Spatial and temporal mining method using GPS data

Xiaolei Li*

School of Electronic and Information, Ningbo Dahongying University, Ningbo, Zhejiang, 315175, China

Received 1 March 2014, www.cmnt.lv

Abstract

Geographic information has spawned many novel Web applications where global positioning system (GPS) plays important roles in bridging the applications and end users. Learning knowledge from users' raw GPS data can provide rich context information for both geographic and mobile applications. However, so far, raw GPS data are still used directly without much understanding. Spatial-temporal data analysis plays an important role in many applications, including transportation infrastructure, border security and inland security. To analyse the moving patterns of vehicles on a road network, a measure for determining the similarity of vehicle trajectories with respect to space and time has to be defined. Although previous research has addressed the trajectory similarity problem, most of the studies focus on Euclidian distance instead of network distance. This paper deals with the variations in applying a spatial-temporal similarity measure with given Points of Interest (POI) and Time of Interest (TOI), treating spatial similarity as a combination of structural and sequence similarities that is evaluated using the techniques of dynamic programming. The similarity set thus formed will be used by the remote database to broadcast trigger-based messages to participating vehicles in a neighbourhood for future route- and information-sharing activities. The performance of the scheme is evaluated using experiments on standard real-life data.

Keywords: K-means, clustering algorithm, error rate, iteration, reduction, stable

1 Introduction

The increasing pervasiveness of location-acquisition technologies, such as GPS and GSM network, is leading to the large collection of spatial-temporal datasets. Such datasets have supported a variety of novel Web applications, in which locality and mobility usually connect to one another closely. For instance, people can tag user-generated contents like photos with locations; trace their outdoor mobility; and use location-based services. Recently, a branch of GPS-track-sharing applications using Web maps appeared on the Internet. In this category of Web applications, people can record their travel routes using a GPS-equipped device and then share travel experiences among each other by publishing these GPS tracks in a Web community. GPS-track-sharing offer a more fancy and interactive approach than text-based articles to better express people's travel experiences, which provide users with valuable references when planning a travel itinerary.

However, so far, these applications require people either to manually label their own trajectories or to use raw GPS data such as GPS coordinates and timestamps without much understanding. Neither of these methods is optimal to the development of such applications. Actually, users become easily frustrated by the additional data labelling effort, and then give up uploading their data. Moreover, people intend to understand an individual's mobility, and learn information about user behaviours as well as user intentions behind the raw data. Being an important kind of human behaviour, transportation modes, such as walking, driving, and taking a bus, can enrich their mobility with knowledge and provide pervasive computing systems with more contexts.

In recent years, information technology has significantly penetrated surface transportation. The transportation environment is embedded with various mobile sensors, including on-board GPS receivers, sensors mounted on public transportation vehicles and pedestrian cell phones. These sensors continuously generate spatial-temporal data and enable applications such as vehicle tracking and environmental monitoring. Studying people and vehicle movements within a certain road network is both interesting and useful, especially if it can be used to understand, manage and predict the traffic flows. By studying the massive flow of traffic data as a trajectory, the traffic flow can be monitored and traffic-related patterns can be discovered. The development of Intelligent Transportation Systems (ITS) allows better monitoring and traffic control to optimize traffic flow.

Advances in social networking and location-based services are increasingly creating new, sophisticated mechanisms that can foster a seamless integration of information among travellers to provide alternatives and support sustainable economic and social policies. All of these issues raise challenges to the dissemination of information, such as safety, traffic, entertainment, service and content, in a vehicular network. In all of the above cases, a major challenge arises from the fact that the relevance of information changes with time and with the location of the vehicle. The quality of data is measured by their spatial-temporal relevance, which indicates the probability that the resource will still be available when the vehicle reaches it. For example, in the case of the

