Chapter 3
A Survey on Spatiotemporal and Semantic Data Mining

Quan Yuan, Chao Zhang, and Jiawei Han

Abstract The wide proliferation of GPS-enabled mobile devices and the rapid development of sensing technology have nurtured explosive growth of semantics-enriched spatiotemporal (SeST) data. Compared to traditional spatiotemporal data like GPS traces and RFID data, SeST data is multidimensional in nature as each SeST object involves location, time, and text. On one hand, mining spatiotemporal knowledge from SeST data brings new opportunities to improving applications like location recommendation, event detection, and urban planning. On the other hand, SeST data also introduces new challenges that have led to the developments of various techniques tailored for mining SeST information. In this survey, we summarize state-of-the-art studies on knowledge discovery from SeST data. Specifically, we first identify the key challenges and data representations for mining SeST data. Then we introduce major mining tasks and how SeST information is leveraged in existing studies. Finally, we provide an overall picture of this research area and an outlook on several future directions of it. We anticipate this survey to provide readers with an overall picture of the state-of-the-art research in this area and to help them generate high-quality work.

Keywords Spatiotemporal data • Semantic data • Data mining techniques

3.1 Introduction

With the wide proliferation of GPS-enabled mobile devices and the rapid advance of sensing technology, recent years are witnessing a massive amount of semantics-rich spatiotemporal data accumulated from various sources. For example, on social media platforms like Twitter, millions of geo-tagged tweets are created every day, where each geo-tagged tweet consists of a timestamp, a location, and short text

Q. Yuan (✉) • C. Zhang • J. Han
Department of Computer Science, University of Illinois at Urbana-Champaign,
201 N Goodwin Ave, Urbana, IL 61801, USA
e-mail: qyuan@illinois.edu; czhang82@illinois.edu; hanj@illinois.edu

M. Behnisch, G. Meinel (eds.), *Trends in Spatial Analysis and Modelling*,
Geotechnologies and the Environment 19, DOI 10.1007/978-3-319-52522-8_3

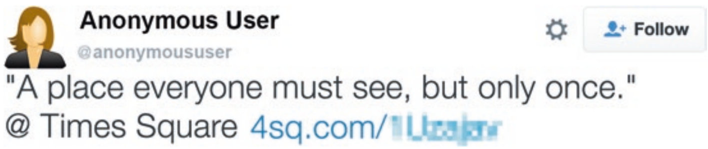

Fig. 3.1 An example of geo-annotated tweet. User name and url link are anonymized for privacy preservation

(Fig. 3.1). For another example, mainstream search engines (e.g., Google, Bing) are continuously collecting queries from GPS-enabled mobile devices. These queries are also associated with timestamps and locations as metadata.

Compared to traditional spatiotemporal data like GPS traces and RFID data, semantics-enrich spatiotemporal (abbreviated as SeST onwards) data is multidimensional in nature. A typical SeST object involves three different data types (location, time, and text) and thus provide a unified where-when-what (three W) view of people's behaviors. As such, the prevalence of SeST data brings new opportunities to spatiotemporal knowledge discovery and opens doors to improving a lot of real-life applications. Consider mobility understanding as an example. While traditional GPS trace data can reveal how an individual moves from one location to another, the SeST data allows us to go beyond that and understand what activities the individual does at various locations. Such semantics-level information is essential in terms of capturing people's mobility patterns, improving applications e.g. location prediction, advertising targeted users, and urban planning.

While SeST data sheds light on improving a wide variety of real-life applications, it is by no means trivial to fully unleash its power. Compared with knowledge discovery in traditional spatiotemporal data, mining SeST data introduces a handful of new challenges:

- **How to integrate diverse data types?** SeST data involves three data types: location, time, and text. Considering the distinct representations of these data types (continuous or discrete) and the complicated correlations among them, it is difficult to effectively integrate them for spatiotemporal knowledge discovery.
- **How to overcome data sparsity?** Unlike intentionally collected tracking data, most SeST data is low-sampling in nature. Take geo-tagged tweets as an example, a user is unlikely to report her activity at every visited location. Such data sparsity makes it challenging to apply classic data mining and machine learning techniques.
- **How to extract useful knowledge from noisy data?** Text in SeST data is usually short and noisy. For example, a geo-tagged tweet contains no more than 140 characters, and most geo-tagged Instagram photos are associated with quite short text descriptions. Moreover, existing studies have revealed that about 40% social media posts are just pointless babbles (Kelly 2009). Still, it is nontrivial to make use of such noisy and incomplete text data to acquire useful knowledge.
- **How to handle large-scale SeST data to build scalable and efficient systems?** Many spatiotemporal applications (e.g., local event detection, location prediction)

requires the back-end system to deal with large-scale SeST data and to respond to users' needs in a timely manner. Since practical SeST data comes in a massive volume, how to develop efficient techniques to handle such big SeST data remains challenging.

