# A Survey on Recommendations in Location-based Social Networks

Jie Bao, University of Minnesota
Yu Zheng, Microsoft Research Asia
David Wilkie, University of North Carolina
Mohamed F. Mokbel, University of Minnesota

Recent advances in position localization techniques have fundamentally enhanced social networking services, allowing users to share their locations and location-related content, such as geo-tagged photos and notes. We refer to these social networks as location-based social networks (LBSNs). Location data both bridges the gap between the physical and digital worlds and enables a deeper understanding of user preferences and behavior. This addition of vast geospatial datasets has stimulated research into novel recommender systems that seek to facilitate users' travels and social interactions. In this paper, we offer a systematic review of this research, summarizing the contributions of individual efforts and exploring their relations. We discuss the new properties and challenges that location brings to recommendation systems for LBSNs. We present a comprehensive survey of recommender systems for LBSNs, analyzing 1) the data source used, 2) the methodology employed to generate a recommendation, and 3) the objective of the recommendation. We propose three taxonomies that partition the recommender systems according to the properties listed above. First, we categorize the recommender systems by the objective of the recommendation, which can include locations, users, activities, or social media.Second, we categorize the recommender systems by the methodologies employed, including content-based, link analysis-based, and collaborative filtering-based methodologies. Third, we categorize the systems by the data sources used, including user profiles, user online histories, and user location histories. For each category, we summarize the goals and contributions of each system and highlight one representative research effort. Further, we provide comparative analysis of the recommendation systems within each category. Finally, we discuss methods of evaluation for these recommender systems and point out promising research topics for future work. This article presents a panorama of the recommendation systems in location-based social networks with a balanced depth, facilitating research into this important research theme.

Categories and Subject Descriptors: H.2.8 [**Database Management**]: Database Applications, Data mining, Spatial databases and GIS

General Terms: Survey

Additional Key Words and Phrases: Location-based Social Networks, Recommender Systems, Social Networks, Location-based Services, Location recommendations, Friend recommendations, Community discoveries, Activity recommendations, Social media recommendations

**ACM Reference Format:**
Bao, J., Zheng, Y., Mokbel, M. 2012. Recommendations in Location-based Social Networks. ACM TIST. V, N, Article A (January YYYY), 30 pages.
DOI = 10.1145/0000000.0000000 http://doi.acm.org/10.1145/0000000.0000000

## 1. INTRODUCTION

With millions of users, social networking services like Facebook and Twitter have become some of the most popular Internet applications. The rich knowledge that has accumulated in these social sites enables a variety of recommendation systems for new friends and media.

Author's addresses: Jie Bao and Mohamed F. Mokbel, Department of Computer Science & Engineering, University of Minnesota; Email: {baojie, mokbel}@cs.umn.edu
Yu Zheng, Microsoft Research Asia, No. 5 Danling, Haidian District, Beijing, China; Email: yuzheng@microsoft.com

DOI 10.1145/0000000.0000000 http://doi.acm.org/10.1145/0000000.0000000

Recently, advances in location-acquisition and wireless communication technologies have enabled the creation of location-based social networking services, such as Foursquare, Twinkle, and GeoLife [Zheng et al.2008;Zheng et al.2010b;Zheng et al. 2009c]. In such a service, users can easily share their geospatial locations and location-related content in the physical world via online platforms. For example, a user with a mobile phone can share comments with his social network about a restaurant at which he has dined on an online social site. Other users can expand their social networks using friend suggestions derived from overlapped location histories. For instance, people who constantly hike on the same mountain can be put in contact.

The location dimension bridges the gap between the physical world and the digital online social networking services, giving rise to new opportunities and challenges in traditional recommender systems in the following aspects:

(1) *Complex objects and relations:* A location is a new object in location-based social networks (LBSNs), generating new relations between users, between locations, and between users and locations. New recommendation scenarios, like location and itinerary recommendations, can be enabled using this new knowledge, and traditional recommendation scenarios, such as friend and media recommendation, can be enhanced. However, doing so requires new methodologies for generating high-quality recommendations.
(2) *Rich knowledge:* A location is one of the most important components defining a user's context. Extensive knowledge about a user's behavior and preferences can be learned via their location history [Ye et al. 2009]. The huge volume of location-related data generated by users improves the likelihood that social opinions, e.g., the most favorite dish in a restaurant or the most popular activity at a point of interest, can be accurately assessed by recommendation systems.

These opportunities and challenges have been tackled by many new approaches to recommendation systems, using different data sources and methodologies to generate different kinds of recommendations. In this article, we provide a survey of these systems, and the publications proposing them, with a systematic review on over fifty articles published over the last four years in the major journals, conferences, and workshops, including KDD, WWW, Ubicomp, ACM SIGSPATIAL, LBSN, RecSys, ACM TIST and VLDB. For each publication, we analyze 1) what a produced recommendation is (i.e., the objective of a recommendation), 2) the methodology employed to generate a recommendation, and 3) the data source it used. According to these three aspects, we propose three taxonomies to respectively partition the recommender systems. This survey presents a panorama of the recommendations in location-based social networks with a balanced depth, facilitating research into this rising topic. The contributions of this article are detailed as follows:

— We distinguish LBSNs from conventional social networks and define their unique properties, challenges, and opportunities.
— We categorize the major recommender systems for LBSNs in three taxonomies, organized by data sources, methodologies, and recommendation objectives. In each category, we summarize the goals and contributions of each system. In addition, we highlight one representative system in each category, providing a more in-depth view of the methodology.
— We summarize the major methods for evaluating the recommendations in LBSNs .
— We point out promising research directions in LBSN recommendation systems, paying special attention to directions that result from the analysis and synthesis of the different recommendation system categories.

The rest of the paper is organized as follows: In Section 2, we provide an overview of location-based social networks. We then propose taxonomies for existing recommendation systems for LBSNs in the three subsequent sections. In Section 3, we propose a taxonomy organized by objective of the recommendations. In Section 4, we propose a taxonomy organized by the methodology of the recommendation system. In Section 5, we propose a taxonomy organized by the data source used by the recommendation systems. In Section 6, we summarize the major methods for evaluating a recommendation in an LBSN. In Section 7, we present potential future research directions and discuss how they relate to the existing recommendation systems. Finally, in Section 8 we present our concluding remarks.

## 2. OVERVIEW

In this section, we first present a formal definition of location-based social networks. After that, we summarize the unique properties of locations as data and discuss the new challenges they bring to recommendation systems for LBSNs.

### 2.1. Concepts of Location-Based Social Networks

A social network is an abstract structure comprised of individuals connected by one or more types of relations, such as friendships, common interests, and shared knowledge. A social networking service is a participatory digital representation of real-world social networks. These services reflect their users actual social networks, but also enhance those networks and enable their growth by allowing users to share ideas, activities, events, and interests.

The addition of location data strengthens the connection between the social networking services and the real-world social networks. Zheng proposed a formal definition for these location-based social networks [Zheng 2011; 2012]:

"*A location-based social network (LBSN) does not only mean adding a location to an existing social network so that people in the social structure can share location-embedded information, but also consists of the new social structure made up of individuals connected by the interdependency derived from their locations in the physical world as well as their location-tagged media content, such as photos, video, and text. Here, the physical location consists of the instant location of an individual at a given timestamp and the location history that an individual has accumulated in a certain period. Further, the interdependency includes not only that two persons co-occur in the same physical location or share similar location histories but also the knowledge, e.g., common interests, behaviors, and activities, inferred from an individual's location (history) and location-tagged data*."

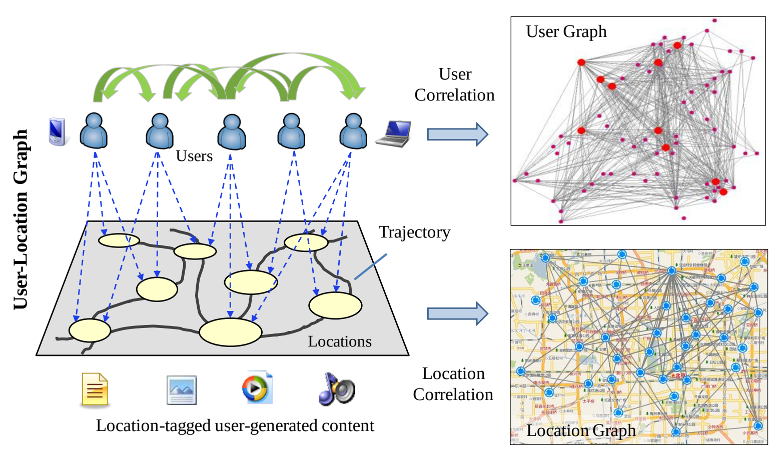



Fig. 1. Concept of location-based social networks.

Figure 1 gives an overview of location-based social networks, in which the addition of locations creates new relations and correlations. Based on this new information, we can build three graphs in location-based social networks: a location-location graph, a user-location graph, and a user-user graph.

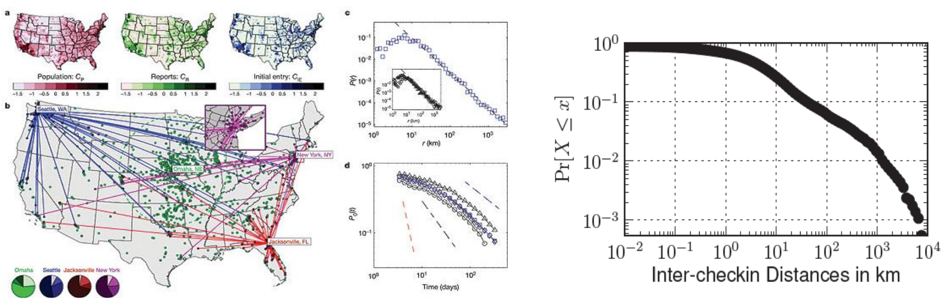
— *Location-location graph.* In the location-location graph (shown in the bottom-right of Figure 1), a node is a location and a directed edge between two locations indicates that some users consecutively visited the locations. The weight associated with an edge represents the strength of the correlation between the two locations.
— *User-location graph.* In the user-location graph (shown in the left of Figure 1), there are two types of nodes, users and locations. An edge starting from a user and ending at a location indicates that the user has visited the location, and the weight of the edge can indicate the number of visits.
— *User-user graph.* In the user-user graph (shown in the top-right of Figure 1), a node is a user and an edge between two nodes represents two relations. One relation is the original connection between two users in an existing social network. The other relation is the new connection derived from the users' locations, e.g., two users may be connected if they have visited the same location, or similar types of places. The latter connection, initially inferred from a user's location history, can be translated to the former through a recommendation mechanism. In other words, we can recommend users to an individual based on the inferred location-based connection. Once the individual accepts the recommendation, the relationship switches from the second category to the first.

The existing location-based social networking services can be classified into three major groups:

— *Geo-tagged-media-based.* Geo-tagging services enable users to add a location label to media content such as text, photos, and videos generated in the physical world. The tagging can occur passively when the content is created or can be added explicitly by the user. Users can view their content in the geographic context in which it was created (on a digital map or in the physical world using a mobile phone). Representative websites of such location-based social networking services include Flickr, Panoramio, and Geo-twitter. Though a location dimension has been added to these social networks, the focus of these services is still on the media content. That is, location is used only as a feature to organize and enrich the media content while the connections between users are based on the media itself.
— *Point-location-based.* Applications like Foursquare and Google Latitude encourage people to share their current locations, such as restaurants or museums. In Foursquare, points and badges are awarded for checking in at venues. The individual with the most number of check-ins at a venue is crowned "Mayor." With the real-time location of users, an individual can discover friends (from her social network) around their physical location to enable social activities in the physical world, e.g., inviting people to have dinner or go shopping. Users can also add tips to venues that other users can read, which serve as suggestions of things to do, see, or eat at the location. With this kind of service, a venue (point location) is the main element determining the connections between users, while user-generated content such as tips and badges are associated with point locations.
— *Trajectory-based.* In a trajectory-based social networking service, such as Bikely, SportsDo, and Microsoft GeoLife, users record both point locations and the route connecting the point locations. These services tell users' basic information, such as distance, duration, and velocity, about a particular trajectory, but they also show users' experiences, represented by tags, tips, and photos along the trajectories. In short, these services provide "how and what" information in addition to "where and when." Other users can reference these experiences (e.g. travel) by browsing or replaying the trajectory on a digital map or in the real world with a GPS-enabled phone.

