Trace Analysis and Mining for Smart Cities: Issues, Methods, and Applications

Gang Pan, Guande Qi, Wangsheng Zhang, Shijian Li, and Zhaohui Wu, Zhejiang University
Laurence Tianruo Yang, Huazhong University of Science and Technology and St. Francis Xavier University

ABSTRACT

Traces of moving objects in a city, which depict lots of semantics concerning human mobility and city dynamics, are becoming increasingly important. In this article, we first give a brief introduction to trace data; then we present six research issues in trace analysis and mining, and survey the state-of-the-art methods; finally, five promising application domains in smart cities are discussed.

INTRODUCTION

The smart city is critical for sustainable urban development. It could alleviate many critical problems accompanying the current overwhelming urbanization process, for example, traffic jams, environment pollution, and natural resource limits. A smart city features the utilization of information and communication technology (ICT) infrastructure, human resources, social capital, and environment resources for economic development, social/environmental sustainability, and high quality of human life.

From a technological view, smartness means understanding, learning, and self-awareness. Analysis and mining of sensed data from dynamic cities is an inevitable step toward making a city smart. Traces provide important information on the mobility of moving objects (e.g., humans and vehicles). It is becoming easily available in indoor and outdoor environments due to the prevalence of the global positioning system (GPS) and other localization technologies. A trace, generated by a moving object, is usually described by a temporal sequence of spatial points with their timestamps. It is simple, but conveys underlying information on people and cities, such as traffic, human activity, and social events. Trace analysis and mining can exact and reveal inherent information or knowledge about a city and its people. It will benefit broad applications, such as transportation, urban planning, public health, public security, and commerce.

A general framework of trace mining for smart cities is illustrated in Fig. 1. There are massive moving objects in a city (e.g., humans, animals, vehicles, cargo, goods, and trash). Various kinds of ubiquitous sensors in the city can collect different types of traces of them. The trace data can then be input for mining to characterize knowledge about mobility, people, and the city. Finally, applications can exploit the knowledge mined to make them smarter in different domains of a smart city.

TRACE DATA

How can trace data of massive moving objects be collected in a city? Currently, many sensors and devices could be used to detect and report location information. According to the types of sensors and devices, the main sources of trace data can be grouped into four categorizations:

Mobile devices: Mobile devices, such as phones and pads, are becoming ubiquitous. Such portable devices could provide diverse location data with the help of GPS, WiFi, GSM, and Bluetooth. Usually, trace data of mobile devices approximately reflects traces of their owners.

Vehicles: Nowadays more and more vehicles are equipped with GPS devices for navigation services. A vehicle's GPS traces depict not only the trajectory of the vehicle itself, but also that of its driver and passengers. Traces of private and public vehicles usually have different privacy requirements.

Smart cards: Bank cards and transportation cards are two kinds of typical smart cards in a city. Each consumption activity with a smart card is associated with a swiping machine, whose location is usually fixed.

Floating sensors: An object equipped with a localization module can report traces of itself, such as trash tracking [1] and traces of cargos with radio frequency identification (RFID) tags. These floating sensors enable trace data collection for various kinds of moving objects.

The semantic meaning of trace data depends on localization technology. The commonly used localization methods for generating trace data include GPS, WiFi, GSM, Bluetooth, and RFID.

Technology	Data	Reference	Expression	Accuracy	Coverage
GPS	Geographic coordinate	Absolute	Physical	1–5 meters (95–99%)	Outdoors
WiFi	Access point ID + signal strength or local coordinate	Relative	Symbolic/ physical	1–20 meters	< 100 meters from an access point
Cell Tower	Cell tower ID + signal strength or geographic coordinate	Relative/ absolute	Symbolic/ physical	50–200 meters in cities	Cell coverage. 5–30km from a cell tower.
Bluetooth	Device ID	Relative	Symbolic	Sensing range of Bluetooth	5–10 meters for Class 1; 20–30 meters for Class 2
RFID	Reader's ID/position	Relative/ absolute	Symbolic/ physical	Sensing range of RFID	1 meters for passive RFID; 100 meters for active RFID