Because of the large potential of SeST data in improving various spatiotemporal applications as well as the unique challenges in fully unleashing the power of SeST data, mining SeST data has attracted a lot of research attention from communities like data mining, civil engineering, transportation, and environmental science. SeST data mining has potentially great impact on a variety of fields such as sociology, epidemiology, psychology, public health, etc. We notice several review works on spatial and spatiotemporal data mining (Cheng et al. 2014; Shekhar et al. 2015), but a systematic summarization on state-of-the-art techniques for mining SeST data is still an untouched topic. In this survey, we summarize recent research studies on knowledge discovery from SeST data. Specifically, we introduce data representations, key research problems, methodologies, and future directions. We anticipate this survey to provide an over- all picture of this area, which can help the community better understand the cutting edge and generate quality research results.

The organization of this survey is as follows. In Sect. 3.2, we survey the datasets and the representations of spatial, temporal and semantic information used in existing studies, and introduce the major approaches to SeST data mining. Then, in Sect. 3.3 we review the major tasks that are widely studied in existing SeST data mining works. Important directions for future research are discussed in Sect. 3.4. In the end, Sect. 3.5 summarizes the article.

3.2 Framework of Mining the SeST Data

Mining the SeST data is a general process of acquiring, integrating, analyzing, and mining semantics-enriched spatiotemporal data. In this section, we overview the data sources and representations of SeST data, and then introduce the major approaches to mining SeST knowledge.

3.2.1 Data Sources

Various types of SeST data are used in existing studies. In this section, we list the major data sources and introduce their properties.

- **User generated content.** With the development of social media websites and GPS technology, a great quantity of user generated content has been accumulated, which involves spatial, temporal and semantic information. Examples are social media posts such as Tweets and Facebook statuses, reviews in crowd-sourced review based social networks (e.g., Yelp, Dianping), check-ins in location-based social networks (e.g., Foursquare), events in event-based social networks and travelogues.

- **Survey study data.** Some organizations collect the mobility behaviors of users via survey studies. Representative survey data includes MIT Reality Mining,[1] American Time Use Survey,[2] and Puget Sound Regional Council Household Activity Survey.[3] In the survey data, each visit of an individual involves location, visiting time, and semantics describing the activity (working, shopping, etc.) or visiting purpose. Sometimes, survey data also contains demographic information of individuals, such as age, gender, job, etc. This enables us to study the correlations between user mobility and their demographics.
- **GPS trajectories.** Trajectories, consisting of a series of coordinates-timestamps information, are used to unveil people's mobility. As visits are passively collected, people often need to extract stay points as the locations which a user visited rather than passed by. The stay points are extracted based on other evidence, such as the mobility range in a session and the stay time. Semantic information, such as location categories and descriptions, are often extracted from external data sources, such as Wikipedia, gazetteers, land use around cell towers, etc.
- **Query logs and browsing histories.** As an increasing number of cellphones are 3G/4G enabled, more people search information and browse webpages on the go. As a result, query logs and browsing data are associated with geographic coordinates, representing the current surrounding of the users. The metadata reveal spatial and temporal information, and the text content can be used as the semantic information.
- **News feeds and blogs.** A large amount of news feeds and blogs have location and time information, and thus such data can be also viewed as SeST data. In these datasets, spatial and temporal information can be either collected from metadata, or exacted from the content parsed by some natural language processing (NLP) tools. The text content carries rich and often high-quality semantics.

Among these datasets, the publicly available user generated content are often of large quantity but low quality, in terms of sparsity (some users may only have few posts), noise (users write text in free style), and incompleteness (observations are available only if the users actively submit them). In contrast, survey data is of much higher quality, but they are expensive to get, and the lengths of observations are often short (ranging from 2 to 100 days). Trajectory data often has a reasonable and stable sampling rate, but it is hard to collect and additional steps are needed to extract stay-points and to infer the semantics. Query logs are not publicly available. Only search companies such as Google, Bing, Yahoo! can access such data. News articles have many constraints due to the natures. For example, it cannot be used to analyze user behaviors. In practical data mining tasks, researchers may exploit multiple data source. For example, various data, such as Tweets, News feeds, survey data, and webpages is used to forecast events in a city (Sect. 3.3.4).

[1] http://realitycommons.media.mit.edu/realitymining.html

[2] www.bls.gov/tus/

[3] https://survey.psrc.org/web/pages/home/

3.2.2 Data Representations

In the literature, spatial, temporal, and semantic information can be represented in various forms.