### 2.2. Influence of Locations in Social Networks

Users' location histories contain a rich set of information reflecting their preferences, once the patterns and correlations in the histories has been analyzed [Eagle and Pentland 2009]. Research into location histories found that the distribution of locations often fit a power law, i.e. the closer locations have a much higher probability of being visited, e.g., [Couldry 2004; Brockmann et al. 2006; Jiang et al. 2008]. In [Brockmann et al. 2006], the authors study the location histories of marked currency as it circulates (shown in Figure 2a). They collect a total of 20,540 trajectories throughout the United States. The authors investigate the probability $P(r)$ of finding a traversal distance $r$ within a number of days. A total of 14,730 (that is, a fraction $Q$ = 0.71) secondary reports occurred outside a short range radius $L_{min}$ = 10 km. The distribution shows power-law behavior $P(r)\, r^{(1+\beta)}$ with an exponent $\beta = 0.59 \pm 0.02$. Recent investigations found similar patterns in users' location histories in LBSNs. For example, [Noulas et al. 2011] studies a large point-location data set collected from Foursquare that reveals several patterns: a user's activities are different during the weekdays and weekends, and the spatiotemporal patterns of users' check-ins fit the power law distribution. They found that 20% of the user's check-ins occur within a distance of 1 km, 60% occur between 1 and 10 km, 20% occur between 10 km and 100 km, and a small percentage extend to distances beyond 100 km. Analysis such as the above, coupled with investigations into user and location correlations and patterns, provide clues of user preferences that can guide recommendation systems.




(a) human mobility study via money circulation [Brockmann et al. 2006]

(b) Human mobility study via user check-ins [Noulas et al. 2011]

Fig. 2. Location Influences in LBSNs.

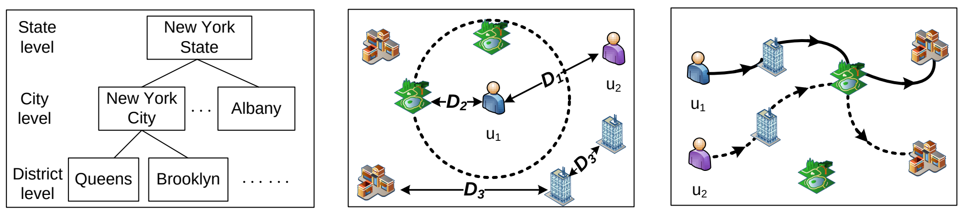
### 2.3. Unique Properties of Locations

Location information brings the following three unique properties to LBSNs, as shown in Figure 3,:

**Hierarchical.** Locations span multiple scales: for example, a location can be as small as a restaurant or as big as a city. Locations with different granularities form a hierarchy, where locations on a lower tiers refer to smaller geographic areas. For example, a restaurant belongs to a neighborhood, the neighborhood belongs to a city, the city belongs to a county, and so on (see Figure 3a). Different levels of location granularity imply different location-location graphs and user-location graphs, even given the same location histories of users. These hierarchical relationships need to be considered as, for example, users who share locations at a lower level (such as a restaurant) likely have a stronger connection than those who share locations at a higher level (such as living in the same city). This hierarchical property is unique in LBSNs, as it does not hold in an academic social network, where a conference never belongs to others.

**Measurable Distances.** Connecting the physical world to a LBSN leads to three new geospatial distance relations, the distance between different users' locations (shown as D1 in Figure 3b), the distance between a user and a location (shown as D2 in Figure 3b), and the distance between two locations (shown as D3 in Figure 3b). According to the first law of geography posed by Waldo Tobler [Tobler 1970], "*everything is related to everything else, but near things are more related than distant things*", we propose that distance affects an LBSN in the following three ways. 1) The user-user distance influences the similarity between users. For example, users with a history of visiting nearby locations are more likely to have similar interests and preferences [Li et al. 2008; Xiao et al. 2010], and users who live close to each other are more likely to be friends [DeScioli et al. 2011]. 2) The user-location distance influences the likelihood a user will be interested in a location. For instance, users in Foursquare visit restaurants close to their homes more frequently than others [Levandoski et al. 2012]. 3) The location-location distance affects the correlations between locations. For example, shopping malls are often placed close to each other [Ye et al. 2011c].

**Sequential ordering.** Subsequent visits by a user to two locations creates a relation with a chronological ordering. For instance, the two users in Figure 3c share a location visiting pattern. From the time of each visit, we can create an ordering which may indicate some similarities between their preferences [Zheng et al. 2009d] or may imply traffic conditions [Tang et al. 2010].




(a) Location Hierarchy Property (b) Location Distance Property (c) Location Sequential Property

Fig. 3. Unique Properties of Locations.

### 2.4. Challenges to Recommendations in LBSNs

While creating new opportunities for LBSNs, the unique properties of locations also bring new challenges such as 1) location context awareness, 2) the heterogeneous domain, and 3) the rate of growth.

*2.4.1. Location Context Awareness.* Recommender systems in LBSNs need to consider how the current location of a user, the location history of the user, and the location histories of other users influences what recommendation to make.

**The Current Location of a User.** A user's current location plays a vital role in generating recommendations in LBSNs due to the following three reasons.

First, a user's current location can be represented on different levels of granularity (the hierarchical property of locations). Choosing a proper granularity for the recommendation scenario is important and challenging. For instance, we should use a fine granularity when recommending restaurants to a user, while a relatively coarse granularity (like in a city or state) for local news recommendations.

Second, the distance property of locations implies that people are more likely to visit nearby locations than distant ones. However, the quality of a location (like a restaurant) is also important for recommendation-making. Ranking a recommendation based on both the user-location distance and the quality of a location is non-trivial. Further, a location indicates a spatial constraint for generating recommendations, but also influences user preferences. For example, beaches might be given a high recommendation rank to a user traveling to Hawaii, even though the user prefers sporting events more than beaches typically. The same user may be more interested in seeing the status of her friends living in Hawaii. An additional challenge is that fine grain location needs to be taken into account quickly: users often access LBSNs via mobile devices that frequently update their location information. Addressing this requires efficient algorithms to generate recommendations quickly.

Third, due to the sequential property of locations, a user's current location affects future travel decisions. For instance, the majority of people visiting Tiananmen Square will subsequently travel to the Forbidden City, or a dessert or drink recommendation may be appropriate after visiting certain restaurants. Discovering these sequential relations and incorporating them into recommendations presents subtle challenges.

**The Historical Locations of the User.** Earlier works, e.g., [Eagle and Pentland 2006; Eagle et al. 2009], have suggest that a user's historical behaviors is a powerful indicator of the user's preferences. A user's historical locations accumulated in an LBSN (e.g., check-ins and geo-tagged photos) reflect more accurately a user's experiences, living patterns, preferences and interests than the user's online behaviors [Zheng and Zhou 2011]. However, it is non-trivial to model a user's location history due to the hierarchy, distance, and sequential properties of locations. Moreover, learning a user's personal preferences from the user's location history is very challenging for the following reasons. 1) As users do not share their locations everywhere, a full set of a user's location history does not exist. Learning a user's preferences from sparse location data is challenging. 2) A user's preferences span multiple kinds of interests, such as shopping, cycling, and arts, rather than consisting of binary decisions, e.g., a set of 'like or dislike' statements. 3) A user's preferences have hierarchies and granularity, such as "Food" $\rightarrow$ "Italian food" $\rightarrow$ "Italian pasta". 4) A user's preferences are constantly evolving (and location dependent).

**The Location Histories of Other Users.** Location histories generated by other users in LBSNs make up the social opinion, which is one of the most important information bases for making recommendations. To extract social opinions from the location histories, however, we are faced with the following two challenges. 1) It is difficult to design a model to consistently represent different users' distinct locations and make these location histories comparable and computable. 2) Users have different degrees of knowledge about different geospatial regions. For instance, local experts of a town are more likely to find high quality restaurants and shopping malls. As a result, weighting different users' data according to their experiences and knowledge is useful when inferring social opinions from the massive user-generated and location-related data. Further, the knowledge of a user is region-related and changes over the granularity of a location. A travel expert in New York City might have less knowledge of Seattle. Likewise, people who are shopping experts in one district of a city might not be the most knowledgeable of the city as a whole. Effectively and efficiently inferring social opinions with respect to users' knowledge of different regions is a difficult problem.

*2.4.2. Heterogeneous Domain.* The graph representing an LBSN is heterogeneous, consisting of at least two types of nodes (user and location) and three types of edges (user-user, location-location, and user-location). Alternatively, we can say there are at least three tightly associated graphs that model an LBSN (as mentioned in Section 2.1). If an LBSN is trajectory-centric, trajectories can be regarded as another type of node in the social network.

A location is not only an additional dimension of information about the user, but also an important object in the LBSN. Inferring the similarity or correlation between two objects in a heterogeneous graph must incorporate the information from related nodes of other types. For instance, determining the connection between two users in an LBSN needs to involve the user-location and location-location relations. A location shared by two users could be evidence of similarity, or it could simply indicate that a location is very popular. Only careful analysis can determine which case holds, and to what extent it should influence the strength of the connection between the users.

*2.4.3. The Rate of Growth.* Location-based social networks evolve at a faster pace than traditional social networks in both social structure and properties of nodes and links. Though academic social networks are also heterogeneous, with authors, conferences, and papers, they evolve at a much slower speed than LBSNs do. For example, adding new links in an LBSN is much easier than it is in a academic social network as visiting a new location is easier than publishing a paper. Further, the properties of nodes and links in a LBSN evolve more quickly than those of academic social networks. A user can become a travel expert in a city after visiting many interesting locations over several months, while a researcher needs years before becoming an expert in a research area. The rate of growth and evolution in LBSNs raise the standard of scalability, efficiency, and updating strategy demanded of recommender systems.

Table I. Comparison of three social networks.

| | Location Awareness | Heterogeneous Environments | Evolving Speed |
|---|---|---|---|
| Academic Social Networks | | √ | Slow |
| General Online Social Networks | | | Fast |
| Location-Based Social Networks | √ | √ | Fast |

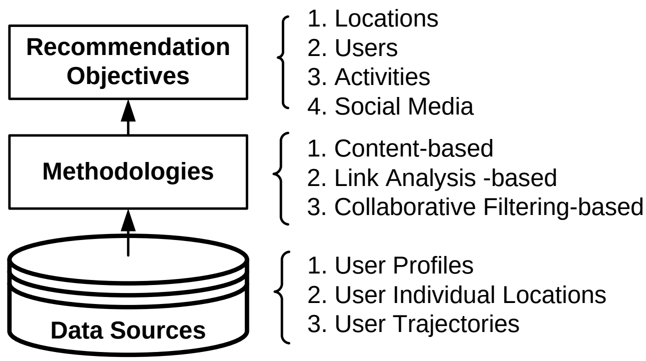



Fig. 4. An overview of recommendation system categories in LBSNs.

We summarize the differences among different types of social networks, e.g., academic networks, such as DBLP, general online social networks, such as Facebook, and location-based social networks, like Foursquare and GeoLife, in Table I. LBSNs present novel opportunities and challenges given the unique properties of locations, the heterogeneous structure of a network, and their high rate of growth and evolution.

### 2.5. Structure of The Paper

To provide a comprehensive survey on recommendations in LBSNs, we studied over forty related publications from the major conferences and journals from 2008 to 2011, as summarized in Table II.

For each publication, we study: 1) what is being recommended (i.e. the objective), 2) the methodology employed to generate the recommendation, and 3) the data source used. Based on these three aspects, we propose three taxonomies to partition these recommendation systems for LBSNs. Following the framework shown in Figure 4, we further detail the three taxonomies taxonomy as follows.

Table II. Statistics on literatures related to Recommendations in LBSNs.

| | Names | 2008 | 2009 | 2010 | 2011 |
|---|---|---|---|---|---|
| Conferences | WWW | 0 | 2 | 3 | 2 |
| | MDM | 1 | 1 | 1 | 1 |
| | KDD | 0 | 0 | 1 | 4 |
| | ACM-GIS | 1 | 1 | 2 | 3 |
| | UbiComp | 0 | 0 | 4 | 1 |
| | LBSN | N/A | 3 | 3 | 5 |
| | RecSys | 0 | 0 | 2 | 1 |
| Journals | VLDB | 0 | 0 | 2 | 0 |
| | ACM-TIST | 0 | 0 | 1 | 1 |
| | ACM TWEB | 0 | 0 | 0 | 1 |
| | PUC | 0 | 0 | 0 | 1 |
| Total Numbers | | 2 | 7 | 19 | 20 |

Table III. Summary of the existing recommendation systems in Location-Based Social Networks.

| | **Objectives** | | | | **Methodologies** | | | **Data Sources** | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Social Media | Location | User | Activity | Content based | Link Analysis | CF | User Profile | Individual Locations | User Trajectories |
| Sandholm, [2011] | √ | | | | | | √ | | √ | |
| Levandoski, [2012] | √ | √ | | | | | √ | | √ | |
| Park, [2007] | | √ | | | √ | | | √ | | |
| Horozov, [2006] | | √ | | | | | √ | | √ | |
| Ye, [2010] | | √ | | | | | √ | | √ | |
| Chow, [2010] | | √ | | | | | √ | | √ | |
| Ye, [2011c] | | √ | | | | | √ | | √ | |
| Tai, [2008] | | √ | | | | √ | | | √ | |
| Yoon, [2010] | | √ | | | | √ | | | | √ |
| Cao, [2010a] | | √ | | | | √ | | | | √ |
| Ye, [2011b] | | √ | | | √ | | | | √ | |
| Cao, [2010b] | | √ | | | √ | | | | | √ |
| Venetis [2011] | | √ | | | √ | | | | √ | |
| Zheng, [2009] | | √ | √ | | | √ | | | | √ |
| Zheng, [2011] | | √ | √ | | | | √ | | | √ |
| Li, [2008] | | | √ | | | | √ | | | √ |
| Hung, [2009] | | | √ | | | | √ | | | √ |
| Xiao, [2010] | | | √ | | √ | | √ | | | √ |
| Ying, [2010] | | | √ | | √ | | √ | | √ | |
| Scellato, [2011b] | | | √ | | | √ | | | √ | |
| Zheng, [2010] | | √ | | √ | | | √ | | √ | √ |
| Symeonidis, [2011] | | √ | | √ | | | √ | | √ | |

**Recommendation objective.** Four types of recommendations are common in LBSNs: 1) location recommendations, which suggest stand-alone locations (e.g., POIs and regions) or sequential locations (such as travel routes and sequences) to a user; 2) user recommendations, which suggest popular users (like local experts and opinion leaders), potential friends (i.e., who share similar interests and preferences), or communities, which a user may wish to join due to shared interests and activities; 3) activity recommendations, which refer to activities that a user may be interested taking into consideration the user's interests and location; 4) social media recommendations, which suggest social media, such as photos, videos, and web contents, to the user taking into account the location of a user and the location metadata of the social media.