Table 1. *Comparison of several popular localization technologies.*

We compare them in Table 1 according to the four properties of trace data [2]:
- **Reference**, which defines whether the data is an absolute location (e.g., 30°12'45''N, 120°10'19''E) or a relative location (e.g. the identity of a moving Bluetooth detected by a phone)
- **Expression**, which describes whether the trace is a physical location (e.g. 30°12'45''N, 120°10'19''E) or a symbolic location label (e.g., Zhejiang University, Hangzhou)
- **Precision**, the precision of each location in the trace data
- **Coverage**, the valid range of the method

Research Issues

Trace data implies the underlying patterns and laws of mobility in human society. It is also embedded with rich information on humans (e.g., human activities, social events, and social relationships) and cities (e.g., the semantics of regions and the dynamics of a city). Meanwhile, trace information may raise privacy issues. This article classifies the research issues of trace mining into the following six categories (Fig. 2).

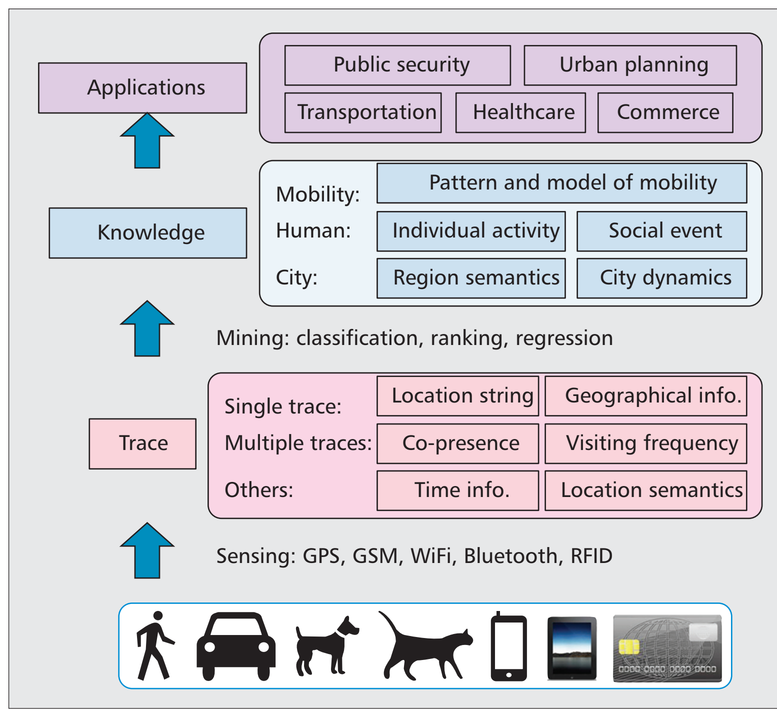

Figure 1. *General framework of trace mining for smart cities.*

Mobility: Patterns, Models, and Prediction

A single individual's trajectory appears as a random motion but actually has inherent patterns. For example, the step size of an individual's movement is power-law distributed [3]. For multiple individuals, measurement of their characteristic distances such as radius of gyration can also be approximated with a truncated power-law. Understanding those patterns that govern human mobility is a fundamental research issue. However, existing work mostly focus on low-level patterns such as step size and orientation.

The absence of a generally accepted mobility model to explain those patterns is a major challenge for mobility analysis. For human mobility, various data sets (e.g., bank note and mobile phone data) seem to agree on the hypothesis that human mobility can be approximated with the Levy flight model or continuous-time random walk (CTRW) model. However, such approximations may conflict with empirical data in details such as the number of significant locations and visiting frequency. This contradiction may be due to missing two generic mechanisms in traditional models: exploration and preferential return. A conceptual model that has the flexibility to make future extension to include new patterns is of crucial importance.