- **Spatial information.** A pair of latitude and longitude is the most representative form of spatial information of a target location. Other representations include lines (e.g., road segments) and polygons (e.g., universities, parks). At the meantime, there are also a lot of studies that index locations by identifiers, such as venues (point-of-interest, POI), cities, grids, etc.
- **Temporal information.** Temporal information can be represented as either continuous or discrete variables. The continuous representation mostly denote time as a real-value offset (e.g., timestamp) with regard to a specific starting time. For the discrete representation, one can choose the appropriate granularity (e.g., hour, day, week, month) depending on the specific tasks. Sometimes temporal information is implicitly modeled as the order of visits (e.g., trajectory mining in Sect. 3.3.6).
- **Semantic information.** Semantic information can be modeled as a categorical variables or plain text, and it can be associated with locations, users, and visits. For example, the semantics of a user could be her jobs, her hobbies extracted from Facebook profiles. For a location, we can get its semantics like category (e.g., hotel, restaurant, airport), and descriptions. We can also get the semantics of a user's visit at a location, which could be a Yelp review or the text content of a geo-annotated tweet. Some studies go one step further beyond text, and extract additional knowledge, such as named entities, emotions, etc., using text mining or NLP techniques. The usage of semantics depends on both specific tasks and data availability.

Figure 3.2 shows the graph representation of SeST data. In this figure, users, POIs and visits are associated with text, and visits and POIs have time and category as metadata, respectively. Based on the time of visits, we can recover the trajectory of a user (e.g., dashed line in Fig. 3.2 for user u_4). In addition, friend links among users may be available in social media data. Although some datasets may have additional objects, e.g., events, the majority of SeST data can be modeled as subgraphs of the figure.

3.2.3 Approaches

Different studies exploit spatial, temporal and semantic information in different ways, which can be categorized into four approaches.

- Some studies use the three types of information independently, and build models for each of them. Then, the results coming out of different models are combined by certain strategies. For example, in order to recommend POIs to users

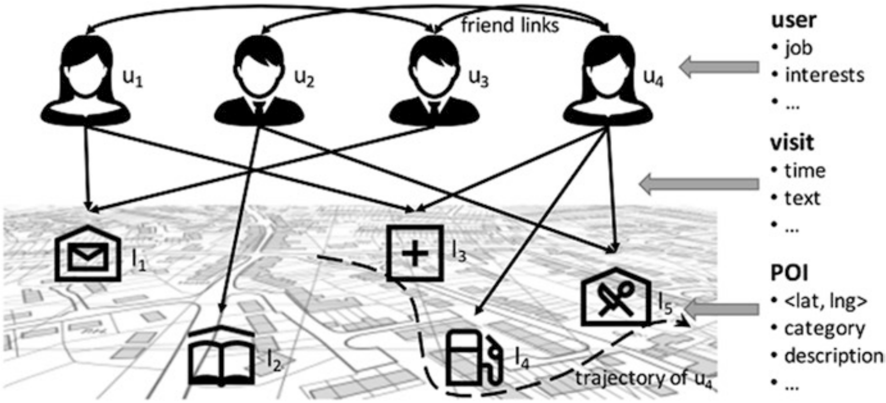

Fig. 3.2 The graph representation of SeST data

- (Section 3.3.2), some researchers build separate models to estimate a target user's spatial, temporal and semantic preference scores of a candidate POI, then these three scores are combined into a final preference score by linear interpolation. Other studies build different classifiers for different types of information and employ co-training strategy to boost the classification performance.
- Some studies extract features from SeST data to train supervised models, such as regression (regression tree, lasso) and classification (support vector machine SVM, Maximum Entropy). This approach is often adopted when a number of features are available, and the target task can be modeled as a supervised or semi-supervised problem, such as quantity prediction (Sect. 3.3.5) and POI typing (Sect. 3.3.8).
- The three types of information are also used as observations for unsupervised models, such as factorization models (tensor factorization, singular value decomposition SVD), graphical models (latent Dirichlet process LDA, hidden Markov model HMM, conditional random field CRF), and graph models (random walk with restart RWR, diffusion model). This approach is often used when the interactions between different information are clear and relatively straightforward to model.
- For some specific tasks, the three types of information are used as optimization constraints or filtering criteria. For example, to plan a trip (Sect. 3.3.7), traveling duration and location category are often modeled as the constraints of an optimization problem. For a second example, many studies on event forecasting do not use the location information in the model. Instead, the location information (e.g., regions, cities) is often used to separate data inside the area of interest from outside.

