DBCACF: A Multidimensional Method for Tourist Recommendation Based on Users' Demographic, Context and Feedback

Maral Kolahkaj
Department of Computer Engineering, Karaj Branch, Islamic Azad University, Karaj, Iran
maral.kolahkaj@kiau.ac.ir
Ali Harounabadi *
Department of Computer Engineering, Central Tehran branch, Islamic Azad University, Tehran, Iran
a.harounabadi@iauctb.ac.ir
Alireza Nikravanshalmani
Department of Computer Engineering, Karaj Branch, Islamic Azad University, Karaj, Iran
nikravan@kiau.ac.ir
Rahim Chinipardaz
Shahid Chamran University of Ahvaz, Ahvaz, Iran
chinipardaz_r@scu.ac.ir

Received: 28/Jan/2019 Revised: 20/Apr/2019 Accepted: 06/May/2019

Abstract

By the advent of applications in the web 2.0 such as social networks which allow the users to share media, numerous opportunities have been provided for the tourists to recognize and visit attractive and unfamiliar Areas-of-Interest (AOIs); however, finding the appropriate areas based on user's preferences is very difficult due to several issues such as huge amount of tourist areas, the limitation of the visiting time, etc. In addition, the available methods have yet failed to provide accurate tourist's recommendations based on geo-tagged media because of several problems such as considering two users with different habits as the same, and ignoring user's information. Therefore, in this paper, a method called "Demographic-Based Context-Aware Collaborative Filtering" (DBCACF) is proposed to investigate the mentioned problems. DBCACF considers personal and side information in combination with the users' feedbacks to overcome the limitations of collaborative filtering methods in dealing with multi-dimensional data. In addition, a new asymmetric similarity measure is proposed in order to overcome the limitations of symmetric similarity methods. The experimental results on Flickr dataset indicated that the use of personal and side information and the addition of proposed asymmetric scheme to the similarity measure could significantly improve the obtained results compared to other methods which used only user-item ratings and symmetric measures. In particular, our method based on the Cosine similarity measurement has provided a better performance (0.34 for Precision and 0.38 for F-score) as compared to our method based on the Pearson similarity measure over data sparsity and cold-start problems.

Keywords: Decision Support Systems; Data Mining; Context-aware Recommendation; Geo-tagged Photo; Asymmetric Similarity.

1. Introduction

By the advent of the web and its relevant applications, new opportunities have been provided for tourists that they can use travel books, personal blogs, and online services such as travel guides, maps and the like [1]; however, the tourists are dealing with problems to find out relevant information about their request by the exponential data growth in the web called "information overload" [2]. Thus, the users have to choose their favorite items from billions of the objects on the web. Obviously, the evaluation of all of these items is impossible by a user [3]. Therefore, the Recommender System (RS) as a type of Decision Support Systems (DSSs) is designed to filter the information overload and present the relevant results. The RS aims to predict and to recommend the ratings and items which the users are interested in visiting [4-7].

Today, in tourism industry, the users prefer a system which is able to recommend the tourism areas based on their own interests. Several RSs have become increasingly popular to present the tourism services such as Tripadvisor, travelblog, and so on. Although the RSs help the users to deal with the information overload, most existing RSs apply the Collaborative Filtering (CF) and Content-Based (CB) methods to compute the preference between a user and an unvisited item [8]. In location RSs the user-location matrix is highly sparse with numerous missing entries, because users have only visited a very small proportion of attractive areas. These methods are suffering from several problems such as low quality recommendations, low accuracy recommendations and unreliability problems [9,10] due to using only two dimensional data (user-item) [11], considering two users with different habits as the same, data sparsity, cold start condition, lack of the benefit data which are utilized in

* Corresponding Author

recommendation process, etc. [12]. The cold-start problem occurs due to an initial lack of ratings for new users who have not rated any item or new items which have not been rated by any user; therefore, it becomes impossible to make reliable recommendations. On the other hand, sparsity problem occurs when the number of users who have rated items is too small compared to the number of items, hence the recommender system cannot generate any recommendations if there is no overlap in ratings with the target user.

