

Spatio-semantic user profiles in location-based social networks

Soha Mohamed¹ · Alia I. Abdelmoty¹ 

Received: 17 December 2016 / Accepted: 8 May 2017
© The Author(s) 2017. This article is an open access publication

Abstract Knowledge of users' visits to places is one of the keys to understanding their interest in places. User-contributed annotations of place, the types of places they visit, and the activities they carry out, add a layer of important semantics that, if considered, can result in more refined representations of user profiles. In this paper, semantic information is summarised as tags for places and a folksonomy data model is used to represent spatial and semantic relationships between users, places, and tags. The model allows simple co-occurrence methods and similarity measures to be applied to build different views of personalised user profiles. Basic profiles capture direct user interactions, while enriched profiles offer an extended view of users' association with places and tags that take into account relationships in the folksonomy. The main contributions of this work are the proposal of a uniform approach to the creation of user profiles on the Social Web that integrates both the spatial and semantic components of user-provided information, and the demonstration of the effectiveness of this approach with realistic datasets.

Keywords Geo-folksonomy · User profiles · Location-based social networks

The work in this paper is based on and extends the conference paper 'Uncovering User Profiles in Location-Based Social Networks'. The Eighth International Conference on Advanced Geographic Information Systems, Applications, and Services—GEOProcessing 2016.

✉ Alia I. Abdelmoty
AbdelmotyAI@cardiff.ac.uk

Soha Mohamed
AlySA@cardiff.ac.uk

¹ School of Computer Science and Informatics, Cardiff University, Cardiff, Wales, UK

1 Introduction

Users of location-based social networks (LBSN) declare where they go (or check-in) and which places are of interest to them (by tagging or leaving tips). Both these spatial and semantic traces are equally useful in understanding people's relationships with place. Whereas spatial tracks can be analysed to determine the frequency of visits and favourite places, semantic interactions can give clues to the sort of activities people carry out in the places they visit and the experiences they share there. Combining both the explicit spatial association to place and the implicit semantics of interaction with place provides a unique opportunity for in-depth understanding of both places and users.

So far, previous works have studied data produced from LBSN from the point of view of enhancing the services provided by these networks, namely, for point of interest (POI) recommendations. There, the question of concern is to find places of interest to a user based on their history of visits to other places and their general interaction with the social network. Most works relied mainly on the spatial dimension of user data [1], with some works more recently exploring the relevance of the social and content data dimensions on these networks [2]. However, data dimensions are normally treated separately, or their outputs are combined in fused models.

In this paper, both semantic and spatial interactions of users are used to project distinct and complementary views of personalised user profiles. Both explicit place affordance; the sort of services offered in a place as denoted by its place type, and implicit place affordance; encapsulated in reference to activities in place annotations, are used in building semantic user profiles. Collective user spatial and semantic interactions with places are used to create profiles for geographic places that in turn provide further enrichment to individual user profiles. In comparison with previous works

in the area of recommendations, LBSN data are treated as folksonomies of users, places, and tags. User annotations in the form of tips, their interaction with places, in the form of check-ins, as well as general place properties, namely, place categories and tags, are analysed concurrently to extract relations between the three elements of the folksonomy. Simple co-occurrence methods and similarity measures are used to compute direct and enriched user profiles.

The proposed approach offers a uniform framework for presenting different views of user profiles, using their direct interactions with the social network or extended with a holistic view of other users' interaction with the network in different regions of geographic space. The homogenous treatment of the different data dimensions allows for the derivation and evaluation of different views of user profiles and offers a flexible and systematic approach to considering different attributes of users and places when building user profiles on LBSN. Previous works addressing POI recommendation used matrix factorisation techniques to handle the multiple data dimensions but did not consider the use of the range of content data as used in this paper. Realistic data from the LBSN Foursquare are used to demonstrate the approach, and evaluation results show its potential value. Whereas it can be argued that the proposed profiles present different views of the user data, and that their effectiveness may depend on the characteristics of the dataset considered and the application context, results show that enriched user profiles can offer potentially more accurate views, than direct profiles, of user's spatial or semantic preferences.

The rest of the paper is organised as follows. Section 2 provides an overview of related works. In Sect. 3, the geofolksonomy data model is described and is used in Sect. 4 to define different types of user profiles. In Sect. 5, an evaluation experiment is presented and its results are discussed. The paper concludes in Sect. 6 with an overview of future work.

2 Related work

This paper considers research questions in the area of user and place modelling in LBSN. Works on modelling user data in LBSN mainly consider two problems; (a) place (or point of interest) recommendation and (b) user similarity calculation. Different types of data are used by different approaches, namely, geographic content, social content as well as textual annotations made by users. Also, different methods are used in analysing the data, for example, distance estimations for geographic data modelling and topic modelling for annotation data analysis.

In the area of POI recommendation, works range from generic approaches that use the popularity of places [3] to recommendation methods that are based on user's individual preferences [4]. A useful survey of these approaches can

be found in [5]. Based on check-in data gathered through Foursquare, Noulas and Mascolo [6] exploit factors such as the transition between types of places, mobility between venues and spatio-temporal characteristics of user check-in patterns to build a supervised model for predicting a user's next check-in. Ye et al. [4] investigated the geographic influence with a power-law distribution. The hypothesis is that users tend to visit places within short distances of one another. Other works considered other distance distribution models [7]. Gao et al. [8] considered a joint model of geo-social correlations for personalised POI recommendation, where the probability of a user checking into a new POI is described as a function of correlations between user's friends and non-friends close to and distant from a region of interest. Liu et al. [9] approached the problem of POI recommendations by proposing a geographic probabilistic factor model that combines the modelling of geographic preference and user mobility. Geographic influence is captured through the identification of latent regions of activity for all users of the LBSN reflecting activity areas for the entire population and mapping the individual user mobility over those regions. Their model is enhanced by assuming a Poisson distribution for the check-in count which better represents the skewed data (users visiting some places one time, while other places 100s of times). While providing some useful insights for modelling the spatial dimension of the data, the above works do not consider the semantic dimension of the data.

Correlations between geographic distance and social connections were noted in [2, 10]. Techniques of personalised POI recommendation with geographic influence and social connections mainly study these two elements separately and then combine their output together within a fused model. Social influence is usually modelled through friend-based collaborative filtering [4, 11, 12] with the assumption that a user tends to be friends with other users who are geographically close to him or would want to visit similar places to those visited by his friends. Ying et al. [13] proposed to combine the social factor with individual preferences and location popularity within a regression-tree model to recommend POIs. The social factor corresponds to similar users, users with common check-ins to the user in question. In this paper, we also use this factor when extending user profiles to represent places of interest within the region of user activity.

More recently, the importance of content information for POI recommendation was recognised. Two types of content can be considered, attributes of places and user-contributed annotations. Place categories are normally used as an indication of user activity; thus, a user visiting a French restaurant would be considered as interested in French food, etc. User annotations in the form of tips and comments are analysed collectively to extract general topics to characterise places or to extract collective sentiment indications about the place. Examples of works that considered place categories are [14–

17]. In [14, 15], the latent Dirichlet allocation (LDA) model was used to represent places as a probability distribution over topics collected from tags and categories or comments made in a place and, similarly, aggregate all tips from places a user has visited to model a user's interest. Aggregation was necessary, as terms associated with a single POI are usually short, incomplete, and ambiguous. [16] on the other hand modelled topics from tweets and reviews from Twitter and Yelp and assumed that the relations between user interests and location are derived from the topic distributions for both users and locations. In [17], a probabilistic approach is proposed that utilises geographic, social, and categorical correlations among users and places to recommend new POIs from historical check-in data of all users. In this paper, we also model user's association to place through the place's relation to tags but add the influence of other users relations in the place to the equation.

Aiming at improving the effectiveness of location recommendation, Yang et al. [18] proposed a hybrid user POI preference model by combining the preference extracted from check-ins and text-based tips which were processed using sentiment analysis techniques. Sentiment analysis is an interesting type of semantics which we do not consider in this work, but can be incorporated in future work.

Studying user similarity from LBSN data is useful; as information available about users, their locations and activities are considered to be sparse. User similarities can be exploited to predict types of activities and places preferred by a user based on those of users with similar preferences. So far, most works on user similarity mainly focused on structured data, e.g. geographic coordinates, or semi-structured data, e.g. tags and place categories. Geo-social metrics are proposed in [19], where spatial degree centrality and spatial closeness centrality metrics are introduced to leverage the geographic influence of use collocation. Recently, Lee and Chung [20] presented a method for determining user similarity based on LBSN data. While the authors made use of check-in information, they relied on the hierarchy of location categories supplied by Foursquare in conjunction with the frequency of check-ins to determine a measure of similarity. Mckenzie et al. [15] suggest exploring unstructured user-contributed data, namely tips provided by users. A topic modelling approach is used to represent users' interests in places. Venues (places in Foursquare) are described as a mixture of a given number of topics, and topic signatures are computed as a distribution across venues. User similarity can then be measured by computing a dissimilarity metric between users' topic distribution. Their method of modelling venues is interesting, but it limits the representation of user profiles, where profiles are based on generated topics derived from collective user annotation on places. Thus, individualised association of users with the place is somewhat ignored. In contrast to the above approach, our model

does not assume constraints on the number of topics represented by the tags but combines the individual's association with both tags and place in the creation of user profiles.