***Corresponding author** e-mail: xiaoleilizj@yeah.net

dissemination of service information (e.g., the availability of parking spaces), the temporal and spatial delivery of information must be carefully considered. In the moving object database, which will have the complete history of the vehicle's trajectory, a spatial-temporal trajectory similarity search process will identify vehicles with very similar movement patterns with respect to space and time. Based on such a similarity set, useful information can be disseminated among neighbouring vehicles in the same similarity set containing extensive applications (e.g., knowing customer preferences, re-routing and scheduling of further travel, and so on). One advantage of this system is that because the broadcast message came from an official central database, it will be more authentic, and thus the nodes can be confident in the messages they have received. However, when internal nodes share information based on the central broad-cast, the system has to go through a confident management process to ensure the accurate dissemination of information. This paper relates how the spatial-temporal similarity measure of vehicle trajectory will help to better disseminate information in vehicular networks for data mining applications.

In this article, we aim to automatically infer transportation modes, including driving, walking, taking a bus, and riding a bicycle, from raw GPS logs based on supervised learning. It is a step toward recognizing human behaviour and understanding user mobility for pervasive computing systems. Also, it is a step toward improving local/mobile applications on the Web and enhancing the connection between mobility and locality by mining knowledge from raw GPS data with minimal user efforts. The contributions of this work lie in the following three areas.

First, we propose a change point-based segmentation method. This method aims to partition each GPS trajectory into separate segments of different transportation modes, while maintaining a segment of one mode as long as possible. In addition, this segmentation method is capable of enhancing the reliability of our methodology facing the variable traffic conditions.

Second, from each segment, we identify a set of sophisticated features, such as direction change rate, velocity change rate, and stop rate. These features have few correlations with the velocity, hence are not affected by differing traffic conditions. These set of features can also be extended to other pervasive computing systems aiming to recognize human behaviour and understand user mobility.

Third, we conduct a graph-based post processing algorithm to further improve the inference performance. In this algorithm, we mine the common sense constraints of the real world and typical user behaviours on a location from user-generated GPS logs. Therefore, we are able to leverage this location-constrained knowledge as probabilistic cues, while maintaining our methodology being in-dependent of an additional database of road networks or points of interests.

Overall, the advantages of our method over the related works include two parts.

1) Our method is independent of other sensor data like GSM signal and heart rate, and map information, for example, road networks and bus stops, etc. Thus, it is generic to be deployed in a broad range of Web applications.

2) The model learned from the dataset of some users can be applied to infer GPS data from others; that is, it is not a user-specific model.

2 Related Works

In conventional approaches, the location information of a moving vehicle is expressed as a geometric coordinate (x, y) in two-dimensional space. However, Lee et al. (2007) express location information using both the hierarchical administrative district and road network in one-dimensional space, more accurately fitting the real world. For instance, if a moving object is in a building at latitude of 125.58 and longitude of 37.34, then it can be expressed as a set of fields according to an administrative district, such as the city, road name, or block (e.g., Seoul, Main Road, 165th block). The dimension reduction of spatial-temporal data management (Abraham and SojanLal, 2008) discusses two algorithms for binary encoding pro-cess: one for location encoding and one for converting a position represented as geometric coordinate into an equivalent binary string. Because the proposed similarity scheme is based on this encoding method, these basic algorithms are briefly discussed below.

Many recent studies regarding data management in vehicular networks and related topics discuss the VANET, while very few focus on the processing of spatial-temporal information. In the case of traffic and service information, Xu et al. (2004) suggest aging information with time and distance. Information is opportunistically pulled from neighbouring vehicles as a target vehicle moves in a given area. Such information needs to have time and location stamps. Because either the vehicle moves farther away from where the information is relevant or time elapses, the information is aged and eventually purged. This enables vehicles to maintain up-to-date information without taxing memory and other resources. Xu et al. and Wu (2004) discuss the propagation and survival of information in time and space. Because of the spatial-temporal relevance of information, a piece of information tended not to propagate beyond a specific boundary. In time, a given piece of information would propagate very quickly until it reached a maximum number of copies, at which point it would also rapidly decline. Wolfson et al. (2005) discuss the dissemination of resource discovery (e.g., parking spaces) using spatial-temporal information. Each disseminated report represents information about a spatial-temporal event, such as the availability of a parking slot or a cab request. Reports are disseminated by a peer-to-peer broadcast paradigm in which an object periodically broadcasts the reports it carries to encountered objects. The authors evaluated the value of resource information in terms of how much time is saved when using the information to discover a resource.

