Journal of Computational Information Systems 10: 16 (2014) 6947–6956
Available at http://www.Jofcis.com

GPS Trajectory Mining: a Survey ⋆

Haibiao LIN, Jianming LV, Can YANG*, Miaoyi DENG, Kaitao WANG, Xintong WANG

School of Computer Science and Engineering, South China University of Technology, Guangzhou 510006, China

Abstract

With the huge amount of GPS trajectories collected by GPS devices, GPS trajectory mining has become a fertile area for research, which covers several research communities such as database, information retrieval and AI. In this paper, we survey the area of GPS trajectory mining and present a global view of the key steps in the mining procedure. Specifically, we focus on the basic mining models, common preprocessing mechanisms and some interesting algorithms in this field.

Keywords: GPS; Trajectory Mining; Algorithms

1 Introduction

In recent years, due to the continuing improvements in location-acquisition technologies especially GPS, large amounts of historical trajectory data have become available for emerging applications [1-4] such as urban planning, spatio-temporal data mining, location-based services and behaviour prediction. Lots of useful information have been achieved by analysing these trajectories.

Typically, a GPS trajectory consists of a sequence of the spatio-temporal points, each of which is formatted as a tuple $(latitude, longitude, timestamp)$. Due to measurement errors and varieties of sampling rate, the collected trajectory is usually not very precise. Some recent researches [5, 6] present data cleaning methods to achieve better results. Besides, plenty of efficient mapping methods [7, 8, 9] are proposed to align GPS trajectories to road networks accurately. All of above works are classified as *GPS trajectories preprocessing* in this paper.

Based on the preprocessing of GPS trajectories, more researches are focused on applications of mining the trajectories, such as route planner [10-16], hot route finder [17, 18, 19, 5, 20, 21, 22],

⋆The work described in this paper was supported by the grants from National Natural Science Foundation of China (61300221), the Fundamental Research Funds for the Central Universities (2014ZZ0038), the Comprehensive Strategic Cooperation Project of Guangdong Province and Chinese Academy of Sciences (2012B090400016) and the Technology Planning Project of Guangdong Province (2012A011100005).

*Corresponding author.
Email address: cscyang@scut.edu.cn (Can YANG).

DOI: 10.12733/jcis11188
August 15, 2014

traffic flow analysis [23, 24], etc. Especially, the researches [22, 21, 25] to find the most desirable path, play a crucial role in smart navigation applications.

The GPS mining research is a converging research area from several research communities such as database, information retrieval and AI research communities. This paper presents an attempt to put the related researches in a more structured way. Since this is a huge and very dynamic research area, there are undoubtedly some omissions in our coverage.

The rest of the paper is organized as follows. Section 2 presents some definitions and gives a basic model of GPS trajectory mining. GPS trajectories preprocessing methods including mapping and smoothing methods are shown in Section 3. Then some GPS trajectory mining methods are introduced in Section 4. Finally, we conclude the paper with Section 5.

2 Basic Model of GPS Trajectory Mining

For clarity, we will firstly introduce some common concepts mentioned in the GPS trajectory mining mechanisms. Then we will discuss a basic model to explain the whole procedure of trajectory mining including preprocessing and main mining algorithms.

In the real world, a moving object's trajectory is continuous and usually called *original route* [6]. However, it is hard to acquire the continuous trajectories through existing positioning techniques and store them in a real database. Thus an alternative feasible solution taken in GPS-enabled services is to store only a set of sampled positions of a trajectory, which is formatted as a tuple $T_r =< p_1, p_2, ..., p_i, ... >$. p_i is a GPS sampled spatial-temporal point, which is a tuple $< t, l, s, e >$. Here t is the timestamp when the data is collected, l is the geographic location of the object, s means the instantaneous speed of the object, and e contains some extra features. This kind of sampled trajectories are called *raw GPS trajectories* in the research [6]. Variety of sampling strategies are adopted in different devices to achieve trajectories. The most widely used ones are time-based sampling, distance-based sampling, turning-based sampling and prediction-based sampling.

