

Article

Interest Aware Location-Based Recommender System Using Geo-Tagged Social Media

Basma AlBanna *, Mahmoud Sakr, Sherin Moussa and Ibrahim Moawad

Department of Information Systems, Faculty of Computer and Information Sciences, Ain Shams University, Cairo 11566, Egypt; mahmoud.sakr@cis.asu.edu.eg (M.S.); sherinmoussa@cis.asu.edu.eg (S.M.); ibrahim_moawad@cis.asu.edu.eg (I.M.)

* Correspondence: basma.hassan@cis.asu.edu.eg; Tel.: +20-010-100-11967

Academic Editors: Georg Gartner, Haosheng Huang and Wolfgang Kainz

Received: 23 October 2016; Accepted: 7 November 2016; Published: 19 December 2016

Abstract: Advances in location acquisition and mobile technologies led to the addition of the location dimension to Social Networks (SNs) and to the emergence of a newer class called Location-Based Social Networks (LBSNs). While LBSNs are richer in their model and functions than SNs, they fail so far to attract as many users as SNs. On the other hand, SNs have large amounts of geo-tagged media that are under-utilized. In this paper, we propose an Interest-Aware Location-Based Recommender system (IALBR), which combines the advantages of both LBSNs and SNs, in order to provide interest-aware location-based recommendations. This recommender system is proposed as an extension to LBSNs. It is novel in: (1) utilizing the geo-content in both LBSNs and SNs; (2) ranking the recommendations based on a novel scoring method that maps to the user interests. It also works for passive users who are not active content contributors to the LBSN. This feature is critical to increase the number of LBSN users. Moreover, it helps with reducing the cold start problem, which is a common problem facing the new users of recommender systems who get random unsatisfying recommendations. This is due to the lack of user interest awareness, which is reliant on user history in most of the recommenders. We evaluated our system with a large-scale real dataset collected from foursquare with respect to precision, recall and the f-measure. We also compared the results with a ground truth system using metrics like the normalized discounted cumulative gain and the mean absolute error. The results confirm that the proposed IALBR generates more efficient recommendations than baselines in terms of interest awareness.

Keywords: location-based social networks; geo-tagged media; location-based services; user interests; recommender systems

1. Introduction

Worldwide accessibility to the Internet is a defining phenomenon of the present time that reshapes the world as we currently know it. A key component in this contribution is the social media that comes in many forms, including blogs, photo-sharing platforms, chat applications and last, but not least, Social Networks (SNs). The number of SN users is expected to reach 2.5 billion by 2018 [1], which is around a third of the Earth's entire population. "A social network can be defined as an abstract structure comprised of individuals connected by one or more types of relations such as friendships, common interests and shared knowledge" [2]. With the advances of location acquisition and wireless communication technologies, another dimension has been added to the existing SNs, which is the location dimension, bridging the gap between the physical world and the social networking services. SNs got affected by this new dimension in two main ways:

1. Geo-enrichment of SNs, enabling users to share geo-tagged content in the form of check-ins, geo-tagged photos or notes.
2. There appeared a new breed of social networks called Location-Based Social Networks (LBSNs) where locations are the root elements of the network structure. Hence, a location can be a stand-alone entity of the network having relationships with other locations and users.

While SNs allow only users to be entities, LBSNs allow both users and locations. Consequently, the relationships in SNs are only user-user, whereas the relationships in LBSNs can be: user-user, user-location and location-location. In an SN, a user-user relationship is derived from a real-world connection between two users [3] (e.g., friendship, co-worker, family). In LBSNs, a user-user relationship can additionally indicate a connection derived from the user locations (e.g., co-location, common interest, activities). In LBSNs, locations can relate to users (e.g., visit, work, study) and to other locations (e.g., neighborhood, containment, correlation). Because locations are entities on their own in LBSNs, concepts like distance, hierarchical containment and sequential ordering are considered as distinguishable features of LBSNs; which place them apart from SNs. A clear distinction between those two types of social networks is illustrated in Figure 1.

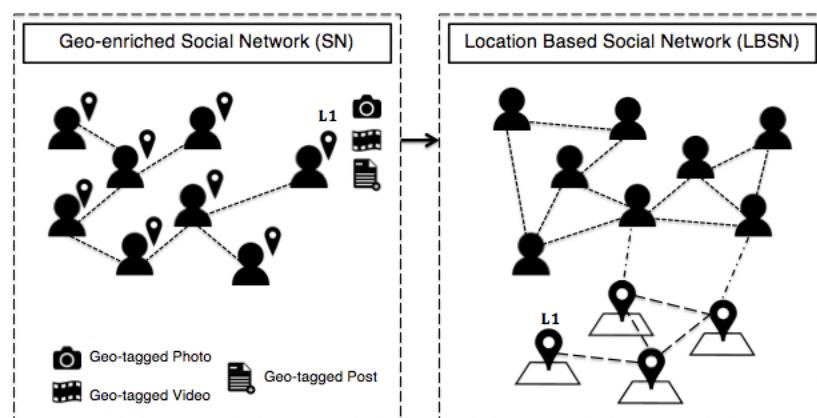


Figure 1. Social Network (SN) and Location-Based SN (LBSN) network structure.

While LBSNs are richer than SNs in terms of the network structure and the range of functionality that can be provided, LBSNs failed to attract as many users as SNs. Based on the literature in [4–8], we created Figure 2, which compares the growth rate of Instagram, the fastest growing SN, with Foursquare, the fastest growing LBSN. According to [9] in 2012, while Instagram has added 10 million users in 10 days, Foursquare added 15 million users in a year.

A main reason behind this gap is because SNs are designed to provide value not only for active users, but also for passive users (i.e., users who do not often generate content or engage with their network, but instead, act as observers). For example, according to [10], only 42% of Facebook users are active users who engage with their network and contribute with content, while the remaining 58% are passive users. The main services of LBSNs (e.g., Foursquare) are broadcasting user locations to their network and recommending places based on proximity and query match. The value of these services depends on the engagement of the user, either through input in the form of venue tips and ratings or having enough friends to share location updates.

On the one hand, SNs have more users and more content than LBSNs. However, on the other hand, LBSNs, due to their inherent spatial capabilities, have a larger capacity to understand the users and their context, allowing for a wider range of applications. Furthermore, LBSNs have a better understanding of user preferences, behaviors, interesting locations, travel sequences, event detection, etc. In this paper, our research was motivated by three main findings:

1. SNs, due to the large user base, have vast amounts of geo-tagged content, which is under-utilized because of the lack of spatial processing capabilities. Precisely, geo-tagged media in photo-sharing SNs (e.g., Instagram) are always associated with tag annotations by which, if exploited, the location content will be enriched significantly. They also provide what can be called “collective wisdom”, i.e., identifying interesting items/activities at a location based on people’s previous experiences at that location.
2. LBSNs, despite their spatial processing capabilities, fail to attract a large user base, because they provide less value for passive users. They also rely heavily on user profiles, user location history and user trajectories to provide recommendations, adding to the severity of the cold start problem. Therefore, the quality of personalized recommendations is not reliant on the technology alone, but also on user activities and properties.
3. In all previous related works, factors affecting personalized location recommendation are limited to user demographics, location proximity, user mobility and location history. However, a specific user interest at any point of time was not addressed before. For example, a user wants a list of restaurants serving the best burger in town on the condition that they are within a specified geo-range. To fulfill this request, the user profile and user location history will not be as important as identifying which restaurants serve the best burger. In other words, it is not only about ranking locations, but also about ranking interesting items/activities in locations.