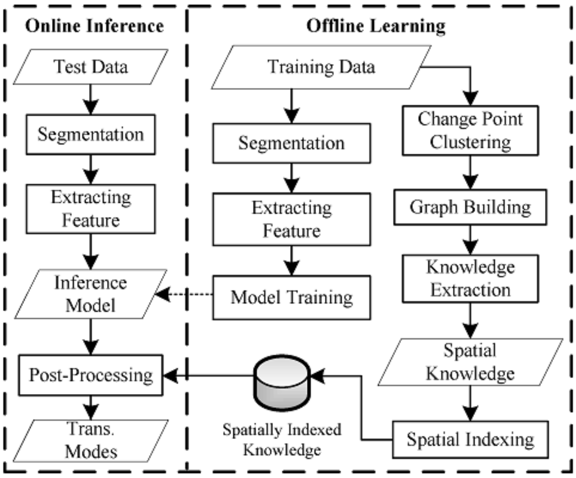

FIGURE 1 Architecture of our approach

The efficiency and reliability of information dissemination will be impacted by how vehicles are clustered when they form and leave groups, how apart they are under different traffic conditions, the density of vehicles and so on. Naumov et al. (2006) simulated realistic vehicular traces, which form trajectories. The authors demonstrated that the realistic traces are noticeably different than those in widely used mobility models. Szczurek et al. (2010) provide another promising approach to the dissemination of spatial-temporal information, such as the current traffic condition of a road segment or the availability of a parking space. Ranking becomes critical in this situation by enabling the most important information to be transmitted under the bandwidth constraint. Delot et al. (2010) proposed a vehicle-to-vehicle communication system for data sharing in vehicular networks using the concept of Encounter Probability to share information. The objective of this paper is to facilitate the dissemination of information between vehicles when they meet each other, taking into account the relevance of the data to the drivers and the type of event (e.g., available parking spaces, obstacles in the road and information relative to the coordination of vehicles in emergency situations) in the network.

Recently, some works (e.g., Delot et al., 2011) have addressed the problem of processing queries in a highly dynamic vehicular network in order to share information between vehicles. Queries interesting in pervasive and mobile computing environments (Vargas-Solar et al., 2010) are usually location-dependent (Ilarri et al., 2010). In vehicular networks, vehicles usually receive data from its neighbours and decides whether they are relevant enough to be stored in a local data cache. Then, a query processor can use the data to retrieve relevant data for the driver. Delot et al. (2011)discuss the challenges and possible solutions of multi-scale query processing techniques to exploit, at the mobile device's level, different access modes (e.g., wireless networks, push, pull) and various data sources (e.g., locally cached data, data stored by vehicles nearby and remote Web services) to provide users with results for their queries. Wolfson and Xu (2010) discuss various research issues related to the methods of routing, navigation, and tracking in transportation networks (the spatial-temporal database) that may involve multiple modes (e.g., train, bus, private car and bicycle).

3 The proposed spatial-temporal mining method

The algorithm for mapping locations in binary code is a recursive procedure, which will successively divide the entire region into sub-regions, map each district into a two-dimensional space and assign a binary string to each district. The total conversion time increases almost linearly with the number of two-dimensional locations, and more than seven million locations can be pro-cessed per second. The characteristics of the binary encoding scheme discussed above, which is used as a baseline in the proposed similarity scheme, are the following: It will be easy to determine the lowest common administrative district by extracting the longest common prefix of a given set of binary strings, and a district containing a set of lower districts can be represented by the range of binary string (i.e., county ''A'' will be represented by the range [00000, 00111]). These advantages make it easy to address the whole country, whole district or a single road by identifying the common prefix in the binary string representing the location of the object. For example, an object's location on the road may be encoded as 001010000110010, where first two bits denote the country, the next four bits denote the district, the following five bits denote the road and the final four bits denote the relative location on the road. Thus, if the first two bits in all of the locations are 00, they all belong to the same country.