**Recommendation system methodology.** We categorize the major methodologies used by the recommendation systems in LBSNs into the following three groups: 1) content-based recommendation, which uses data from a user's profile (e.g., age, gender, and preferred cuisines) and the features of locations (such as categories and tags associated with a location) to make recommendations; 2) link analysis-based recommendation, which applies link analysis models, e.g., hypertext induced topic search (HITS) and PageRank, to identify experienced users and interesting locations; and 3) collaborative filtering (CF) recommendation, which infers a user's preferences from historical behavior (such as from a location history).

**Data sources used.** Recommendation systems in LBSNs can take advantages of various data sources such as: 1) user profiles, which explicitly specify a user's age, gender, interests, preferences, etc.; 2) user geo-located content, which includes a user's ratings of visited locations, geo-tagged content, check-ins, etc.; and 3) user trajectories, consisting of sequential locations contained in a user's GPS trajectories.

Table III provides an overview of some representative publications in regard to the three aspects mentioned above. For instance, Zheng et al. [Zheng et al. 2009b] recommend interesting locations and local experts in a city to users based on user location histories in a form of GPS trajectories using a HIST-based link analysis method.

## 3. CATEGORIZATION BY OBJECTIVES

Location-based social networks open new recommendation possibilities, including locations. In this section, we categorize the existing recommendation systems in LBSNs based on the items they recommended, 1) locations, including the stand-alone locations and traveling routes, 2) users, including expert users, friends recommendation, and community discovery, 3) activities, and 4) social media.

### 3.1. Location Recommendations

As location recommendation is a very broad topic, in this paper, we only focus on location recommendations in the context of social networking. Figure 5 gives an overview of the existing location recommendation system in LBSNs. These systems can be divided into two groups by the objective of their recommendation: 1) stand-alone location recommendation systems, which provide a user with individual locations, such as restaurants or cities, that match their preferences, and 2) sequential location recommendation systems, which recommend a series of locations (e.g., a popular travel route in a city) to a user based on their preferences and their constraints, such as in time and cost. As shown in Figure 5, each type of location recommendation system can be further categorized based on the data sources used.

#### *3.1.1. Stand-alone Location Recommendations.*

The stand-alone location recommendation systems have been a focus of recent research, including the development of multiple prototype systems, e.g.,[Chow et al. 2010;Park et al.2007; Takeuchi and Sugimoto 2006;Yang et al. 2008;Ye et al. 2010;Zheng et al.2010a;Zheng et al.2010c; Zheng and Xie 2011;Zheng et al.2009b]. We can further subdivide and categorize the stand-alone location recommender systems based on the data sources used, as follows.

**User profiles.** These location recommendation systems suggest locations by matching the user's profile against the location metadata, such as description and semantic text and tags. The system proposed in [Park et al. 2007] matches user's profile data – including age, gender, cuisine preferences, and income – against the price and category of a restaurant using a Bayesian network model. In [Ramaswamy et al. 2009], the authors focus on enabling location recommendation on low-end devices capable only of voice and short text messages (SMS). Their approach focuses on using a user's address and 'social affinity', social connections implied by a user's address book, to make recommendations. The social affinity computation and spatiotemporal matching techniques in the system are continuously tuned through the user feedback. In [Kodama et al. 2009], the authors select location candidates using semantic data and make a final recommendation using a skyline operator [Borzsony et al. 2001] that takes into account both the price and the distance of the candidate

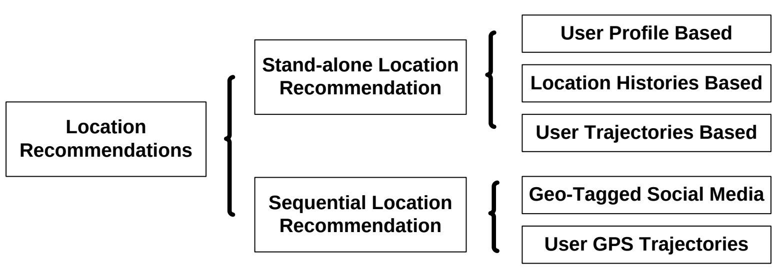



Fig. 5. Location Recommendations in LBSNs.

locations. In [Sheng et al. 2010], the authors propose a method to recommend whole regions, which have geometric extents and points of interest (POI).

Additionally, another branch of ongoing research aims to extract the features of locations, which can be later used in a profile-based location recommendation system. For example, In [Lu et al. 2011], the authors explore the spatial and temporal relationships among individual points within trajectories to identify the subsequences related to the user's preferred activities and assign to them a semantic meaning. [Ye et al. 2011a] proposes a method to extract location features based on the temporal distributions of users' check-ins. [Ye et al. 2011b] extends this work by considering two additional aspects, 1) a set of explicit patterns, including the total number of check-ins, the total number of unique visitors, the maximum number of check-ins by a single visitor, the distribution of check-in times in a week, and the distribution of check-in times in a 24-hour interval, and 2) implicit relatedness, which captures the correlations between locations in check-in behavior.

**User location histories.** A user's location history includes a) their online rating history of locations (e.g., hotels and restaurants) and b) their the check-in history in location-based social networking systems. Using users' location histories, as described above, for making recommendations has advantages over relying solely on profile data as location histories also capture the ratings from the other users. It therefore improves the quality of recommendation by ignoring poorly-reviewed locations that otherwise match user's profile.

Many online web services, e.g. Yelp and Yellowpage, allow users to explicitly express their preferences for locations using ratings. Using these ratings, a body of research, e.g., [Chow et al. 2010; Horozov et al. 2006; Ye et al. 2010; Del Prete and Capra 2010], proposes location recommendation systems using Collaborative Filtering (CF) models that give personalized recommendations for locations that take into account other users' ratings. The intuition behind these methods is that a user will share location preferences with similar users. Most of the CF-based location recommender systems undertake three discrete operations, 1) similarity inference, which calculates a similarity score between users based on their historical ratings, 2) candidate selection, which selects a subset of candidate locations using the user's current location, and 3) recommendation score predication, which predicts the rating a user would give to a location. For example, motivated by the observation that "*people who live in the same neighborhood are likely to visit the same local places*", [Horozov et al. 2006] uses the historical ratings from users living close to the user's query location, which significantly reduces the number of users in the user similarity matrix and thus reduces the computational cost of the recommendation. Similarly [Ye et al. 2010] suggests that solely using the ratings of a user's friends is more efficient and just as effective as using the ratings generated by the top-$k$ most similar users. The authors present a set of experiments showing that a user's friends share more preferences than strangers. In [Chow et al. 2010], the authors focus on the efficiency of a recommendation system that makes location-based recommendations continuously as a user changes their location. In [Del Prete and Capra 2010], the authors present a decentralized mobile recommendation service designed for pervasive environments. In [Ye et al. 2011c], the authors use user check-ins to study the effects of the CF-model, geographical distance, and social structures in making location recommendations. The authors find that geographical distance has the largest impact in their model. In [Shi et al. 2011] the authors propose a personalized location recommendation system based on a category-regularized matrix, which is constructed from the user location histories. The location recommendations consider both the user's preferences as well as a category-based location similarity. [Bao et al. 2012] identifies the three key components in a location recommendation system, a) the user's current location, which constrains the location candidates, b) the user's location histories, which reflect the user's preferences, and c) the location histories from the other users (including local experts), which is considered as the social context.

***Representative Research.*** [Ye et al. 2011c] presents a recommendation system that integrates a) the user's preferences, which are extracted from the check-in history, b) the user's social connections, which are measured by the user's distance to other users in the social network, and c) the geographic distance between the user and the candidate locations. As a result, the probability $S_{i,j}$ of a location $l_j$ to be visited by the user $u_i$ can be estimated using the following equation:

$$S_{i,j} = (1 - \alpha - \beta) \times S^u_{i,j} + \alpha \times S^s_{i,j} + \beta \times S^g_{i,j} \quad (1)$$

where the two weighting parameters $\alpha$ and $\beta$ $(0 \leq \alpha + \beta \leq 1)$ denote the relative importance of social influence and geographical influence compared to user preference. Here $\alpha$ = 1 implies that $S_{i,j}$ depends completely on social influences, $\beta$ = 1 implies that $S_{i,j}$ depends completely on geographical influences, and $\alpha = \beta = 0$ implies that $S_{i,j}$ depends only on user preference.

The authors explore the effect of the different factors in two large data sets from Foursquare and Whrrl. They found their model allowed high precision and recall. Further, they observed that a) geographical influences had a greater impact on the probability of a user visiting a location than did social influences, b) *Random Walk and Restart* may not be suitable for POI recommendations in LBSNs as close social network connections still exhibit significantly different location preferences, and c) the insufficient number of visitors to many locations limits some Collaborative Filtering approaches.

**User trajectories.** Compared to stand-alone check-in data, user-generated trajectories contain a richer set of information, such as the visiting sequence between locations, the path traveled, and the duration of stay at each location. As a result, trajectory data can be used to more accurately estimate a user's preferences. Examples of recommendation systems using trajectory data include [Leung et al. 2011; Takeuchi and Sugimoto 2006; Zheng et al. 2009a; 2009b; Lian and Xie 2011]. In particular, Zheng, et al [Zheng et al. 2009a; 2009b] propose a recommendation framework to find expert users and interesting locations by mining GPS trajectory data. In [Cao et al. 2010a], the authors extend the previous work to consider location-location relations as well as location-user relations. In [Leung et al. 2011] the authors propose a dynamic clustering algorithm in a collaborative location recommendation framework that takes advantage of user classes.

***Representative Research.*** In [Zheng et al. 2009b;Zheng et al.2010c], the authors extend the hypertext induced topic search (HITS) model to extract interesting locations and experienced users using two approaches, 1) dividing the geographical space into a Tree-based Hierarchical Graph (TBHG), and 2) assigning scores to each user and location that indicate the popularity of the location and the travel experience of the user. Figure 6 gives an example of a TBHG structure, in which the multiple layers on the right side of the figure represent the location clusters at different levels of granularity, and the tree structure on the left describes the relationships between the clusters on each level. The intuition behind the score assignment in (2) is that the more experienced users should be better able to recommend interesting locations, while the interesting locations are likely to be accessed by more experienced users, as shown in Figure 7.

In this model, a user's visit to a location is modeled as an edge from the user to the location. Thus, a user is a 'hub' if they have visited many locations, and a location is an 'authority' if it has been accessed by many users. Further, the user's travel experience and a location's interest have a mutually reinforcing relationship. Based on this relationship, a ranking of experienced users and interesting locations can be derived from the model using the following equations:

$$a^l_{ij} = \sum_{u_k \in U} v^k_{jk} \times h^k_{lq} \quad (2)$$

$$h^k_{lq} = \sum_{c_{ij} \in c_{lq}} v^k_{ij} \times a^l_{ij} \quad (3)$$

where the subscripts $ij$ implies that the quantity $x_{ij}$ is of the $i^{th}$ level of the $j^{th}$ cluster in the TBHG, $h^k_{ij}$ represents the $k^{th}$ user's experience, $a^l_{ij}$ represents the location interest, and $c_{lq}$ is $c_{ij}$'s parent

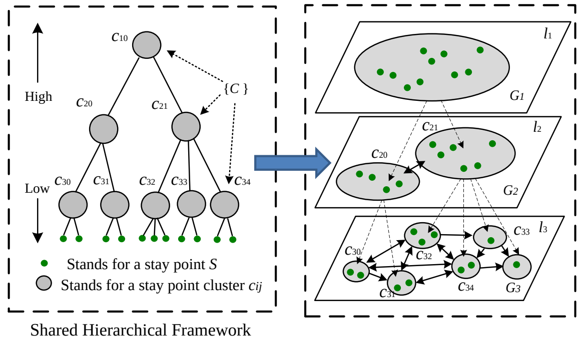



Fig. 6. Tree-based hierarchical graph.