In modelling users on LBSN, several user characteristics as well as user relationships may be considered. For example, users' personal characteristics such as age, gender, and cuisine preferences were used in [21], and social affinity was considered in [22,23]. User's history of online activity can also be collected, for example, search history; history of map browsing and spatial searching logs [24–26], place reviews and ratings [27–29], as well as explicit interaction on LBSN, by tagging and commenting on places [30,31]. In this work, users' location tracks are considered as the primary source of user–place relationships, as these represent explicit interaction with geographic places, normally recording actual visits to places. These tracks are also associated with explicit semantics of tagging and tipping and thus form a useful basis for considering both the spatial and semantic aspects of the user profiles. Other attributes can always be added and considered in a similar way.

3 Geo-folksonomy model

The location-based social networking platform, Foursquare, was used as our source of data. It holds a large number of crowdsourced venues (>65 million places) from a user population estimated recently to around 55 million users. As the application defines it, a venue is a user-contributed 'physical location, such as a place of business or personal residence'. Foursquare allows users to check into a specific venue, sharing their location with friends, as well as other online social networks, such as Facebook or Twitter. Tips on a specific venue normally describe a recommendation, experience, or activity performed in the place.

In Foursquare, a place is related to one or more place categories, where a three-level hierarchy of categories is maintained. Ten main categories form the root of the hierarchy: arts and entertainment, college and university, event, food, nightlife spots, outdoors and recreation, professional and other places, residence, shops and services, and travel and transport. These are then classified to 525 subcategories on the lowest level in the hierarchy. For example, the categories 'Christmas Market' and 'Conference' are subclasses of the class 'Event' and the categories 'American Restaurant' and 'Asian Restaurant' are subclasses of the class 'Food'. The distribution of subcategories in Foursquare is shown in Table 1.

In this work, we use a folksonomy data model to represent user–place relationships and derive tag assignments from users' actions of check-ins and annotation of venues [32]. In particular, tags are assigned to venues in our data model in two scenarios as follows.

Table 1 Distribution of place categories in the Foursquare dataset

Category name	Subcategories count
Arts and entertainment	45
College and university	35
Event	7
Food	127
Nightlife spot	19
Outdoors and recreation	65
Professional and other places	68
Residence	4
Shop and service	111
Travel and transport	41

1. A user's check-in results in the assignment of place categories associated with the place as tags annotated by this user. Thus, a check-in by user u in place r with the categories (represented as keywords) x , y , and z , will be considered as an assertion of the form $(u, r, (x, y, z))$. This, in turn, will be transformed into a set of triples $\{(u, r, x), (u, r, y), (u, r, z)\}$ in the folksonomy.
2. A user's tip in the place also results in the assignment of place categories as tags, in addition to the set of keywords extracted from the tip. Thus, in the above example, a tip by u in r with the keywords (t_1, \dots, t_n) , will be considered as an assertion of the form $(u, r, (x, y, z, t_1, \dots, t_n))$ and is in turn transformed into individual triples between the user, place, and tags in the folksonomy.

Figure 1 depicts the overall process of user profile creation. The process starts with data collection of user tracks and tip data that are then processed to extract users, places, and tags and their associated properties. This step also includes data pre-processing and cleaning and involves the following sequence of steps:

1. Removal of special characters. All non-alphanumeric characters are removed from tags. For instance, the tag *Cardiff&* is changed to *Cardiff*.
2. Filtering of all tags that are less than 3 characters in length.
3. Filtering of tags that represent URLs.
4. Filtering of stop words. A list of 116 stop words, published by Microsoft,¹ is used.
5. Removal of duplicate tags. Duplicates are removed in such a way as to preserve the relations between place resources and users.

After pre-processing, tips are tokenised into words and stored as tags. A tag resolution stage makes use of the

¹ [http://msdn.microsoft.com/en-us/library/bb164590\(v=vs.80\).aspx](http://msdn.microsoft.com/en-us/library/bb164590(v=vs.80).aspx).

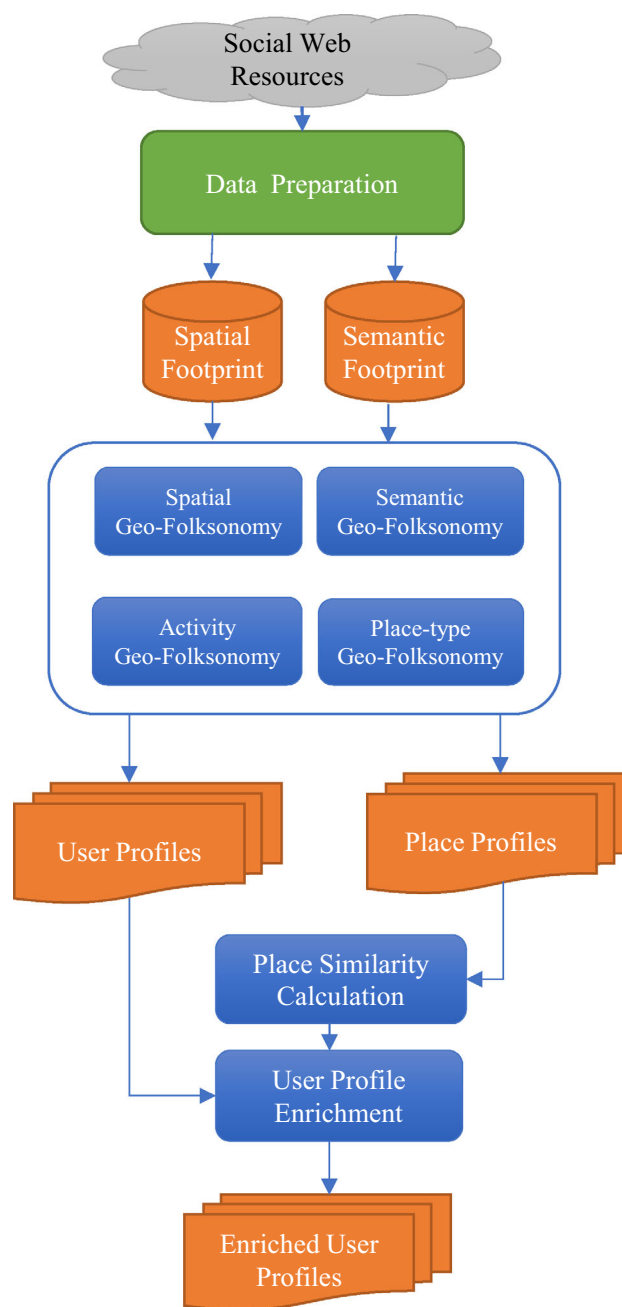


Fig. 1 Overview of the process of user profile creation in the proposed system

WordNet lexicographer (a lexical database for the English language). There are 44 lexicographer files that can be used to classify a word into a suitable category. Table 2 shows some examples of verb and noun categories in WordNet that are used to identify reference to human activity, as described later in the paper.

A data modelling stage is then carried out and results in the creation of a geo-folksonomy, to represent the relationships between the data elements. A geo-folksonomy can be defined

Table 2 Example subset of WordNet lexicographer files

Name	Contents
noun.act	Nouns denoting acts or actions
noun.cognition	Nouns denoting cognitive processes and contents
noun.communication	Nouns denoting communicative processes and contents
noun.event	Nouns denoting natural events
verb.competition	Verbs of fighting, athletic activities
verb.consumption	Verbs of eating and drinking
verb.creation	Verbs of sewing, baking, painting, performing
verb.emotion	Verbs of feeling
verb.motion	Verbs of walking, flying, swimming

as a quadruple $\mathbb{F} := (U, T, R, Y)$, where U, T, R are finite sets of instances of users, tags, and places, respectively, and Y defines a relation, the tag assignment, between these sets; that is: $Y \subseteq U \times T \times R$ [33,34].

A geo-folksonomy can be transformed into a tripartite undirected graph, which is denoted as folksonomy graph $G_{\mathbb{F}}$. A geo-folksonomy graph $G_{\mathbb{F}} = (V_{\mathbb{F}}, E_{\mathbb{F}})$ is an undirected weighted tripartite graph that models a given folksonomy \mathbb{F} , where: $V_{\mathbb{F}} = U \cup T \cup R$ is the set of nodes, $E_{\mathbb{F}} = \{\{u, t\}, \{t, r\}, \{u, r\} | (u, t, r) \in Y\}$ is the set of edges, and a weight w is associated with each edge $e \in E_{\mathbb{F}}$. The weight associated with an edge $\{u, t\}$, $\{t, r\}$ and $\{u, r\}$ corresponds to the co-occurrence frequency of the corresponding nodes within the set of tag assignments Y . For example, $w(t, r) = |\{u \in U : (u, t, r) \in Y\}|$ corresponds to the number of users that assigned tag t to place r .

Different versions of the geo-folksonomy are created in this stage. A process of tag filtering and refinements is carried out to expose different types of place-related semantics, namely place types and place-related activities as shown in Fig. 1. Place type and activities are important semantics used in place ontologies. For example, the National Mapping Agency of Great Britain, the Ordnance Survey, maintains an ontology of buildings and places² and employs a *has-purpose* relationship to document the services offered by different place types. In this work, we make use of the Foursquare categories as place types in the geo-folksonomy model. Tags representing human activ-

² Building and place ontology: <http://www.ordnancesurvey.co.uk/oswebsite/ontology/>.

