Trajectory Data Mining: Overview of Applications

Anjali Sahu[1], Pallavi N.[2]
[1, 2] Dept of Computer Science & Engg
[1, 2] Atria Institute of Technology

Abstract- *The advances in location-acquisition and mobile computing techniques have generated massive spatial trajectory data, which represent the mobility of a diversity of moving objects, such as people, vehicles, and animals. Many techniques have been proposed for processing, managing, and mining trajectory data in the past decade, fostering a broad range of applications. To maintain such streaming data of location updates from mobile clients, conventional approaches such as time-based regular location updating and distance-based location updating have been used. This paper provides a survey on various different approaches towards understanding of trajectory data in storage systems, and also generalizes the other applications implemented in this domain.*

Keywords- trajectory data, data mining, mobile computing, pattern identification

I. INTRODUCTION

In dynamical systems, a trajectory is the set of points in state space that are the future states resulting from a given initial state. In a discrete dynamical system, a trajectory is a set of isolated points in state space. In a continuous dynamical system, a trajectory is a curve in state space. [1]Internet enabled mobile devices are primary sources for obtaining very large volumes of trajectory data that capture the movements of different types of objects such as people, vehicles, animals and vessels. The increasing pervasiveness of location acquisition technologies has enabled collection of very large trajectory datasets for different types of moving objects. Useful patterns discovered from the movement behaviour of moving objects are very valuable and forms a trajectory knowledgebase, and much useful to variety of real time applications. [2]

Trajectory data are collected from various sources. One of the most common types is generated by GPS-equipped vehicles. Besides, other kinds of trajectories probably come from smart phones, online check-in data, geo-tagged messages or media in social networks, RFID readers, and so on. Consequently, moving objects can be human beings, animals, vehicles, and even natural phenomena (e.g., hurricanes).There exist a wide spectrum of applications driven and improved by trajectory data mining, such as path discovery, location/destination prediction, movement behaviour analysis for individual or a group of moving objects, making sense of trajectories and other applications of urban service. These applications significantly benefit the common people, commercial organizations and government agencies. [3]

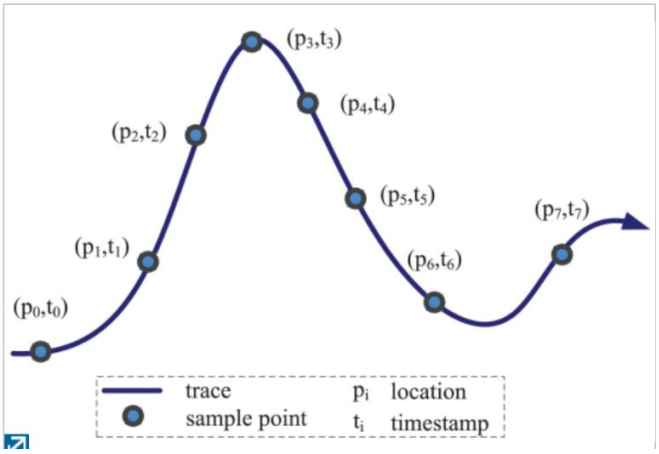

Fig. 1.1. A trajectory is generated by sampling from a continuous trace.

It is unavoidable to extract movement patterns from trajectory data because of many reasons. Finding trajectory patterns is very useful in learning interactions between moving objects. Varieties of trajectory patterns have been proposed in the literature of trajectories. Trajectory data are ubiquitous in the real world and there is an ever increasing interest in performing data analysis over trajectory data.

II. FRAMEWORK OF TRAJECTORY DATA MINING

Firstly, trajectory data are generated by various moving objects and collected from multiple data sources. Trajectory mining techniques are presented with five components, i.e., pre-processing, data management, query processing, trajectory data mining tasks, and privacy protection. Under the circumstances, trajectory data mining has become an increasingly important research theme, attracting the attention from numerous areas, including computer science, sociology, and geography.