3.3 Spatiotemporal and Semantic Data Mining Tasks

Many data mining tasks on spatial, temporal, and semantic information have been studied in the literature, the majority of which, however, only exploit at most two dimensions of the three. Recent studies exploit all of the three dimensions, and the integration of space, time, and semantics provides new possibilities of data mining. In this section, we review several most popular SeST data mining tasks, which can be organized as in Fig. 3.3.

3.3.1 Prediction

Prediction aims to infer the candidate dependent variable for a target variable. Under the SeST scenario, the target variable could be user, visit, social media post, etc., while in most studies the dependent variable is location. Representative tasks include next movement prediction for users, home location inference for users, POI inference for geographic coordinates, location estimation for tweets or photos (Hauff and Houben 2012), etc. In most existing studies, the semantic information is the POI category or the text associated with the records. In this section, we take the former two tasks as examples and review existing studies.

Next-place prediction aims at predicting the next place a user is about to visit based on her current location or recent moves, e.g., suppose an office lady just visited a bank branch after work, where is she going to visit next? Next-place prediction is of great importance to user mobility modeling as well as advertisement. Most initial studies only exploit location and time information to construct trajectories, and then predict the next place based off either frequent trajectories of massive people or the user herself. Many recent studies attempt to use semantic information as additional evidence to estimate user mobility preference, where the semantics can be POI categories or text. One thread of works is to extract the frequent trajectory patterns over POI categories. Suppose the pattern *office bank restaurant* is popular in the database, then we can predict that the next place the lady is going to visit is a restaurant. Based on her current location and time, we can select a specific restaurant as the prediction result. Instead of extracting trajectory patterns, we can also infer the transitions between HMM latent states from users' traces for prediction, where each latent state (e.g., office state, home state) defines a semantic topic

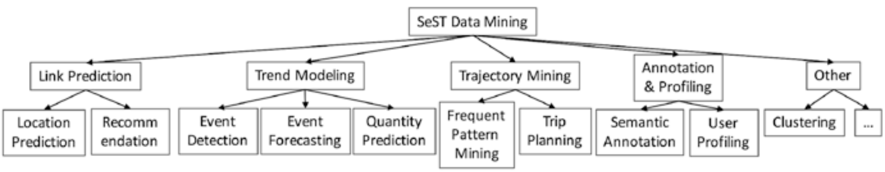

Fig. 3.3 An overview of SeST data mining tasks

(e.g., working), a time range (day-time), and a geographic area (CBD) (Zhang et al. 2016). Another thread is to model the task as a supervised learning problem, in which features like temporal frequencies of categories are extracted to rank candidate POIs by building ranking models.

Home inference is to estimate the home location (e.g., city, state) of a user based on her historical records, such as social media posts, friends, IP address, etc. Home inference is important because users' home information is essential to many tasks such as event detection, personalized recommendation, advertisement, but only few people disclose their home locations on social networks. Pioneering studies exploit social links and user generated text. Suppose a boy has many Facebook friends in New York City, and he has posted a lot about *Net Knicks* and *Bronx*, then he is likely to live in NYC as well. Recent studies use the temporal information to extract the spatiotemporal correlations between text content and locations (Yamaguchi et al. 2014), or extract users' temporal tweeting behavior. For example, if a user posts a tweet right after an earthquake, we can infer that the home of the user should be close to the location of the earthquake. As another example, New Yorkers are more likely to post tweets at 7:00 pm EDT, whereas people who live in Log Angles may tweet less because they are still at work (4:00 pm PDT) in California.

3.3.2 Recommendation

The goal of recommender systems is to suggest new items that a target user might be interested in. While many traditional recommender systems are built on explicit numeric feedbacks (e.g., movie ratings), most recommendation tasks under SeST scenarios deal with implicit binary feedback data, e.g., whether a user will visit a place. Representative tasks include the recommendations of POIs, events, entities (Zhuang et al. 2011), short messages, etc., where the first two tasks are introduced as examples in this section.

POI recommendation aims to suggest unvisited POIs to users based on users' preference. This task has caught a lot of research attention because it can not only help users explore new places but also has great commercial value in advertising. Pioneering studies use spatial and temporal information for recommendation based on the assumptions that users tend to visit their nearby places (Ye et al. 2011b), and a user's preference over POIs is influenced by time (Yuan et al. 2013), e.g., visiting libraries in the morning and bars at night. Some recent works exploit semantic information such as POI categories, check-in text, and reviews to better estimate users' preference implicitly. A straightforward strategy is to recommend the POIs belonging to the categories that the target user visited most. For example, if a user went to many Italian restaurants, then it is safe to continue recommending Italian restaurants to her. We can also infer users' preference transitions over POI categories for recommendation. For example, if a user just visited a restaurant, then we can recommend a theater to her if the pattern *restaurant theater* is frequent in the training set. Rather than using semantics implicitly, we can explicitly take the target user's specific requirements (e.g., *cheesy pizza and spaghetti*) as input, and recommend POIs that best match the target user's semantic profile, mobility behavior and the requirements (Yuan et al. 2015).

