist in the general movement data, like the prorogation(traffic congestion, disease) and periodicity(species migration, traffic congestion). The behavioral patterns more focus on patterns under the specific context or for the specific object, like the fighting and foraging.

Primitive and compound patterns: The generic patterns can be further classified into primitive and compound patterns, according to the number of variable parameters. The primitive patterns is the simplest form of movement with only one changing parameter. On the other hand, the compound patterns involves the changing of multiple parameters. Further more, the Primitive patterns can be classified by the spatial temporal features, which includes spatial patterns, temporal patterns and spatial temporal patterns. While all the compound patterns involve both spatial and temporal features.

Many similar taxonomy also exist, the commonality of them is that all of these work are based on the information that can be extracted from movement data.

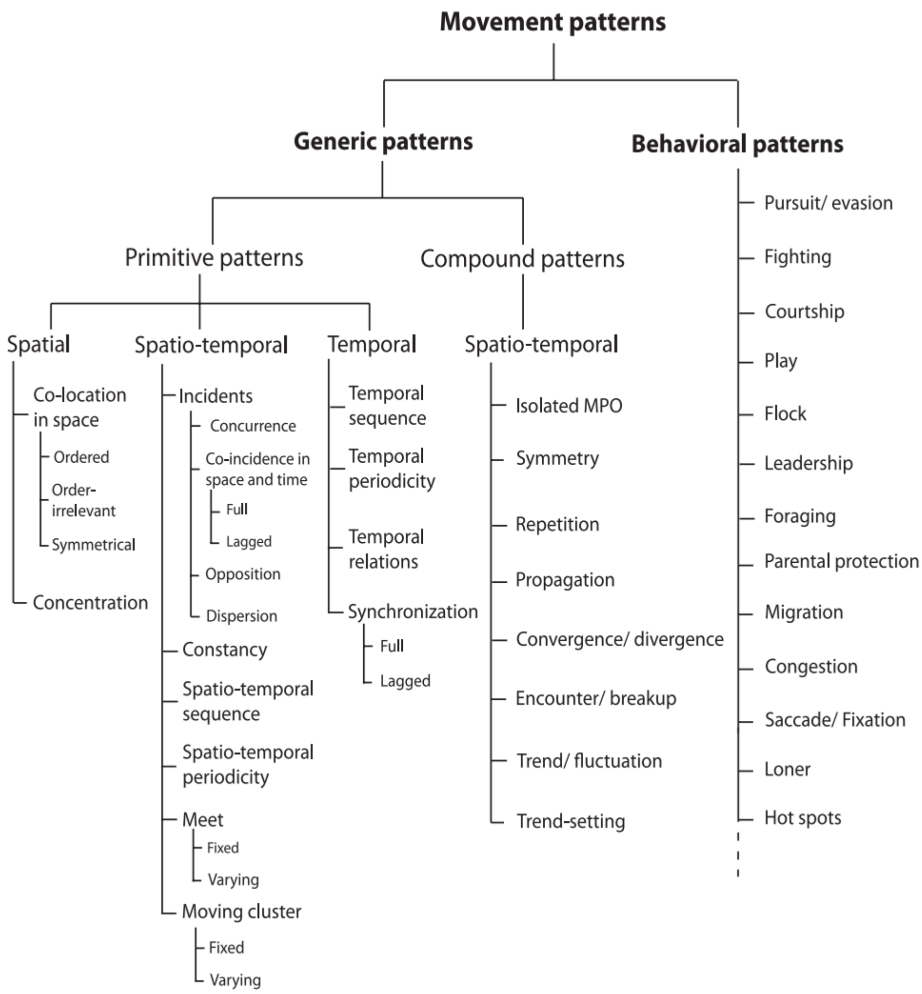

Figure 2.1: Dodge's taxonomy of movement patterns.

2.2 Taxonomy by Andrienko et al

Another category of classification is based on the study object or some high-level application problems. In this section we introduce a recent work from Andrienko et al.(Andrienko et al., 2017).

In this survey,the previous work are divided into four large categories: "Data", "Movement and Transportation Infrastructure", "Movement and Behavior", "Modeling and Planning". However, since the "Data" section is an summary of the features and transformations of movement data, which works as fundamental of other sections, we will not list this section into the taxonomy.

Movement and Transportation Infrastructure: This category focus on the human movements with the transportation infrastructure, including vehicles and pedestrians along the transportation routes, as well as the people through in the public transportation facilities like Subway, bus and train. This category can be further divided into sub-branches including:

- Details of Individual Movements
- Variety of Taken Routes
- Movement Dynamics Along a Route
- Details of Individual Movements
- Linking Origins to Destinations
- Collective Movement Over a Territory
- Events
- Contextualizing Movement
- Impacts and Risks

Movement and Behavior: This category focus on the behavior behind the movement, in addition to the mobility itself, this category more focus on the reasoning of specific movement patterns, like the interest, activities, etc. This category is further divided into three following divisions:

- Use of Transport
- Mass Mobility
- Peoples Activities and Interests

Modeling and planning: This category focus on the analytics of traffic modeling and transportation planning. This includes the derivation of models from data, applications of forecasting and simulation, transportation scheduling.