The layer of trajectory data mining techniques is organized as follows. Pre-processing attempts to improve quality of trajectory data and to partition trajectories into sub-trajectories for further processing. Data management solves

the problem of storing a huge amount of trajectory data in an efficient and scalable manner. Query processing aims to retrieve appropriate data from the underlying storage system efficiently. The component of trajectory data mining tasks summarizes several important types of mining tasks. Protecting privacy of users with privacy-preserving techniques is an essential problem throughout these four components above and thus it can be combined with any component.

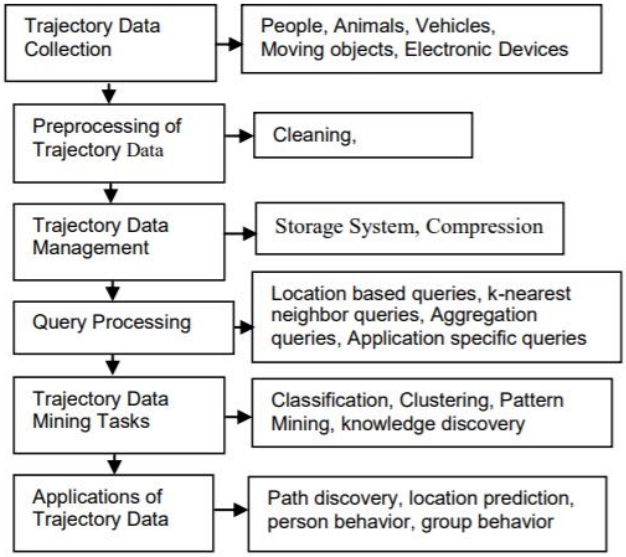

Fig. 2.1: A general framework of data mining process

There exist a wide spectrum of applications driven and improved by trajectory data mining and these applications significantly benefit the common people, commercial organizations and government agencies. Trajectory data mining techniques are directly related to the collected trajectory data and the requirements of the applications. Trajectory data mining techniques are divided into six categories [4]:

2.1. Trajectory data collection

Based on the technology by which they are recorded, mobility data is available in different forms: GPS (global positioning system), GSM (global system for mobile communications), and geo-social network based trajectory data. Other forms are; RFID (radio frequency identification) based and Wi-Fi based data. GPS based data is composed of temporally ordered sequences of geographic coordinates recorded by a GPS-enabled device carried by the moving object. [5]

2.2. Pre-processing of Trajectory Data

Trajectory data pre-processing includes trajectory data cleaning and it is the most important and prerequisite requirement for any trajectory data mining technique. The performance of trajectory data mining decreases as the number of outliers increases in the trajectory data.. Main goal of trajectory data cleaning is to find and remove outliers at the beginning itself. Uncertain trajectory data sets must be used with the help of other rules and constraints. Trajectory data completion is a way of improving the quality of uncertain trajectories before actually usage of them in any trajectory data mining.

2.3. Trajectory data management

How to store a huge amount of trajectory data is a fundamental problem in trajectory data mining.As some location points in a trajectory are often redundant, trajectory compression algorithmsare promising to reduce storage requirements and communication loads. A compression approach often makes a trade-off between compression ratio and maximum error. Generally, the higher the compression ratio, the poorer the quality of compressed trajectory data. Typical approaches are line generation and delta compression.

A dynamic indexing structure is needed to maintain tremendous amount of trajectory databases. B+ tree is the state-of-the-art indexing data structure for relational databases but it is one dimensional. R+ tree, R* tree and R tree are spatial indexing data structures.

2.4. Query processing

Retrieving data from an underlying storage system is a crucial operation. The objective of retrieval is to find appropriate data efficiently.A location-based query attempts to find trajectories that are close to all query locations where the query is a small set of locations with or without a specific order constraint.