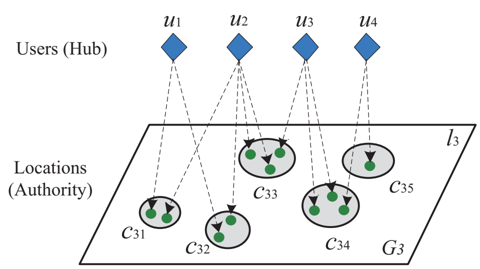


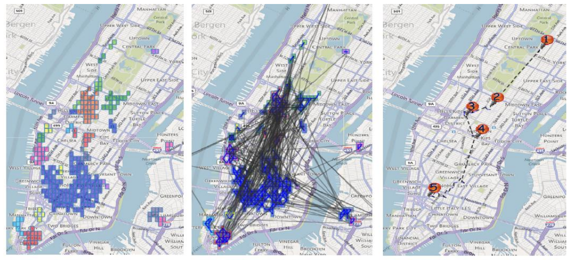

Fig. 7. The HITS inference model.

node on the $l^{th}$ level. The rating is local, as the system rates user experience and location interest at every level of the TBHG, which is consistent with the intuition, for example, that a very experienced user in New York may not have any idea of the interesting locations in Beijing. The authors use this model to extract the top $n$ most interesting locations and the top $k$ most experienced users in a given region using a power iteration method.

*3.1.2. Sequential Location Recommendations.*

Sequential location recommendations can have more complex objectives. For example, a suggested location path could maximize the number of interesting places visited while minimizing travel time or energy consumption. From a user's location history, one can infer how a users preferences for locations are correlated [Zheng and Xie 2010]. A number of sequential location recommendation systems have been proposed based on either users' geo-tagged social media posts [Tai et al. 2008; Lu et al. 2010; Lian and Xie 2011; Wei et al. 2012; Liu et al. 2011] and users' GPS trajectories [Doytsher et al. 2011; Chang et al. 2011; Ge et al. 2011; Ge et al. 2010; Yoon et al. 2010; 2011; Zheng and Xie 2011].

*1) Mining Geo-tagged social media.* A user's geo-tagged social media content can be used as a knowledge base for making sequential location recommendations, e.g., as done in [Arase et al. 2010; Hao et al. 2010]. In [Tai et al. 2008] the authors use association rule mining [Agrawal and Srikant 1994; Han et al. 2000] and sequence mining [Han et al. 2004; Srikant and Agrawal 1996] over sequences of locations extracted from geo-tagged photos. Based on the user's historical visiting pattern, the system creates an itinerary of scenic locations to visit that are popular among other users. Using a vast amount of geo-tagged photos collected from Panoramio, the authors of [Lu et al. 2010] propose a Travel Route Suggestion algorithm to suggest customized travel plans that take into account the time spent at each location, the total travel time, and user preferences. In [Yin et al. 2011a], the authors propose a trip recommendation method that focuses on ranking trajectory pat-



(a) Regions (b) Routable graph (c) Top-1 route

Fig. 8. Construct Popular Routes.

terns mined from uploaded photos. In [Lian and Xie 2011], the authors make use of users' historical visiting patterns, including the type of location, to suggest subsequent locations.

***Representative Research*** In [Wei et al. 2012; Liu et al. 2011], the authors propose the Route Inference framework based on Collective Knowledge (RICK) to construct popular routes from uncertain trajectories. Given a location sequence and a time span, RICK constructs the top-k routes by aggregating uncertain trajectories in a mutually reinforcing way. RICK is comprised of constructing a routable graph and inferring popular routes, as seen in Figure 8. First, RICK constructs a routable graph from uncertain trajectories by aggregating user check-in data. Second, a routing algorithm is used to construct the top-k routes according to a user-specified query.

*2) Mining GPS trajectory.*

GPS trajectories contain a rich set of information, including the duration a user spent at a location and the order of location visits, that can improve sequential location recommendations. In [Doytsher et al. 2011], the authors present a graph model for socio-spatial networks that stores information about frequently traveled routes and implement a route recommendation system using their query language. In [Chang et al. 2011], the authors propose a route recommendation system that takes into account a user's own historically preferred road segments, mined from the user's historical trajectories. The intuition for this approach is that users may feel more comfortable traveling on familiar roads. In [Ge et al. 2011], the authors propose an approach to travel recommendation based on the user's cost constraints, where the travel costs are learned using tour data from professional travel agencies. In [Ge et al. 2010], the authors integrate energy consumption into a mobile recommender system by learning energy-efficient transportation patterns from trajectories.

***Representative Research.*** The itinerary recommendation system [Yoon et al. 2010; 2011; Zheng and Xie 2011] further extended the previous works by incorporating additional constraints, such as 1) a total time constraint on the trip, e.g., a user only has 8 hours for traveling, 2) a destination constraint, which indicates that the user wants to end the trip with a selected location, e.g. a user may need to return to a hotel or the airport, and 3) a constraint on specific ratio metrics, including a) the elapsed time ratio (ETR) between the duration of the recommended trip to the total time constraint, which captures a user's desire to utilize as much available time as possible, b) the stay time ratio (STR) between the amount of time a user stays at location to the amount of time spent traveling between locations, which captures a user's desire to maximize the time in the interesting locations, and c) the interest density ratio (IDR), which is the summation of interest scores for all the locations in the trip over the maximum total interest. Figure 9 shows the architecture of the itinerary recommendation system, containing the following two components:

*Offline model building.* The offline system builds the model used to identify interesting locations and estimate travel times. First, it detects points along the user trajectories at which a user has stayed at a location for some significant duration of time. Next, it clusters these points into interest locations. The duration of a user's stay and the travel time between each location is then computed. Finally, the system infers the interest level based on the HITS model.

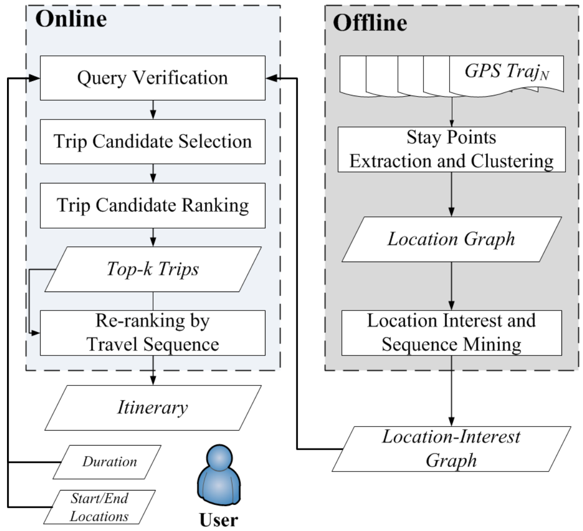



Fig. 9. An overview of itinerary recommendation system.

*Online recommendation.* The online system receives a user's query, including a starting location, a destination, and a time constraint, and returns an itinerary with a sequence of locations. This computation involves three main steps, 1) query verification, which checks the feasibility of the query with the spatial and temporal constraints, 2) itinerary candidate selection, which collects the candidate itineraries based on the HITS model generated in the model building step, and 3) itinerary candidate ranking, which ranks the candidate trips based on the elapsed time ratio, stay time ratio, and interest density ratio.

### 3.2. User Recommendations

User recommendations, which includes popular user discovery [Valente 1996; Burt 1999; Gilbert and Karahalios 2009], friend recommendation [Chen et al. 2009; Backstrom and Leskovec 2011; Roth et al. 2010; Xiang et al. 2010; Yin et al. 2010], and community discovery [Lin et al. 2009; Wiese et al. 2011], have been extensively studied in the context of traditional social networks. The traditional user recommendation approaches are based on the underlying social structure and user interaction patterns. Location-based social networks provide a new way to make user recommendations by also considering users' location histories. Location histories provide rich contextual information and have significant correlations to real social behaviors [Cranshaw et al. 2010]. Several studies reveal that geographical information actually plays a vital role in determining user relationships within social networks. For example, by analyzing the spatial dissemination of new baby names, [Goldenberg and Levy 2009] confirms the importance of geographical proximity, despite the interconnectedness of the Internet era. [Liben-Nowell et al. 2005] shows that at least 2/3 of the friendships in an online social network are determined by the users' locations. [Scellato et al. 2011] analyzes the data collected from a location-based social networking system (Foursquare) and finds that 1) about 40% of the connections are within 100 km, 2) a strong heterogeneity exists across users regarding the spatial distances of connections between their social ties and triads, and 3) gravity models may influence how these social connections are created. Thus, considering users' location histories in an LBSN can improve the effectiveness and efficiency of user recommendations. In this section, we summarize the existing work in user recommendation for location-based social networks, e.g., [Hung et al. 2009; Zheng et al. 2009b; Li et al. 2008; Xiao et al. 2010; Ying et al. 2010; Ying et al. 2011], categorizing each work by its objective, 1) popular user discovery, 2) friend recommendation, or 3) community discovery.

**Popular user discovery.** Traditional approaches to popular user discovery [Valente 1996; Burt 1999] find the opinion leaders in a social networking service by analyzing the node degrees within the information diffusion networks. In LBSNs, we consider 'popular users' to be the users with more knowledge about the locations. Finding experienced users is very important for the recommender systems in LBSNs as these users can provide high quality location recommendations. [Zheng et al. 2009b] finds that a user's traveling experiences are regional, and a user's experience is best determined by considering the qualities of the locations in addition to the number of locations visited. The authors propose a system to identify experienced travelers by applying a HITS inference model over a Tree-Based Hierarchical Graph of users' historical trajectories. [Ying et al. 2011] extends the previous work and proposes four metrics that are used for analysis on EveryTrail (a website for sharing trips). They found that users who share more trajectories get more attention from other users, and users who are popular are more likely to connect to other popular users.

**Friend recommendation.** Traditional friend recommendation systems provide a user with promising potential friends based on their user profiles [Chen et al. 2009; Xiang et al. 2010], the social structure [Doyle and Snell 2000], and the users' interactions [Backstrom and Leskovec 2011; Gilbert and Karahalios 2009; Roth et al. 2010]. Location information in can significantly improve the effectiveness of friend recommendations. The basic intuition is that user location histories reveal preferences, and thus users with similar location histories have similar preferences and are more likely to become friends. Several publications investigate the impact of users' geographical locations on their social relations. For example, a recent study [DeScioli et al. 2011] on MySpace data reveals that users' social connections are highly related to their geographical distances, i.e. that the users living close to each other are more likely to be friends. Moreover, [Backstrom et al. 2010] observes that at medium to long-range distances, the probability of friendship is roughly proportional to the inverse of the distance. However, at shorter ranges, distance does not play as large a role in determining the likelihood of friendships. Similarly, [Scellato et al 2011b] analyzes a large set of data from Gowalla (a location-based social networking system), from which they find that the link prediction space can be reduced by 15 times by focusing on location-friends and friends-of-friends. Based on this observation, they propose a link predication model using supervised learning that considers the users' visited locations. [Yu et al. 2011] builds a pattern-based heterogeneous information network to predict connection probabilities using an unsupervised link analysis model. The connections inside the information network reflect users' geographical histories as well as their social relationships. The connection probability and the friend recommendation score are calculated by a random walk process over the user-location network. Other works, such as [Cho et al. 2011], study the relationship between user movement and friendships through an analysis of mobile phone communications and check-ins. The authors discover that users' short term periodical movement is irrelevant to social structure, but their long distance movement significantly affects their social structure.

A related body of research proposes to measure the similarity between two users from their historical locations and trajectories. [Li et al. 2008] presents a user similarity algorithm that builds a tree-based hierarchical graph of locations. A user's detailed trajectory is abstracted as a set of sequentially visited locations. Based on a sequence matching algorithm that takes into account location hierarchies, the system finds users with similar traveling patterns. [Xiao et al. 2010] extends the user similarity approach by considering the available semantic information for each location, such as its tags and categories. This allows connections between users who have different geographic behaviors, e.g., living in different cities, but share similar semantic behaviors, i.e. they go to the same types of locations. For this approach, the authors transform users' trajectories int location histories with category information. Similarity scores between users are calculated by matching their maximal traveling sequences at different spatial granularities. [Ying et al. 2010] expands on the use of location semantic information. Their framework consists of four phases,

1) semantic trajectory transformation, which converts a user trajectory into a sequence of locations with semantic data, such as parks and schools; 2) maximum semantic trajectory pattern mining, which applies the sequential pattern mining algorithm to each user's trajectory to find the most frequent sequence, 3) semantic similarity measurement, which computes a similarity score between users maximum semantic trajectories, and 4) potential friend recommendation, which uses the constructed user similarity matrix to suggest potential friends.