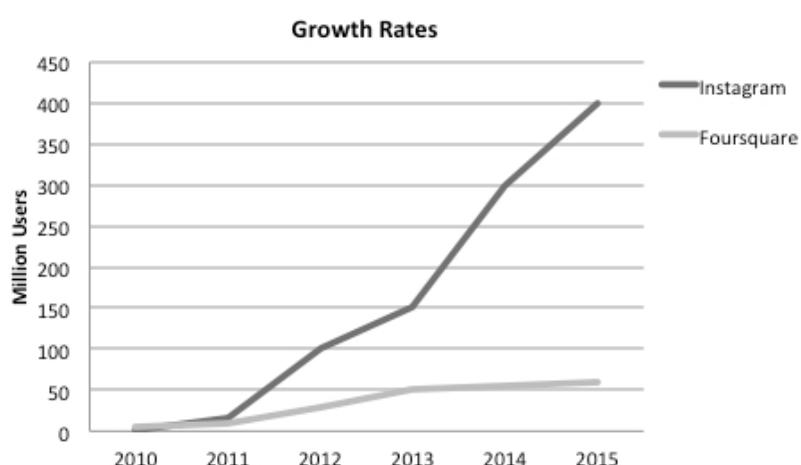


Figure 2. Growth rates of Instagram vs. Foursquare.

Therefore, we propose joining SNs with LBSNs in a way that uses the advantages of both. In this way, we will enrich the LBSNs’ content with the geo-tagged content of SNs, which will allow the following benefits: (1) the utilization of the geo-tagged content of SNs will be maximized when it is made available to LBSNs, because of their spatial processing capabilities; (2) based on this enriched content in LBSNs, personalized interest-aware recommendation can be offered for both active and passive users. Thus, it will attract the pool of passive users that can still capture value from LBSNs with minimal behavioral changes and contribution.

In this paper, we illustrate and examine these benefits by proposing a novel Interest-Aware Location-Based Recommender (IALBR). This recommender is built upon an LBSN and uses geo-tagged content generated from an SN. This paper’s contributions can be summarized as follows:

1. Linking SNs and LBSN for interest-aware recommendation: joining geo-tagged media annotations in SNs to their corresponding locations in LBSNs.
2. Interest-aware recommendation and scoring: the IALBR supports recommendations (e.g., restaurants) matching a very specific user interest (e.g., French fries). This recommendation is done based on a new scoring methodology that takes into consideration both the popularity of an interest in the location and the popularity of the location.

The rest of this paper is organized as follows: Section 2 summarizes the related work. Section 3 gives an overview of the proposed IALBR system while showcasing the underlying architecture, data model and system modules. Section 4 demonstrates the application scenario. Section 5 evaluates the proposed IALBR along with the associated experiments and results conducted to examine the proposed system. Section 6 contains a further discussion, and Section 7 concludes the paper and presents the future work.

2. Related Work

Based on the framework proposed in [2], there are four types of recommenders in LBSNs: (1) the location recommender, which suggests stand-alone locations (e.g., Points of Interests (POIs)) or sequential locations (e.g., travel routes); (2) the user recommender, which suggests expert users, potential friends or communities; (3) the activity recommender, which suggests activities that could be of interest to the user; (4) the social media recommender, which suggests social media content. The scope of this research falls in the first type, location recommender. Within this type, the proposed framework in [2] categorized the related works based on the data sources used, as follows: (1) user profile-based; (2) user location history-based; (3) user trajectory-based; and we added to this categorization the (4) location information-based and the (5) hybrid recommenders.

1. User profile-based: These are recommenders that suggest locations by matching location semantics or metadata with user profiles. For example, [11] matches user profile data (e.g., gender, income and age) with the price and category of a restaurant using a Bayesian network model. Others took into consideration the proximity of the candidate location [12] or supported low-end devices [13] by focusing on using the user address and social affinity. In [14–16], the authors tried to explore spatial and temporal relations between locations within trajectories by extracting features and identifying correlations and sub-sequences related to the user's preferred activities.
2. User location history-based: These are recommenders that use user location history either through the ratings they give to locations they visited or through their check-in history. For example [17–22] provided personalized recommendations for locations taking into account other user's rating using Collaborative Filtering (CF) models. Hence, this improved the quality of recommendation by ignoring poorly-reviewed locations that could match a user profile in the user profile-based recommenders. The CF-based recommendation was done by first identifying similar users, using their rating history, and then selecting a subset of candidate locations based on the user's location and, finally, predicting the rating a user would give to this location. The work in [23] also developed a random walk-based model that recommended new venues by learning from both social ties and check-ins simultaneously. Furthermore, the authors in [24] used check-in data to model the probability of a user's check-in on a location as a Multi-center Gaussian Model (MGM). This was done through fusing matrix factorization with geographical and social influence. The authors in [25] also leveraged check-in data for location recommendation through using weighted matrix factorization that incorporates spatial clustering. Recently, [26] introduced a personalized location recommendation approach that is based on identifying correlations between POIs through capitalizing on historical check-in data of users.
3. User trajectory-based: These are recommenders that estimate user preferences based on the sequence of visited locations, the path traveled and the stay time at locations. Example systems include [27–33], which proposed recommendation frameworks, which suggest interesting locations by mining GPS trajectory data. The authors in [34] proposed GTAG, a time-aware location recommender that took into account both geographic and temporal influences through exploiting large volumes of user spatiotemporal trajectory data. Recently, [26] developed a new model for location recommendation called LOREthat extracted sequential patterns from the historical check-in data of all users to use them in predicting the location that users might be interested in visiting.

4. Location information-based: These are recommenders that recommend POIs based on location information represented in POI properties (e.g., spatial location), user-generated content (e.g., tips), user geo-tagged media, etc. Among those recommenders is ORec [35], an opinion-based recommendation framework that recommends POIs based on user opinions found in venue tips. In [36], the authors used Implicit-feedback Content-aware Collaborative Filtering (ICFF) for location recommendation to incorporate semantic content (e.g., geo-tagged tweets) without negative sampling (i.e., user's negative preferences). Furthermore, the authors in [37] proposed Geo-SAGE, a geographical sparse additive generative model for POI recommendation, which exploited and integrated geographical and social influence, temporal effect and content information of POIs for recommendation. Furthermore, [38,39] proposed systems that exploit geo-tagged media on photo-sharing applications to identify semantically meaningful tourist locations. They applied collaborative filtering to obtain tourist's preferences in one city, from their public geo-tagged photos, and used it for personalized travel recommendation in another city.
5. Hybrid: These are recommenders that integrate the above categories of data sources for location recommendation. For example, the authors in [40] used a probabilistic factor model that takes into account user preferences by counting user check-ins (user location history), geographical influence (location information) and user mobility behaviors (user trajectory) to provide personalized recommendation of places. The authors in [41] developed a recommendation framework that capitalized on POI information, represented as location properties, user interests and sentiment detection, to improve POI recommendation in LBSNs. Another example can be found in [42], which proposed a recommender system that mines large-scale sets of geo-tagged media (location information) to identify the most popular landmarks, as well as the best travel routing between the landmarks (user trajectory) to be used for road-based travel recommendation. A similar, but more personalized framework was proposed in [43] where large-scale sets of photographer geo-tagged media (location information) were mined to identify routes (user trajectory) across interesting landmarks, using them to develop probabilistic photographer behavior models for personalized travel route recommendation.

Our proposed LBSN recommender falls under the location information-based category and differs from the related works in two ways: (1) exploiting geo-tagged media: to the best of our knowledge, the proposed approach is the first to leverage the huge amount of geo-tagged content existing in photo-sharing SNs to extract the most popular items existing in a specific location to be used for location recommendation in LBSNs; (2) interest granularity: the other recommenders are either too generic, having recommendations based on public opinions on locations or travel routes extracted from GPS trajectories in a city, regardless of location categories (i.e., restaurants, parks, shopping malls, etc.); or less generic (more personalized); or recommenders that filter candidate locations based on categories (i.e., restaurants). However, none of them considered specific interest-based recommendation (e.g., the user wants to eat a burger instead of just going to a nearby restaurant). Furthermore, the proposed recommender supports high-level and generic queries (e.g., bakery) and low-level queries (e.g., cheesecake) within a user-specified geo-spatial range.