Table 3 Example activities in the dataset used

Activity class	Example
noun.act	Move, tour, soccer, game, dining, sightseeing
verb.competition	Play, win, score, golf, attack, defeat

ities are identified by matching against WordNet categories: ‘noun.act’ and ‘verb.(competition, consumption, contact, creation, motion)’. Table 3 shows some examples of activities extracted from tags in the dataset. A sample of activity tags and their related place types, as extracted from the Foursquare dataset, is shown in Fig. 2.

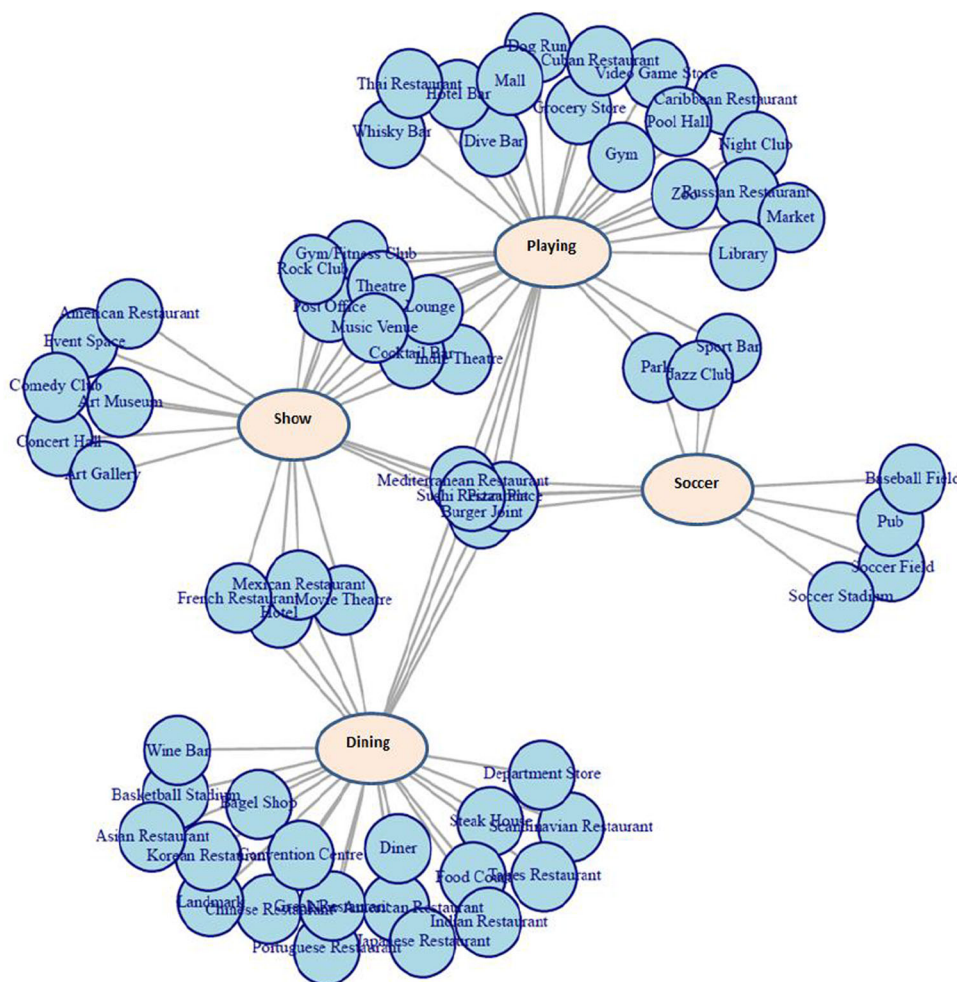
4 User modelling strategies

The user–place data collected on the LBSN are multidimensional and dynamic. Users’ interaction with geographic places is not-uniform. Some users are very active and frequently record their check-ins, while others check-in occasionally. Some users provide many annotations to places, while others are more casual annotators, etc. Thus, the approach proposed here offers different methods of describing the user–place relationships. The effectiveness of the profiles cannot be absolutely compared as it is likely to depend on the nature of the dataset considered and the density of the different data dimensions represented.

Thus, in this section, a uniform approach is proposed for building different types of user profiles by considering all the data dimensions. A spatial profile represents the association of a user with the places he visits. A semantic profile describes his association with the concepts he uses to annotate the places he visits. A combined spatio-semantic profile is a customised view of both the previous profiles that projects the user’s specific interest in particular types of places or place-related activities.

A user profile is built in stages. Starting with a basic profile that utilises direct check-in and annotation histories, a user profile is then extended by computing the relationship between places and concepts derived from the collective behaviour of other users in the dataset. A basic profile represents actual interactions with places, while the extended profile describes ‘recommended’ associations given overall interactions between users, places, and concepts in the dataset. We are able to model such interactions separately in the extended profile by controlling the similarity function used to create the profile. For example, we can focus on

Fig. 2 A sample of the activities and their related place types in the dataset



modelling the types of places visited by the user or take into account visit behaviour of other users whose profiles overlap with the user, as discussed below.

4.1 Basic user profiles

Definition 1 (*Spatial user profile*) A spatial user profile $P_R(u)$ of a user u is deduced from the set of places that u visited or annotated directly.

$$P_R(u) = \{(r, w(u, r)) | (u, t, r) \in Y, w(u, r) = |\{t \in T : (u, t, r) \in Y\}|\}$$

$w(u, r)$ is the number of tag assignments, where user u assigned some tag t to place r through the action of checking-in or annotation. Hence, the weight assigned to a place simply corresponds to the frequency of the user reference to the place either by checking-in or by leaving a tip.

We further normalise the weights so that the sum of the weights assigned to the places in the spatial profile is equal to 1. We use \overline{P}_R to explicitly refer to the spatial profile where the sum of all weights is equal to 1, with $\overline{w}(u, r) =$

$\frac{|\{t \in T : (u, t, r) \in Y\}|}{\sum_{i=1}^n \sum_{j=1}^m |\{t_i \in T : (u, t_i, r_j) \in Y\}|}$, where n and m are the total number of tags and resources, respectively. More simply, $\overline{w}(u, r) = \frac{N(u, r)}{N_T(u)}$, where $N(u, r)$ is the number of tags used by u for resource r , while $N_T(u)$ is the total number of tags used by u for all places. Figure 3 shows a sample of places visited in an example spatial user profile of user: ‘user164’.

Correspondingly, we define a semantic, tag-based, profile of a user; $P_T(u)$ as follows.

Definition 2 (*Semantic user profile*) A semantic user profile $P_T(u)$ of a user u is deduced from the set of tag assignments linked with u .

$$P_T(u) = \{(t, w(u, t)) | (u, t, r) \in Y, w(u, t) = |\{r \in R : (u, t, r) \in Y\}|\}$$

$w(u, t)$ is the number of tag assignments where user u assigned tag t to some place through the action of checking-in or annotation.

\overline{P}_T refers to the semantic profile where the sum of all weights is equal to 1, with $\overline{w}(u, t) = \frac{N(u, t)}{N_R(u)}$, where $N(u, t)$ is the number of resources annotated by u with t and $N_R(u)$ is the total number of resources annotated by u . Figure 4 shows



Fig. 3 A spatial profile for a sample user: ‘user164’

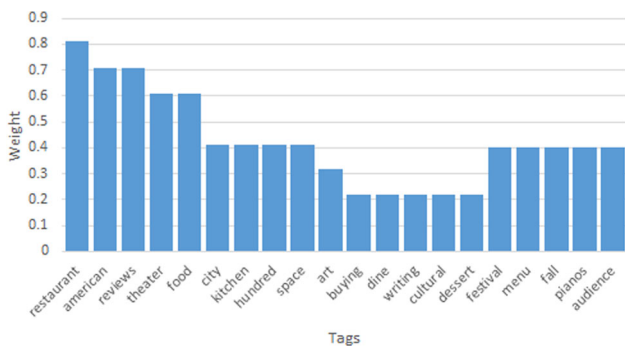


Fig. 4 Semantic profile for a sample user: ‘user164’

the top 20 tags stored in the semantic user profile of user: ‘user164’.

4.2 Spatio-semantic user profiles

Spatio-semantic user profiles of a user u are profiles that represent the user’s interests in geographic places using both their geographic location and other semantic properties, such as their type, the kind of services they provide, or the activities that take place in them. Here, we define two possible versions of such profiles: a place type-oriented profile and an activity-oriented profile. A process of restructuring the geo-folksonomy needs to be carried out to map the relationships between users and tags to place types and activities.