Despite these various mobility models, prediction of human mobility remains a challenge. Although researchers have found a 93 percent potential predictability in human mobility [4], transient prediction as simple as predicting the next visited places with knowledge of historical traces has poor results. To achieve better short-range temporal fidelity, it is necessary to consider the periodic modulations and spatial correlations that constrain human mobility when making predictions.

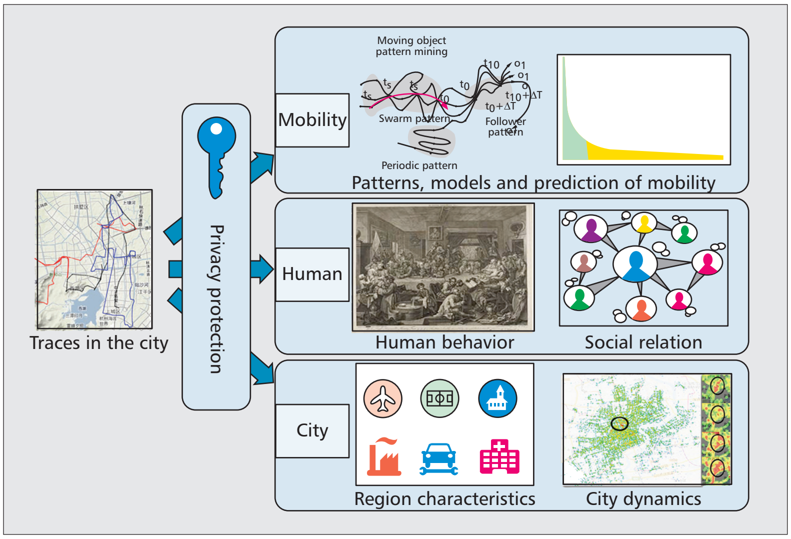

Figure 2. *Research issues of trace mining for smart cities.*

Human Behavior: From Location to Activity and Event

Human behavior is an individual's interaction with an environment or other individuals. Low-level human activities, such as walking, sitting, and standing, may be detected from trace data. Social events, which are sets of human behaviors, including epidemics or organized crime, often accompany the mobility of a crowd.

A major challenge for recognizing human behavior is mining high-level semantics from low-level activities. We should be able to answer questions such as where he/she has visited, who he/she has contacted, what he/she has done, and where he/she will go, based on his/her trace data. Existing machine learning techniques to recognize human behavior are still limited in specific context for specific behaviors; sophisticated inference algorithms in certain contexts cannot adapt to new situations.

With the knowledge of human behaviors, we can reconstruct the whole picture of a city and find social events under very unusual conditions, such as earthquakes and tsunamis, and social events in daily life, such as popular sports games and unexpected traffic congestion.

Social Relation: Trace-Based Social Analysis

A social relation refers to a relationship among individuals. It includes direct and indirect interactions, ranging from low-level face-to-face interaction to high-level interactions like reading books or following traditions. Similar to online social network data, trace data implies social relations among individuals living in a city [5]. For example, frequent co-presence, which means the occurrence of two or more people together in the same place and time, indicates that those people may have some social relation such as classmates or colleagues.

Trace data is spatially closer to the real world than online data because it reflects an individual's locations in the real world. Co-presence is not the only trace-based social relation we can define; people visiting similar places can also be viewed to have social relations even if they do not know each other. A major challenge for trace-based social analysis is to find potential social relation in trace data and construct a corresponding social network. Besides a traditional social network, which involves individuals, we may construct a social network between locations via location correlation. This makes trace-based social analysis more flexible.

Region Characteristics: Semantics and Significance

City regions are diverse in their characteristics such as semantic functions and land usage types [6]. Connection of regional characteristics with traces exists since region characteristics partly result from human mobility. For example, the population of a scenic region during a holiday is more than that on a workday; the activity intensity of a bar region at night is higher than that in day time. However, quantitative analysis of city region characteristics is difficult. First, data quality may be low due to the imperfect data collecting infrastructure; second, the analysis methods devoted to region characteristics analysis are few.