3.1 ARCHITECTURE OF OUR APPROACH

As shown in Figure 1, the architecture of our approach includes two parts, offline learning and online inference. In the offline learning section, on one hand, we first partition GPS trajectories into segments based on change points and extract features from each segment. Then, the features and corresponding ground truths are employed to train a classification model for online inference. On the other hand, using a density-based clustering algorithm, we group the change points detected from all users' GPS logs into clusters. Subsequently, a graph based on these clusters and user-generated GPS trajectories is built.

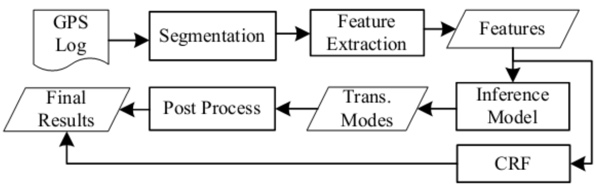

FIGURE 2 Procedure of inferring transportation mode

From this graph we can mine some location-constrained knowledge, such as the probability distribution of different transportation modes on each edge. The knowledge can be employed as probabilistic cues to improve the inference performance in the post processing. In addition, a spatial index is built over the detected spatial knowledge to enhance the processing efficiency. In the online operation, when a GPS trajectory comes, like the offline training process, we first partition it into segments and extract the same features from each segment. Second, given the features, the generative inference model will predict the transportation mode of each segment in a probabilistic manner. Third, given the probabilities of a segment being different transportation modes, a post-processing algorithm is used to improve the inference accuracy by leveraging the spatial knowledge mined from the training data. Finally, the transportation mode with maximum posterior probability will be selected as the ultimate result.

3.2 INFERENCE STRATEGY

As shown in Figure 2, when a GPS Log file comes, first, we divide the GPS track into trips and then partition each trip into segments by change points. Then, we extract the features from each segment and send these features to the inference model. Two alternative ways are considered when we attempt to learn a user's transportation mode. In one way, we regard the segments of GPS tracks as independent instances. General classifiers like Decision Tree are employed to perform inference. After the inference, a post-processing, which takes the transition probability between different transportation modes into account, is implemented to improve the prediction accuracy. In the other way, GPS data are deemed as a kind of sequential data. Conditional random field (CRF) [13], a framework for building probabilistic models to segment and label sequence data, is leveraged to perform the inference. Since the conditional probabilities between different transportation modes have been considered in the CRF graphical model, in this way, we do not take the post-processing. In the inference, the mode of transportation can take four different values including Bike, Bus, Car and Walk. At the same time, we do not discriminate driving private car from taking taxi. Both of them are deemed as Car.

3.3 SPATIAL KNOWLEDGE EXTRACTION

Figure 3 illustrates the four steps toward mining spatial knowledge from users' GPS logs. The knowledge includes a change point-based graph and the probability distribution on each edge of the graph. First, given GPS logs with labeled ground truths, we can get the special points consisting of change points and the start/end points of each GPS trajectory. These special points were subsequently grouped into several nodes (clusters) using a density-based clustering algorithm. The reasons we prefer to use density-based clustering instead of agglomerative methods, such as K-Means, lie in two aspects.