Taking the raw GPS trajectories as input, the GPS trajectory mining model consists of two basic phases: offline analysing and online querying as illustrated in Fig. 1. In the offline phase, there are two main sub-stages including trajectory preprocessing and mining procedure. In the online phase, applications provide different services to users based on the mining results.

As we know, GPS records usually contain some useless information which makes the original record dataset large and redundant. Furthermore, due to the errors caused by data sampling and encryption in GPS navigation services, many GPS records are not precise and generally needed to be aligned with road networks or some points of interest (POI) before trajectory mining. Thus a trajectory preprocessing procedure (stage 1 in Fig. 1) is necessary to filter useless information and perform map-matching [8, 9]. The detailed preprocessing procedure will be presented in Section 3.

After the preprocessing stage, we can obtain the GPS trajectories which contain a sequence of tuples about an object's spatio-temporal information including time, location, speed, etc. By analysing these processed trajectories, some important information can be retrieved at the mining stage (stage 2 in Fig. 1), which can be applied for different application services such as hot region detection, traffic outliers detection and smart navigation. All above trajectory mining procedures run offline and provide knowledge information to the users through the online querying system

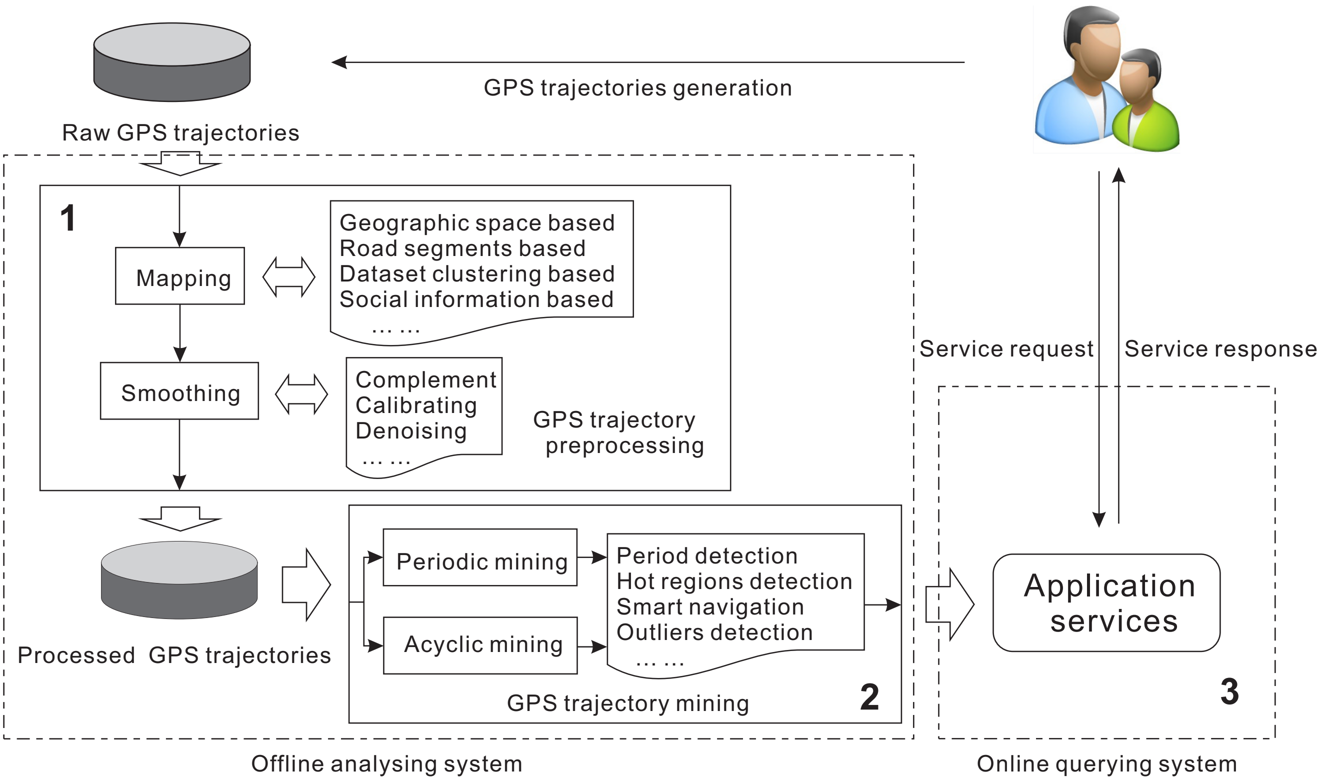

Fig. 1: The basic model of GPS mining process

(stage 3 in Fig. 1). The detail will be introduced in Section 4.