3. IALBR System Design

This section first introduces the architecture of the IALBR system, then it presents the used data model, followed by a formal description of the IALBR system modules. The main concept behind the IALBR, which is built upon LBSNs, is that venues in LBSNs get enriched with the corresponding geo-tagged content from SNs. In other words, we join the venue information in both LBSN and SN. As a result, we create a dataset that is rich in tags and tips about venues. This dataset is then used as the basis for user interest queries. We also propose a novel scoring method for ranking the query results.

3.1. System Architecture

IALBR is composed of three main modules: (1) interest tags mapping module; (2) interesting venues and expert user inferring module; and (3) interest-aware recommender module. It is also composed of several data stores containing data about users, venues, interests and check-ins. IALBR is designed to be an extension to existing LBSNs that will be joined to SNs to enrich their venue information. It provides a recommender that complements the location-based services of existing LBSNs by making them more interest aware through capitalizing on the richness of the geo-tagged content in SNs. Furthermore, the addition of the interest tags dimension to the current LBSN user and location dimension creates a new horizon of applications that support diverse levels of preference granularity. Figure 3 depicts the generic architecture of IALBR.

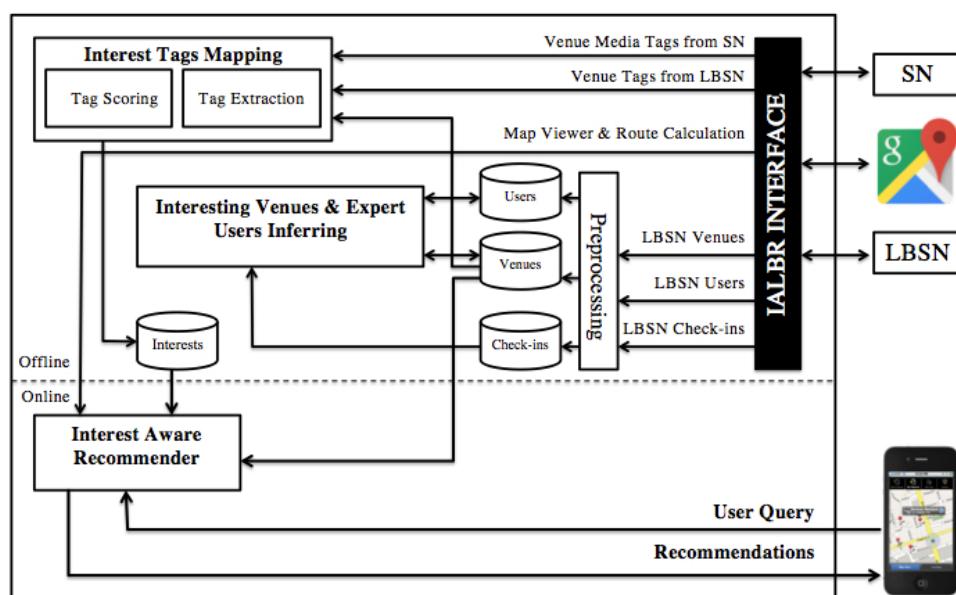


Figure 3. Interest-Aware Location-Based Recommender system (IALBR) system architecture.

The way of joining the venue information in both SN and LBSN varies according to the used SN and LBSN. Several criteria can be employed to improve the join accuracy (e.g., spatial proximity, name similarity, tag similarity, etc). Some existing SNs and LBSNs maintain common keys for venues. Instagram and Foursquare, for instance, maintain such keys. In this work, we used them to illustrate and implement our system. The IALBR activities can be summarized into two main groups:

1. Offline modeling: This starts by crawling Instagram geo-tagged media corresponding to Foursquare venues using Instagram API functions and venue unique keys, common to Foursquare and Instagram. It then extracts the geo-tagged media tags representing interests in order to calculate interest scores based on the frequency of occurrence of those tags. It also extracts the interest tags annotating this venue on Foursquare. Finally, an average-weighted score is calculated for each interest tag based on its frequency of occurrence in both the LBSN and the SN. Figure 4 shows geo-tagged media extracted from a venue called “Teds” on Instagram. Each photo contains tags annotating it. In this example, the interest tags resembling these photos are: #strawberry, #mojito, #moltencake, #chocolatesouffle, #macncheese, #macroni and #cheese.

Another important offline process is done at the interesting venues and expert user inferring module, where users get hub scores and venues get authority scores using check-in information. A user having a high hub score is a user who visits many interesting venues (i.e., venues having high authority scores), and a venue having a high authority score is a venue visited by many expert users (i.e., users having high hub scores).

2. Online recommendation: This starts when the IALBR receives a user query asking for venues matching an interest (e.g., burger) within a specific geo-range (e.g., 10 km). A power score p is calculated for venues matching this interest and are within the predetermined geo-range. This power score is calculated based on the interest tag score resulting from the interest tags mapping module and the venue's authority score resulting from the interesting venues and expert user inferring module. Finally, it returns back a ranked list of venue recommendations.



Figure 4. A sample of Instagram geo-tagged media at a venue.

3.2. Data Model

A venue in an LBSN (e.g., Foursquare) is described as:

$$v_i = \langle n, c, lon, lat, a(v_i) \rangle$$

This means that a venue v_i has a name n , belongs to category c , existing at a longitude lon and a latitude lat . We further enrich this venue entry by an authority score $a(v_i)$ indicating the popularity of this venue. This score is computed using the Hypertext-Induced Topic Search (HITS) method [32]. An example of a venue would be presented as: ("Walter Foods", "American Restaurant", -73.95861, 40.71341, 5.89).

Similarly, let U be the set containing all users of the same LBSN. A user can be modeled as follows:

$$u_i = \langle n, hc, h(u_i) \rangle$$

where a user u_i has a name n and home city hc . We enrich the user entry by a calculated hub score $h(u_i)$. This score is calculated using the method in [32], as well. A higher hub score indicates that the user is experienced and that his/her tags are more trusted. An example of a user would be presented as: ("Thomas Rossi", "Washington, DC", 2.97).

Users in U perform check-ins in venues V . This check-in history is obtained from the LBSN. We group the check-ins by the user and represent the history of a user as:

$$c_i = (u_i, \langle (v, t) \rangle)$$

meaning a user $u_i \in U$ has an ordered list of pairs (v, t) , where $v \in V$ is the venue and $t \in T$ is a time-stamp. An example of check-ins for a user would be presented as: ("Thomas Rossi", < ("Amrit", 17 January 2013 9:02:30), ("TUM", 17 January 2013 9:05:51), ("Marineplatz", 17 January 2013 9:05:51)>). Let C be the set of such entries for all users in U .

Finally, we compute the list of interests τ . Let τ be the set containing all interests extracted from the descriptive tags of geo-tagged media in a selected SN (e.g., Instagram) and extracted from the tags annotating all venues in the LBSN. An interest can be modeled as follows:

$$\tau_i = (interest_i, \langle (v, s) \rangle)$$

This means that an interest with the name n_i is represented in the form of an ordered list of pairs (v, s) , where each pair contains a venue v_j and a score $s_j \in R$, which is the weight of the interest τ_i at venue v_j calculated in the interest mapping module. An example would be (“Burger”, < (“Walter Foods”, 64), (“Burger King”, 58), (“ShaknShack”, 42)>).

3.3. Interest Tags Mapping Module

This module fills in the list τ by collecting data from both the LBSN and the SN. For all of the venues coming from the LBSN, it extracts the corresponding descriptive tags from both the SN and the LBSN. Every $v_i \in V$ has both tags and tips in the LBSN. Tags are keywords that best describe the venue, in an LBSN. Tips (one of the most popular interaction mechanisms in LBSNs, like Yelp, Foursquare, whrrl, etc.) are user posts that correspond to the venue. This scoring method applies for both of them. Typically, LBSNs extract tags by analyzing the tips. In the SN, these descriptive tags exist in the form of annotations on geo-tagged social media content. We are only interested in the geo-tagged media that are related to the set of venues V , coming from the LBSN.