***Representative Research.*** [Zheng et al. 2011] further extends the user similarity measure framework presented in [Li et al. 2008] by considering the sequences of locations at different spatial granularities. The authors propose a new sequence matching algorithm that divides the location sequences and considers the popularity of each visited locations separately. The newly proposed framework, referred to as a hierarchical-graph-based similarity measurement (HGSM, shown in Figure 10), is proposed to model each individual's location history and measure the similarity between each user. This similarity is based on the users' location histories and is measured using three factors, 1) the shared sequence of users' movements, i.e. the longer the sequence of similar visitations shared by two users, the more similar the two users, 2) the baseline popularity of the locations, e.g. two users visiting a location less traveled might be more correlated than others visiting a popular location, and 3) the hierarchy of geographic spaces, i.e. the finer the granularity of geographic regions shared by two individuals, the more similar these two individuals.

**Community discovery.** Traditional approaches to community discovery often cluster users with either spectral clustering [Mishra et al. 2007; Von Luxburg 2007; Li et al. 2012] or tensor factorization [Lin et al. 2009] based on the social structure (see [Getoor and Diehl 2005] for a detailed survey). With the availability of location information, community discovery in LBSNs can be extended to discover user communities with similar location preferences. For example, [Hung et al. 2009] clusters users based on their traveling patterns, which are mined from their trajectories. First, the authors extract each user's frequently visited locations. They then apply a distance based clustering algorithm to discover communities within the social networks. This computation includes 1) constructing profiles, consisting of a probability suffix tree (PST) for each user describing the

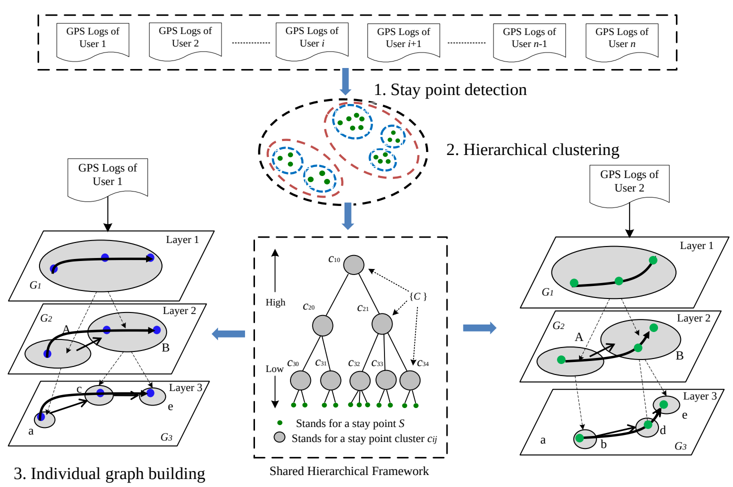



Fig. 10. Hierarchical graph modeling individual location history.

frequency of location visits, 2) measuring the distance between profiles, and 3) identifying communities using a clustering algorithm.

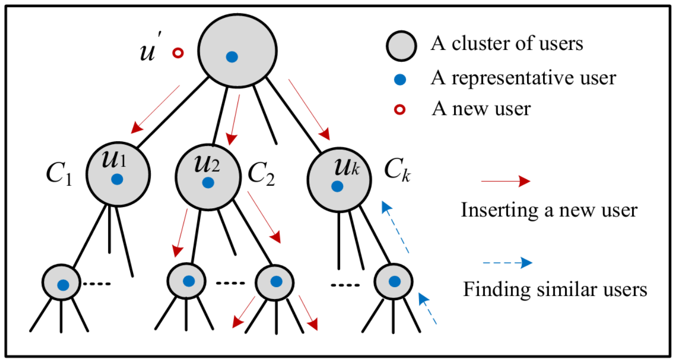



Fig. 11. Hierarchical graph modeling individual location history.

***Representative Research.*** [Xiao et al. 2012] presents an example of this line of research. They hierarchically cluster users into groups by clustering according to the similarity measure proposed in [Xiao et al. 2010]. Consequently, as depicted in Figure 11, they can build a hierarchy of user clusters, where a cluster denotes a group of users sharing some similar interests, at different levels of similarity. The clusters on the higher layers stand for big communities in which people share some high-level interests, e.g. sports. The clusters occurring at the lower layers denote people sharing some narrower interests, e.g. hiking a particular mountain.

### 3.3. Activity Recommendations

An activity recommendation in an LBSN is an information retrieval operation of one or more activities that are appropriate for a query location. For example, sightseeing, boating, and jogging could be recommended for the Olympic Park of Beijing. A list of possible activities at a location can be obtained directly from user-labeled tags or inferred from users' location histories and the semantic data attached to each location.

*3.3.1. Individual Inference-based Approaches.* A user's activity at a certain location can be inferred from the user's geo-tagged social media data and the POI dataset. For example, [Yin et al. 2011b] studies the distributions of some geographical topics (like beach, hiking, and sunset) from the geo-tagged photos acquired from Flickr. [Pozdnoukhov and Kaiser 2011] studies a large set of geo-tagged tweets to explore the spatial-temporal distribution of the topical content. The authors show that the topics, and thus activities, are often geospatially correlated. [Huang et al. 2010] proposes a method to automatically detect activities using the spatial temporal attractiveness (STPA) of points of interest (POI). By comparing the sub-trajectories contained in each POI's STPA, the authors show that most likely activities and their durations can be discovered. The accuracy of this method depends on the POIs and trajectories having accurate arrival time, duration, spatial accuracy, as well as other background factors.

*3.3.2. Collaborative Learning-based Approaches.* One shortcoming of individual inference-based approaches is that they have difficulty dealing with data sparsity, which can be a common occurrence in LBSNs as some users may have a limited location history and some locations may receive few visitors. An alternative approach based on collaborative learning uses information from all users to discover activities. This idea was first proposed in [Zheng et al. 2009d], which extracts the location semantics from GPS data and uses it in conjunction with user profile data to identify

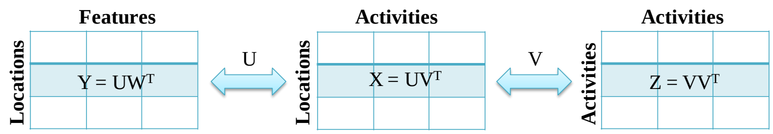



Fig. 12. Collaborative location-activity leaning model.

activities. The system exploits the connections between the user activities and profiles in a joint learning process. Further, [Zheng et al. 2010c] proposes a new model for location and activity histories using a user-location-activity rating tensor. Their system uses this model to provide location-specific activity recommendations. [Zheng et al. 2012] proposes a new algorithm that uses a ranking-based collective tensor and matrix factorization model. Separately, [Symeonidis et al. 2011] extends the previous work by using the Higher Order Singular Value Decomposition (HOSVD) technique to perform dimensionality reduction and semantic analysis. As more data is accumulated by their system, it uses incremental solutions to update a tensor that includes users, locations and activities.

***Representative Research.*** [Zheng et al. 2010c] provides location and activity recommendations in LBSNs to answer two questions for the tourists, 1) where to go for activities such as sightseeing or dining in a large city and 2) what activities are available at specific locations, e.g. if someone visits the Bird's Nest in Beijing Olympic park, what can they do there? The major challenge is due to data sparsity, as users in the system have very limited histories. To this end, the authors propose a collaborative-based approach to extract the features for the locations. Three matrices are constructed as the data model, as shown in Figure 12:
*Location-activity matrix.* A user can log an activity in order to associate it with a point in a trajectory. For example, in Foursquare, users can associate content with venues to share with their friends. The specification of both activity and location in this social media enables the authors to study the correlation between locations and activities and to construct a location-activity matrix. Ideally, the activities associated with a location can be discovered from the location-activity matrix. However, the matrix is typically very sparse as the amount of user-added content is dwarfed by the number of locations. To address this, the paper uses the location-feature and activity-activity matrices to infer missing items in the location-activity matrix, as shown in Figure 12.
*Location-feature matrix.* This matrix connects locations and categories (such as restaurants, cafes, and bars) based on the intuition that locations of the same category are likely to have the same activity possibilities. In this matrix, a location may include multiple categories (or features). For example, a mall would include shops, movie theaters, and cafes. The matrix is built from a POI database, in which each POI is associated with a set of properties such as, name, address, GPS coordinates, and categories.
*Activity-activity matrix.* This matrix models the correlations between different activities. From this, the authors infer the likelihood of an activity being performed at a location given that a user has performed some other activity. The paper suggests two ways to determine these correlations, (1) by mining the user-created content and (2) by using the number of search engine results for the activity terms (if the user-content is insufficient).

After the system constructs the three matrices, a filtering approach is applied to train the location-activity recommendation system using collective matrix factorization [Singh and Gordon 2008]. An objective function, shown in Equation 4, is defined to infer the missing values. This function is iteratively minimized using gradient descent.

$$L(U,V,W) = \frac{1}{2} \parallel I \circ (X - UV^T) \parallel_F^2 + \frac{\lambda_1}{2} \parallel Y - UW^T \parallel_F^2 + \frac{\lambda_2}{2} \parallel Z - VV^T \parallel_F^2 + \frac{\lambda_3}{2} (\parallel U \parallel_F^2 + \parallel V \parallel_F^2 + \parallel W \parallel_F^2) \quad (4)$$

Where $\| \cdot \|_F$ denotes the Frobenius norm. $I$ is an indicator matrix with its entry $I_{ij} = 0$ if $X_ij$ is missing, $I_{ij} = 1$ otherwise. The operator "$\circ$" denotes the entry-wise product. As shown in Figure 12, the authors propagate the information among $X_{m \times n}$, $Y_{m \times l}$ and $Z_{n \times n}$ by requiring the matrices to share the low-rank matrices $U_{m \times k}$ and $V_{n \times k}$. The first three terms in Equation 4 control the loss in matrix factorization, and the last term controls the regularization over the factorized matrices to prevent over-fitting. From the final location-activity matrix, the top $k$ values are suggested as activities for the location.

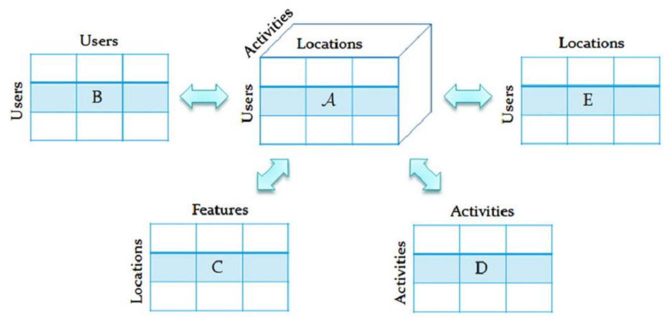



Fig. 13. Personalized Collaborative location-activity leaning model.

One limitation of the proposed activity recommendation approach is that it can not provide personalized recommendations for the users that take into account each user's preferences. Therefore, [Zheng et al. 2010a] extends the approach to create a personalized activity recommendation system which includes user-user and user-location matrices. Specifically, the authors model the user-location-activity tensor $\mathcal{A}$ under the factorization framework and use additional information to address the data sparsity issue. Figure 13 illustrates the new tensor model. Data scarcity results in missing entries in tensor $\mathcal{A}$ that must be filled. In addition to the location-features, activity-feature, and activity-activity matrices used in the previous system, the matrix $B \in \mathbb{R}^{m \times m}$, which encodes the user-user similarities, and the matrix $E \in \mathbb{R}^{m \times n}$, which models the user's location visiting preferences, are added to the computation. Finally, to fill the entries in tensor $\mathcal{A}$, model-based methods are applied [Srebro et al. 2003; Singh and Gordon 2008] to decompose the tensor $\mathcal{A}$ with respect to each tensor entity.

### 3.4. Social Media Recommendations

Social media recommendation aims to provide users with suggestions of photos, videos, or other web content they might like. Using location information in LBSNs can improve both the effectiveness and efficiency of traditional social media recommendations. Several works in spatial keyword searching for web content show the effectiveness of this pairing, e.g., [Chen et al. 2009; Zhang et al. 2009; Cao et al. 2010b; Cao et al. 2011; Bouidghaghen et al. 2011].

[Mokbel et al. 2011] analyzes the rating data from MovieLens [MovieLens ] and finds that people at different locations have different preferences. For example, users from Minnesota are more interested in crime and war movies, while users from Florida are more interested in fantasy and animation movies. Location-aware image ranking algorithms have been proposed to increase the relevance of the search results, e.g., [Arase et al. 2009; Kawakubo and Yanai 2011]. [Silva 2011] improves the quality of the image tags using a recommender system to automatically infer and suggest candidate location tags. [Daly and Geyer 2011] discovers events using both social and location information.

The efficiency of recommendation systems can be significantly improved by using location data to prune out irrelevant information. [Scellato et al. 2011a] improves the efficiency of content delivery networks using a novel caching mechanism based on geographic location. [Sandholm and Ung 2011] builds a real-time recommendation system for online web content using a collaborative filtering method to make more diverse and personalized recommendations within a geographical area. [Levandoski et al. 2012] proposes a novel location-aware recommendation framework, LARS, to exploit users' ratings of locations using a technique that uses the distance of querying users to influences recommendations.