Definition 3 (Place type-oriented user profile) A place type-oriented user profile describes the association between user u and place types in the folksonomy. The strength of the

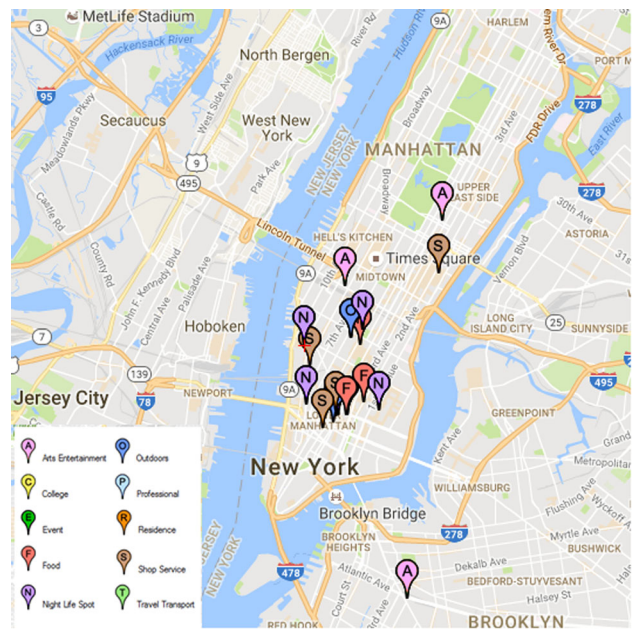


Fig. 5 Place type-oriented profile of a sample user: ‘user164’

association with a place type (weight) is computed as the sum of all weights of places in the user spatial profile that belong to that place type. Let the set of place types for a place r be denoted as r_C .

$$P_C(u) = \{(c, w(u, c)) | (r, w(u, r)) \in P_R(u) \wedge c \in r_C, w(u, c) = \{\sum_{i=1}^n w(u, r_i) : (r_i, w(u, r_i)) \in P_R(u) \wedge c \in r_{iC}\}$$

where places r_1, \dots, r_n are all the places in the spatial profile $P_R(u)$ and $w(u, c)$ is the sum of the weights for all places with place type c in $P_R(u)$.

$\overline{P_C}$ refers to the place type profile where the sum of all weights is equal to 1, with $\overline{w}(u, c) = \frac{N(u, r_C)}{N_R(u)}$, where $N(u, r_C)$ is the number of resources annotated by u whose place type is c and $N_R(u)$ is the total number of resources annotated by u . Figure 5 shows a sample of place type in the user profile of user: ‘user164’.

Definition 4 (Activity-oriented user profile) An activity-oriented user profile describes the association between user u and the human activities carried out in places in the folksonomy. It can be regarded as a restriction of the spatial profile $P_R(u)$ by tags representing human activities and thus describes the user’s association with places in his profile that are annotated by human activities.

Let \mathbb{A} be a set of tags representing all possible human activities. Let $T(a) \subseteq \mathbb{A}$ be a subset of all tags in the folksonomy that correspond to human activities. Let $\mathbb{F}_u = (T(a)_u, R_u, I_u)$ of a given user $u \in U$ be the restriction of the geo-folksonomy \mathbb{F} to u , such that, $T(a)_u$ and R_u are finite sets of

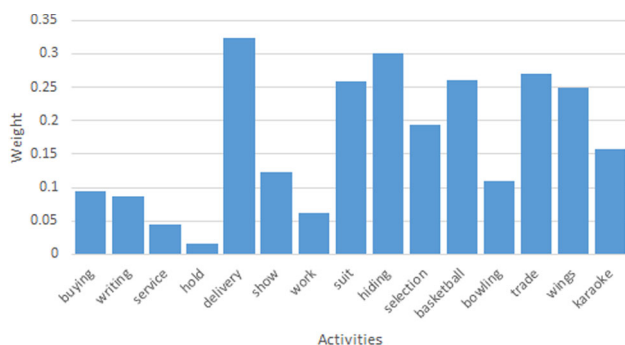


Fig. 6 Tags denoting human activity in the user profile of user ‘user164’

tags, representing human activities, and places, respectively, that are referenced from tag assignments performed by u , and I_u defines a relation between these sets: $I_u := \{(a, r) \in T(a)_u \times R_u \mid (u, a, r) \in Y\}$.

An activity-oriented user profile $P_A(u)$ of a user u is deduced from the set of tag assignments made for place r by u .

$$P_A(u) = \{([r, a], w_u([r, a])) \mid (a, r) \in I_u, w_u([r, a]) = |\{a \in T(a)\} : (a, r) \in I_u|\}$$

where $w([r, a])$ is how often user u assigned tag a to place r .

$\overline{P_A}$ is the spatio-semantic profile where the sum of all weights is equal to 1, with $w_u([r, a]) = \frac{N(u, [r, a])}{N_{RT}(u)}$, where $N(u, [r, a])$ is the number of times u annotate r with a , and $N_{RT}(u)$ is the total number of activity tags assigned by u for r . Figure 6 shows the activity tags in the profile of user: ‘user164’ (approximately 5% of all tags).

4.3 Basic place and tag profiles

So far, the basic user profile provides only a limited view of the user association with places and concepts derived directly from captured data. Basic profiles reduce the dimensionality of the folksonomy space by considering only 2 dimensions at a time, user–place and user–tag, leading to a loss of correlation information between all three elements. Users profiles can be extended to represent possible latent relationships in the dataset, by identifying places (respectively, tags) that are similar to those in the basic profile. Similarity between places (respectively, tags) is measured through the collective actions of other users of check-ins and annotations.

To compute tag–tag similarity, profiles for tags are first defined through the places they are used to annotate. Thus, a place-based tag profile ($P_R(t)$) of a tag t is a weighted list of places r that are annotated by t . That is, $w(r, t)$ is determined by the number of users’ check-ins and tips that resulted in assigning t to r in the geo-folksonomy. Similarity between

tags is defined as the cosine similarity between their place-based tag profiles as follows.

$$\text{CosSim}(t_1, t_2) = \frac{|P_R(t_1) \cap P_R(t_2)|}{\sqrt{|P_R(t_1)| \cdot |P_R(t_2)|}}$$

On the other hand, similarity between places is defined by measuring the similarity of their tag-based and user-based profiles. Let $P_T(r)$ and $P_U(r)$ be the tag-based place profile and user-based place profile for place r (defined in a similar manner to user profiles above). Conceptually, a tag-based place profile is a description of the place by the tags assigned to it and a user-based place profile is an account of users’ visits to the place. Cosine similarity between tag-based place profiles ($\text{CSim}_{\text{tag}}(r_1, r_2)$) and between user-based place profiles ($\text{CSim}_{\text{user}}(r_1, r_2)$) constructs a tag-oriented ranking and user-oriented ranking, respectively. These similarity rankings can be aggregated using the Borda method [35] to compute a generalised similarity score between two places as shown in the following equation.

$$\text{PlaceSim}(r_1, r_2) = \gamma * \text{CSim}_{\text{tag}}(r_1, r_2) + (1 - \gamma) * \text{CSim}_{\text{user}}(r_1, r_2)$$

where $0 \leq \gamma \leq 1$ is a parameter that determines the balance of importance given to similarity scores from $P_T(r)$ and $P_U(r)$. Conceptually, similarity between two places is a function of the overlap between their tag assignments only (for $\gamma = 0$), a measure of their common visitors only (for $\gamma = 1$), or both (for γ between 0 and 1).

4.4 Enriched user profiles

We extend the basic user profiles by the information extracted from the computation of tag and place similarity above. Enriched user profiles therefore present a modified view of how users are associated with places and reflect collective user behaviour on the LBSN.

Definition 5 (*Enriched spatial user profile*) An enriched spatial user profile $\hat{P}_R(u)$ of a user u is an extension of the basic profile by places with the highest degree of similarity to places in $\overline{P_R}(u)$. Let R_u be the set of all places in $\overline{P_R}(u)$ and w_i is the weight associated with place i in the profile.

$$\hat{P}_R(u) = \{ \langle r_i, w_i \rangle \mid w_i = \begin{cases} w_i, & \text{if } r_i \in R_u \\ w_i * \text{Max}(\text{PlaceSim}(r_i, r_j)), & \forall (r_i \in \{R - R_u\} \wedge r_j \in R_u) \end{cases} \}$$

We compute the maximum similarity of the 10 most similar places in the dataset for every place in the basic user profile and use the highest similarity score as the weight for the new place in the enriched user profile. The process of building

- 1: procedure SPATIAL ENRICHMENT ($P_R(u), \gamma$)
- 2: for all place (r_i, w_i) in Spatial-Profile $P_R(u)$ do
- 3: Compute $PlaceSim(r_i, \gamma): [r_j, sim_j]$
- 4: for all $j = 1$ st to 10thdo
- 5: $w_j = w_i * sim_j$
- 6: add $\langle r_j, w_j \rangle$ to $P_R(u)$
- 7: end for
- 8: end for
- 9: return $\hat{P}_R(u)$
- 10: end procedure

Fig. 7 Algorithm for building the enriched user profiles



Fig. 8 Spatial user profile for user 'user164'

the enriched spatial profile from place similarity with γ as an input is shown in Fig. 7.

Figures 8 and 9 show the spatial profile and the enriched spatial profiles for user 'user164', respectively. $\gamma = 0.5$ was used in the place similarity equation of the enriched profile. The size of the dots in the figures represents the weight of the place in the profile.

Definition 6 (*Enriched semantic user profile*) An enriched tag-based user profile $\hat{P}_T(u)$ of a user u is an extension of the basic profile by tags with the highest degree of similarity to tags in $\overline{P_T(u)}$. Let T_u be the set of all tags in $\overline{P_T(u)}$ and w_i is the weight associated with tag i in the profile.

$$\hat{P}_T(u) = \{ \langle t_i, w_i \rangle \mid w_i = \begin{cases} w_i, & \text{if } t_i \in T_u \\ w_i * \text{Max}(\text{Sim}(t_i, t_j)), & \forall (t_i \in \{T - T_u\} \wedge t_j \in T_u) \end{cases} \}$$

A similar algorithm to that of enriching the spatial user profiles is used for choosing the tags and weights.