3 GPS Trajectories Preprocessing

This section mainly describes the GPS trajectories preprocessing phase including mapping and smoothing techniques.

3.1 GPS trajectories mapping

In the raw trajectories, there exist lots of redundant information and measurement errors, which can bring serious side effects on the data accuracy of trajectory mining. The Map-matching methods [7, 8, 9] are typical preprocessing mechanisms to clean trajectories. These methods construct a *reference graph* containing a set of *reference points*, each of which is a fixed spatial location in the space depending on some partition strategies. The reference graph can be considered a simplified version of real road networks. Each raw GPS trajectory is converted to a simplified trajectory which only contains reference points. It is important to construct a reasonable reference points dataset to reduce the loss of information after converting.

Actually, there are a lot of division strategies to construct the reference point dataset with different resolutions and density distributions. Here we will introduce some typical methods, including geographic space division, road segments division, dataset clustering division and social information division as follows.

Geographic space division Geographic space division [6, 21] is to divide a whole geographic space into uniform grid cells, and take the centroids of the cells as reference points. An obvious advantage of this method is that we can quickly divide a digital map into cells and get the reference points without any extra road network information or trajectory dataset.

Road segments division Road segments division [9, 8] is to divide the whole map into different road segments and use each road segment number as a reference point. In this method, a road network information should be achieved first.

Dataset clustering division Dataset clustering division [5, 21] is to select a large set of historical GPS points, then cluster these points and use the clustering center points as the reference points. This method aims to find out the top-n (n means the predetermined number of clustering centers) highest density points according to different trajectories. The drawback of this method is that it is less stable than other methods due to the dynamic property of clustering.

Social information division Social information division [6] is also a method to determine the reference points according to extra social information. In real life, some semantic locations is well known such as restaurants, hotels, shopping centers, etc. These locations are usually called *Landmarks* or *POIs*, which are stable in terms of their locations and are also frequently visited by people. The research [6] adopts the POIs as reference points. Furthermore, some density-based clustering methods are also proposed to merge nearby POIs into a single reference point to reduce computing complexity.

Based on the reference points, some methods are presented to map GPS trajectory points to reference points and they can be classified into two categories: local/incremental methods [25, 26, 27] and global methods [28, 29, 30, 31, 8, 9, 6].

The local/incremental methods map the GPS points to reference points one by one, and only consider the current or neighbour positions while performing mapping. These methods can run fast, but sensitive to the decrease of sampling frequency. The global methods aim to match the entire trajectory with road networks based on some distance/similarity measures of trajectories to achieve more precise results. Recently, some researches [31, 6, 9, 8] develop statistical models to improve the mapping precision by incorporating historical information. Specifically, in [31] a Hidden Markov Model based method is presented to model topological constraints of road networks. Similarly, the researches [6, 9, 8] find out the global optimum routes by constructing a probability state space of trajectories.

3.2 GPS trajectories smoothing

Besides trajectories mapping mentioned above, some smoothing mechanisms [7, 29, 32, 33, 6] are proposed to make trajectories more reasonable by reducing outliers and adding missing points.

Specifically, Brakatsoulas et al. [29] incorporate the average Frechet distance to reduce the side effects caused by outliers. Lou et al. [8] leverage both the spatial topological structures of road networks and some temporal/speed constraints of trajectories to reduce outliers. Moreover, Chen et al. [5] present a direction smoothing method to alleviate the noise of position fluctuation.

In order to solve the problem of missing points caused by low sampling rate, Chen et al. [5] conduct linear interpolation to complement points to keep trajectories continuous. Su et al. [6]

complement missing points of a low resolution trajectory according to some similar trajectories which have high sampling rate. In this research, a calibration system based on spatial geometry and historical statistics is proposed to complement trajectories by using reference points.