A pre-processing task has to be done to keep only the tags that reflect interests. For example, an Indian restaurant containing geo-tagged media tags like #foodie #love #favourite #indian #restaurant #beryani will be reduced to #indian #restaurant #beryani after the pre-processing. This might be done by removing non-indicative tags, like stop words, the name of the place, etc. These remaining tags are what we call *interestin* τ .

Following the pre-processing, we proceed by constructing the list $\langle v, s \rangle$ for every *interest*. That is, we construct the list of venues fulfilling this interest along with scores that indicate how well the interest is met in this venue. This score s is a weighted average score that combines two scores $\omega(v, interest)$ and $\varphi(v, interest)$.

The score $\omega(v, interest)$ calculated in Equation (1) indicates the popularity of the interest in a venue using only the LBSN’s tips, where $tips_v$ denotes tips corresponding to a venue v , containing an *interest*. For example, “I love this place” is not considered a tip in our scoring method, while “This place makes a great burger” will be counted as a tip indicating an *interest* (burger). Thus, the final score will be a percentage of venue tips, which is coming from the LBSN and containing the given *interest*.

$$\omega(v_i, interest_j) = \frac{tips_{v_i} \text{ containing } interest_j}{tips_{v_i}} \quad (1)$$

The score $\varphi(v, interest)$ calculated in Equation (2) indicates the popularity of the interest in a venue using only the SN geo-tagged media, where $media_v$ denotes geo-tagged media items corresponding to a venue v having annotations containing an *interest*. Therefore, the final score will be a percentage of venue geo-tagged media items, coming from the SN, containing the given *interest*.

$$\varphi(v_i, interest_j) = \frac{media_{v_i} \text{ annotated by } interest_j}{media_{v_i}} \quad (2)$$

Finally, for every $interest_i$ in τ , the average weighted score s_{ij} of $interest_i$ in every venue v_j meeting this interest is calculated using Equation (3).

$$s_{ij} = \frac{tips_{v_i} * \omega(v_j, interest_i) + media_{v_j} * \varphi(v_j, interest_i)}{tips_{v_j} + media_{v_j}} \quad (3)$$

3.4. Interesting Places and Expert User Inferring Module

The Hypertext-Induced Topic Search (HITS) [32] is used to infer interesting venues and experienced users. Figure 5 illustrates the main idea of the HITS-based inference model. In this model, a user visit to a venue is regarded as a directed link from the user to that venue. Thus, a user is a hub if he/she has check-ins in many venues, and a venue is an authority if it has been checked-in by many users. Following that, a scoring algorithm is applied to the proposed model, which gives users and venues scores based on those edges. The assigned scores to each user and venue indicate the popularity of the venue and the travel experience of the user.

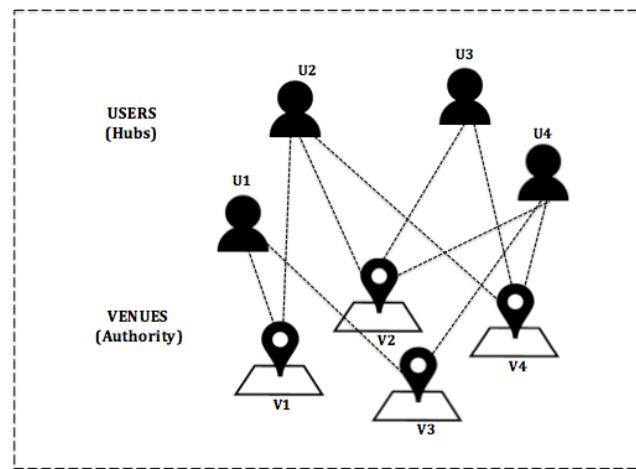


Figure 5. Hypertext-Induced Topic Search (HITS)-based inference model.

This interesting places and expert user inferring module of IALBR uses a HITS-based inference model [32] to identify experienced users and interesting venues to support our interest-aware recommendation. Given the set C containing the history of check-ins for all users U , for every venue $v_i \in V$, an authority score $a(v_i)$ is calculated. This score indicates how popular this venue is, based on the hub scores of the users who checked-in this venue. It also calculates for every user $u_i \in U$ a hub score $h(u_i)$ based on the authority scores of the venues visited by this user. Since there is a mutual reinforcement relationship between the hub users and the authority venues, they are both calculated using the power iteration method shown in Equations (4) and (5).

$$h(u_i) = \sum_{v_j \in V} \sigma_j * a(v_j) \quad (4)$$

$$a(v_i) = \sum_{u_j \in U} \lambda_j * h(u_j) \quad (5)$$

where σ is the number of check-ins of user u_i in v_j and λ is the number of check-ins of user u_j in v_i .

3.5. Interest-Aware Recommender Module

This module is responsible for receiving interest-aware location-based queries and returning back a ranked list of venues having matching interests. A user query can be modeled as follows:

$$q = \langle u_i, lng, lat, interest, range \rangle$$

An example of a user query would be presented as: $q = \langle Thomas\ Rossi, -6.98, 42, "Fries", 1\ mile \rangle$. The recommendations are made using our interest-aware recommender algorithm (i.e., demonstrated in Algorithm 1) having as data the set of all interests τ calculated in the interest tags mapping module

and the set of all venues V enriched by the venue authority scores $auth(v)$ calculated in the interesting places and expert user inferring module. It retrieves a ranked list of the venues meeting the user $interest$ and within the specified geo-range. Their ranking is based on the power score calculated in Equation (6).

$$p = s_{ij} * a(v_j) \quad (6)$$

where s_{ij} is the average weighted score of an $interest_i$ at a venue v_j . Its value is less than one; while $a(v_j)$ is the authority score of venue v_j calculated using HITS. This way, the power score p of an $interest_i$ will be a fraction of the total authority of the venue v_j , determined by the popularity of the $interest_i$ in it, i.e., the higher the value of p , the higher is the popularity of the $interest_i$ in v_j .

Algorithm 1: Interest aware recommender.

```

input : (1) User geographic location  $loc$ ; (2) recommendation geo-range  $r$ ; (3) user interest  $i$ 
output: A sorted list  $L$  of venue recommendations.

begin
  Select the venues  $V'$  within the specified range  $r$  from the user location  $loc$ 
  Get the venues  $V''$  from  $V'$ , which are venues containing user interest  $i$  given the set  $\tau$ 
  for  $v_j \in V''$  do
    Get  $s_{ij}$  //average weighted score of an  $interest_i$  at a venue  $v_j$ 
    Get  $a(v_j)$  //authority score of  $v_j$ 
     $L[j] \leftarrow$  multiply  $s_{ij}$  and  $a(v_j)$  //power score of  $v_j$ 
  end
   $L \leftarrow Sort(L)$ 
  Sort venues based on power score return  $L$ 
end
  
```

4. System Implementation

We implemented the proposed IALBR as an android mobile application on top of an LBSN (Foursquare) and joined it with an SN (Instagram) using their API functions. Foursquare has been chosen since it is the largest LBSN to date, while Instagram is the fastest growing SN having a huge user base contributing large amounts of geo-tagged media, which can be directly mapped to Foursquare's venues. Additionally, user-generated geo-tagged media on Instagram support preference granularity; for example: when a user uploads a photo on Instagram while having dinner at an Indian restaurant, the tags associating this photo will be something like #biryani #indianfood #amrit #berlin, i.e., food item, food category, restaurant and city. Moreover, interest tags extracted from geo-tagged media are mostly positive tags, i.e., when a user uploads a photo of a food item they tried at a restaurant, it is because they like it. This conclusion is derived from [44], a research study that examined the behavior of common users on Instagram. We are using PostgreSQL, since it has an advanced geo-spatial handling. Moreover, we integrated the application with Google Maps for the map view and direction features.