Fig. 9 Enriched spatial user profile for user 'user164' with $\gamma = 0.5$

5 Experiments and results

As previously pointed out, the different types of user profiles that can be produced from the geo-folksonomy reflect the focus on the different data dimensions. The density of the users' interaction in the dataset, e.g. the frequency of their check-ins and their co-location in geographic places, will affect the quality of the user profiles produced. Hence, it is not possible to compare the effectiveness of individual profiles in an absolute manner. Spatial profiles may be more effective for one dataset, while the semantic profiles may be more effective for another. In this work, we demonstrate the differences between the profiles by considering a sample of a realistic dataset in a recommendation context. The experimental set-up considered two main factors in building the geo-folksonomy: the number of users and the frequency of use of the LBSN, as discussed below.

5.1 Dataset

Approximately 10 months of check-in data in New York City were collected from Foursquare between April 2012 and February 2013. This data consist of 227,428 anonymised user check-ins, with venue ids, venue category, longitude and latitude of venues and time stamps of check-ins. The data were then used to recursively extract venue-related tips (tip id, text, and time stamp) and subsequently all venues for users related to the tips collected.

In total, 604,924 tips were collected for 167,786 users in 36,940 venues. Time stamps of the tip data range from January 2009 to June 2015. Figure 10 shows the number of places versus the number of users in the collected dataset. As the figure shows, about 94% of the users visited less than 10 places and about 3% of users visited 11 to 20 places and the remaining 3% visited 21 to 400 places (Fig. 11).

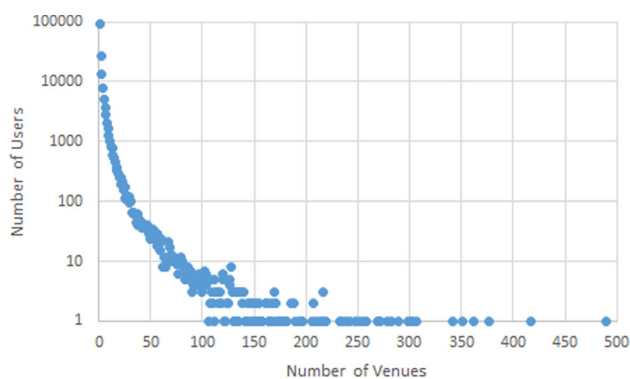


Fig. 10 Number of users and venues visited in the dataset

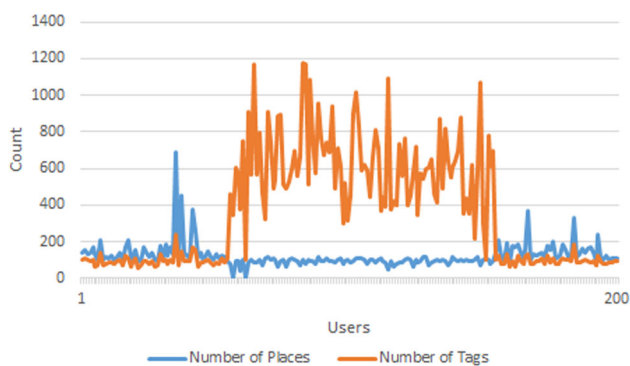


Fig. 11 Number of distinct places and tags for each user in the dataset

Experiments were carried out using a sample of 200 users with a high frequency of check-ins and co-location rate. In choosing the sample set of users, a balanced ratio of users with check-in and tipping data was used. Also, a balanced distribution of visited venues was chosen. All the 8 basic and enriched profiles were created for each user. The experiment was first applied on a set of 20 users and then extended to a set of 200 users in the dataset. The results of the two groups are presented below.

Table 4 shows some summary statistics of the sample datasets used for both groups of users, and Fig. 12 shows the distribution of users across the main place categories of Foursquare.

5.2 Experimental set-up

The experiment took the form of place (and tag) top-*N* recommendation problem using the different constructed user profiles based on the users profiles cosine similarities and sought to establish how well the profiles reflect the user spatial and semantic characteristics when using the LBSN.

We use recall@*N*, precision@*N*, and F1@*N* as our success measures, where *N* is a predefined number of places (or tags) to be recommended. Recall measures the ratio of correct recommendations to the number of true places (or tags)

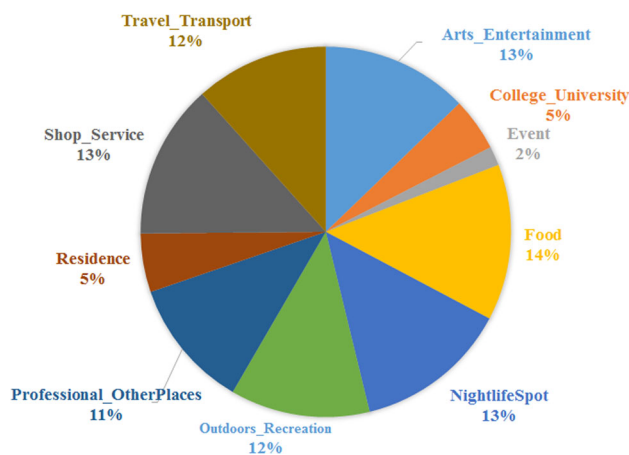


Fig. 12 Distribution of place and users by place type in the dataset

Table 4 Statistics of the two datasets used in the experiment

	20 Users	200 Users
Number of venues	2,041	10,988
Total number of check-ins	4,212	50,584
Total number of tips	942	10,469
Total number of distinct tags	3,357	13,396
Total number categories	317	495
Total Number of relationships	17,955	165,453

of a test check-in or tip record, whereas precision measures the ratio of correct to false recommendations made. Recall and precision are given by

$$\text{Recall} = \frac{TP}{TP + FN}$$

$$\text{Precision} = \frac{TP}{TP + FP}$$

where true positives (TP) is the number of correct place (or tags) recommended, false positives (FP) is the number of wrong recommendations, and false negatives (FN) is the number of true place (or tags) which were not recommended. F1 measure combines both recall and precision:

$$F1 = \frac{2 * \text{precision} * \text{recall}}{\text{precision} + \text{recall}}$$

The values of TP, FP, FN are determined by randomly splitting the users into two sets, the training set and the testing set. Multifold cross-validation was used to ensure a fair partitioning between test data and training data. Data were split 90% for training and 10% for testing, and the process was repeated 5 times to create fivefold, and the mean of the performance was reported. Evaluation of the basic spatial, semantic, and spatio-semantic profiles is first carried out, followed by an evaluation of the enriched profiles.

5.3 Evaluation of basic user profiles

Place type-based user profiles Figure 13 shows the place type distribution for users in the dataset. The average number of place types per user is half the average number of places annotated by the user. This leads to more dense, but smaller matrices in the geo-folksonomy (user–place type, place type–tag and user–tag matrices).

Figure 14a, b shows the precision and recall values for evaluating both spatial and place type-based user profiles. As expected, results show that the coarser relationships in the

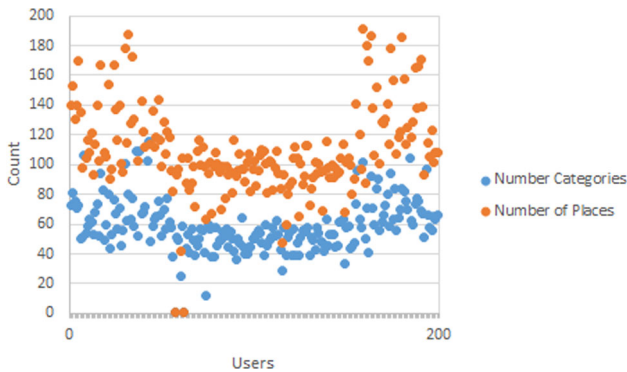
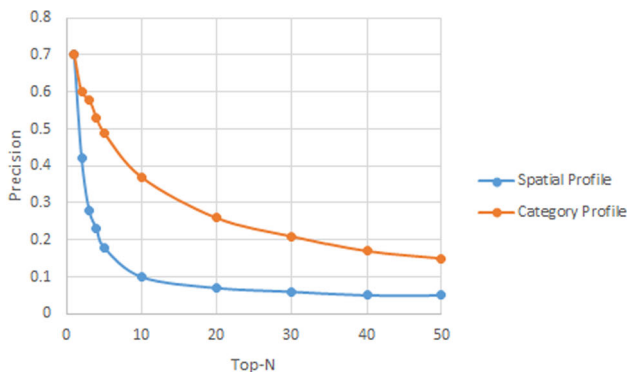
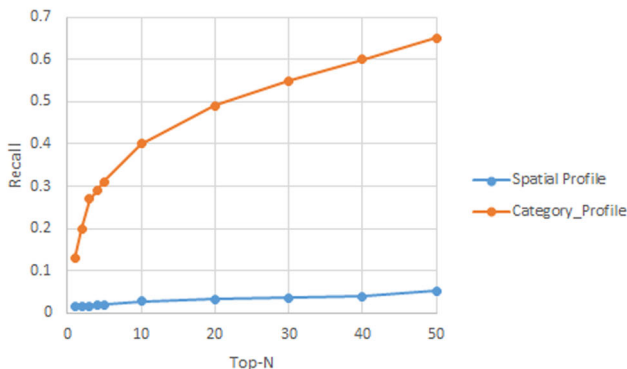


Fig. 13 The number of distinct categories and places for each user



(a) Precision



(b) Recall

Fig. 14 Spatial versus place type-oriented user profiles

place type geo-folksonomy improve the values of precision and recall.