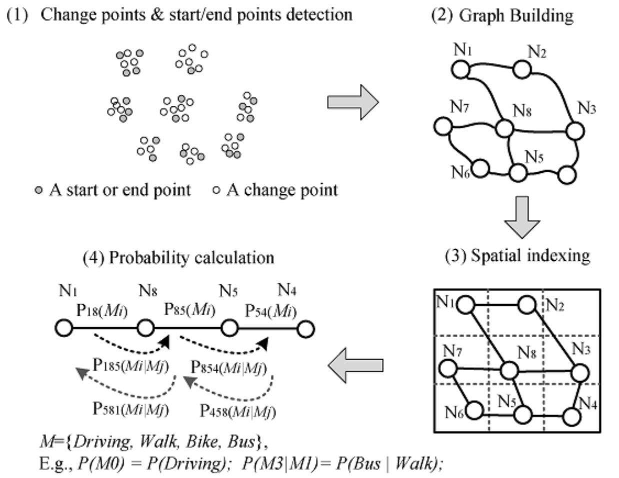

FIGURE 3 Mining spatial knowledge from GPS logs

First, a density-based approach is capable of detecting clusters with irregular structures, which may stand for bus stops or parking places. Second, it can discover popular places where most people change their transportation modes while removing sparse change points representing places with low access frequency.

Second, with the GPS trajectories from multiple users' GPS logs, we can construct an undirected graph. In such a graph, a node represents a cluster of the special points mentioned above, and an edge denotes users' transitions between two nodes. Here, we do not differentiate various trajectories with similar start/end points; that is, all the trajectories passing two graph nodes are regarded as similar trajectories.

Third, we build a grid-based spatial index over the graph to improve the efficiency of accessing the information of each node and each edge. The space covered by the graph is partitioned into many disjoint grids. Then, the graph nodes falling in different grids are associated with the grids. Therefore, when a new GPS trajectory comes to be inferred, we only need to match the special points detected from the trajectory against the graph nodes pertaining to the grids where these special points falling in. Of course, this step is optional unless the scale of the GPS dataset is quite large.

Fourth, we are able to calculate the probability distribution of different transportation modes on each edge. For instance, as depicted in the fourth step of stands for the likelihood of the event that people take buses on the edge between node 1 and node 8. Further, the conditional probability be-tween different transportation modes can also be computed based on the graph, for example, (Bus |Walk) represents the transition probability from Walk to Bus between edge 18 and edge 85. In other words, it denotes the likelihood of the event that a user walks from node 8 to node 5 based on the observed occurrence that the user takes a bus from node 1 to node 8.

Such previously mentioned knowledge is promising in improving the inference accuracy due to the following reasons:

1) This implies people's typical behaviors among different places. The clusters of change points represent the places many people change their transportation modes. Usually, these places could be bus stops, parking lots, and railway stations. We can take into account user behavior among these nodes as probabilistic cues when we infer other trajectories passing these two nodes.

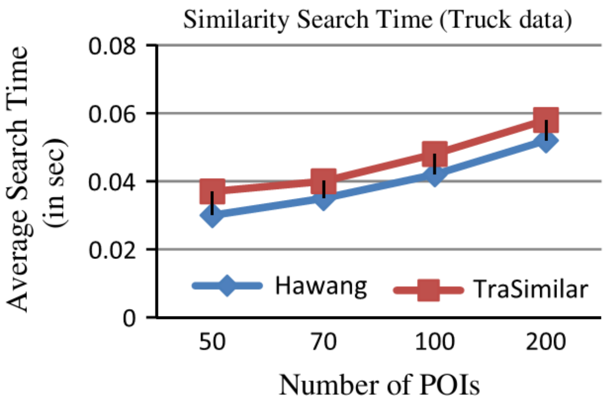

FIGURE 4 Comparison of the average search times for the truck dataset

Naturally, for example, if most users take a bus between two nodes, we can suggest that the two nodes could be bus stops, and the edge between them could have a very high probability of being a bus line.