4 GPS Trajectory Mining Algorithms

In this section, we list some researches on GPS trajectory mining in recent years. We try to offer a global view of the variety of representations, processes, methods, and applications in this field.

So far the research of GPS trajectory mining can be roughly classified into two categories: periodic pattern mining and acyclic pattern mining. In periodic pattern mining, researches are focused on finding some cyclical phenomena and patterns. On the contrary, the acyclic pattern mining pays attention to discovering non-periodic frequent patterns, e.g., detecting hot regions and retrieving special routes.

Based on the trajectories mining algorithms, some special patterns can be retrieved to support trajectory classification, prediction and intelligent decision.

4.1 Periodic pattern mining

In general, periodicity can be found everywhere among GPS trajectories and it reveals the regularity of objects' behaviours. Recently, some periodic pattern mining techniques [34-38] have been proposed. Specifically, Yang et al. [34-36] propose a series of works to deal with variations of periodic pattern mining. They try to discover different kinds of patterns such as asynchronous patterns in [34], surprising periodic patterns in [35] and patterns with gap penalties in [36].

Above periodic pattern mining methods are all based on some global minimum support threshold, and output a large set of patterns with similar periods. However, in real life, most of observed movements are generated from multiple interleaved periodic behaviours. In order to discover different periodic behaviours of moving objects, Li et al. [39] combine Fourier transform and autocorrelation methods to retrieve the values of multiple periods, and present a probabilistic model to characterize the periodic behaviours.

4.2 Acyclic pattern mining

While mining GPS trajectories, lots of acyclic information can be retrieved such as hot regions, optimal path, etc. Here we will illustrate three kinds of representative research about acyclic mining.

4.2.1 Hot region discovering

Hot regions are those areas where many moving objects stay for a variety of reasons. Identifying hot regions of these moving objects is essential for many intelligent services.

Some researches [17] detect hot regions by clustering GPS trajectory points based on spatial density. Liu et al. [40] propose a novel mobility-based clustering approach, which uses the speed information to infer the crowdedness of moving objects. Lu et al. [41] explore both spatial and temporal relationships among trajectory points to extract hot regions. In this research, a sequential clustering approach based on spatial-temporal density is proposed to discover semantic regions from individual trajectories. Furthermore, a shared nearest neighbour (SNN) based clustering algorithm is also presented to group similar semantic regions. More over, Yuan et al. [42] propose a framework (named DRoF) to discover regions of different functions in a city, such as educational areas, entertainment areas, and regions of historic interests.

4.2.2 Road segment outliers mining

The detection of road segment outliers in GPS trajectory dataset is an important research direction in the data mining community. Work et al. [23] propose a method to monitor highway traffic condition and estimate traffic states. In their research, a Kalman filtering approach is proposed to estimate the velocity field on the highway. Through analysing the velocity, traffic monitoring can be implemented and road segment outliers can be detected. Similarly, Hunter et al. [43] forecast future traffic conditions by monitoring real time traffic flows on some road segments. In their research, an expectation maximization algorithm is taken to learn the historical traffic conditions and applied to detect road segment outliers. Liu et al. [40] propose a mobility-based clustering algorithm to identify road segment outliers caused by traffic jam. This research is based on the assumption that vehicles are driven fast in a sparse region, while moving slowly in crowded areas for security concerns.

However, the relationships, especially causal interactions among detected traffic outliers is still not investigated in above researches. To solve this problem, Liu et al. [24] construct outlier causality trees based on temporal and spatial properties of detected outliers. Frequent substructures of the causality trees reveal not only recurring interactions among outliers, but also potential flaws in the design of existing traffic networks.

4.2.3 Transportation navigation

Navigation services aim to help users to find some special routes between two locations, which has shortest distance, costs shortest time, passes through least road segments, or is visited by most people.

Chen et al. [5] investigate the problem of discovering the Most Popular Route (MPR) between two locations by observing the historical travelling behaviours of many users. MARKOV train is adopted here to calculate the probability of each route from the source to the target location, and the route with maximum probability is chosen as MPR. Based on above research, Luo et al. [22] design a algorithm to find out the most frequent path (MFP) during user-specified time periods. The MFPs retrieved by this method have following optimal properties: 1) *suffix-optimal*: any suffix of an MFP is also an MFP; 2) *length-insensitive*: the selection of MFPs is independent of the length of paths; 3) *bottleneck-free*: MFP does not contain infrequent edges.