A principal way to understand city region characteristics is to analyze hotspots, which are critical regions in a city with significantly high and regular activity intensity. Research on hotspots includes hotspot discovery, representation and visualization of activities in hotspots, mining high-level semantics of hotspots, and so on.

City Dynamics: Expression, Visualization, and Evaluation

Trace data can be viewed as a kind of sampling of a city and help to better understand a city's evolution. It reflects urban dynamic information in many fields such as energy consumption, traffic flow, epidemic spread, and urban growth.

Representation, visualization, and evaluation are three key steps to analyze city dynamics and develop smart city applications. A major challenge for those steps is cross-domain inference, which comes from the heterogeneous nature of trace data. There are many kinds of trace sources (traces of mobile devices, vehicles, smart cards, and floating sensors) and localization systems (GPS, Wi-Fi, GSM, Bluetooth, and payment records). Some of them may be composed of symbolic location names, while others may contain locations denoted by relative distance away from a landmark. Researchers face multisource and multimodal trace data. A joint context inference framework with full consideration of all contexts is crucial.

Privacy: Disclosure and Protection

Traces, especially personal traces, contain semantics about individual preferences, social relations, and physical locations. The disclosure of such information can cause consequences "from the uncomfortable creepiness of being watched to unwanted revelations of a person's activities to actual physical harm" [7].

Personal identity disclosure could happen in collecting, publishing, and utilizing trace data. First, localization techniques may record user or device ID and cause risks. Location by GPS is more secure than GSM, WiFi, Bluetooth, and RFID because centric servers do not need to know device IDs. Second, personal identity could be inferred from published locations, in spite of having been removed/ anonymized from the trace data records. Third, traces may expose unwanted privacy information to personalized services and applications. In such case, an anonymizing proxy is to be trusted to store, manage, and protect user locations, and to communicate between applications and users. The challenges are to keep fidelity of data for applications meanwhile protecting privacy.

Methods for Trace Analysis and Mining

Techniques for mining trace data depend on specific research problems. We roughly divide the mining techniques into five classes: clustering, classification, ranking, regression, and physical statistical modeling. Some example problems and their methods are shown in Table 2.

Clustering

Clustering is for finding hotspots, querying similar traces, mining moving patterns, and finding objects with similar traces. It is the task of grouping a set of objects so that objects in the same group are most similar. Clustering in trace analysis usually has two types of objects: points and traces. The main difference between them is that points are in the same space such as the physical world (of spatial points) or feature space (of features), while traces, which are sequences of spatial points, cannot be put into the same space directly because of different lengths (dimensions).

The most important problem in clustering is the similarity (or dissimilarity) measure for points or traces. For objects that lie in the same space, typical similarity measures are Euclidean norm, cosine similarity, and Pearson correlation. For traces, similarity may be computed with length of common subsequence and string edit distance.

Based on similarity measures, lots of clustering methods have been presented:

- Hierarchical methods recursively aggregate small and similar clusters into a larger one, or divide a large cluster of dissimilar objects into smaller ones.
- Partitioning methods find the best way to partition objects into a pre-set number of clusters.
- Density-based methods assume thatobjects of each cluster are drawn from a certain probability distribution and fit the distribution of objects with some mixed distribution.

Classification

Classification could be used to recognize individual activity, social event, and region semantics. It is a technique that identifies the class to which each sample belongs, with a model learned from a training set of samples with their class labels. For example, when recognizing individual activity from trace data, each trace is a sample, and each type of activity is a class. Most classification algorithms have three stages: feature extraction, training procedure, and test procedure. The features extracted and the models used for training are most important for classification.

Features extracted from trace data can be categorized into four types according to the views of traced data.

- When trace data is viewed as a sequence of locations, a string feature is extracted to represent mobility.
- When trace data is treated as discrete points or curves on the map space, a geometrical and geographical feature is retrieved.
- When traces are aggregated according to time and space, a time and frequency domain feature can be calculated.
- With the location name, a semantic feature could be inferred to recognize human behavior and region semantics.