2.3 Taxonomy design based on modeling

Dodge's taxonomy is based on the patterns that can be extracted from the movement data; while Andrienko's taxonomy focuses on the applications. After reviewing the research work in the past tens of years, we have found there are three types data modeling methods, which are spatial points, origin-destination and trajectory. These three modeling methods differ in the granularity of movement and are respectively utilized in different data or applications.

Need modify: taxonomy shouldn't include the interactions and aggregation

In spatial temporal analysis, aggregation works as a very important step, but the aggregation methods are quite different under different modeling methods. We plan to introduce the aggregation techniques as the start of each subcategory. The overall structure taxonomy is described as follows:

Mobility based on spatial points: In this category, the mobility will be analyzed through the separated spatial points. It should be noted that in this category, the movement of individual mover will not be tracked, the spatial points can reflect the overall spatial distribution of movers, and the distribution of multiple time will indicate the overview movement. We first introduce the aggregation techniques that will be used in the spatial points. Then we classify the previous work based on the entity to be analyzed, including movers and events.

Mobility based on origin-destination: In this category, the movement among origin and destination will be analyzed. We first introduce the aggregation, query and interaction based on origin-destination. Then we classify this category into single origin/destination and multiple origin/destination, which could be further adapted to the different application problem.

Mobility based on trajectory: In the category, the movement records are collected in a finest-grained level and provide a way to conduct the fine-grained tasks, like individual monitoring. In addition to the aggregation techniques and interaction techniques, we further classify the research about trajectory into massive analysis and individual analysis.

CHAPTER 3

MOBILITY BASED ON SPATIAL POINTS

In this section, we will introduce the mobility analytics based on spatial points. To make the spatial points different from the the other two modeling methods, here we will not track the movement of single movers. Under this modeling, the overall movement will be generated according to the comparison of the spatial distribution in different time range.

3.1 Aggregation techniques for spatial points

Bibliography

Andrienko, G., Andrienko, N., Bak, P., Keim, D., & Wrobel, S. (2013a). *Visual analytics of movement*. Springer Science & Business Media.

Andrienko, G., Andrienko, N., Bosch, H., Ertl, T., Fuchs, G., Jankowski, P., & Thom, D. (2013b). Thematic patterns in georeferenced tweets through space-time visual analytics. *Computing in Science & Engineering*, *15*(3), 72–82.

Andrienko, G., Andrienko, N., Chen, W., Maciejewski, R., & Zhao, Y. (2017). Visual analytics of mobility and transportation: State of the art and further research directions. *IEEE Transactions on Intelligent Transportation Systems*.

Chae, J., Thom, D., Jang, Y., Kim, S., Ertl, T., & Ebert, D. S. (2014). Public behavior response analysis in disaster events utilizing visual analytics of microblog data. *Computers & Graphics*, *38*, 51–60.

Dodge, S., Weibel, R., & Lautenschütz, A.-K. (2008). Towards a taxonomy of movement patterns. *Information visualization*, *7*(3-4), 240–252.

Guo, H., Wang, Z., Yu, B., Zhao, H., & Yuan, X. (2011). Tripvista: Triple perspective visual trajectory analytics and its application on microscopic traffic data at a road intersection. In *Pacific Visualization Symposium (PacificVis), 2011 IEEE*, pp. 163–170. IEEE.

Itoh, M., Yokoyama, D., Toyoda, M., Tomita, Y., Kawamura, S., & Kitsuregawa, M. (2016). Visual exploration of changes in passenger flows and tweets on mega-city metro network. *IEEE Transactions on Big Data*, *2*(1), 85–99.

Lampe, O. D., & Hauser, H. (2011). Interactive visualization of streaming data with kernel density estimation. In *Pacific Visualization Symposium (PacificVis), 2011 IEEE*, pp. 171–178. IEEE.

Liu, D., Weng, D., Li, Y., Bao, J., Zheng, Y., Qu, H., & Wu, Y. (2017). Smartadp: visual analytics of large-scale taxi trajectories for selecting billboard locations. *IEEE transactions on visualization and computer graphics*, *23*(1), 1–10.

Willems, N., Van De Wetering, H., & Van Wijk, J. J. (2009). Visualization of vessel movements. In *Computer Graphics Forum*, Vol. 28, pp. 959–966. Wiley Online Library.

Wood, J., Dykes, J., & Slingsby, A. (2010). Visualisation of origins, destinations and flows with od maps. *The Cartographic Journal*, *47*(2), 117–129.

Wood, J., Slingsby, A., & Dykes, J. (2011). Visualizing the dynamics of london's bicycle-hire scheme. *Cartographica: The International Journal for Geographic Information and Geovisualization*, *46*(4), 239–251.

Wu, W., Xu, J., Zeng, H., Zheng, Y., Qu, H., Ni, B., Yuan, M., & Ni, L. M. (2016). Telcovis: Visual exploration of co-occurrence in urban human mobility based on telco data. *IEEE transactions on visualization and computer graphics*, *22*(1), 935–944.

Zhao, J., Forer, P., & Harvey, A. S. (2008). Activities, ringmaps and geovisualization of large human movement fields. *Information visualization*, *7*(3-4), 198–209.