2) The probability on each edge implies constraints of the real world. For instance, buses only take passengers on at bus stops, cars are left in parking lots, and cars and buses only travel on streets, etc. This knowledge mined from multiple users' GPS logs take advantages of the location constraints while keeping our method independent of an additional map database. In such a way, it is not necessary to match each GPS trajectory against the road network. Mean-while, we do not need to maintain a database of bus stops, railway stations and parking lots.

4 Experimental Results

This experimental setup focuses on vehicle trajectories in which two common characteristics are assumed. First, vehicle trajectories typically follow road networks (i.e., they are not free movements in two-dimensional space). Second, vehicle positions are measured at a reasonably good temporal resolution (e.g., one GPS measurement per minute). Many existing vehicle trajectory datasets satisfy the above resolution requirement, such as the truck data set and the INFATI data derived from the INFATI Project (Jensen http://www.infati.dk/uk). Both of these datasets were used in this research, resulting in one GPS measurement for every 30 s.

In this study's experiments, the performance and scalability of similarity search techniques have been evaluated. All experiments have been conducted on a Intel Core 2 Duo machine running Windows XP with 2 GB of RAM and a 320 GB.

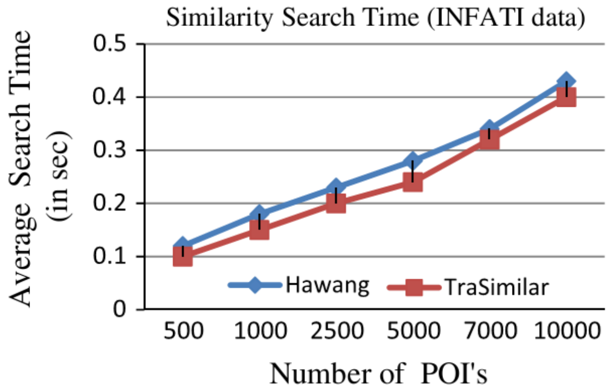

FIGURE 5 Comparison of the average search times for the INFATI dataset

SATA2 16-MB hard drive. The performance of the similarity search technique of the proposed algorithm is measured by comparing the average search time with that of the Hawang algorithm for different numbers of POIs, ranging from 50 to 300 for small datasets (truck data) and from 500 to 3000 for large datasets (INFATI data). Scalability is measured in terms of how the search time grows with respect to an increase in the number of POIs. The truck dataset consists of 50 trucks delivering concrete to several construction places around the Athens metropolitan area in Greece for 33 different days. The data set has 276 trajectories and 112,203 GPS points (approximately one GPS measurement for every 30 s in most trajectories). This study's approach can be used to analyse other vehicle trajectory datasets with similar temporal resolutions, such as the Milan data set. The structure of each record in truck data set is as follows: {obj-id, traj-id, date (dd/mm/yyyy), time (hh:mm:ss), lat, lon, x, y}, where (lat, lon) is in the WGS84 reference system and (x, y) is in the GGRS87 reference system.

As shown in Figure 4, the experimental results confirm that the average search time of the proposed spatial-temporal similarity algorithm increases linearly with an increase in the number of POIs, supporting the fact that the algorithm is scalable. As both of the graphs are similar in terms of average search time, the search performance of TraSimilar is con-firmed to be similar to that of the earlier algorithm even when the proposed spatial similarity extends the structural and sequence similarity measures.

The similarity search performance was tested with a large GPS log dataset (INFATI) of car movements to consider a large number of input trajectories and a POI number greater than 300. The INFATI data are derived from the INFATI Project, which is an intelligent speed-adaptation project carried out by a team of researchers at the Department of Development and Planning, Aalborg University. For each car that delivered data, the INFATI data contains one file with GPS log data.. As shown in Figure 5, the experimental results for a large number of POIs confirm the scalability and search performance of the algorithm in agreement with the results for a smaller number of POIs. The results of the experiment also shows that the TraSimilar search performance is better than that of the Hawang method when there is a larger number of POIs. The authors claim that this is due to the proposed method's advantage of reduced dimensions when representing location and trajectory data in the binary-encoded scheme.