However, the above research [5] needs high resolution trajectories to cluster raw trajectories, and the research [22] depends on an extra precise road network to construct the reference points. To solve these problems, Wei et al. [21] propose a new route inference framework based on Collective Knowledge (RICK), which constructs the popular routes from uncertain trajectories. Firstly, they explore the spatial and temporal characteristics of low sampling uncertain trajectories to construct a routable graph. Then a robust routing algorithm is proposed to construct the top-k

routes according to user specified queries.

Recently, some researches [11, 44, 13, 3, 12, 14, 16] focus on finding fastest path based on statistics of historical GPS trajectories. Especially in [14, 16], Yuan et al. propose a cloud based system for computing fast driving routes by analysing the trajectories of a large number of GPS-equipped taxis and GPS-enabled phones. In this paper, a Variance-Entropy-Based Clustering approach is devised to estimate the distribution of travel time between two landmarks in different time slots. Based on the same experimental dataset, Yuan et al. [15] also propose an recommendation algorithm to provide taxi drivers with optimal parking places, which are more likely to pick up passengers to maximize profit. A probabilistic model is constructed in this method to estimate the profit of different parking places for particular drivers. This recommend system is also able to suggest some locations to passengers where they can easily find vacant taxis.

5 Conclusion

In this paper we survey the research in the area of GPS trajectory mining in recent years. We point out the basic model of GPS trajectory mining. We also introduce the GPS trajectory preprocessing and mining categories and then situate some of the research with respect to these categories.

Moreover, GPS trajectory mining also brings some new challenges to traditional data mining algorithms due to the big volume and fast updating property of trajectory data. Thanks to the cloud computing techniques, GPS trajectory mining can be more efficient and popularly deployed in the future.

References

[1] Yu Zheng, Like Liu, Longhao Wang, and Xing Xie. Learning transportation mode from raw gps data for geographic applications on the web. WWW '08, pages 247-256, New York, NY, USA, 2008. ACM.

[2] Yu Zheng, Lizhu Zhang, Xing Xie, and Wei-Ying Ma. Mining interesting locations and travel sequences from gps trajectories. In 18th international conference on World wide web, pages 791-800, New York, NY, USA, 2009. ACM.

[3] Jing Yuan, Yu Zheng, Xing Xie, and Guangzhong Sun. Driving with knowledge from the physical world. In 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '11, pages 316-324, New York, NY, USA, 2011. ACM.

[4] Bin Yang, Nicolas Fantini, and Christian S. Jensen. ipark: Identifying parking spaces from trajectories. In 16th International Conference on Extending Database Technology, EDBT '13, pages 705-708, New York, NY, USA, 2013. ACM.

[5] Zaiben Chen, Heng Tao Shen, and Xiaofang Zhou. Discovering popular routes from trajectories. In 27th IEEE International Conference on Data Engineering, ICDE '11, pages 900-911, 2011.

[6] Han Su, Kai Zheng, Haozhou Wang, Jiamin Huang, and Xiaofang Zhou. Calibrating trajectory data for similarity-based analysis. In ACM SIGMOD International Conference on Management of Data, pages 833-844, 2013.

[7] MohammedA. Quddus, WashingtonYotto Ochieng, Lin Zhao, and RobertB. Noland. A general map matching algorithm for transport telematics applications. GPS Solutions, 7(3): 157-167, 2003.

[8] Yin Lou, Chengyang Zhang, Yu Zheng, Xing Xie, Wei Wang, and Yan Huang. Map-matching for low-sampling-rate gps trajectories. In 17th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems, pages 352-361, New York, NY, USA, 2009. ACM.

[9] Jing Yuan, Yu Zheng, Chengyang Zhang, Xing Xie, and Guang-Zhong Sun. An interactive-voting based map matching algorithm. In 11th International Conference on Mobile Data Management, pages 43-52, 2010.