The models for training in classification algorithms are usually geometrical or probabilistic:

- A geometrical model considers the spatial distribution of features from different classes. Typical geometrical models are linear discriminant analysis, a neural network, and a support vector machine.
- A probabilistic model represents the probabilistic dependence between features and categories, such as naïve Bayes, the hidden Markov model, a Bayesian network, and a conditional random field.

Ranking

Ranking is used to find the most desired/undesirable traces, moving objects, and regions, for example, recommending the most popular places. It is formally defined as, given a set of samples and their properties, designate a range with partial order relation and use a model to map the samples to the range for rendering the ordered list of these samples. Ranking is widely used for evaluating mobility patterns, finding important persons from location-based social networks, and recommending places.

Two important issues for ranking with traces are finding relevant properties and designating a range of measures with partial order. Relevant properties of traces for ranking can be gotten from different views of traces (e.g., location string, geometric curve, geographic route, sequence of semantic labels, co-presence, and time information).

Designating a range with partial order relies on measuring a mobility pattern, including gain and loss (time, money, and other resources) of a trace, and disparity among traces. For example, ranking the best route for taxi drivers would consider the income of a historical trace.

Designating a range with partial order relies on measurement of a mobility pattern, including gain and loss (time, money, and other resources) of a trace, and disparity among traces. For example, ranking the best route for taxi drivers would consider the income of a historical trace.

Regression

Regression is useful for measuring trace-based social relationships and fitting a mobility model. It is a technique that builds a continuous function between a dependent variable and one or more independent variables. Regression relies on loss function to learn a best model with least error.

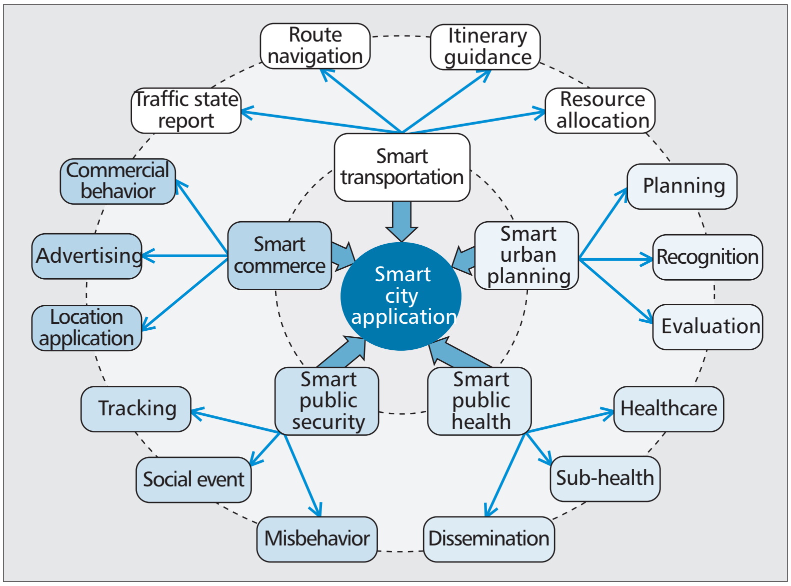

Figure 3. *Applications of trace mining for smart cities.*

The dependent variable is defined according to the specific problem, for example, relationship strength for building social relations. The independent variables can be defined based on characteristics of traces (e.g., co-presence frequency). Usually, the loss function is the divergence between the predicted value from regression and the true value. The regression model may be linear or nonlinear, depending on the empirical assumption of the relationship between variables.

Physical Statistical Modeling

The availability of massive traces inspires the understanding of human mobility. Physical statistical modeling is a technique that builds and evaluates statistical models to depict physical laws or patterns of human mobility with characteristics from concrete step length, number of visiting places, visiting frequency, and order to abstract visiting patterns.

The modeling process is to discover the physical phenomenon in a statistical way. Many laws of nature are of statistical origin, and statistical physics has been successfully applied to modeling. Typical models for human mobility are power law and continuous time random walk. The model is then trained by regression using real-world data and evaluated by goodness of fit.