A range query always gives two bounds - lower and upper bounds. There exist many applications of trajectory data management that involves range queries. One example for range query is to find all trajectories that contain a set of cities in the path trajectory such that starting from the 9th city to 18th city. Range queries are particularly suitable for handling uncertain trajectories. [6]

Aggregate queries are useful to find statistical details of trajectories stored in the trajectory databases. Total or average number of persons travelled on a specific road segment in a given time interval is an example for aggregate query. Average velocity, maximum and minimum velocity of

a set of vehicle trajectories on a particular road segment is another aggregate query.

2.5. Trajectory data mining tasks

Most important and frequently used data mining tasks are: classification, clustering, pattern mining, knowledge discovery, frequent pattern mining and so on. Trajectory data mining tasks are classified as:

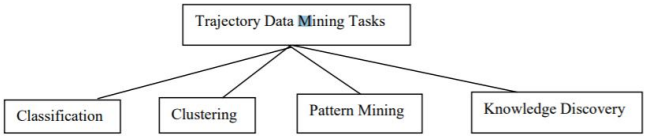

Fig. 2.2. The most frequent tasks in data mining

Pattern mining is to analyse mobility patterns of a moving object or multiple moving objects together. There are various types of patterns, such as gathering/group patterns, sequential patterns and periodic patterns. Knowledge that is mined in trajectory data mining may be used in many places and in many applications.

Classification of trajectories is to build a model on training data and then to apply the trained model to predict the labels of test trajectories. Main focus is on a classification problem and introduce duration information to boost prediction accuracy. Privacy protection is one of the important tasks in trajectory data mining. [7]

III. APPLICATIONS OF TRAJECTORY DATA MINING

Potential applications of data mining are retailing, banking, credit card management, insurance, telecommunications, telemarketing and human resource management. There exist a many applications of similarity search on spatio-temporal trajectories. A suitable distance function is used for finding similarity between uncertain trajectories. Trajectory clustering algorithms for moving objects are very useful in finding traffic jams, important location identification, and facilities available at a particular location at particular time etc.

Some of the applications of trajectory data mining are Route Recommendation, Animal Migration, Transportation Management, Real time Traffic Information Details of Transport Organization, and Tourism. Some of the useful trajectory data mining applications are path discovery, shortest path discovery, individual behaviour prediction, group behaviour prediction, location prediction, service prediction and so on. Applications of trajectory data mining can be classified based on domain of application as follows:

3.1. Path Discovery

Path discovery, also called route discovery, is to find at least one path that satisfy a predefined objective given a source and a destination. Routes must be derived based on a specific road network. Furthermore, geographical locations in numerical style in trajectories should be matched to a map in order to derive candidate paths or path segments. Historical trajectories on the road network provide valuable intelligence to estimate, compare and even construct candidate routes.

3.2. Movement Behavior Analysis

A set of Trajectory may represent movement behaviour of a single person or movement behaviour of a group of persons. Movement behaviour of moving objects is represented in the form of trajectories and then analysed and processed before taking crucial decisions. Finding movement patterns is very important in order to understand human behaviour clearly. Predicting human behaviour accurately under emergency is a crucial issue for disaster alarming, disaster management, disaster relief and societal reconstruction after disasters. Human mobility behaviour can be studied from spatial, temporal and social aspects. [8]

3.3. Location/Destination Prediction

Location based services (LBSs), also called location-aware services, are increasingly beneficial to people in urban areas. It has been revealed that human mobility is extraordinary regular and thus predictable. Many location based applications require location prediction or destination prediction to send advertisements to targeted consumers, to recommend tourist spots or restaurants, or to set destinations in navigation systems.

Destination prediction is closely related to path discovery. If an ongoing trip matches part of a frequent route in a dataset of historical trajectories, the destination of the frequent route is possibly the destination of the ongoing trip. However, there exist a few constraints in real world scenarios.

3.4. Urban Service

The knowledge discovered after applying trajectory data mining techniques is useful in many ways for urban area people. Large scale trajectory data are collected by using electronic techniques. People living in a city need the knowledge to assist their decision on buying or renting a house, choosing a job. Meanwhile, the knowledge helps urban planners to make decisions on future development of the city

and to estimate effects of previous policies. There exist a great many studies to facilitate interpretation of raw trajectory data.