5 Conclusion

Our approach consists of three parts: a change point based segmentation method, an inference model and a post-processing algorithm based on conditional probability. The similarity set thus formed will be used by the remote database to broadcast trigger-based messages to participating vehicles in a neighbourhood for future route and information-sharing activities. The performance of the scheme is evaluated using experiments on standard real-life data.

In this paper, by using knowledge mined from raw GPS data, we aim to improve geographic applications on the Web and build closer connections between locality and mobility. The knowledge we gained as well as the connections enable more novel applications and improve user experience in a variety of tasks. An approach has been proposed to automatically learn transportation mode from raw GPS data. The inferred transportation mode can help Web users more deeply understand their own experience while better sharing other users' knowledge. It also enables context-aware computing based on a user's present transportation mode and creation of innovative user interface for Web users. The proposed approach is independent of other information and devices. Therefore, it is universal to be performed in both mobile devices and servers.

References

[1] Tietbohl Palma A, Bogorny V, Kuijpers B, Otavio Alavares L 2008 A clustering based approach for discovering interesting places in trajectories *Proceedings of SAC'08* 429-32

[2] Andrienko G, Andrienko N 2008 Spatio-temporal aggregation for visual analysis of movements *Proceedings of IEEE Symposium on Visual Analytics Science and Technology (VAST '08)* 51-8

[3] Chang J W, Bista R, Kim Y C, Kim Y K 2007 Spatio-temporal Similarity measure algorithm for moving objects on spatial networks *Proceedings of ICCSA* 1165-78.

[4] Chen, L, Özsu M T, Oria V 2005 Robust and fast similarity search for moving object trajectories *Proceedings of ACM SIGMOD* 491-502

[5] Cuevas-Vicenttin V 2008 Towards multi-scale query processing *Proceedings of IEEE 24th International Conference on Data Engineering (ICDE'08) Workshops* 137-44

[6] Delot T, Cenerario N, Ilarris S 2010 Vehicular event sharing with a mobile peer-to-peer architecture *Transportation Research Part C* **18**(4) 584-98

[7] Delot T, Ilarri S, Thilliez M, Vargas-Solar G, Lecomte S 2011 Multi-scale query processing in vehicular networks *Journal of Ambient Intelligence and Humanized Computing* 2(3) 213-26

[8] Gunduz S, Özsu M T 2003 A web page prediction model based on click-stream tree representation of user behavior *Proceedings of the 9th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* 535-40

[9] Hwang J R, Kang H Y, Li K J 2005 Spatio-temporal Similarity Analysis between Trajectories on Road Networks *Perspectives in Conceptual Modeling Lecture Notes in Computer Science* **3770** 280-9

[10] Hwang J R, Kang H Y, Li K J 2006 Searching for Similar Trajectories on Road Networks using Spatio-Temporal Similarity *Advances in Databases and Information Systems Lecture Notes in Computer Science* **4152** Springer-Verlag LNCS 282-95

[11] Ilarri S, Mena E, Illarramendi A 2010 Location-dependent query processing: where we are and where we are heading *ACM Computing Survey* **42**(3) 1-73

[12] Leontiadis I, Mascolo C 2007 Opportunistic spatio-temporal dissemination system for vehicular networks *Proceedings of the 1st International Mobisys Workshop on Mobile Opportunistic Networking (MobiOpp'07)* 39-46

[13] Salhi I, Cherif M O, Senouci S M 2009 A New Architecture for Data Collection in Vehicular Networks *Proceedings of IEEE/ICC'09* Dresden DE 1-6

[14] Li Q, Zheng Y, Xie X, Chen Y, Liu W, Ma W Y 2008 Mining user similarity based on location history *Proceedings of GIS 08* 1-10

[15] Naumov V, Baumann R, Gross T 2006 An evaluation of inter-vehicle ad hoc networks based on realistic vehicular traces *Proceedings of ACM Mobihoc06* 108-19