Event recommendation aims at recommending local events (e.g., a BBQ party) in event-based social networks (e.g., Meetup[4]) for users to participate in. Initial studies on event recommendation mainly focus on spatial, social and semantic information, based on the assumption that a user tend to participate in events that (1) held close to her, (2) topically attractive to her, and (3) many of her friends also took part in. However, time is also a factor that needs to be taken into consideration because users can only join an event if she is available at that time. To exploit time, several methods (Pham et al. 2015) have been proposed under the frameworks of RWR or ranking models, assuming user tends to attend events held at similar times (e.g., time of a day and day of a week) of the events she attended before.

3.3.3 Event Detection

Event detection is to detect unusual semantic trends that are temporally spiking. Pioneering studies focus on temporal and semantic information to detect global events, e.g., stock market fluctuations and presidential elections, while recent studies exploit the spatial information to detect local events from geo-annotated data, where local event is defined as *something that happens at some specific time and place* (Lee 2012), such as a basketball game or a terrorist attack. Different from the global ones, the local events should be bursty in terms of both location and time. For example, an unlarge number of tweets are talking about *explosion* at Istanbul airport indicate there is an local event terrorist bombing at the airport. The can be detected either by monitoring the changes of spatial and temporal distributions of semantics (Chen and Roy 2009), or comparing the predicted count of tweets generated by regression models with the actual count of tweets for each region (Krumm and Horvitz 2015). Some studies are designed to detect specific types of events such as earthquakes and traffic congestions, in which task-specific evidence is utilized, e.g., the change of massive drivers' routing behavior on road network for traffic anomaly detection (Pan et al. 2013).

3.3.4 Event Forecasting

Event forecasting aims to predict whether an event will happen in the near future. The forecast results make it possible for individuals, organizations and government to prepare for potential crisis in advance. Early studies produce forecasts via either supervised models or time series evolution models that use the temporal and

[4] http://www.meetup.com/

semantic information. The spatial information makes it possible to forecast local events. Existing studies on local event forecasting are domain-specific, i.e., they can detect a specific type of events, such as civil unrest and disease outbreaks. They assume local events can be predicted by monitoring some indicative features, such as keywords counts, tweet cascades, extended vocabulary, etc., extracted from various data sources. For example, if a large portion of tweets in a city are talking about *protest* and *march*, but the portion is small in other cities, there is likely an civil unrest in the city. To forecast whether an event will happen, we can either estimate the development stages (e.g., emerging, uprising, peak, etc.) by monitoring the tweet stream (Zhao et al. 2015), or build regression or classification models on the extracted features.

3.3.5 Quantity Prediction

The availability of SeST information enables us to predict quantity of event or objects based on current observations. Representative tasks include popularity prediction, air quality prediction, traffic volume prediction, etc.

Popularity prediction aims to predict the number of objects adoptions (e.g., hashtags, topics) at specific time in social media, and try to answer the questions like *how many times the hashtag #brexit will be discussed tomorrow in twitter?* Predicting hashtag popularity is important to the identification of commercial and ideological trends. It has been shown that the content of the hashtag (e.g., character length, number of words), the locations mentioned in the tweets, the social network topology (e.g., the number of followers of users who used the hashtag) and the counts of the hashtags in each time intervals are all important features to train a regression model for popularity prediction (Tsur and Rappoport 2012).

Some studies focus on air quality prediction for city regions. This is a challenging task because only a limited number of air quality monitor stations are available in a city, and the air quality depends on various factors, such as meteorology, traffic, land use, etc. This follows our intuition that the air quality in an industry region with high traffic speed is likely to be worse than that in a university. Several supervised or semi supervised models (Zheng et al. 2013) have been developed to predict air-quality based on various spatiotemporal features such as human mobility, traffic speed, the categories of POIs within a region (e.g., factories, parks), etc. Similar strategies are employed to predict the traffic volume in different road segments, where users' activity such as shopping and leisure is an important consideration.

3.3.6 Frequent Pattern Mining

Spatiotemporal frequent pattern mining aims to extract patterns that frequently occur in the given spatiotemporal database. While classic studies on frequent pattern mining in spatiotemporal data can uncover the regularity of people's spatiotemporal movements, the availability of SeST data adds semantics to them and enables us to discover interpretable patterns. Broadly speaking, frequent spatiotemporal patterns in SeST data can be classified into two categories: frequent event patterns, mobility patterns.