## 4. CATEGORIZATION BY METHODOLOGY

Although traditional recommendation systems have been successful by using community opinions, e.g., inventories in Amazon [Linden et al. 2003] and news from Google [Das et al. 2007], incorporating location information requires novel approaches. In this section, we categorize the major methodologies used by recommendation systems in location-based social networks as being based on: 1) content, 2) link analysis, or 3) collaborative filtering.

### 4.1. Content-based Recommendations

Content-based recommendation systems, such as [Park et al. 2007; Ramaswamy et al. 2009], match user preferences, discovered from users' profiles, with features extracted from locations, such as tags and categories, to make recommendations. These systems require accurate and structured information for both the user profiles and the location features to make high quality recommendations.

The major advantages of the content-based approach that such a system is robust against the cold start problem for both new users and locations. As long as the newly added user or location has the appropriate descriptive content, they can be handled effectively. However, content-based recommendation systems have many drawbacks in regard to LBSNs: 1) content-based recommendation systems do not consider the aggregated community opinions (inferred from users), which may result low quality recommendations, and 2) content-based recommendation systems require that the structured information for both users and locations be created and maintained, which can be costly, especially in LBSNs in which the majority of the contents (i.e., user profiles and location tags) are generated by the users.

### 4.2. Link analysis-based Recommendations

Link analysis algorithms, e.g., PageRank [Page et al. 1999] and Hypertext Induced Topic Search (HITS) [Chakrabarti et al. 1998; Kleinberg 1999], are widely used to rank the web pages. These algorithms extract high quality nodes from a complex network by analyzing the structure. In LBSNs, there are interconnected networks of different types, e.g., user-user, user-location, and location-location networks. [Zheng et al. 2009b] extends the HITS algorithm for discovering experienced users and interesting locations in an LBSN. In their system, each location is assigned a popularity score, and each user is assigned a hub score, which indicates their travel expertise. Based on a mutually reinforcing relationship, a ranking of expert users and interesting locations is computed. Similarly, [Raymond et al. 2011] extends a random walk-based link analysis algorithm to provide location recommendation.

The advantages of link analysis-based methodologies are that 1) they take into account the user's experiences when making recommendations and amplify ratings from experienced users, and 2) they are robust against the cold start problem. However, they have a major drawback: they can only provide generic recommendations for all users, which overlooks users' personal preferences.

### 4.3. Collaborative Filtering-based Recommendations

Collaborative filtering (CF) is widely used in conventional recommendation systems [Adomavicius and Tuzhilin 2005]. The intuition in extending the CF model for recommendations in LBSNs is that a user is more likely to visit a location if it is preferred by similar users. The CF approach used by recommender systems in LBSNs consists of three processes: 1) candidate selection, 2) similarity inference, and 3) recommendation score predication.

**Candidate Selections.** The first step of CF-based recommendation systems is to select a subset of candidate nodes to reduce the computational overhead. The traditional CF-based recommendation algorithms limit use the most similar users (or locations, activities, etc.) as the candidates. CF-based recommender systems in LBSNs can also use geographic bounds and associations to constrain the candidate selection process. A spatial range can be computed to prune candidate locations, e.g., [Chow et al. 2010]. [Horozov et al. 2006] selects candidate users by considering only individuals who live near the user's querying location. Non-geographic criteria can also be used. In [Ye et al. 2010], the authors select candidates by considering user preference and social influence, but also geographic influence modeled as a power-law probabilistic model.

**Similarity Inferences.** Similarities between users (or locations, activities, etc.) are inferred from users' ratings and location histories in LBSNs. The traditional CF models can be divided into two subgroups: 1) user-based models, such as [Herlocker et al. 1999], that use similarity measures between each pair of users; and 2) item-based models, such as [Lemire and Maclachlan 2005], that use similarity measures between each pair of items (media content, activities, etc.). The following equation demonstrates a simple user similarity computation for user $u$ and $u'$ using the Cosine correlation function in a user-based CF model:

$$\mathit{UserSim}(u,u') = \frac{\sum_{o\in\mathcal{O}} r(o,u)\times r(o,u')}{\sqrt{\sum_{o\in\mathcal{O}} r(o,u)^2}\sqrt{\sum_{o\in\mathcal{O}} r(o,u')^2}} \tag{5}$$

where $r(o,u)$ is the rating user $u$ gives to each object $o$ in the set of all objects $O$. Many of the existing recommendation systems in LBSNs, e.g., [Chow et al. 2010; Horozov et al. 2006; Ye et al. 2010; Del Prete and Capra 2010], provide location recommendations based on the distribution of user's ratings over their visited locations using the above equation.

Similarity inference between users (and locations etc.) can also be done by analyzing the pattern of location co-visitation. Recently, systems have been proposed that use the number of visitations (e.g., tips and check-ins) at locations as an implicit rating of the location, e.g., [Takeuchi and Sugimoto 2006; Shi et al. 2011]. Location similarity can also be captured using sequential relations [Li et al. 2008] or semantic similarities [Xiao et al. 2010].

**Recommendation Score Predication.** Finally, CF systems predict a recommendation score for each object $o_i$ (locations, social media, etc.) in the candidate set. These scores are calculated from ratings given by the set of users ($U$) and the similarity measures between individual users. The following equation gives an example of a recommendation score computation:

$$\mathit{RecScore}(o_i,u) = \frac{\sum_{u_j\in\mathcal{U}'} \mathit{UserSim}(u,u_j)\times r(o_i,u_j)}{\sum_{u_j\in\mathcal{U}':r(o_i,u_j)>0} |\mathit{UserSim}(u,u_j)|} \tag{6}$$

The advantages of the collaborative filtering models are that 1) they do not need to maintain well structured descriptions of items (locations, activities, etc.) or users, and 2) they take advantage of community opinions, which provide high quality recommendations. However, CF models also suffers from several drawbacks: 1) when data is sparse, e.g. the number of user ratings is low, the user-item (location, etc.) rating matrix is very sparse and the collaborative filtering model fails to make effective recommendations; 2) due to the large number of users and items in the systems, the similarity model construction process is very time consuming, presenting a scalability challenge that is exacerbated by the rapid growth and evolution of LBSNs, and 3) the CF model deals poorly with the cold start problem, providing recommendations for new users or new items in the system.

## 5. CATEGORIZATION BY DATA SOURCES

In this section, we summarize the different types of data sources used in recommendation systems for LBSNs, including 1) user profiles, 2) user online histories, and 3) user location histories.

### 5.1. User Profiles

As in the conventional social networks, LBSN users maintain profiles that may include democraphic data, interests, and preferences. Such profile information is used by many content-based recommender systems, e.g., [Park et al. 2007], to recommend locations based on the location's categories, user generated tags, etc. Other research, e.g. [Ye et al. 2011a; 2011b], focuses on improving the accuracy of the location tags and categories by extracting user activity patterns for each location.

### 5.2. User Online Histories

Users' online histories come in three main classes, user ratings, user interaction patterns, and user search histories. Users in LBSNs may leave explicit ratings for locations to express their opinions, just as they can in traditional recommender systems. User ratings in LBSNs are associated with locations and can be used to find similar users or similar locations, e.g., [Chow et al. 2010; Horozov et al. 2006; Ye et al. 2010]. User interaction patterns in LBSNs include user tags and commenting patterns. The user interaction patterns are used for friend recommendation and community discovery systems, e.g. as in [Gilbert and Karahalios 2009; Xiang et al. 2010]. User search histories include map browsing histories and spatial searching logs. By accumulating such information, recommendation systems can estimate the community's knowledge and preferences, e.g., [Weakliam et al. 2005; Ballatore et al. 2010; Venetis et al. 2011].

### 5.3. User Location Histories

A user location history is a record of a user's previously visited locations accumulated in an LBSN, including for example check-in data and trajectories. A user's location history can be a more accurate data source to study the user's behaviors and preferences as it records where users actually go, rather than what they list as preferences. Location histories can also be used for friend recommendation. For example, when two users share the location history sequence or stay similar amounts of time at a same location, it provides evidence that the users share preferences and interests.

## 6. METHODS OF EVALUATION

Recommendation systems in LBSNs have typically used two methods to evaluate the effectiveness of their recommendations, 1) user studies and 2) precision and recall ratios.

**User Studies.** To conduct a user study of a recommendation system, the researchers invite multiple subjects to use the recommender system and evaluate its performance, e.g., [Zheng et al. 010c]. For each recommendation task, the subjects need to evaluate the top-k recommendations suggested by the recommendation system.

To create a baseline for evaluation, researchers aggregate all the feedback provided by the subjects to create an ideal ranking list. As recommendations are based on result rankings, the normalized discounted cumulative gain (nDCG) [Manning et al. 2008] is used to measure the effectiveness of the recommendation list. nDCG is also commonly used in information retrieval to measure search engine performance. A higher nDCG value means that more relevant items appear first in the results list.

**Precision and Recall Ratios.** Precision and recall ratios are also used to evaluate the effectiveness of recommendations in LBSNs, e.g.,[Ye et al. 2011c], [Bao et al. 2012]. To use this evaluation method, a user's location history is divided into two parts, 1) the location history generated within a

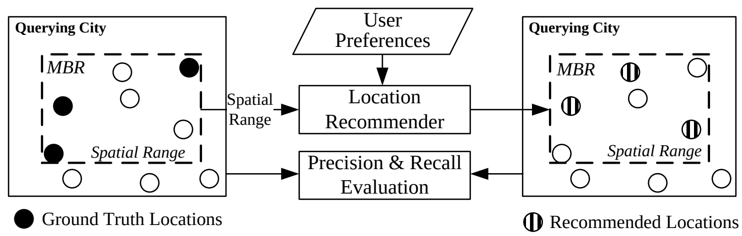



Fig. 14. Evaluate Recommendation using Precision and Recall Ratios.

query area, which is used as ground truth, and 2) the rest of the user's location history, which is used as a training set to learn the user's preferences and build the recommendation model. The system is then evaluated by whether it can suggest those sites within the querying region that the user has actually visited based on the training data (the location history outside of the query region).

For example, in the left part of Figure 14, the black dots are the venues the user visited. A system trained with data outside the query region (the dotted square) recommends the venues illustrated by the striped dots in the right part of Figure 14. Using the black dots as ground truth, recall and precision can be calculated.

$$precision = \frac{number\ of\ recovered\ ground\ truths}{total\ number\ of\ recommendations} \tag{7}$$

$$recall = \frac{number\ of\ recovered\ ground\ truths}{total\ number\ of\ ground\ truths}. \tag{8}$$

This evaluation measurement may be pessimistic as, for example, a user may still prefer a location even if the user has not yet visited it.

## 7. FUTURE WORK

Although many recommendation systems have been proposed in LBSNs, there are still many open questions and challenges to be addressed. In this section, we summarize potential research directions to improve the effectiveness and efficiency of recommendation systems in LBSNs.

### 7.1. Effectiveness of Recommendations

To improve their effectiveness, recommendation systems need more accurate estimations of user preferences and social knowledge. Potential paths to achieve this include 1) using diverse data sources, 2) integrating and hybridizing different types of recommendation methodologies, and 3) increasing context awareness.

**Diverse Data Sources.** Most recommendation systems in LBSNs currently use only one type of the data source to make recommendations. However, there are many different types data in LBSNs, e.g., users' friendships, online interactions, and user location histories. Moreover, many of the data sources are related and may mutually reinforce each other. By considering more diversified data sources, more effective recommendations can be provided. For instance, the user online interactions, social structures, and location histories are all very relevant to friend recommendation. If two users have more online interactions, are close in the social structure, and have overlapped location histories, these users are likely to be compatible. A friend recommender system that can consider all these factors will make higher quality friend recommendations.

**Hybrid Methodologies.** The recommendation methodologies used in the existing recommendation systems each have their own drawbacks. For example, in collaborative filtering based recommendation systems, data sparsity and cold starts are challenging problems. Link analysis-based recommendation systems avoid these problems, but only provide generic recommendations that ignore users' personal preferences. By integrating CF and link analysis-based techniques, a hybrid recommendation system could overcome the weaknesses of both.

**Context Awareness.** Current recommendation systems in LBSNs use a user's history to extract preferences. However, the user's context is currently ignored. A context aware recommender system in LBSNs would need to consider 1) user context, including static attributes like income, profession, and age, as well as dynamic attributes include current user location, mood, and status, (e.g., at home or in meeting) and 2) environmental context, including information about the surrounding environment, e.g. the current time, weather, traffic conditions, events, etc.

### 7.2. Efficiency of Recommendations

Recommendations in LBSNs can be computationally costly, especially given the frequency with which users add new location data and content.

**User Mobility.** Users in LBSNs interact with the services using mobile devices and want up-to-date recommendations based on their current location. However, processing continuous recommendation requests as multiple individual requests is inefficient as many redundant computations are undertaken between the consecutive recommendation queries. To address this, more advanced recommendation algorithms are required that leverage prior computations to reduce the cost of continuous recommendation requests.