[16] Orenstein J A, Merrett T H 1984 A class of data structures for associative searching *Proceedings of 3rd ACM SIGACT-SIGMOD-SIGART Symposium on Principles of Database Systems* Waterloo Ontario 181-90

[17] Krishnamurthy P 2008 Information dissemination and information assurance in vehicular networks: a survey *Proceedings of ACM Schools Conference 08*

[18] Abraham S, Sojan-Lal P 2008 Trigger based security alarming scheme for moving objects on road networks *Proceedings of IEEE International Conference on Intelligence and Security Informatics (ISI 08) LNCS* Springer-Verlag Berlin 92-101

[19] Abraham S, Sojan-Lal P 2010 Trajectory similarity of network constrained moving objects and applications to traffic security *Proceedings of Pacific Asia International Workshop on Security Informatics (PAISI 10)* Springer-Verlag Berlin LNCS 31-43

[20] Sakurai Y, Yoshikawa M, Faloutsos C 2005 FTW: Fast Similarity Search under the Time Warping Distance *PODS 05* 326-37

[21] Lee S, Park S, Kim W 2007 An efficient location encoding method for moving objects using hierarchical administrative district and road network *Information Sciences* **177** 832-43

[22] Shim C B, Chang J W 2003 Similar sub-trajectory retrieval for moving objects in spatio-temporal databases *Proceedings of the 7th EECADIS* 308-22

[23] Szczurek P, Xu B, Lin J, Wolfson O 2010 Spatio-temporal information ranking in VANET applications *Inaugural Issue of the International Journal of Next-Generation Computing (IJNGC)* **1**(1) 52

[24] Theodoridis, 2007. R-Tree Portal. <http://www.rtreeportal.org> (viewed 28.02.07)

[25] Tiakas E, Papadopoulos A N, Nanopoulos A, Manolopoulos Y 2006 Trajectory similarity search in spatial networks *Proceedings of the 10th IDEAS* 185-92

[26] kas E, Papadopoulos A N, Nanopoulos A, Manolopoulos Y, Stojanovic D, Djordjevic-Kajan S 2008 Searching for similar trajectories in spatial networks *Journal of Systems and Software* **82**(5) 772-88

[27] Vargas-Solar G, Ibrahim N, Collet C, Adiba M, Petit J-M, Delot T 2010 Pervasive Computing and Communications Design and Deployment: Technologies, Trends and Applications *IGI Global*

[28] Abraham S, Sojan-Lal P 2012 *Spatio-temporal similarity of network-constrained moving object trajectories using sequence alignment of travel locations* Transportation Research Part C **23** 109-23

[29] Cherfaoui V, Denouex T, Cherfi Z L 2008 Distributed data fusion: application to confidence management in vehicular networks *Proceedings 11th International Conference of Information Fusion 08* 1-8
[30] Vlachos M, Gunopulos D, Kollios D 2002 Robust similarity measures of mobile object trajectories *Proceedings of the 13th International Workshop on DEXA* 721-8
[31] Weinan W, Osmar R Z 2002 Clustering web sessions by sequence alignment *Proceedings of the 13th International Workshop on Database and Expert Systems Applications* 394-8
[32] Wolfson O, Xu B 2010 Spatio-temporal databases in urban transportation *Bulletin of the IEEE Computer Society Technical Committee on Data Engineering* 18-25
[33] Wolfson O, Xu B, Yin H, Rishe N 2005 Resource discovery using spatio-temporal information in mobile ad-hoc networks *Proceedings of the 5th International Workshop on Web and Wireless Geographical Information Systems W2GIS'05* 129-35

Authors

Xiaolei Li, born in December, 1981, Ningbo, Zhejiang Province, China

Current position, grades: lecturer at the School of Electronic and Information, Ningbo Dahongying University, China.
University studies: Master's degree in Information Engineering.
Scientific interest: information engineering and data mining-algorithms.
Publications: 12 papers.