[10] L. Fu, D. Sun, and L.R. Rilett. Heuristic shortest path algorithms for transportation applications: State of the art. Computers & Operations Research, 33(11): 3324-3343, 2006. Part Special Issue: Operations Research and Data Mining.

[11] E. Kanoulas, Yang Du, Tian Xia, and Donghui Zhang. Finding fastest paths on a road network with speed patterns. In 22nd International Conference on Data Engineering, ICDE '06, pages 10-10, 2006.

[12] Bolin Ding, Jeffrey Xu Yu, and Lu Qin. Finding time-dependent shortest paths over large graphs. In 11th International Conference on Extending Database echnology: Advances in Database Technology, EDBT '08, pages 205C216, New York, NY, USA, 2008. ACM.

[13] Ugur Demiryurek, Farnoush Banaei-Kashani, and Cyrus Shahabi. A case for time-dependent shortest path computation in spatial networks. In 18th SIGSPATIAL International Conference on Advances in Geographic Information Systems, GIS '10, pages 474-477, New York, NY, USA, 2010. ACM.

[14] Jing Yuan, Yu Zheng, Chengyang Zhang, Wenlei Xie, Xing Xie, Guangzhong Sun, and Yan Huang. T-drive: driving directions based on taxi trajectories. In 18th SIGSPATIAL International Conference on Advances in Geographic Information Systems, pages 99-108, New York, NY, USA, 2010. ACM.

[15] Jing Yuan, Yu Zheng, Liuhang Zhang, XIng Xie, and Guangzhong Sun. Where to find my next passenger. In 13th international conference on Ubiquitous computing, pages 109-118, New York, NY, USA, 2011. ACM.

[16] Jing Yuan, Yu Zheng, Xing Xie, and Guangzhong Sun. T-drive: Enhancing driving directions with taxi drivers' intelligence. IEEE Transactions on Knowledge and Data Engineering, 25(1): 220-232, 2013.

[17] Fosca Giannotti, Mirco Nanni, Fabio Pinelli, and Dino Pedreschi. Trajectory pattern mining. In 13th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pages 330C339, New York, NY, USA, 2007. ACM.

[18] Xiaolei Li, Jiawei Han, Jae-Gil Lee, and Hector Gonzalez. Traffic density-based discovery of hot routes in road networks. volume 4605 of Lecture Notes in Computer Science, pages 441-459. Springer Berlin Heidelberg, 2007.

[19] Gyz Gidfalvi and TorbenBach Pedersen. Mining long, sharable patterns in trajectories of moving objects. GeoInformatica, 13(1): 27-55, 2009.

[20] Hechen Liu, Ling-Yin Wei, Yu Zheng, M. Schneider, and Wen-Chih Peng. Route discovery from mining uncertain trajectories. In 11th IEEE International Conference on Data Mining Workshops, ICDMW '11, pages 1239-1242, 2011.

[21] Ling-Yin Wei, Yu Zheng, and Wen-Chih Peng. Constructing popular routes from uncertain trajectories. In 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pages 195-203, 2012.

[22] Wuman Luo, Haoyu Tan, Lei Chen, and Lionel M. Ni. Finding time period-based most frequent path in big trajectory data. In 2013 ACM SIGMOD International Conference on Management of Data, SIGMOD '13, pages 713-724, 2013.

[23] D.B. Work, O.-P. Tossavainen, S. Blandin, A.M. Bayen, T. Iwuchukwu, and K. Tracton. An ensemble kalman filtering approach to highway traffic estimation using gps enabled mobile devices. In 47th IEEE Conference on Decision and Control, CDC '08, pages 5062-5068, 2008.

[24] Wei Liu, Yu Zheng, Sanjay Chawla, Jing Yuan, and Xie Xing. Discovering spatio-temporal causal interactions in traffic data streams. In 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '11, pages 1010-1018, 2011.

[25] S.S. Chawathe. Segment-based map matching. In Intelligent Vehicles Symposium, 2007 IEEE, pages 1190-1197, 2007.

[26] Joshua S Greenfeld. Matching gps observations to locations on a digital map. In National Research Council (US). Transportation Research Board. Meeting (81st: 2002: Washington, DC). Preprint CD-ROM, 2002.