**Frequent User Updates.** Users in LBSNs can be very active. They visit many locations over short time spans, which adds information related to their preferences at a high rate. It is very inefficient to re-compute the user preferences and user similarities every time a user undertakes a new activity. As a result, new recommendation techniques are required to efficiently address the update frequency in LBSNs.

## 8. CONCLUSION

Motivated by the prevalence of location-based social networks and the importance of recommendation systems, we have provided a systematic survey of the related recent research. We studied over 50 papers published in the last five years at over 10 major conference and in journals, such as KDD, WWW, RecSys, UbiComp, ACM SIGSPATIAL LBSN, ACM TIST, and ACM TWEB. We provided categorizations of existing systems in regard to their data sources, their methodologies, and their recommendation objective. This survey presents a panorama of this research with a balanced depth and scope. Further, this survey serves as a tutorial, introducing the concepts, unique properties, challenges, representative solutions and systems, evaluation methods, and future work for recommendation systems in LBSNs.

## REFERENCES


ADOMAVICIUS, G. AND TUZHILIN, A. 2005. Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. *Knowledge and Data Engineering, IEEE Transactions on 17,* 6, 734–749.

AGRAWAL, R. AND SRIKANT, R. 1994. Fast algorithms for mining association rules. In *Proc 20th Int Conf Very Large Data Bases VLDB*. Vol. 1215. Citeseer, 487–499.

ARASE, Y., XIE, X., DUAN, M., HARA, T., AND NISHIO, S. 2009. A game based approach to assign geographical relevance to web images. In *Proceedings of the 18th international conference on World wide web*. ACM, 811–820.

ARASE, Y., XIE, X., HARA, T., AND NISHIO, S. 2010. Mining people's trips from large scale geo-tagged photos. In *Proceedings of the international conference on Multimedia*. ACM, 133–142.

BACKSTROM, L. AND LESKOVEC, J. 2011. Supervised random walks: predicting and recommending links in social networks. In *Proceedings of the fourth ACM international conference on Web search and data mining*. ACM, 635–644.

BACKSTROM, L., SUN, E., AND MARLOW, C. 2010. Find me if you can: improving geographical prediction with social and spatial proximity. In *Proceedings of the 19th international conference on World wide web*. ACM, 61–70.

BALLATORE, A., MCARDLE, G., KELLY, C., AND BERTOLOTTO, M. 2010. Recomap: an interactive and adaptive map-based recommender. In *Proceedings of the 2010 ACM Symposium on Applied Computing*. ACM, 887–891.

BAO, J., ZHENG, Y., AND MOKBEL, M. 2012. Location-based and preference-aware recommendation using sparse geo-social networking data. In *ACM SIGSPATIAL*.

BORZSONY, S., KOSSMANN, D., AND STOCKER, K. 2001. The skyline operator. In *Data Engineering, 2001. Proceedings. 17th International Conference on*. IEEE, 421–430.

BOUIDGHAGHEN, O., TAMINE, L., AND BOUGHANEM, M. 2011. Personalizing mobile web search for location sensitive queries. In *Mobile Data Management (MDM), 2011 12th IEEE International Conference on*. Vol. 1. IEEE, 110–118.

BROCKMANN, D., HUFNAGEL, L., AND GEISEL, T. 2006. The scaling laws of human travel. *Nature 439,* 7075, 462–465.

BURT, R. 1999. The social capital of opinion leaders. *The Annals of the American Academy of Political and Social Science 566,* 1, 37–54.

CAO, X., CONG, G., AND JENSEN, C. 2010a. Mining significant semantic locations from gps data. *Proceedings of the VLDB Endowment 3,* 1-2, 1009–1020.

CAO, X., CONG, G., AND JENSEN, C. 2010b. Retrieving top-k prestige-based relevant spatial web objects. *Proceedings of the VLDB Endowment 3,* 1-2, 373–384.

CAO, X., CONG, G., JENSEN, C., AND OOI, B. 2011. Collective spatial keyword querying. In *Proceedings of the 2011 international conference on Management of data*. ACM, 373–384.

CHAKRABARTI, S., DOM, B., RAGHAVAN, P., RAJAGOPALAN, S., GIBSON, D., AND KLEINBERG, J. 1998. Automatic resource compilation by analyzing hyperlink structure and associated text. *Computer Networks and ISDN Systems 30,* 1-7, 65–74.

CHANG, K.-P., WEI, L.-Y., PENG, W.-C., AND YEH, M.-Y. 2011. Discovering personalized routes from tejaectories. In *GIS-LBSN*.

CHEN, J., GEYER, W., DUGAN, C., MULLER, M., AND GUY, I. 2009. Make new friends, but keep the old: recommending people on social networking sites. In *Proceedings of the 27th international conference on Human factors in computing systems*. ACM, 201–210.

CHEN, Y., WANG, W., LIU, Z., AND LIN, X. 2009. Keyword search on structured and semi-structured data. In *Proceedings of the 35th SIGMOD international conference on Management of data*. ACM, 1005–1010.

CHO, E., MYERS, S., AND LESKOVEC, J. 2011. Friendship and mobility: User movement in location-based social networks. *ACM SIGKDD*.

CHOW, C.-Y., BAO, J., AND MOKBEL, M. F. 2010. Towards Location-Based Social Networking Services. In *The 2nd ACM SIGSPATIAL International Workshop on Location Based Social Networks*.

COULDRY, N. 2004. *Mediaspace: Place, scale and culture in a media age*. Routledge.

CRANSHAW, J., TOCH, E., HONG, J., KITTUR, A., AND SADEH, N. 2010. Bridging the gap between physical location and online social networks. In *Proceedings of the 12th ACM international conference on Ubiquitous computing*. ACM, 119–128.

DALY, E. AND GEYER, W. 2011. Effective event discovery: using location and social information for scoping event recommendations. In *Proceedings of the fifth ACM conference on Recommender systems*. ACM, 277–280.

DAS, A., DATAR, M., GARG, A., AND RAJARAM, S. 2007. Google news personalization: scalable online collaborative filtering. In *Proceedings of the 16th international conference on World Wide Web*. ACM, 271–280.

DEL PRETE, L. AND CAPRA, L. 2010. differs: A mobile recommender service. In *Mobile Data Management (MDM), 2010 Eleventh International Conference on*. IEEE, 21–26.

DESCIOLI, P., KURZBAN, R., KOCH, E., AND LIBEN-NOWELL, D. 2011. Best friends. *Perspectives on Psychological Science 6,* 1, 6.

DOYLE, P. AND SNELL, J. 2000. Random walks and electric networks. *Carus mathematical monographs 22*.

DOYTSHER, Y., GALON, B., AND KANZA, Y. . 2011. Storing routes in socio-spatial networks and supporting social-based route recommendation. In *Proceedings of the 3nd ACM SIGSPATIAL International Workshop on Location Based Social Networks*. ACM.

EAGLE, N. AND PENTLAND, A. 2006. Reality mining: sensing complex social systems. *Personal and Ubiquitous Computing 10,* 4, 255–268.

EAGLE, N. AND PENTLAND, A. 2009. Eigenbehaviors: Identifying structure in routine. *Behavioral Ecology and Sociobiology 63,* 7, 1057–1066.

EAGLE, N., PENTLAND, A., AND LAZER, D. 2009. Inferring friendship network structure by using mobile phone data. *Proceedings of the National Academy of Sciences 106,* 36, 15274.

GE, Y., LIU, Q., XIONG, H., TUZHILIN, A., AND CHEN, J. 2011. Cost-aware travel tour recommendation. In *Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 983–991.

GE, Y., XIONG, H., TUZHILIN, A., XIAO, K., GRUTESER, M., AND PAZZANI, M. 2010. An energy-efficient mobile recommender system. In *Proceedings of the 16th ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 899–908.

GETOOR, L. AND DIEHL, C. 2005. Link mining: a survey. *ACM SIGKDD Explorations Newsletter 7,* 2, 3–12.

GILBERT, E. AND KARAHALIOS, K. 2009. Predicting tie strength with social media. In *Proceedings of the 27th international conference on Human factors in computing systems*. ACM, 211–220.

GOLDENBERG, J. AND LEVY, M. 2009. Distance is not dead: Social interaction and geographical distance in the internet era. *arXiv preprint arXiv:0906.3202*.

HAN, J., PEI, J., AND YAN, X. 2004. From sequential pattern mining to structured pattern mining: A pattern-growth approach. *Journal of Computer Science and Technology 19,* 3, 257–279.

HAN, J., PEI, J., AND YIN, Y. 2000. Mining frequent patterns without candidate generation. In *ACM SIGMOD Record*. Vol. 29-2. ACM, 1–12.

HAO, Q., CAI, R., WANG, C., XIAO, R., YANG, J., PANG, Y., AND ZHANG, L. 2010. Equip tourists with knowledge mined from travelogues. In *Proceedings of the 19th international conference on World wide web*. ACM, 401–410.

HERLOCKER, J., KONSTAN, J., BORCHERS, A., AND RIEDL, J. 1999. An algorithmic framework for performing collaborative filtering. In *Proceedings of the 22nd annual international ACM SIGIR conference on Research and development in information retrieval*. ACM, 230–237.

HOROZOV, T., NARASIMHAN, N., AND VASUDEVAN, V. 2006. Using location for personalized poi recommendations in mobile environments. In *SAINT*. 124–129.

HUANG, L., LI, Q., AND YUE, Y. 2010. Activity identification from gps trajectories using spatial temporal pois' attractiveness. In *Proceedings of the 2nd ACM SIGSPATIAL International Workshop on Location Based Social Networks*. ACM, 27–30.

HUNG, C., CHANG, C., AND PENG, W. 2009. Mining trajectory profiles for discovering user communities. In *Proceedings of the 2009 International Workshop on Location Based Social Networks*. ACM, 1–8.

JIANG, B., YIN, J., AND ZHAO, S. 2008. Characterizing human mobility patterns in a large street network. *arXiv preprint arXiv:0809.5001*.

KAWAKUBO, H. AND YANAI, K. 2011. Geovisualrank: a ranking method of geotagged imagesconsidering visual similarity and geo-location proximity. In *Proceedings of the 20th international conference companion on World wide web*. ACM, 69–70.

KLEINBERG, J. 1999. Authoritative sources in a hyperlinked environment. *Journal of the ACM (JACM) 46,* 5, 604–632.

KODAMA, K., IIJIMA, Y., GUO, X., AND ISHIKAWA, Y. 2009. Skyline queries based on user locations and preferences for making location-based recommendations. In *Proceedings of the 2009 International Workshop on Location Based Social Networks*. ACM, 9–16.

LEMIRE, D. AND MACLACHLAN, A. 2005. Slope one predictors for online rating-based collaborative filtering. *Society for Industrial Mathematics*.

LEUNG, K., LEE, D., AND LEE, W. 2011. Clr: a collaborative location recommendation framework based on co-clustering. In *Proceedings of the 34th international ACM SIGIR conference on Research and development in Information*. ACM, 305–314.

LEVANDOSKI, J., SARWAT, M., ELDAWY, A., AND MOKBEL, M. 2012. Lars: A location-aware recommender system. In *IEEE International Conference on Data Engineering*.

LI, Q., ZHENG, Y., XIE, X., CHEN, Y., LIU, W., AND MA, W. 2008. Mining user similarity based on location history. In *Proceedings of the 16th ACM SIGSPATIAL international conference on Advances in geographic information systems*. ACM, 34.

LI, Y., ZHANG, Z., AND BAO, J. 2012. Mutual or unrequited love: Identifying stable clusters in social networks with uni-and bi-directional links. *Algorithms and Models for the Web Graph*, 113–125.

LIAN, D. AND XIE, X. 2011. Learning location naming from user check-in histories. In *ACM SIGSPATIAL*. ACM.

LIBEN-NOWELL, D., NOVAK, J., KUMAR, R., RAGHAVAN, P., AND TOMKINS, A. 2005. Geographic routing in social networks. *Proceedings of the National Academy of Sciences of the United States of America 102,* 33, 11623–11628.

LIN, Y., SUN, J., CASTRO, P., KONURU, R., SUNDARAM, H., AND KELLIHER, A. 2009. Metafac: community discovery via relational hypergraph factorization. In *Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 527–536.

LINDEN, G., SMITH, B., AND YORK, J. 2003. Amazon. com recommendations: Item-to-item collaborative filtering. *Internet Computing, IEEE 7,* 1, 76–80.

LIU, H., WEI, L., ZHENG, Y., SCHNEIDER, M., AND PENG, W. 2011. Route discovery from mining uncertain trajectories. In *Data Mining Workshops (ICDMW), 2011 IEEE 11th International Conference on*. IEEE, 1239–1242.

LU, C., LEI, P., PENG, W., AND SU, I. 2011. A framework of mining semantic regions from trajectories. In *Database Systems for Advanced Applications*. Springer, 193–207.

LU, X., WANG, C., YANG, J., PANG, Y., AND ZHANG, L. 2010. Photo2trip: generating travel routes from geo-tagged photos for trip planning. In *Proceedings of the international conference on Multimedia*. ACM, 143–152.