Applications for Smart Cities

How could mining trace data help smart cities? Information about mobility is helpful for traffic systems; learning human behavior and social relations can benefit public health, security, and commerce; characterizing regions and cities is critical for urban planning. As shown in Fig. 3, trace data mining could help a smart city from at least five aspects, discussed below.

Smart Transportation

In trace data analysis, vehicles could be viewed as floating sensors. They sense and report explicit traffic information such as traffic conditions, road maps, transportation supply and demand, along with implicit information and knowledge, such as traffic accidents, driving strategy, and route navigation. Based on such information, potential applications of smart transportation could help drivers, passengers, and administration.

Monitoring of traffic dynamics. Vehicle traces could be used to monitor traffic dynamics, for example, construction of road map, monitoring of real-time traffic flow, prediction of possible traffic jams, discovery of traffic accidents, and alarms of improper driving behaviors.

Intelligent navigation services. They integrate traffic condition information to recommend the best route, given various measures other than distance, such as time (least time-consuming route) and energy (most energy-saving route).

Guidance for itinerary. It suggests popular spots, hotels, and shopping malls for tourists or citizens. Moreover, the possible time consumed on the way is considered to recommend the best travel route between locations.

Dynamic dispatch of public transportation resources. Given the current and historical human flow data, intelligent systems can predict human flow dynamics in the near future. Furthermore, it could assess whether the public transportation supply could satisfy the future demand and give suggestions on allocating or dispatching public transportation resources to balance the supply and demand.

Smart Urban Planning

Urban planning is a technical and political process concerned with the control of land use and design of urban environments, which can benefit from trace data [15] analysis and mining. For instance, the visiting frequency, such as the number of pick-ups in a region of vehicles on a road, and of patients waiting outside a hospital, is related to land use [6]. In detail, trace data could help urban planning in many ways:

Guiding urban planning. Trace data could guide urban planning in solving two problems:

- How many infrastructures are required for a region
- How to distribute these infrastructures in a region

Visiting frequency is directly related with demands of infrastructure. Moving patterns could answer the second question.

Monitoring of urban land use. Previous research has revealed the potential relationship between visiting frequency and the region's social function [6]. Detailed information and reporting of urban land use could be inferred from massive trace data of individuals.

Evaluating urban planning. According to the visiting frequency to infrastructures, we could measure whether current city planning is sufficient or excessive for service demands. For example, the number of vehicles on roads could be used to evaluate traffic planning [15].

Smart Public Health

Public health could be improved by mining traces of patients, susceptible individuals, and unhealthy people. Patients' traces could help to monitor their social and physical activities relat-

Problem	Method	Instances
Laws of mobility	Physical statistical model	Power law for step length [3], continuous-time random walk for mobility model, and regularity of mobility based on entropy analysis or visiting pattern [4]
Mining moving behavior	Clustering	Finding similar trajectory [8], detecting frequent traces [9]
Route recommendations	Ranking	Navigation [10] and itinerary guidance
Activity recognition	Classification	Recognition of transportation behavior [11]; daily living activities, abnormality detection [12]
Social event recognition	Classification	Recognition of five events: cinema, family, music, performing arts, and sports [13]
Correlation of social networks	Regression	Correlation between reported friendship and co-presence [5]
Recognition of region semantics	Classification	Recognizing social function of regions [6] and POI labeling
Predicition in urban transportation	Regression	Predicting amount of passengers, supply of public bicycles, and traffic conditions [14]

Table 2. *Samples of typical research problems and their methods.*

ed to body condition. Traces of susceptible individuals are critical for understanding and controlling disease dissemination. The physical activity of unhealthy individuals, which can be inferred from traces, could be used for persuasive health. In detail, these applications are:

Monitoring behaviors of patients. With help of patients' traces, an intelligent healthcare system could collect important information on patients' behavior (e.g., frequency of going to restrooms).