[27] C. Wenk, R. Salas, and D. Pfoser. Addressing the need for map-matching speed: Localizing global curve-matching algorithms. In 18th International Conference on Scientific and Statistical Database Management, pages 379-388, 2006.

[28] Helmut Alt, Alon Efrat, Gunter Rote, and Carola Wenk. Matching planar maps. In 14th Annual ACM-SIAM Symposium on Discrete Algorithms, SODA '03, pages 589-598, Philadelphia, PA, USA, 2003. Society for Industrial and Applied Mathematics.

[29] Sotiris Brakatsoulas, Dieter Pfoser, Randall Salas, and Carola Wenk. On map-matching vehicle tracking data. In 31st International Conference on Very Large Data Bases, VLDB '05, pages 853-864. VLDB Endowment, 2005.

[30] Huabei Yin and O. Wolfson. A weight-based map matching method in moving objects databases. In 16th International Conference on Scientific and Statistical Database Management, pages 437-438, 2004.

[31] O. Pink and B. Hummel. A statistical approach to map matching using road network geometry, topology and vehicular motion constraints. In 11th IEEE International Conference on Intelligent Transportation Systems, ITSC '08, pages 862-867, 2008.

[32] Kai Zheng, Yu Zheng, Xing Xie, and Xiaofang Zhou. Reducing uncertainty of low-samplingrate trajectories. In 28th IEEE International Conference on Data Engineering, ICDE '12, pages 1144-1155, 2012.

[33] Michel Bierlaire, Jingmin Chen, and Jeffrey Newman. A probabilistic map matching method for smartphone gps data. Transportation Research Part C: Emerging Technologies, 26(0): 78-98, 2013.

[34] Jiong Yang, Wei Wang, and P.S. Yu. Mining asynchronous periodic patterns in time series data. IEEE Transactions on Knowledge and Data Engineering, 15(3): 613-628, 2003.

[35] Jiong Yang, Wei Wang, and Philip S. Yu. Infominer: Mining surprising periodic patterns. In Proceedings of the 7th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '01, pages 395-400, New York, NY, USA, 2001. ACM.

[36] Jiong Yang, Wei Wang, and P.S. Yu. Infominer+: mining partial periodic patterns with gap penalties. In IEEE International Conference on Data Mining, ICDM '02, pages 725-728, 2002.

[37] Nikos Mamoulis, Huiping Cao, George Kollios, Marios Hadjieleftheriou, Yufei Tao, and David W. Cheung. Mining, indexing, and querying historical spatiotemporal data. In 10th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '04, pages 236-245, New York, NY, USA, 2004. ACM.

[38] Huiping Cao, Nikos Mamoulis, and D.W. Cheung. Discovery of periodic patterns in spatiotemporal sequences. IEEE Transactions on Knowledge and Data Engineering, 19(4): 453-467, 2007.

[39] Zhenhui Li, Bolin Ding, Jiawei Han, Roland Kays, and Peter Nye. Mining periodic behaviors for moving objects. In 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '10, pages 1099-1108, 2010.

[40] Siyuan Liu, Yunhuai Liu, Lionel M. Ni, Jianping Fan, and Minglu Li. Towards mobility-based clustering. In 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '10, pages 919-928, 2010.

[41] Chun-Ta Lu, Po-Ruey Lei, Wen-Chih Peng, and Ing-Jiunn Su. A framework of mining semantic regions from trajectories. In Database Systems for Advanced Applications, pages 193-207, 2011.

[42] Jing Yuan, Yu Zheng, and Xing Xie. Discovering regions of different functions in a city using human mobility and pois. In 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '12, pages 186-194, 2012.

[43] T. Hunter, R. Herring, P. Abbeel, and A. Bayen. Path and travel time inference from gps probe vehicle data. Proc. Neural Information Processing Systems, 2009.

[44] Hector Gonzalez, Jiawei Han, Xiaolei Li, Margaret Myslinska, and John Paul Sondag. Adaptive fastest path computation on a road network: A traffic mining approach. In 33rd International Conference on Very Large Data Bases, VLDB '07, pages 794-805. VLDB Endowment, 2007.