The above result can further be demonstrated when further clustering of place types is used. Figure 15 shows the metrics when place types are clustered to only the top eight parent place categories defined in Foursquare. The top- k values of 1, 2, 4, and 8 (where 8 is the maximum number of place types in this case) are shown.

Activity-oriented user profiles Here we evaluate a semantic activity-oriented profile—describing the users’ association with tags marked as human activities, against the general semantic profile—describing the user’s association with all tags in the folksonomy. In this case, the matrices in the activity geo-folksonomy are smaller (small percentage of tags are activity tags), but not denser.

Predictive accuracy metrics; mean average precision (MAP), root-mean-square error (RMSE) and mean average error (MAE) are used here to measure the extent a recommender system can predict users’ ratings. MAP is the mean of the precision score after each relevant user is retrieved for different top- N values.

$$MAP = \frac{\sum_1^N p@n}{N}$$

MAE is the deviation of the prediction from the true value and is defined as,

$$MAE = \frac{1}{|K|} \sum_{i,j \in K} |r_{i,j} - \hat{r}_{i,j}|$$

$$= \frac{FP + TN}{FP + FN + TP + TN}$$

where K is the set of all user-item pairings (i, j) for which we have a predicted rating $\hat{r}_{i,j}$ and a known rating $r_{i,j}$ and which was not used to learn the recommendation model. RMSE is similar to MAE, but penalises larger errors more strongly than MAE.

$$RMSE = \sqrt{\frac{\sum_{i,j \in K} (r_{i,j} - \hat{r}_{i,j})^2}{|K|}}$$

Figure 16 shows the average values of these metrics for the activity-oriented profiles and the general semantic profiles in the folksonomy. As expected, the semantic profile produces overall better results. This is mainly due to the reduction in the size of the folksonomy based on the activity tags. The denser semantic user profile produces a higher degrees of precision, as shown in the improved MAP measure in Fig. 16.

Fig. 15 Comparison of metrics for different views of the geo-folksonomy: spatial, place type, and parent place type user profiles

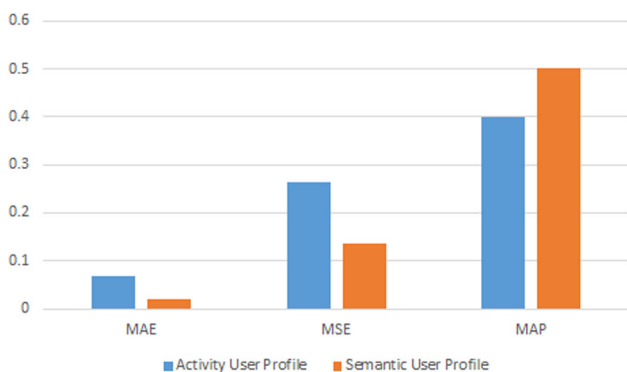
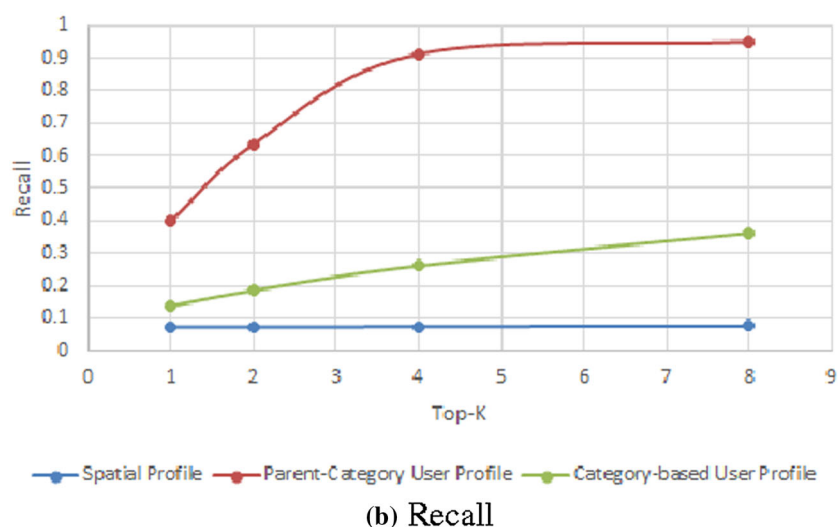
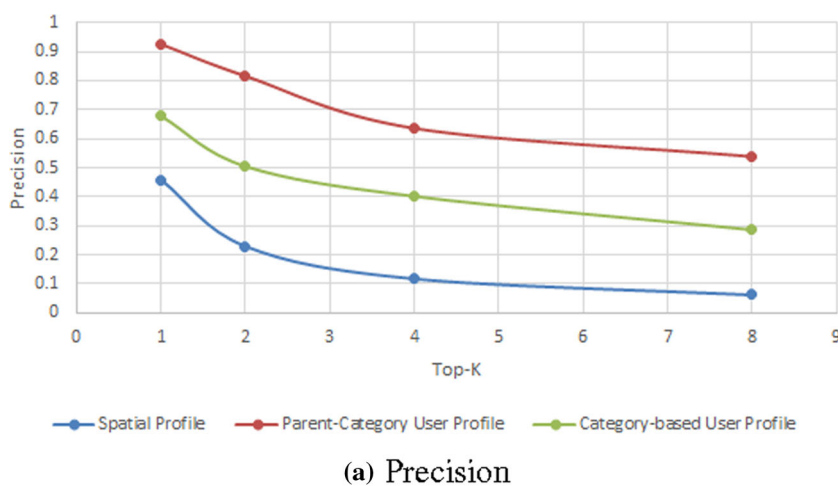


Fig. 16 Semantic and activity-oriented user profiles

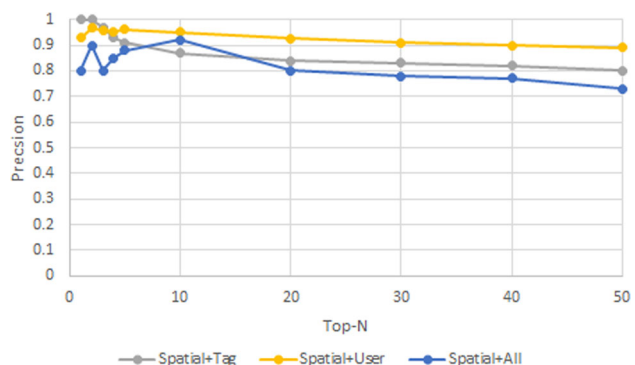


Fig. 17 Precision values for the top-N place recommendations with enriched spatial profiles

5.4 Evaluation of enriched profiles

Here similar experiments are carried out to test the enriched profiles. Different versions of the enriched spatial profiles, using different place similarity measures, were created, (a) using $\gamma = 0$ (to represent enrichment with place-tag sim-

ilarity only), (b) using $\gamma = 1$, (to represent enrichment with place-user similarity only), and (c) using $\gamma = 0.5$ for an aggregated view of both effects. Figures 17, 18, and 19 show the precision, recall, and F1 measure for the different profile, Enriched-(Spatial+Tag) for $\gamma = 1$, Enriched-(Spatial+User) for $\gamma = 0$, and Enriched-(Spatial+All) for

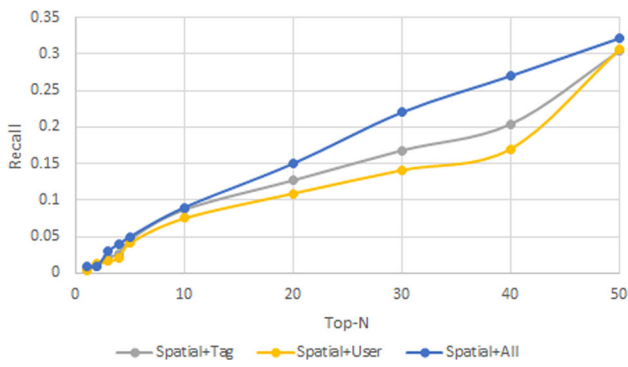


Fig. 18 Recall values for the top-*N* place recommendations with enriched spatial profiles