Event pattern mining aims to extract frequently co-occurring, cascade, and sequential event patterns from historical SeST data. Co-occurring patterns are events that frequently happen at the same time, e.g., the pattern {*morning, breakfast, at home} → {read news}* detected from smartphone context data; cascade patterns are partially ordered subsets of events located together and occurring serially, e.g., the event *bar closing* leads to subsequent event *assault*, and the two together result in *drunk driving* (Mohan et al. 2012); sequential patterns consist of a series events that happen usually in order, e.g., the disease transmission pattern *bird → mosquito → human being*.

In addition to finding frequent event patterns, researchers have also utilized SeST data to mine frequent movement patterns and tried to understand people's mobility regularity. For example, we can extract sequential patterns over location groups from semantic trajectory databases, where locations in each group are close in distance and consistent in semantic categories (Zhang et al. 2014). For example, a frequent sequential pattern in London could be {House of Parliament, Westminster Abbey} → {*Hyde Park, Regent's Park*}. The former set contains historic sites, while the latter set contains parks. The locations in each set are close to each other.

3.3.7 Trip Planning

The goal of trip planning is to construct a series of locations as travel route for target users. Intuitively, individuals may have limited budget and time for a trip, and different users may have different preferences over places of interests. For example, suppose a girl has only 1 day to travel in London, and she is interested in historic and cultural sites, we should construct a sequence of places within the city, such as *the House of Parliament*, *the British Museum*, *Tower Bridge*, etc., instead of *Wimbledon Tennis Court* or *Windsor Castle* because of they cannot fulfill the girl's interests and the travel time is too long. To incorporate such information, most existing studies (Brilhante et al. 2015) model the trip planning task as an optimization problem by selecting a series of POIs that can meet constraints on time, expense, categories, etc.

3.3.8 Semantic Annotation and Profiling

Semantic annotation aims to infer semantics (categories, descriptions, etc.) for objects (POIs, regions, user visits, trajectories, etc.). Semantic annotation is of great importance to many applications. For example, about 30% of POIs in Foursquare

are lacking any meaningful textual descriptions (Ye et al. 2011a). Annotating these POIs with categories can facilitate both place exploration for users and recommendation services for businesses.

To archive semantic annotation, it is important to exploit spatial, temporal and semantic jointly. To take POI annotation as an example, it is observed that the categories of POIs visited by the same user at the same time are similar. In addition, visitors' demographic information (e.g., age and gender) and the surrounding business are both good indicators of the POI category: a student may stop by a restaurant for lunch at noon because of a break between two classes, and the restaurant is close to other restaurants and grocery stores. Classification models are effective in combining spatiotemporal and semantic features for category estimation. Similarly, the function of a region can be inferred from various SeST evidence such as human mobility and POI categories (Yuan et al. 2012). There are also studies on visit annotation, which uncover the visiting purpose by both static and dynamic features: on the one hand, the visiting purpose is related to the static features such as POI category and region's land use that are invariant to time; on the other hand, the purpose is also influenced by dynamic local events, such as sport games or festival celebrations (Wu et al. 2015). Consider a man who visits Oracle Arena in Oakland on June 19 2016. Then the purpose of his visit might be watching NBA Finals Game 7. Recent studies use static and dynamic features to infer the visiting purpose and achieve satisfactory performance.

Profiling is to characterize entities by spatiotemporal or semantic data. Several papers are published to profile users, POIs and words. Among them, user profiling received the most research interest. Different methods profile users from different aspects, such as individuals' frequent routines (Farrahi and Gatica-Perez 2011), spatiotemporal mobility and topic preference such as movie, hiking (Yuan et al. 2015), and demographic information such as age, gender, marital status (Zhong et al. 2015). It has been shown that the spatial, temporal and semantic information of users' visiting records are all important evidence for profiling.

3.3.9 Clustering

Clustering, which aims to group objects such that objects in the same group are more similar to each other than those in other groups, is a fundamental task in data mining. The availability of spatial, temporal and semantic information enables us to better estimate the relatedness between objects and form clusters. The detected clusters can not only provide a high-level summary of the whole data, but also are important to a number of tasks, such as community detection, frequent pattern mining, recommendation, event detection, next-place prediction, etc. Several methods have been proposed to cluster objects such as tags, hashtags, tweets, users, trajectories, etc. For different objects, the SeST information is used in different ways. For example, to cluster tags or hashtags, the co-occurrence between two objects is an important measure. In other cases, however, the co-occurrence itself is not enough to estimate the relatedness. For example, in Flickr photos, the tags *The Statue of Liberty* and *Times Square* are seldom used together, but the two tags are highly correlated because both of them refer to two famous landmarks in New York City. Thus, in addition to co-occurrence, spatial and temporal features are extracted to cluster tags, based on the assumption that related concepts should have similar spatial and temporal distributions (Zhang et al. 2012). For another example, trajectories can be clustered not only based on locations and visiting orders, but also based on the semantics, e.g., location categories. Now, we can identify groups of objects such as individuals and taxi drivers based on the behaviorally-driven markers of individual and collective movement (Liu et al. 2013).