When the user logs into the application, the homepage shown in Figure 6a appears to the user. The user types an interest in the "Search Interest" text field and selects a geo-range (in kilometers) for venue recommendations. For example, if a user selects the range "20 km" and "burger" as an interest and then clicks the "Go" button, the query is executed, and a number of venues matching the user's interest appears, as shown in Figure 6b.

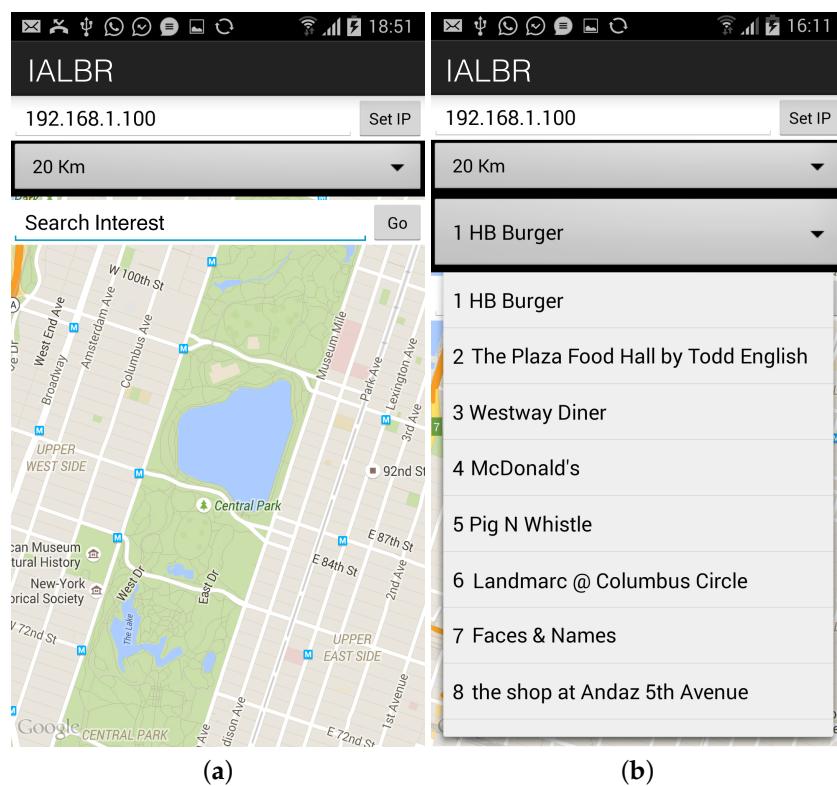


Figure 6. IALBR mobile application. (a) Home page; (b) user query results.

5. Experimental Evaluation

In this section, we describe the settings of the experiments, including the used dataset, the evaluation methods and the baseline approach. The experimental results are then reported.

5.1. Experimental Settings

In order to validate the proposed IALBR, we carried out various experiments based on the large-scale dataset provided in [45,46], which is collected from Foursquare. It has 35,023 venues, 10,377 tips and 392,247 check-ins generated by 17,463 users in NYC. For the venues in this dataset, we crawled 641,079 geo-tagged media items from Instagram. This has resulted in extracting 217,511 tags that represent interests associated with these venues.

Our experiments aim at measuring the following aspects of our approach:

1. Utility of recommendations.
2. Ranking of recommendations.

To measure these aspects, we conducted 10 interest queries simulating various user interests from a fixed location in NYC (latitude = 40.7678052451765 and longitude = −73.98199327290058). The user requests venue recommendations matching his/her interest within a radius of 20 km. The queries varied in the interest level, for example, we queried “cheesecake” (low level) and “bakery” (higher level). The queries also had varying interest types ranging from health and wellness like “gym” and “running” to food items like “sushi” and “burger”.

5.2. Experimental Results

5.2.1. Utility of Recommendations

In this experiment, the baseline approach is the current Foursquare recommender before its enrichment with Instagram’s geo-tagged media annotations. It will recommend venues within the

20-km range, having venue tags and tips matching the interest query; whereas our IALBR will recommend venues within the 20-km range having matching interest tags in both Foursquare's tags and tips and Instagram's geo-tagged media items. The venues resulting from both approaches are given power scores p (see Equation (6)), and they are ranked based on this.

Table 1 shows the summary of the results after performing ten interest queries using both approaches. It shows the total number of results and the percentage of relevant results per query (precision at ten). We used Yelp as a ground truth of relevancy by considering a venue to be a relevant venue only if it includes positive reviews of the query item. The table also shows if the best match (i.e., the venue having the highest power score p) was among the results of both approaches.

Table 1. Comparison between the results of IALBR and Foursquare.

Query Item	Foursquare (Baseline)			IALBR		
	Total No. of Results	Relevant Results (%)	Best Match Found	Total No. of Results	Relevant Results (%)	Best Match Found
Bagel	1	100%	No	10	100%	Yes
Burger	4	100%	Yes	10	100%	Yes
Pizza	7	100%	Yes	18	72%	Yes
Sushi	2	100%	Yes	10	80%	Yes
Cheesecake	0	0%	No	11	100%	Yes
Running	0	0%	No	12	58%	Yes
Bakery	3	100%	No	5	80%	Yes
Gym	6	100%	Yes	14	43%	Yes
Organic	1	100%	Yes	12	75%	Yes
Seafood	0	0%	No	21	95%	Yes
Average	3	70%	50%	13	80%	100%

On average, Foursquare queries resulted in around three venues per query with a precision of 1.0 and had the best match 50% of the time; while IALBR's queries resulted in around 13 venues per query with a precision of 0.8 and included the best match 100% of the time. This means that for every query, IALBR added on average 10 new venues for the user out of which, eight are relevant venues that are very likely to match the user interest. Furthermore, IALBR was able to answer all of the interest queries, while Foursquare failed to answer 30% of the queries.

Table 2 shows the precision Pr , recall R and F-measure F of the 10 queries at different thresholds N , where $N = 5, 10, 15$. We did this experiment to identify the threshold yielding the best F-measure (i.e., the best weighted average of the precision and recall). It is a single value that indicates the overall utility of the recommendation list. The precision, recall and F-measure were calculated using Equations (7)–(9), respectively.

$$Pr = \frac{\text{Relevant Retrieved}}{\text{Relevant Retrieved} + \text{Irrelevant Retrieved}} \quad (7)$$

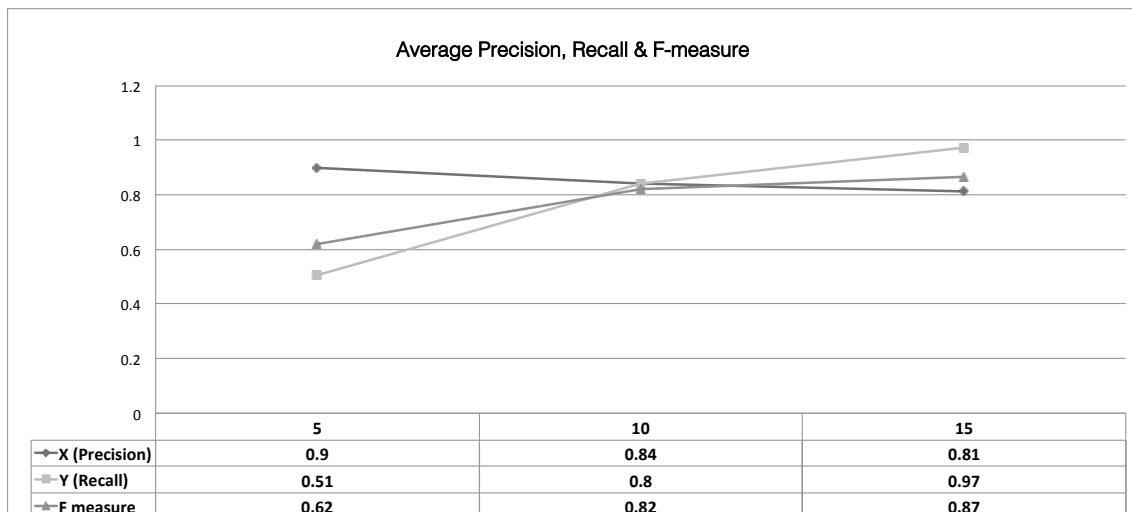
$$R = \frac{\text{Relevant Retrieved}}{\text{Relevant Retrieved} + \text{Relevant Not Retrieved}} \quad (8)$$

$$F = \frac{2 * Pr * R}{Pr + R} \quad (9)$$

Figure 7 shows the average recall, precision and F-measure for the 10 queries, indicating that $N = 15$ achieves the highest F-measure (0.87). It also shows that the highest average precision is achieved at $N = 5$, and the highest average recall is achieved at $N = 15$.