Controlling dissemination of epidemics. Many epidemics are disseminated partly by physical proximity or contact. Such contact is generated by movement and co-presence. Massive trace data could discover information on co-presence and density of individuals in a physical space, which is critical to understanding and controlling epidemic dissemination.

Reducing health problems. Obesity and other chronic conditions are becoming increasingly common, and are highly correlated with daily human behaviors. For example, they may be caused by long work hours with little exercise. Trace data of individuals could help to make people healthier by monitoring and encouraging exercise.

SMART PUBLIC SECURITY

Public security is closely related to human trace data since public security is about individual activity and social events in most cases. Human traces could be used to infer lots of human behaviors, such as occurrences of social events, abnormal gatherings of people, misbehavior of individuals, movement of criminals and the lost, and behaviors of people in disaster. Based on the human and crowd behavior information mined from trace data, public security could be enhanced in many aspects, for instance:

Detecting misbehavior of individuals. Most people have routine routes every day, for example, a repetitive pattern of travel between home and work on weekdays. Abnormal traces could hint at potential misbehavior of individuals.

Monitoring social events and crowd behavior. Gathering and movement of crowds are often accompanied by abnormal human flows, which could be detected and monitored with massive trace data.

Searching and tracking critical people. Predicting traces of the lost and people in disaster could help to locate them. Traces of suspected criminals could infer actual criminals, find gangs, and detect insecure communities.

SMART COMMERCE

Commerce can be improved by trace mining. For example, we could place commercial advertisements in the most visited places. Trace data, especially check-in information, is also associated with cash flow and business affairs. The main applications of trace data for smart commerce may include:

Location-based services. The most popular commercial application inspired by location systems is location-based service. It includes location-based social networking services like Foursquare, location-based recommendation services, driving navigation, and itinerary guidance.

Smart advertising based on human flows. Advertisement is important for companies to grow their business. Where and when to put advertisements, which is critical for advertising, depends on the flow of potential customers. Trace data contains massive information of such flows and could be used to improve advertising strategy.

Improving shopping services by knowledge of consuming behaviors. Payment records reflect interests of customers. Moreover, customers' traces reflect their routine consuming behavior. For example, customers' traces in shopping malls can depict how long they stay at a goods shelf.

CONCLUSION

Traces are a sampling of dynamics of moving objects in the temporal and spatial dimensions. Analysis and mining of trace data is becoming a very promising way to discover the underlying knowledge on human activities and city dynam-

Traces are a sampling of dynamics of moving objects in the temporal and spatial dimensions. Analysis and mining of trace data is becoming a very promising way to discover the underlying knowledge of human activities and city dynamics.

ics. It helps well understanding of human, society, and cities. Trace mining can be exploited in a wide range of potential applications to make a city smarter. Still, it remains challenging to aggregate heterogeneous and limited trace data into a unified view of complex and dynamic real-world city life.

ACKNOWLEDGMENTS

This work is partly supported by the National Key Basic Research Program of China (2013CB329504) and Qianjiang Talent Program of Zhejiang (2011R10078). Dr. Z. Wu is the corresponding author.