3.4 Future Directions

Although mining SeST data has been gaining much research attention in recent years, many remaining challenging issues call for new and effective solutions. We outlook some important directions for future research in mining SeST data.

- **Deeper understanding of semantics.** While various techniques have been proposed to incorporate semantics information into the process of mining SeST for useful knowledge, the modeling of the semantics is still built upon simple models e.g., bag of keywords. More accurate methods, e.g., phrase mining (Liu et al. 2015), named entity recognition and typing (Ren et al. 2015), sentiment analysis, etc., are necessary so as to capture the intrinsic semantics more accurately.
- **Managing and integrating multiple data sources.** Current research for mining SeST data mostly consider only one data source. It is interesting and important to integrate the data from different sources (e.g., social media, sensor data) to extract valuable evidence in various aspects.
- **Interactive exploration of SeST data.** In many real-life applications, it is not easy to determine the data mining techniques and model parameters before-hand. Extracting the most useful knowledge from the given SeST data usually involves extensive model and parameter tuning. Therefore, it has become an urgent need to develop techniques that can support interactive exploration of SeST data.
- **Assisting decision making.** How to discover knowledge from SeST data to aid decision making is a promising direction. For example, the semantic enriched mobility of massive people is at great importance to urban planning such as site selection for a new airport. In addition, how to generate interpretable and explorable knowledge is also a critical problem to facilitate decision making processes.

3.5 Summary

With the prevalence of GPS technology and the development of social networks, a sheer amount of SeST data has been accumulated. It involves additional types of information compared with traditional datasets which have up to two dimensions among location, time, and semantics. The multi-dimensional SeST data bring new opportunities along with new challenges to extract knowledge. In this article, we introduced the major challenges, data sources, information representation, and general mining approaches to SeST data mining. We also reviewed nine important tasks and cutting edge studies. Some promising directions for future work were also discussed. To the best of our knowledge, this is the first survey that focuses on summarizing existing techniques for mining SeST data. We hope this article can provide readers with a high-level and systematic overview of the research on SeST data mining.

References

Brilhante IR, Macedo JA, Nardini FM, Perego R, Renso C (2015) On planning sightseeing tours with tripbuilder. Inf Process Manag 51(2):1–15

Chen L, Roy A (2009) Event detection from flickr data through wavelet-based spatial analysis. In: Proceedings of 18th ACM international conference on information and knowledge management, ACM, pp 523–532

Cheng T, Haworth J, Anbaroglu B, Tanaksaranond G, Wang J (2014) Spatiotemporal data mining. In: Fischer MM, Nijkamp P (eds) Handbook of regional science. Springer, Heidelberg, pp 1173–1193

Farrahi K, Gatica-Perez D (2011) Discovering routines from large-scale human locations using probabilistic topic models. ACM Tran on Intell Syst Technol (TIST) 2(1):3

Hauff C, Houben GJ (2012) Placing images on the world map: a microblog-based enrichment approach. In: Proceedings of 35th ACM SIGIR international conference on research and development in information retrieval. ACM, pp 691–700

Kelly R (2009) Twitter study reveals interesting results about usage 40% is "pointless babble". http://www.pearanalytics.com/blog/2009/twitter-study-reveals-interesting-results- 40- percent-pointless-babble/. Retrieved: January 18, 2012

Krumm J, Horvitz E (2015) Eyewitness: identifying local events via space-time signals in twitter feeds. In: Proceedings of 23rd ACM SIGSPATIAL international conference on advances in geographic information systems. ACM, pp 20:1–20:10

Lee CH (2012) Mining spatio-temporal information on microblogging streams using a density-based online clustering method. Expert Syst Appl 39(10):9623–9641

Liu S, Wang S, Jayarajah K, Misra A, Krishnan R (2013) Todmis: mining communities from trajectories. In: Proceedings of 22nd ACM international conference on information and knowledge management. ACM, pp 2109–2118

Liu J, Shang J, Wang C, Ren X, Han J (2015) Mining quality phrases from massive text corpora. In: Proceedings of 2015 ACM SIGMOD international conference on management of data. ACM, pp 1729–1744