3.5. Making Sense of Trajectories

Raw trajectory data which are in the form of sequence of geographical locations and timestamps fail to make sense to people without semantic description. There exist a great many studies to facilitate interpretation of raw trajectory data.

Unlike semantic trajectory that cannot express movement properties of moving objects, e.g., overspeed, stopover, a method was proposed a partition-and-summarization approach that automatically generates a short human-readable text to describe a trajectory. The approach not only extends expressivity of traditional semantic trajectories but also avoids a challenging problem of storage, processing and transmission of large volume of semantic trajectories.

3.6. Group Behavior Analysis

Moving objects, especially people and animals, sometimes tend to form groups or clusters due to their social behaviour. For instance, movement of a person is affected by not only personal activities, but also social ties with that of the groups he belongs to. Besides, a gathering pattern, as a novel modelling of trajectory patterns, describes movement pattern of a group of moving objects. Examples include celebrations, parades, traffic congestion, large-scale business promotions, protests etc.

Another study also aims at efficiently discovering moving objects which move together. A group is defined as a cluster that at least m moving objects being densely connected for at least a certain duration of time. It is very different from gathering meaning aforementioned. Besides, a sampling-independent approach is proposed to avoid flaws of sampling dependent ones, e.g., convoy, swarm.[9]

IV. CONCLUSION

The current trends and opportunities with trajectory data mining in general and with trajectory data clustering in particular are addressed. The opportunities for the research towards community movement clustering are identified. The high amount of scope and need is there in future to analyse the current data trends to go for better decisions and knowledge of movement behaviours of moving objects obtained by analysing and processing historical trajectories.

Location acquisition technologies generate huge amount of trajectory data. Trajectory data which track traces of moving objects is typically represented by a sequence of time stamped geographical locations. A large amount of applications are created upon mining trajectory data. This paper reviews an extensive collection of existing studies in the proposed framework of trajectory data mining.

REFERENCES

[1] Nykamp DQ, “Trajectory definition.” From Math Insight. http://mathinsight.org/definition/trajectory_dynamical_systems
[2] V. Tanuja & P. Govindarajulu, A Survey on Trajectory Data Mining, International Journal of Computer Science and Security (IJCSS), Volume (10) : Issue (5) : 2016
[3] Zhenni Feng, Yanmin Zhu, A Survey on Trajectory Data Mining: Techniques and Applications, IEEE, Volume: 4, 10.1109/ACCESS.2016.2553681
[4] Namgu Kwon, Jibum Kim, Sunho Lim, Jinseok Chae, Heemin Park, T-Reduce: Route-Aware Mobile Trajectory Data Reduction in Transportation Management Systems, IEEE Access, 10.1109/ACCESS.2018.2852787
[5] Shaojie Qiao, Huidong Jin, Yunjun Gao, Lu-An Tang, Huanlai Xing, Trajectory Data Mining in Distributed Sensor Networks, July 9, 2015, https://doi.org/10.1155/2015/913165
[6] Zheng, Yu. (2015). Trajectory Data Mining: An Overview. ACM Transactions on Intelligent Systems and Technology. 6. 1-41. 10.1145/2743025.
[7] Mazimpaka, Jean Damascene & Timpf, Sabine. (2016). Trajectory data mining: A review of methods and applications. Journal of Spatial Information Science. 13. 61–99.
[8] Oliviu Ghica, Goce Trajcevski, Ouri Wolfson, Ugo Buy, Peter Scheuermann, Fan Zhou, Dennis Vaccaro, Trajectory Data Reduction in Wireless Sensor Networks, International Journal of Next-Generation Computing, Vol. 1, No. 1, July 2010.
[9] Yingqi Xu & Wang-Chien Lee, Compressing Moving Object Trajectory in Wireless Sensor Networks, 14 Mar 2007, International Journal of Distributed Sensor Networks, Volume 3, 2007 - Issue 2: Sensor Networks, Ubiquitous and Trustworthy