Table 2. Precision, recall and F-measure at every N .

Query Item	$N = 5$			$N = 10$			$N = 15$		
	Pr	R	F	Pr	R	F	Pr	R	F
Bagel	1	0.5	0.67	1	1	1	1	1	1
Burger	1	0.4	0.53	1	0.7	0.83	0.7	1	0.84
Pizza	1	0.4	0.56	0.9	0.7	0.78	0.9	0.7	0.78
Sushi	1	0.6	0.77	0.8	1	0.89	0.8	1	0.89
Cheesecake	1	0.5	0.63	1	0.9	0.95	1	1	1
Running	0.6	0.4	0.5	0.6	0.9	0.71	0.6	1	0.74
Bakery	0.8	1	0.89	0.8	1	0.89	0.8	1	0.89
Gym	0.6	0.5	0.55	0.5	0.8	0.63	0.4	1	0.6
Organic	1	0.6	0.71	0.9	1	0.95	0.8	1	0.86
Seafood	1	0.3	0.4	0.9	0.5	0.62	0.7	0.8	0.75

**Figure 7.** Average precision, recall and F-measure.

5.2.2. Ranking of Recommendations

In this experiment, we tried to measure the quality of IALBR's venue ranking by comparing it with the ratings on Yelp. To conduct this experiment, we analyzed the resulting venues from the interest query "burger" and calculated the Mean Absolute Error (MAE), the Discounted Cumulative Gain (DCG) and the Normalized Discounted Cumulative Gain (NDCG) of those venues relative to Yelp. These measures are defined as:

$$DCG = DG_1 + \sum_{i=2}^n \frac{rel_i}{\log_2(i)}, \quad IDCG = IDG_1 + \sum_{i=2}^n \frac{rel_i}{\log_2(i)}, \quad NDCG = \frac{DCG}{IDCG} \quad (10)$$

where rel is the graded relevance of the result at position i . To be able to do this comparison, we had to transform IALBR's scores P into a five-star scale similar to that of Yelp. We adjusted the scale so that it will give a score of five to the venue having the highest p for the interest tag "burger" in our entire dataset and will give a score of one to the venue having the lowest p for the same interest tag.

Another problem we faced in this comparison is that Yelp ratings reflect the overall rating of the venue, not the rating of the interest (e.g., burger) within this venue. IALBR, on the other hand, rates the interests (e.g., burger) in the venue. Therefore, we had to make an assumption that if a venue is popular for serving a specific item, this venue is likely to have a good rating on Yelp. This is why we used Yelp ratings as a relatively acceptable ground truth. Moreover, we could not find an alternative platform that provides item-based ranking.

Table 3 lists IALBR venue results of the interest query “burger” and compares their ratings (after scaling the power scores p) with the venue ratings on Yelp. We achieved an MAE of 1.3. It is the square root of the average of the right-most column (absolute error). A smaller MEA indicates a smaller deviation between our method and the ground truth. Tables 4 and 5 show the DCG and the NDCG values for IALBR, respectively. In Table 4, the venues were ranked by their IALBR p scores, and their Discounted Cumulative Gain (DCG) was calculated using the Yelp rating in Equation (10). In Table 5, we ranked the venues based on their Yelp ratings assuming that this is the ideal ranking, and their Ideal Discounted cumulative Gain (IDG) was calculated using the reordered Yelp ratings, then the Normalized Discounted Cumulative Gain (NDCG) was calculated. The achieved NDCG value (0.99) indicates that the positioning/ranking of the venues based on IALBR’s power scores p is very close to the positioning of venues based on the Yelp rating.

Table 3. IALBR mean absolute error (MAE).

Venue	IALBR P	IALBR Rating	Yelp Rating	Absolute Error
Shake Shack	220	5	4	1
Natureworks	182	4.5	4	0.5
HB burger	60	2	3.5	1.5
The Plaza Food Hall	48	1.5	4	2.5
Westway diner	48	1.5	3	1.5
Taproom	32	1.5	3.5	2
McDonalds	23	1	3	2
Pig N Whistle	15	1	3.5	2.5
Landmarc	11.6	1	3.5	2.5
Faces and Names	5	1	2.5	1.5
Mean Absolute Error				1.3

Table 4. IALBR Discounted Cumulative Gain (DCG).

Venue	Yelp Stars	DG	DCG
Shake Shack	4	4	4
Natureworks	4	4	8
HB burger	3.5	2.2	10.2
The Plaza Food Hall	4	2	12.2
Westway diner	3	1.3	13.5
Taproom	3.5	1.4	14.9
McDonalds	3	1.1	15.9
Pig N Whistle	3.5	1.2	17.1
Landmarc	3.5	1.1	18.2
Faces and Names	2.5	0.8	19

Table 5. IALBR Normalized Discounted Cumulative Gain (NDCG). IDG, Ideal Discounted cumulative Gain.

Venue	Yelp Stars	IDG	IDCG	NDCG
Shake Shack	4	4	4	1
Natureworks	4	4	8	1
The Plaza Food Hall	4	2.5	10.5	0.97
HB burger	3.5	1.8	12.3	0.99
Taproom	3.5	1.5	13.8	0.98
Pig N Whistle	3.5	1.4	15.1	0.98
Landmarc	3.5	1.3	16.4	0.97
Westway diner	3	1	17.4	0.98
McDonalds	3	1	18.4	0.99
Faces and Names	2.5	0.8	19.1	0.99

6. Discussion

In this section, we analyze the IALBR's scoring approach explained in Section 3. This is done by comparing the effectiveness of each component independently. We compare the results of a user query “burger” using four scenarios: (1) LBSN resources alone; (2) LBSN and SN resources together; (3) HITS applied on LBSN; and (4) IALBR, which combines the three. As shown in Table 6, the different scenarios demonstrate how each component by itself is not enough and how effective they became when combined in IALBR.

Table 6. Different scenarios of recommendation.

LBSN	LBSN and SN	HITS	IALBR
Shake Shack	Shake Shack	New York Hilton Midtown	Shake Shack
Landmarc	HB Burger	Natureworks	Nature Works
Taproom	McDonalds	Bierhaus	HB Burger
HB Burger	The Plaza Food Hall	The Plaza Food Hall	The Plaza Food Hall
	Westway Diner	Westway Diner	Westway Diner
	Nature Works	Taproom	Taproom
	Taproom	Landmarc	McDonalds
	Pig N Whistle	Pig N Whistle	Pig N Whistle
	Landmarc	McDonalds	Landmarc
	Faces and Names	HB Burger	Faces and Names

1. Scenario 1, LBSNs:

Let us assume that a Foursquare user is looking for a place to eat a “burger” within a 20-km geo-range. The only way Foursquare can answer this query is to look up the word “burger” in the tags of venues within this geo-range. In our experiment, four recommendations were returned, which demonstrates that LBSN alone offers fewer recommendations. For the same query, IALBR discovered new relevant options with even better p scores (e.g., Nature Works and The Plaza Food Hall).