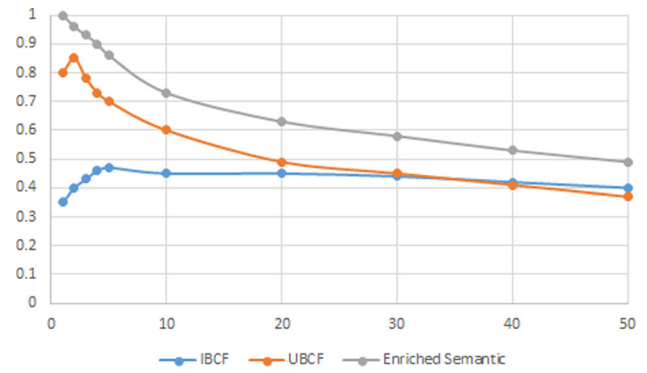
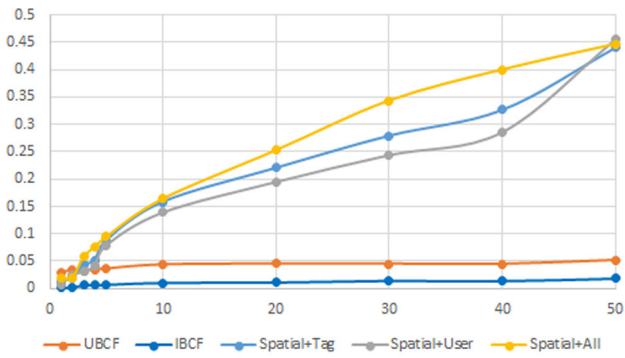
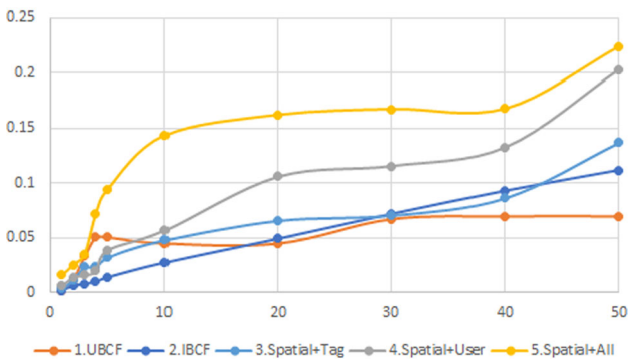


Fig. 20 Precision values for top-*N* tag recommendations for enriched semantic profiles



(a) 200 users



(b) 20 users

Fig. 19 F1 measure values for the top-*N* place recommendations with enriched spatial profiles: **a** for the 200 user dataset, **b** for the 20 user dataset

$\gamma = 0.5$. As can be shown in the figures, enrichment of user profiles results in significant improvements in the precision, recall, and F1 measures compared to the basic spatial profiles in Fig. 14. Both the tag and user enrichments give comparable results. The question of which enrichment to choose, or indeed whether a combined enrichment method is suitable, can be addressed in the context of the datasets considered. The density of the check-ins, and corresponding collocated check-ins, as well as the density of the tags used in places are all factors which are likely to influence the quality of the

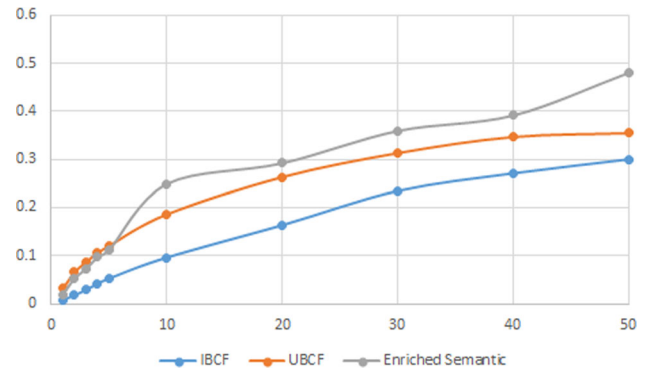
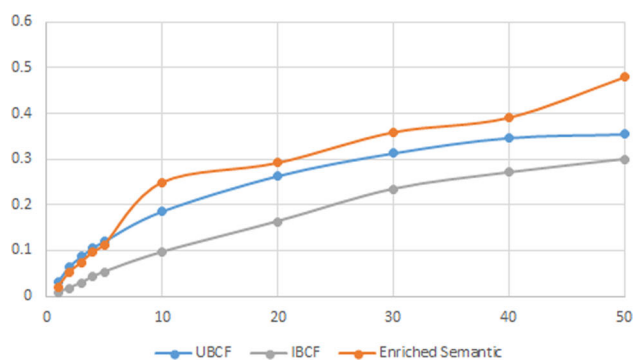


Fig. 21 Recall values for top-*N* tag recommendations for enriched semantic profiles

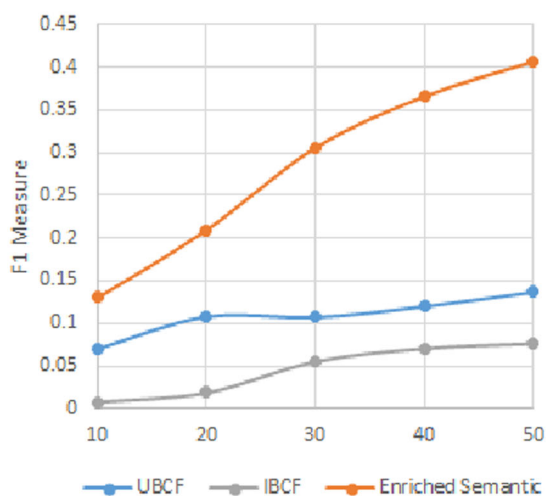
results. Hence, the results derived here can only be seen as general pointers to the possible benefits of using these methods in the context of location-based social networks. Results for the F1 measure are also shown in Fig. 19 for the smaller 20 user sample dataset, where a similar trend in the results can be seen.

Semantic profiles are compared against the item-based collaborative filtering (IBCF) [36] and user-based collaborative filtering (UCBF) [37] approaches. Figures 20, 21, and 22 show the results of the top-10, 20, 30, 40, and 50 tag recommendations using the different methods. As shown in Fig. 20, the enriched semantic profile demonstrates significant improvements with respect to both the traditional approaches. Similar patterns of results from the smaller, 20 user, sample dataset are also shown in the figure. The results are not surprising, since the enriched profiles in both the spatial and semantic cases represent much denser geofolksonomies (with more data in the user–place–tag matrices) in comparison with the raw folksonomies of the basic profiles, on which the IBCF and UCBF methods are applied.

Finally, results were also computed for spatio-semantic profiles. In particular, place type-based profiles were built using activity tags in the folksonomy only. Profiles were then enriched by considering similarity between place types using



(a) 200 users



(b) 20 users

Fig. 22 F1 measure values for the top- N tag recommendations with enriched semantic profiles: **a** for the 200 user dataset, **b** for the 20 user dataset

association with activities only CATREC-Activity, using association with users only CATREC-User and a combination of both CATREC-User+Activity. Results are shown in Tables 5, 6 and 7. A combined enrichment approach seems to produce the best results in this case.

It is worth noting that the improvements in the recommendation results in the case of spatio-semantic profiles may be attributed to the consideration of the inherent coupling [38] between attributes of the place resources. In particular,

a place category is essentially a hierarchical classification of places by their place type. Place type is also inherently related to place activity, for example, a school is associated with concepts of learning and teaching and a hospital is associated with concepts of treatment and cure, etc. [39]. The importance of understanding the similarities of users and items and their relationships and the significance of realising the interdependence between the user and item distributions are proposed by Cao [39]. El-gindy and Abdelmoty [40] describe how geo-folksonomies can be used to extract taxonomies of place types and activities. An in-depth study of such relationships in location-based social networks is the subject of future research.

6 Conclusions

This paper considers the problem of user profiling on location-based social networks. Both the spatial (where) and the semantic (what) dimensions of user and place data are used to construct different views of a user's profile. A place is considered to be associated with a set of tags or labels that describe its associated place types as well as summarise the users' annotations in the place. A folksonomy data model and analysis methods are used to represent and manipulate the data to construct user profiles and place profiles. It is shown how user profiles can be extended from a basic model that describes user's direct links with a place, to enriched profile describing richer views of place data on the social network. The model is flexible and can be adjusted to focus on the spatial and semantic dimensions separately or in combination. Results demonstrate that the proposed methods produce user profiles that are more representative of user's spatial and semantic preferences. In particular, it is noted that a combined treatment of the spatial, social, and semantic aspects of the data can result in richer and more representative profiles. Several open research questions still remain and are subject to future work including the following.

- There is a need to consider different usage pattern for users. Would the methods be effective for users who use the network less frequently?

Table 5 Precision

Top- N	IBCF	UBCF	CATREC-Activity	CATREC-User	CATREC-User+Activity
1	0.102491	0.7	0.905714	0.96	0.963333
2	0.138205	0.42	0.804286	0.88	0.709206
3	0.154188	0.28	0.766667	0.822222	0.659048
4	0.144835	0.23	0.727143	0.796667	0.637579
5	0.139033	0.18	0.68419	0.794667	0.602317
10	0.104344	0.1	0.608095	0.717333	0.459333

Table 6 Recall

Top- <i>N</i>	IBCF	UBCF	CATREC-Activity	CATREC-User	CATREC-User+ Activity
1	0.029784	0.13	0.106253	0.064058	0.18225
2	0.09705	0.2	0.170805	0.105911	0.251613
3	0.1449	0.27	0.227286	0.145425	0.337592
4	0.175471	0.29	0.268756	0.183316	0.427367
5	0.209241	0.31	0.307064	0.225094	0.504571
10	0.288989	0.4	0.504859	0.384251	0.716239

Table 7 F1 measure

Top- <i>N</i>	IBCF	UBCF	CATREC-Activity	CATREC-User	CATREC-User+ Activity
1	0.046155	0.219277	0.190194	0.120103	0.306512
2	0.114028	0.270968	0.281771	0.189066	0.371445
3	0.1494	0.274909	0.350626	0.247138	0.446479
4	0.158688	0.256538	0.392457	0.298049	0.511726
5	0.167061	0.227755	0.423888	0.350817	0.549128
10	0.153327	0.16	0.551689	0.500435	0.559714

- Sentiment is another type of semantics that may be extracted from user annotation. The effectiveness of adding sentiment as well other possible types of implicit semantics can be explored.
- The temporal dimension of the data needs to be studied with the aim of deriving dynamic user profiles that project users' activity and association with place over time.