REFERENCES

[1] A. Boustani *et al.*, "Investigation of the Waste-Removal Chain through Pervasive Computing," *IBM J. Res. Dev.*, vol. 55, no. 1.2, Jan. 2011, pp. 1–11.
[2] J. Hightower and G. Borriello, "Location Systems for Ubiquitous Computing," *Computer*, vol. 34, no. 8, Aug. 2001, pp. 57-66.
[3] D. Brockmann, L. Hufnagel, and T. Geisel, "The Scaling Laws of Human Travel," *Nature*, vol. 439, no. 7075, Jan. 2006, pp. 462–65.
[4] C. Song *et al.*, "Limits of Predictability in Human Mobility," *Science*, vol. 327, no. 5968, Feb. 2010, pp. 1018–21.
[5] N. Eagle *et al.*, "Inferring Friendship Network Structure by Using Mobile Phone Data," *P. Natl. Acad. Sci. USA*, vol. 106, no. 36, July 2009, pp. 15,274–78.
[6] G. Pan *et al.*, "Land-Use Classification Using Taxi GPS Traces," *IEEE Trans. Intell. Transp. Sys.*, vol. 14, no. 1, Mar. 2013, pp. 113–23.
[7] J. Krumm, "A Survey of Computational Location Privacy," *Pers. Ubiquit. Comput.*, vol. 13, no. 6, Aug. 2009, pp. 391–99.
[8] Y. Li, J. Han, and J. Yang, "Clustering Moving Objects," *KDD'04*, Seattle, WA, Aug. 2004, pp. 617–22.
[9] F. Giannotti *et al.*, "Trajectory Pattern Mining," *KDD '07*, San Jose, CA, Aug. 2007, pp. 330–39.
[10] J. Yuan *et al.*, "T-Drive: Enhancing Driving Directions with Taxi Drivers' Intelligence," *IEEE Trans. Knowl. Data Eng.*, vol. 25, no. 1, Jan. 2013, pp. 220–32.
[11] S. Reddy *et al.*, "Using Mobile Phones to Determine Transportation Modes," *ACM T. Sensor Network*, vol. 6, no. 2, Feb. 2010, p. 13.
[12] D. Zhang *et al.*, "iBAT: Detecting

Anomalous Taxi Trajectories from GPS Traces," *Ubicomp '11*, Beijing, China, Sept. 2011, pp. 99–108.
[13] F. Calabrese *et al.*, "The Geography of Taste: Analyzing Cell-Phone Mobility and Social Events," *Pervasive '10*, Helsinki, Finland, May. 2010, pp. 22–37.
[14] P. S. Castro *et al.*, "From Taxi GPS Traces to Social and Community Dynamics: A Survey," *ACM Comp. Survey*, 2013, in press.
[15] Y. Zheng *et al.*, "Urban Computing with Taxicabs," *Ubicomp '11*, Beijing, China, Sept. 2011, pp. 89–98.

BIOGRAPHY

GANG PAN (gpan@zju.edu.cn) received his B.Sc. and Ph.D. in computer science in 1998 and 2004, respectively, from Zhejiang University, Hangzhou, China. He is a professor of the College of Computer Science and Technology, Zhejiang University, China. His research interests include pervasive computing, computer vision, and pattern recognition. He has co-authored more than 100 refereed papers.

GUANDE QI received his B.Sc. in life science from Zhejiang University. He became a Ph.D. student of the Department of Computer Science at Zhejiang University in 2008. His research interests include machine learning and data mining. His research topic is about mining taxi traces for intelligent transportation and social awareness during the Ph.D. program.

WANGSHENG ZHANG is a Ph.D. candidate in the Department of Computer Science at Zhejiang University. He received his M.Sc. in industrial engineering from Tsinghua University in 2010. His research interests are in pervasive computing, social networks and human mobility patterns with focus on the modeling and analyzing of individual trajectory.

LAURENCE T. YANG is a professor in School of Computer Science and Technology at Huazhong University of Science and Technology, Wuhan, China, and in the Department of Computer Science, St. Francis Xavier University, Canada. His research interests include parallel and distributed computing, and embedded and ubiquitous/pervasive computing. He has a Ph.D. in computer science from the University of Victoria, Canada.

SHIJIAN LI received his Ph.D. from the College of Computer Science and Technology, Zhejiang University in 2006. He worked at the Institute Telecom SudParis, France, as a visiting scholar in 2010. His research interests include sensor networks, ubiquitous computing, and social computing. He has published over 40 papers in the above domains. He works as an editor of the *International Journal of Distributed Sensor Networks*, and as a reviewer or PC member of more than 10 conferences.

ZHAOHUI WU received his B.Sc. and Ph.D. in computer science in 1988 and 1993, respectively, from Zhejiang University. He is a professor in the Department of Computer Science at Zhejiang University. His research interests include distributed artificial intelligence, semantic grid, and pervasive computing. He is a standing council member of the China Computer Federation.