2. Scenario 2, LBSNs and SNs:

A user is using an LBSN that exploits SN's geo-tagged media to map interests into venues. This scenario does not employ the HITS methodology for inferring expert users and interesting locations. The user is looking for a place to eat a “burger” within a 20-km geo-range. It recommends venues that frequently have the query item on both the geo-tagged media and the LBSN's tips and tags. Running this query, we were able to recommend 10 places sorted using the average weighted score s explained in Section 3.2, but without taking into consideration if those places are “interesting places” or are usually visited by “expert” users. Comparing these results with the IALBR results, we find that IALBR not only recommends places that have the query interest among the top interests, but also should be quality places (i.e., they are visited frequently by expert users). For example: McDonald's is famous for burgers, so it is likely to appear with a high rank in this scenario. It is however known for medium food quality and not usually visited by expert users, giving it a low authority score. Since the IALBR approach takes into consideration the authority score of the venue, McDonald's gets a lower rank. This shows that IALBR makes the right trade-off between expert's opinion and public opinion regarding a query interest.

3. Scenario 3, HITS:

This is a user who is using an LBSN that exploits SN's geo-tagged media to map interests into venues. It employs the HITS model to infer expert users and interesting locations. If the user is looking for a place to eat a “burger” within a 20-km geo-range, this system recommends interesting “venues”, which also contain the user's interest “burger”. Simulating this system,

we were able to recommend 10 places sorted using the venue authority score $a(v)$ explained in Section 3.3. Comparing the results of this scenario with the results of IALBR, we found that HITS alone provides venues that have very high scores as “interesting” venues visited by “experts” and that include the query item among their tags. The result, however, includes venues that are not of much relevance. For example: Using HITS alone brought “New York Hilton Midtown” as the first option for burger. It is true that you can have a great burger there, but this recommendation was based on the popularity of the venue, not on the popularity of the interest item (i.e., burger) in the venue.

7. Conclusions

This paper presents an Interest-Aware Location-Based Recommender system (IALBR), which uses geo-tagged social media content to provide venue recommendations to a user query at a specific position within a predetermined geo-range. The recommendations are based on the popularity of the query item in the venue and the popularity of this place calculated using the HITS-based inference model. This kind of recommendation supports high-level queries (e.g., bakery) and low-level queries (e.g., bagel). It leverages the huge amount of geo-tagged content existing in social media in order to extract the most frequent ones and uses them for location recommendation and interest matching. We evaluated our proposed system using multiple experiments based on data collected from Foursquare, having 35,023 venues, 10,377 tips and 392,247 check-ins generated from 17,463 users in NYC. For the venues in this dataset, we crawled 641,079 geo-tagged media items from Instagram to extract 217,511 tags representing interests associated with those venues. Foursquare and Instagram are meant to be examples for LBSNs and SNs, without loss of generality, because many other LBSNs and SNs share similar features. The main contribution of our approach lies in: (1) providing interest-based venue recommendations having the ability to respond to interests of different types (e.g., activity or food item) and of varying abstraction levels (e.g., generic food category or specific food item); (2) the information about venues in a given LBSN is enriched with geo-tagged content collected from a photo-sharing social network.

In terms of future work, we seek to handle complex queries (e.g., question answering, interest synonyms, etc.) as our implementation handles so far only one-word queries. We also intend to join LBSNs with SNs without being limited to pre-existing unique keys between the LBSN (Foursquare) and the SN (Instagram) for this join. This will significantly enrich the locations by including other geo-referenced information in other LBSNs that do not share same keys (e.g., flicker, twitter, etc.). Furthermore, we need to work on identifying the polarity (positive, neutral, negative) of the venue tips and geo-tagged media that contain the queried interest, before calculating the average weighted score of that interest at that venue. Another future research direction includes making the IALBR approach a self-learning approach, i.e., users will rate recommendations given by the system, and this rating will be used in interest scoring.

Author Contributions: B.A. identified the problem and the casual idea of the IALSBN system design. She also designed and implemented the programs for the experiments. M.S. and S.M. formalized the problem and the solution, and devised the IALBSN architecture. B.A. in collaboration with M.S. wrote the manuscript with the review and input of S.S. and I.M.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Radicati, S.; Hoang, Q. *Email Statistics Report, 2011–2015*; The Radicati Group, Inc.: Palo Alto, CA, USA, 2011.
2. Bao, J.; Zheng, Y.; Wilkie, D.; Mokbel, M.F. A survey on recommendations in location-based social networks. *ACM Trans. Intell. Syst. Technol.* **2013**, *19*, 525–565.
3. Zheng, Y. Tutorial on location-based social networks. In Proceedings of the 21st International Conference on World Wide Web (WWW), Lyon, France, 16–20 April 2012.

4. Constine, J. Instagram Hits 300 Million Monthly Users to Surpass Twitter, Keeps It Real With Verified Badges, 2014. Available online: <http://techcrunch.com/2014/12/10/not-a-fad/> (accessed on 6 December 2016).
5. D'Onfro, J. This Chart Shows How Instagram Reached 150 Million Users in Half the Time of Twitter, 2014. Available online: <http://www.businessinsider.com/instagram-growth-chart-2014-2> (accessed on 6 December 2016).
6. Smith, C. Foursquare's New Big Data Initiative Is Going to Help It Thrive, Even as the Check-in Withers, 2014. Available online: <http://www.businessinsider.com/foursquare-surpasses-45-million-registered-users-and-begins-collecting-data-in-new-ways-2-2014-1> (accessed on 6 December 2016).
7. Number of Monthly Active Instagram Users From January 2013 to June 2016 (in Millions), 2016. Available online: <https://www.statista.com/statistics/253577/number-of-monthly-active-instagram-users/> (accessed on 6 December 2016).
8. Weber, J.N.H. Foursquare by the Numbers: 60 m Registered Users, 50 m Maus, and 75 m Tips to Date, 2015. Available online: <http://venturebeat.com/2015/08/18/foursquare-by-the-numbers-60m-registered-users-50m-maus-and-75m-tips-to-date/> (accessed on 6 December 2016).
9. Wilhelm, A. What Instagram's Crazy Facebook Buyout Tells Us About the Value of Foursquare, 2012. Available online: <http://thenextweb.com/insider/2012/04/13/what-instagrams-crazy-facebook-buyout-tells-us-about-the-value-of-foursquare/> (accessed on 6 December 2016).
10. GWI Brand Summary q1 2014. Available online: http://www.media2000.it/wp-content/uploads/2014/04/GWI_Brand_Summary.pdf (accessed on 6 December 2016).
11. Park, M.H.; Hong, J.H.; Cho, S.B. Location-based recommendation system using bayesian user's preference model in mobile devices. In *Ubiquitous Intelligence and Computing*; Springer: Berlin/Heidelberg, Germany, 2007; pp. 1130–1139.
12. Borzsonyi, S.; Kossmann, D.; Stocker, K. The skyline operator. In Proceedings of the IEEE 17th International Conference on Data Engineering, Washington, DC, USA, 2–6 April 2001.
13. Ramaswamy, L.; Deepak, P.; Polavarapu, R.; Gunasekera, K.; Garg, D.; Visweswarah, K.; Kalyanaraman, S. Caesar: A context-aware, social recommender system for low-end mobile devices. In Proceedings of the 10th International Conference on Mobile Data Management: Systems, Services and Middleware (MDM'09), Taipei, Taiwan, 18–21 May 2009; pp. 338–347.
14. Lu, C.T.; Lei, P.R.; Peng, W.C.; Su, J. A framework of mining semantic regions from trajectories. In *Database Systems for Advanced Applications*; Springer: Berlin/Heidelberg, Germany, 2011; pp. 193–207.
15. Ye, M.; Janowicz, K.; Mülligann, C.; Lee, W. 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, Chicago, IL, USA, 1–4 November 2011; pp. 102–111.
16. Ye, M.; Shou, D.; Lee, W.; 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 (KDD'11), San Diego, CA, USA, 21–24 August 2011.
17. 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, San Jose, CA, USA, 3–5 November 2010; pp. 31–38.
18. Horozov, T.; Narasimhan, N.; Vasudevan, V. Using location for personalized poi recommendations in mobile environments. In Proceedings of the 2006 International Symposium on Applications and the Internet (SAINT 2006), Phoenix, AZ, USA, 23–27 January 2006.
19. Del Prete, L.; Capra, L. diffeRS: A mobile recommender service. In Proceedings of the 2010 11th International Conference on Mobile Data Management (MDM), Kansas City, MO, USA, 23–26 May 2010; pp. 21–26.
20. 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, San Jose, CA, USA, 3–5 November 2010; ACM: New York, NY, USA, 2010; pp. 458–461.