MANNING, C., RAGHAVAN, P., AND SCHUTZE, H. 2008. *Introduction to information retrieval*. Vol. 1. Cambridge University Press Cambridge.

MISHRA, N., SCHREIBER, R., STANTON, I., AND TARJAN, R. 2007. Clustering social networks. In *Proceedings of the 5th international conference on Algorithms and models for the web-graph*. Springer-Verlag, 56–67.

MOKBEL, M., BAO, J., ELDAWY, A., LEVANDOSKI, J., AND SARWAT, M. 2011. Personalization, Socialization, and Recommendations in Location-based Services 2.0. In *5th International VLDB workshop on Personalized access, Profile Management and context awareness in Databases (PersDB)*. VLDB.

MovieLens. MovieLens. http://www.MovieLens.org/.

NOULAS, A., SCELLATO, S., MASCOLO, C., AND PONTIL, M. 2011. An empirical study of geographic user activity patterns in foursquare. *ICWSM*.

PAGE, L., BRIN, S., MOTWANI, R., AND WINOGRAD, T. 1999. The pagerank citation ranking: Bringing order to the web. *Stanford Technical Report*.

PARK, M., HONG, J., AND CHO, S. 2007. Location-based recommendation system using bayesian users preference model in mobile devices. *Ubiquitous Intelligence and Computing*, 1130–1139.

POZDNOUKHOV, A. AND KAISER, C. 2011. Space-time dynamics of topics in streaming text. In *Proceedings of the 3rd ACM SIGSPATIAL International Workshop on Location-Based Social Networks*. ACM, 8.

RAMASWAMY, L., DEEPAK, P., POLAVARAPU, R., GUNASEKERA, K., GARG, D., VISWESWARIAH, K., AND KALYANARAMAN, S. 2009. Caesar: A context-aware, social recommender system for low-end mobile devices. In *Mobile Data Management: Systems, Services and Middleware, 2009. MDM'09. Tenth International Conference on*. IEEE, 338–347.

RAYMOND, R., SUGIURA, T., AND TSUBOUCHI, K. 2011. Location recommendation based on location history and spatio-temporal correlations for an on-demand bus system. In *ACM SIGSPATIAL*. ACM.

ROTH, M., BEN-DAVID, A., DEUTSCHER, D., FLYSHER, G., HORN, I., LEICHTBERG, A., LEISER, N., MATIAS, Y., AND MEROM, R. 2010. Suggesting friends using the implicit social graph. In *Proceedings of the 16th ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 233–242.

SANDHOLM, T. AND UNG, H. 2011. Real-time, location-aware collaborative filtering of web content. In *Proceedings of the 2011 Workshop on Context-awareness in Retrieval and Recommendation*. ACM, 14–18.

SCELLATO, S., MASCOLO, C., MUSOLESI, M., AND CROWCROFT, J. 2011a. Track globally, deliver locally: Improving content delivery networks by tracking geographic social cascades. In *Proceedings of the 20th international conference on World wide web*. ACM, 457–466.

SCELLATO, S., NOULAS, A., LAMBIOTTE, R., AND MASCOLO, C. 2011. Socio-spatial properties of online location-based social networks. *Proceedings of ICWSM 11*, 329–336.

SCELLATO, S., NOULAS, A., AND MASCOLO, C. 2011b. Exploiting place features in link prediction on location-based social networks. *ACM SIGKDD*.

SHENG, C., ZHENG, Y., HSU, W., LEE, M., AND XIE, X. 2010. Answering top-k similar region queries. In *Database Systems for Advanced Applications*. Springer, 186–201.

SHI, Y., SERDYUKOV, P., HANJALIC, A., AND LARSON, M. 2011. Personalized landmark recommendation based on geotags from photo sharing sites. *AAAI*.

SILVA, ANA .AND MARTINS, B. 2011. Tag recommendation for georeferenced photos. In *Proceedings of the 3nd ACM SIGSPATIAL International Workshop on Location Based Social Networks*. ACM.

SINGH, A. AND GORDON, G. 2008. Relational learning via collective matrix factorization. In *Proceeding of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 650–658.

SREBRO, N., JAAKKOLA, T., ET AL. 2003. Weighted low-rank approximations. In *MACHINE LEARNING-INTERNATIONAL WORKSHOP THEN CONFERENCE-*. Vol. 20. 720.

SRIKANT, R. AND AGRAWAL, R. 1996. Mining sequential patterns: Generalizations and performance improvements. *Advances in Database TechnologyłEDBT'96*, 1–17.

SYMEONIDIS, PANAGIOTIS .AND PAPADIMITRIOU, A., MANOLOPOULOS, Y., SENKUL, P., AND TOROSLU, I. 2011. Geo-social recommendations based on incremental tensor reduction and local path traversal. In *Proceedings of the 3nd ACM SIGSPATIAL International Workshop on Location Based Social Networks*. ACM.

TAI, C., LIN, L., AND CHEN, M. 2008. Recommending personalized scenic itinerarywith geo-tagged photos. In *Multimedia and Expo, 2008 IEEE International Conference on*. IEEE, 1209–1212.

TAKEUCHI, Y. AND SUGIMOTO, M. 2006. Cityvoyager: An outdoor recommendation system based on user location history. *Ubiquitous Intelligence and Computing*, 625–636.

TANG, K., LIN, J., HONG, J., SIEWIOREK, D., AND SADEH, N. 2010. Rethinking location sharing: exploring the implications of social-driven vs. purpose-driven location sharing. In *Proceedings of the 12th ACM international conference on Ubiquitous computing*. ACM, 85–94.

TOBLER, W. 1970. A computer movie simulating urban growth in the detroit region. *Economic geography 46*, 234–240.

VALENTE, T. 1996. Social network thresholds in the diffusion of innovations. *Social Networks 18,* 1, 69–89.

VENETIS, P., GONZALEZ, H., JENSEN, C., AND HALEVY, A. 2011. Hyper-local, directions-based ranking of places. *Proceedings of the VLDB Endowment 4,* 5, 290–301.

VON LUXBURG, U. 2007. A tutorial on spectral clustering. *Statistics and Computing 17,* 4, 395–416.

WEAKLIAM, J., BERTOLOTTO, M., AND WILSON, D. 2005. Implicit interaction profiling for recommending spatial content. In *Proceedings of the 13th annual ACM international workshop on Geographic information systems*. ACM, 285–294.

WEI, L., ZHENG, Y., AND PENG, W. 2012. Constructing popular routes from uncertain trajectories. In *Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 195–203.

WIESE, J., KELLEY, P., CRANOR, L., DABBISH, L., HONG, J., AND ZIMMERMAN, J. 2011. Are you close with me? are you nearby? investigating social groups, closeness, and willingness to share. In *Proceedings of the 13th international conference on Ubiquitous computing*. ACM, 197–206.

XIANG, R., NEVILLE, J., AND ROGATI, M. 2010. Modeling relationship strength in online social networks. In *Proceedings of the 19th international conference on World wide web*. ACM, 981–990.

XIAO, X., ZHENG, Y., LUO, Q., AND XIE, X. 2010. Finding similar users using category-based location history. In *Proceedings of the 18th SIGSPATIAL International Conference on Advances in Geographic Information Systems*. ACM, 442–445.

XIAO, X., ZHENG, Y., LUO, Q., AND XIE, X. 2012. Inferring social ties between users with human location history. *the International Journal of Ambient Intelligence and Humanized Computing*.

YE, M., JANOWICZ, K., LEE, W., AND MÜLLIGANN, C. 2011a. What you are is when you are: The temporal dimension of feature types in location-based social networks. In *ACM SIGSPATIAL*. ACM.

YE, M., SHOU, D., LEE, W., YIN, P., AND JANOWICZ, K. 2011b. On the semantic annotation of places in location-based social networks. *ACM SIGKDD*.

YE, M., YIN, P., AND LEE, W. 2010. Location recommendation for location-based social networks. In *SIGSPATAIL*. ACM, 458–461.

YE, M., YIN, P., LEE, W., AND LEE, D. 2011c. Exploiting geographical influence for collaborative point-of-interest recommendation. In *Proceedings of the 34th international ACM SIGIR conference on Research and development in Information*. ACM, 325–334.

YE, Y., ZHENG, Y., CHEN, Y., FENG, J., AND XIE, X. 2009. Mining individual life pattern based on location history. In *Mobile Data Management: Systems, Services and Middleware, 2009. MDM'09. Tenth International Conference on*. IEEE, 1–10.

YIN, Z., CAO, L., HAN, J., LUO, J., AND HUANG, T. 2011a. Diversified trajectory pattern ranking in geo-tagged social media. In *SDM*. 980–991.

YIN, Z., CAO, L., HAN, J., ZHAI, C., AND HUANG, T. 2011b. Geographical topic discovery and comparison. In *Proceedings of the 20th international conference on World wide web*. ACM, 247–256.

YIN, Z., GUPTA, M., WENINGER, T., AND HAN, J. 2010. Linkrec: a unified framework for link recommendation with user attributes and graph structure. In *Proceedings of the 19th international conference on World wide web*. ACM, 1211–1212.

YING, J., LU, E., LEE, W., WENG, T., AND TSENG, V. 2010. Mining user similarity from semantic trajectories. In *Proceedings of the 2nd ACM SIGSPATIAL International Workshop on Location Based Social Networks*. ACM, 19–26.

YING, J. J.-C., LEE, W.-C., YE, M., CHEN, T. C.-Y., AND TSENG, V. S. 2011. User association analysis of locales on location based social networks. In *GIS-LBSN*.

YOON, H., ZHENG, Y., XIE, X., AND WOO, W. 2010. Smart itinerary recommendation based on user-generated gps trajectories. *Ubiquitous Intelligence and Computing*, 19–34.

YOON, H., ZHENG, Y., XIE, X., AND WOO, W. 2011. Social itinerary recommendation from user-generated digital trails. *Personal and Ubiquitous Computing*, 1–16.

YU, X., PAN, A., TANG, L., LI, Z., AND HAN, J. 2011. Geo-friends recommendation in gps-based cyber-physical social network. In *2011 International Conference on Advances in Social Networks Analysis and Mining*. IEEE, 361–368.

ZHANG, D., CHEE, Y., MONDAL, A., TUNG, A., AND KITSUREGAWA, M. 2009. Keyword search in spatial databases: Towards searching by document. In *IEEE International Conference on Data Engineering*. IEEE, 688–699.

ZHENG, V., CAO, B., ZHENG, Y., XIE, X., AND YANG, Q. 2010a. Collaborative filtering meets mobile recommendation: A user-centered approach.

ZHENG, V., ZHENG, Y., XIE, X., AND YANG, Q. 2010c. Collaborative location and activity recommendations with gps history data. In *WWW*. ACM, 1029–1038.

ZHENG, V., ZHENG, Y., XIE, X., AND YANG, Q. 2012. Towards mobile intelligence: Learning from gps history data for collaborative recommendation. *Artificial Intelligence*.

ZHENG, V., ZHENG, Y., AND YANG, Q. 2009d. Joint learning user's activities and profiles from gps data. In *Proceedings of the 2009 International Workshop on Location Based Social Networks*. ACM, 17–20.

ZHENG, Y. 2011. Location-based social networks: Users. In *Computing with Spatial Trajectories*, Y. Zheng and X. Zhou, Eds. Springer.

ZHENG, Y. 2012. Tutorial on location-based social networks. In *In International conference on World Wide Web (WWW 2012)*.

ZHENG, Y., CHEN, Y., XIE, X., AND MA, W.-Y. 2009c. GeoLife2.0: A Location-Based Social Networking Service. In *MDM*.

ZHENG, Y. AND XIE, X. 2010. Learning location correlation from gps trajectories. In *Mobile Data Management (MDM), 2010 Eleventh International Conference on*. IEEE, 27–32.

ZHENG, Y. AND XIE, X. 2011. Learning travel recommendations from user-generated gps traces. *ACM Transactions on Intelligent Systems and Technology (TIST) 2,* 1, 2.

ZHENG, Y., XIE, X., AND MA, W. 2010b. Geolife: A collaborative social networking service among user, location and trajectory. *IEEE Data Engineering Bulletin 33,* 2, 32–40.

ZHENG, Y., ZHANG, L., MA, Z., XIE, X., AND MA, W. 2011. Recommending friends and locations based on individual location history. *ACM Transactions on the Web (TWEB) 5,* 1, 5.

ZHENG, Y., ZHANG, L., XIE, X., AND MA, W. 2009a. Mining correlation between locations using human location history. In *Proceedings of the 17th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*. ACM, 472–475.

ZHENG, Y., ZHANG, L., XIE, X., AND MA, W. 2009b. Mining interesting locations and travel sequences from gps trajectories. In *Proceedings of the 18th international conference on World wide web*. ACM, 791–800.

ZHENG, Y. AND ZHOU, X. 2011. *Computing with Spatial Trajectories*. Springer.