Compliance with ethical standards

Conflict of interest On behalf of all authors, the corresponding author states that there is no conflict of interest.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Li, Q., Zheng, Y., Xie, X., Chen, Y., Liu, W., Ma, W.-Y.: Mining user similarity based on location history. In: Proceedings of the 16th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems, p. 34. ACM (2008)
- Gao, H., Tang, J., Liu, H.: Exploring social-historical ties on location-based social networks. In: ICWSM, pp. 114–121 (2012)
- Cao, X., Cong, G., Jensen, C.S.: Mining significant semantic locations from gps data. *Proc. VLDB Endow.* **3**(1–2), 1009–1020 (2010)
- Ye, M., Liu, X., Lee, W.-C.: Exploring social influence for recommendation—a probabilistic generative model approach. arXiv preprint [arXiv:1109.0758](https://arxiv.org/abs/1109.0758) (2011)
- Bao, J., Zheng, Y., Wilkie, D., Mokbel, M.: Recommendations in location-based social networks: a survey. *GeoInformatica* **19**(3), 525–565 (2015)
- Noulas, A., Mascolo, C.: Exploiting foursquare and cellular data to infer user activity in urban environments. In: 14th International Conference on Mobile Data Management, vol. 1, pp. 167–176. IEEE (2013)
- Zhang, J.-D., Chow, C.-Y.: igslr: personalized geo-social location recommendation: a kernel density estimation approach. In: Proceedings of the 21st ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems, pp. 334–343. ACM (2013)
- Gao, H., Tang, J., Liu, H.: gscorr: modeling geo-social correlations for new check-ins on location-based social networks. In: Proceedings of the 21st ACM International Conference on Information and Knowledge Management, pp. 1582–1586. ACM (2012)
- Liu, B., Xiong, H., Papadimitriou, S., Fu, Y., Yao, Z.: A general geographical probabilistic factor model for point of interest recommendation. *IEEE Trans. Knowl. Data Eng.* **27**(5), 1167–1179 (2015)
- Cho, E., Myers, S.A., Leskovec, J.: Friendship and mobility: user movement in location-based social networks. In: Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 1082–1090. ACM (2011)
- Zhou, D., Wang, B., Rahimi, S.M., Wang, X.: A study of recommending locations on location-based social network by collaborative filtering. In: Kosseim, L., Inkpen, D., (eds.) *Advances in Artificial Intelligence, LNAI*, vol. 7310, pp. 255–266. Springer, Heidelberg (2012)
- Wang, H., Terrovitis, M., Mamoulis, N.: Location recommendation in location-based social networks using user check-in data. In: Proceedings of the 21st ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems, pp. 374–383. ACM (2013)
- Ying, J. J.-C., Lu, E. H.-C., Kuo, W.-N., Tseng, V. S.: Urban point-of-interest recommendation by mining user check-in behaviors. In: Proceedings of the ACM SIGKDD International Workshop on Urban Computing, pp. 63–70. ACM (2012)
- Liu, B., Fu, Y., Yao, Z., Xiong, H.: Learning geographical preferences for point-of-interest recommendation. In: Proceedings of

- the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 1043–1051. ACM (2013)
15. McKenzie, G., Adams, B., Janowicz, K.: A thematic approach to user similarity built on geosocial check-ins. In: Vandenbroucke, D., Bucher, B., Crompvoets, J. (eds.) *Geographic Information Science at the Heart of Europe*, pp. 39–53. Springer (2013)
 16. Hu, B., Ester, M.: Spatial topic modeling in online social media for location recommendation. In: *Proceedings of the 7th ACM conference on Recommender systems*, pp. 25–32. ACM (2013)
 17. Zhang, J.-D., Chow, C.-Y.: Geosoca: Exploiting geographical, social and categorical correlations for point-of-interest recommendations. In: *Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 443–452. ACM (2015)
 18. Yang, D., Zhang, D., Yu, Z., Wang, Z.: A sentiment-enhanced personalized location recommendation system. In: *Proceedings of the 24th ACM Conference on Hypertext and Social Media*, pp. 119–128. ACM (2013)
 19. Lima, A., Musolesi, M.: Spatial dissemination metrics for location-based social networks. In: *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*, pp. 972–979. ACM (2012)
 20. Lee, M.-J., Chung, C.-W.: A user similarity calculation based on the location for social network services. In: Yu, J.X., Kim, M.H., Unland, R. (eds.) *Database Systems for Advanced Applications, DASFAA 2011*. LNCS, vol. 6587, pp. 38–52. Springer, Berlin, Heidelberg (2011)
 21. Park, M.-H., Hong, J.-H., Cho, S.-B.: Location-based recommendation system using bayesian user's preference model in mobile devices. In: Indulska, J., Yang, L., Ungerer, L., Cao, J. (eds.) *Proceedings of Ubiquitous Intelligence and Computing*, 4611, pp. 1130–1139. Springer, Berlin (2007)
 22. Ye, M., Janowicz, K., Mülligann, C., Lee, W.-C.: What you are is when you are: the temporal dimension of feature types in location-based social networks. In: *Proceedings of the 19th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, pp. 102–111. ACM (2011)
 23. Ye, M., Shou, D., Lee, W.-C., Yin, P., Janowicz, K.: On the semantic annotation of places in location-based social networks. In: *Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 520–528. ACM (2011)
 24. Weakliam, J., Bertolotto, M., Wilson, D.: Implicit interaction profiling for recommending spatial content. In: *Proceedings of the 13th annual ACM International Workshop on Geographic Information Systems*, pp. 285–294. ACM (2005)
 25. Ballatore, A., McArdle, G., Kelly, C., Bertolotto, M.: Reomap: an interactive and adaptive map-based recommender. In: *Proceedings of the 2010 ACM Symposium on Applied Computing*, pp. 887–891. ACM (2010)
 26. Venetis, P., Gonzalez, H., Jensen, C.S., Halevy, A.: Hyper-local, directions-based ranking of places. *Proc. VLDB Endow.* **4**(5), 290–301 (2011)
 27. Chow, C.-Y., Bao, J., Mokbel, M.F.: Towards location-based social networking services. In: *Proceedings of the 2nd ACM SIGSPATIAL International Workshop on Location Based Social Networks*, pp. 31–38. ACM (2010)
 28. Horozov, T., Narasimhan, N., Vasudevan, V.: Using location for personalized poi recommendations in mobile environments. In: *International Symposium on Applications and the Internet (SAINT'06)*, p. 6. IEEE(2006)
 29. Ye, M., Yin, P., Lee, W.-C.: Location recommendation for location-based social networks. In: *Proceedings of the 18th SIGSPATIAL International Conference on Advances in Geographic Information Systems*, pp. 458–461. ACM (2010)
 30. Gilbert, E., Karahalios, K.: Predicting tie strength with social media. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 211–220. ACM (2009)
 31. Xiang, R., Neville, J., Rogati, M.: Modeling relationship strength in online social networks. In: *Proceedings of the 19th International Conference on World Wide Web*, pp. 981–990. ACM (2010)
 32. Mohamed, S., Abdelmoty, A.: Uncovering user profiles in location-based social networks. In: *GEOProcessing 2016, The Eighth International Conference on Advanced Geographic Information Systems, Applications, and Services*, pp. 14–21. IARIA (2016)
 33. Abel, F.: Contextualization, User Modeling and Personalization in the Social Web. PhD Thesis, Gottfried Wilhelm Leibniz University Hannover, April (2011)
 34. Hotho, A., Jäschke, R., Schmitz, C., Stumme, G.: Information retrieval in folksonomies: search and ranking. Sure, Y., Domingue, J. (eds.) *ESWC 2006*. LNCS, vol. 4011, pp. 411–426. Springer, Heidelberg (2006)
 35. Dwork, C., Kumar, R., Naor, M., Sivakumar, D.: Rank aggregation methods for the web. In: *Proceedings of the 10th International Conference on World Wide Web*, pp. 613–622. ACM (2001)
 36. Sarwar, B., Karypis, G., Konstan, J., Riedl, J.: Item-based collaborative filtering recommendation algorithms. In: *Proceedings of the 10th International Conference on World Wide Web*, pp. 285–295. ACM (2001)
 37. Pazzani, M.J.: A framework for collaborative, content-based and demographic filtering. *Artif. Intell. Rev.* **13**(5–6), 393–408 (1999)
 38. Survey, O.: Ordnance survey buildings and places ontology (osbp). (2009). <http://www.ordnancesurvey.co.uk/oswebsite>
 39. Cao, L.: Non-iid recommender systems a review and framework of recommendation paradigm shifting. *Engineering* **2**, 212–224 (2016)
 40. ElGindy, E., Abdelmoty, A.: Enriching user profiles using geo-social place semantics in geo-folksonomies. *Int. J. Geogr. Inf. Sci.* **28**(7), 1439–1458 (2014)