21. Bao, J.; Zheng, Y.; Mokbel, M.F. Location-based and preference-aware recommendation using sparse geo-social networking data. In Proceedings of the 20th International Conference on Advances in Geographic Information Systems, Redondo Beach, CA, USA, 7–9 November 2012; pp. 199–208.
22. Sarwat, M.; Bao, J.; Eldawy, A.; Levandoski, J.J.; Magdy, A.; Mokbel, M.F. Sindbad: A location-based social networking system. In Proceedings of the 2012 ACM SIGMOD International Conference on Management of Data, Scottsdale, AZ, USA, 20–24 May 2012; pp. 649–652.
23. Noulas, A.; Scellato, S.; Lathia, N.; Mascolo, C. A random walk around the city: New venue recommendation in location-based social networks. In Proceedings of the 2012 International Conference on Privacy, Security, Risk and Trust (PASSAT) and 2012 International Conference on Social Computing (SocialCom), Amsterdam, The Netherlands, 3–5 September 2012; pp. 144–153.
24. Cheng, C.; Yang, H.; King, I.; Lyu, M.R. Fused matrix factorization with geographical and social influence in location-based social networks. In Proceeding of the 26th AAAI Conference on Artificial Intelligence (AAAI'12), Toronto, ON, Canada, 22–26 July 2012; Volume 12, pp. 17–23.
25. Lian, D.; Zhao, C.; Xie, X.; Sun, G.; Chen, E.; Rui, Y. GeoMF: joint geographical modeling and matrix factorization for point-of-interest recommendation. In Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, New York, NY, USA, 24–27 August 2014; pp. 831–840.
26. Zhang, J.-D.; Chow, C.-Y. Spatiotemporal sequential influence modeling for location recommendations: A gravity-based approach. *ACM Trans. Intell. Syst. Technol. (TIST)* **2015**, *7*, 11.
27. Leung, K.W.-T.; Lee, D.L.; Lee, W.-C. 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 Retrieval, Beijing, China, 24–28 July 2011.
28. Takeuchi, Y.; Sugimoto, M. CityVoyager: An outdoor recommendation system based on user location history. In *Ubiquitous Intelligence and Computing SE-64*; Ma, J., Jin, H., Yang, L., Tsai, J.J.-P., Eds.; Springer: Berlin/Heidelberg, Germany, 2006; pp. 625–636.
29. Zheng, Y.; Zhang, L.; Xie, X.; Ma, W.-Y. Mining interesting locations and travel sequences from gps trajectories. In Proceedings of the 18th International Conference on World Wide Web, Madrid, Spain, 20–24 April 2009; pp. 791–800.
30. Zheng, Y.; Zhang, L.; Xie, X.; Ma, W.-Y. Mining Correlation Between Locations Using Human Location History. In Proceedings of the 17th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems, Seattle, WA, USA, 4–6 November 2009; ACM: New York, NY, USA; pp. 472–475.
31. Lian, D.; Xie, X. Learning location naming from user check-in histories. In Proceedings of the 19th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems, Chicago, IL, USA, 1–4 November 2011.
32. Zheng, Y.; Xie, X. Learning travel recommendations from user-generated GPS traces. *ACM Trans. Intell. Syst. Technol.* **2011**, *2*, 1–29.
33. Zheng, V.W.; Zheng, Y.; Xie, X.; Yang, Q. Collaborative location and activity recommendations with GPS history data. In Proceedings of the 19th International Conference on World Wide Web, Raleigh, NC, USA, 26–30 April 2010; pp. 1029–1038.
34. Yuan, Q.; Cong, G.; Sun, A. Graph-based point-of-interest recommendation with geographical and temporal influences. In Proceedings of the 23rd ACM International Conference on Conference on Information and Knowledge Management, Shanghai, China, 3–7 November 2014; pp. 659–668.
35. Zhang, J.-D.; Chow, C.-Y.; Zheng, Y. ORec: An opinion-based point-of-interest recommendation framework. In Proceedings of the 24th ACM International on Conference on Information and Knowledge Management, Melbourne, Australia, 19–23 October 2015; pp. 1641–1650.
36. Lian, D.; Ge, Y.; Zhang, F.; Yuan, N.J.; Xie, X.; Zhou, T.; Rui, A. Content-aware collaborative filtering for location recommendation based on human mobility data. In Proceedings of the 2015 IEEE International Conference on Data Mining (ICDM), Atlantic City, NJ, USA, 14–17 November 2015; pp. 261–270.
37. Wang, W.; Yin, H.; Chen, L.; Sun, Y.; Sadiq, S.; Zhou, X. Geo-SAGE: A geographical sparse additive generative model for spatial item recommendation. In Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Sydney, Australia, 10–13 August 2015; pp. 1255–1264.
38. Majid, A.; Chen, L.; Chen, G.; Mirza, H.T.; Hussain, I.; Woodward, J. A context-aware personalized travel recommendation system based on geotagged social media data mining. *Int. J. Geogr. Inform. Sci.* **2013**, *27*, 662–684.

39. Memon, I.; Chen, L.; Majid, A.; Lv, M.; Hussain, I.; Chen, G. Travel recommendation using geo-tagged photos in social media for tourist. *Wirel. Pers. Commun.* **2015**, *80*, 1347–1362.
40. 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, Chicago, IL, USA, 11–14 August 2013; pp. 1043–1051.
41. Gao, H.; Tang, J.; Hu, X.; Liu, H. Content-aware point of interest recommendation on location-based social networks. In Proceedings of the 29th AAAI Conference on Artificial Intelligence (AAAI'15), Austin, TX, USA, 25–30 January 2015; pp. 1721–1727.
42. Sun, Y.; Fan, H.; Bakillah, M.; Zipf, A. Road-based travel recommendation using geo-tagged images. *Comput. Environ. Urban Syst.* **2015**, *53*, 110–122.
43. Kurashima, T.; Iwata, T.; Irie, G.; Fujimura, K. Travel route recommendation using geotagged photos. *Knowl. Inform. Syst.* **2013**, *37*, 37–60.
44. HosseiniMardi, H.; Li, S.; Yang, Z.; Lv, Q.; Rafiq, R.I.; Han, R.; Mishra, S. A comparison of common users across Instagram and Ask. fm to better understand cyberbullying. In Proceedings of the 2014 IEEE 4th International Conference on Big Data and Cloud Computing (BdCloud), Sydney, Australia, 3–5 December 2014; pp. 355–362.
45. Yang, D.; Zhang, D.; Qu, B. Participatory cultural mapping based on collective behavior data in location based social networks. *ACM Trans. Intell. Syst. Technol.* **2015**, *7*, 30.
46. Yang, D.; Zhang, D.; Chen, L.; Qu, B. NationTelescope: Monitoring and visualizing large-scale collective behavior in lbsns. *J. Netw. Comput. Appl.* **2015**, *55*, 170–180.



© 2016 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC-BY) license (<http://creativecommons.org/licenses/by/4.0/>).