Mohan P, Shekhar S, Shine JA, Rogers JP (2012) Cascading spatio-temporal pattern discovery. IEEE Trans Knowl Data Eng (TKDE) 24(11):1977–1992

Pan B, Zheng Y, Wilkie D, Shahabi C (2013) Crowd sensing of traffic anomalies based on human mobility and social media. In: Proceedings of 21st ACM SIGSPATIAL international conference on advances in geographic information systems. ACM, pp 344–353

Pham TAN, Li X, Cong G, Zhang Z (2015) A general graph-based model for recommendation in event-based social networks. In: Proceedings of 31st IEEE international conference on data engineering (ICDE). IEEE, pp 567–578

Ren X, El-Kishky A, Wang C, Tao F, Voss CR, Han J (2015) Clustype: effective entity recognition and typing by relation phrase-based clustering. In: Proceedings of 21st ACM SIGKDD international conference on knowledge discovery and data mining. ACM, pp 995–1004

Shekhar S, Jiang Z, Ali RY, Eftelioglu E, Tang X, Gunturi V, Zhou X (2015) Spatiotemporal data mining: a computational perspective. ISPRS Int J Geo-Inf 4(4):2306–2338

Tsur O, Rappoport A (2012) What's in a hashtag?: content based prediction of the spread of ideas in microblogging communities. In: Proceedings of 5th ACM international conference on web search and data mining. ACM, pp 643–652

Wu F, Li Z, Lee WC, Wang H, Huang Z (2015) Semantic annotaion of mobility data using social media. In: Proceedings of 24th international conference on world wide web. International World Wide Web Conference Steering Committee, pp 1253–1263

Yamaguchi Y, Amagasa T, Kitagawa H, Ikawa Y (2014) Online user location inference exploiting spatiotemporal correlations in social streams. In: Proceedings of 23rd ACM international conference on information and knowledge management. ACM, pp 1139–1148

Ye M, Shou D, Lee WC, Yin P, Janowicz K (2011a) On the semantic annotation of places in location-based social networks. In: Proceedings of 17th ACM SIGKDD international conference on knowledge discovery and data mining. ACM, pp 520–528

Ye M, Yin P, Lee WC, Lee DL (2011b) Exploiting geographical influence for collaborative point-of-interest recommendation. In: Proceedings of 34th ACM SIGIR international conference on research and development in information retrieval. ACM, pp 325–334

Yuan J, Zheng Y, Xie X (2012) Discovering regions of different functions in a city using human mobility and pois. In: Proceedings of 18th ACM SIGKDD international conference on knowledge discovery and data mining. ACM, pp 186–194

Yuan Q, Cong G, Ma Z, Sun A, Thalmann NM (2013) Time-aware point-of-interest recommendation. In: Proceedings of 36th ACM SIGIR international conference on research and development in information retrieval. ACM, pp 363–372

Yuan Q, Cong G, Zhao K, Ma Z, Sun A (2015) Who, where, when, and what: a nonparametric bayesian approach to context-aware recommendation and search for twitter users. ACM Trans Inf Syst (TOIS) 33(1):2

Zhang H, Korayem M, You E, Crandall DJ (2012) Beyond co-occurrence: discovering and visualizing tag relationships from geo-spatial and temporal similarities. In: Proceedings of 5th ACM international conference on web search and data mining. ACM, pp 33–42

Zhang C, Han J, Shou L, Lu J, La Porta T (2014) Splitter: mining fine-grained sequential patterns in semantic trajectories. Proc VLDB Endowment 7(9):769–780

Zhang C, Zhang K, Yuan Q, Zhang L, Hanratty T, Han J (2016) Gmove: group-level mobility modeling using geo-tagged social media. In: Proceedings of 22ed ACM SIGKDD international conference on knowledge discovery and data mining. ACM

Zhao L, Chen F, Lu CT, Ramakrishnan N (2015) Spatiotemporal event forecasting in social media. In: Proceedings of 15th SIAM international conference on data mining. SIAM, pp 963–971

Zheng Y, Liu F, Hsieh HP (2013) U-air: when urban air quality inference meets big data. In: Proceedings of 19th ACM SIGKDD international conference on knowledge discovery and data mining. ACM, pp 1436–1444

Zhong Y, Yuan NJ, Zhong W, Zhang F, Xie X (2015) You are where you go: inferring demographic attributes from location check-ins. In: Proceedings of 8th ACM international conference on web search and data mining. ACM, pp 295–304

Zhuang J, Mei T, Hoi SC, Xu YQ, Li S (2011) When recommendation meets mobile: contextual and personalized recommendation on the go. In: Proceedings of 13th international conference on ubiquitous computing. ACM, pp 153–162