Social RSs are a novel generation of recommenders which utilizes the social media and the interaction data among the users to present the recommendations [13]. The location-based social networks (LBSNs) with the possibility of sharing geo-tagged photos are the examples of networks that the tourists can use to share their travel experiences by uploading photos, providing ratings and comments [10,14,15]. By considering the social data on the social networks, new challenges are also emerged, such as what data is useful for the recommendation process and how to handle this data in order to provide the relevant and accurate recommendations [12,16]. In LBSNs, locations are encoded with latitude and longitude, which distinguishes locations form other items, such as books, music and movies in conventional recommender systems [8,10]. Often, the previous methods which were based on analyzing travel reports and geo-tagged photos, mined popularity areas and do not consider the users' appropriate attributes and contexts on the social network [17]. Therefore, the aim of this study is to propose Demographic- Based Context- Aware Collaborative Filtering (DBCACF) method by using a weighted hybrid method. Based on DBCACF, three RSs are combined to design a tourist recommender system including Collaborative Filtering RS, Demographic RS and Context-Aware RS. The main benefit of combination of these RSs is to take advantage of each particular RS while overcoming limitations of individual RSs. For example, user's demographic and context information is used to overcome data sparsity and cold start problems. Also, to overcome the limitations of methods that used two dimensional data (user-item) and to obtain more accurate and relevant recommendations, this study proposes a method to provide suggestion to the users by handling multi-dimensional data. Furthermore, an asymmetric similarity measure is proposed to imply difference between two different users. Though the previous studies have used various techniques in their recommender systems, the novelty of this work is in integrating Collaborative Filtering, Context-Aware and Demographic as a hybrid recommender system. The main contributions of this paper are as follows:

- Utilizing the explicit users' demographic information in tourism RS based on geo-tagged photos.
- Combining demographic information with the contextual information in tourism RS based on geo-tagged photos to address the cold-start problem and to alleviate sparsity problem.
- Presenting a novel asymmetric similarity measure which determines the appropriate user's neighbors and accordingly recommends relevant and accurate suggestions to the user.
- Conducting extensive experiments to evaluate the performance of DBCACF on Flickr data set.

The organization of this paper is as follows: Section 2 explains the related works in the field of RSs and tourism recommendation. Section 3 presents the proposed method. Section 4 discusses the implementation and evaluations of the proposed method. Finally, the conclusion and future works are considered in Section 5.

2. Related Work

In social media websites which are able to share photos and videos, the tourists participate in sharing the geo-tagged photos [15]. These social media play an increasingly important role as information sources for travelers [18]. On the other hand, the users are interested in searching for the attraction places [19]. In this case, the time-based and the location-based data from social media can be used to provide the assorted recommendations [3,20]. A large body of research was conducted to present different RS methods [4,9,21-30]. The important related methods to this study that can be used in the tourism based on the geo-tagged photos are discussed as follows.

2.1 Symmetric Similarity Methods

CF methods are the dominant methods in RS algorithms such as user-based and item-based approaches [31]. These methods improve the recommendation process by considering the users' similar interests and reduce the overspecialization problem [32]. Despite the success of the CF methods, the performance of these methods is strongly influenced by data sparsity and cold start problems due to the numerous items on the web [25]. A large number of studies have presented several similarity measures to obtain the nearest neighbors for the user [33-41]. In memory-based collaborative filtering, the traditional similarity functions can be used, such as Cosine [42], the Euclidean distance, the mean square difference [43], symmetric Kullback-Leibler divergence [15], Pearson correlation [44,45], to calculate the similarity measure between two tourists. These functions calculate the similarity measure as symmetrical term and thus the users have the same effect on each other for receiving recommendations. In the methods which only considered the ratings of common items, the ratings of uncommon items are ignored in calculating the symmetric similarity between users. As mentioned, these methods are unreliable in cold start and data sparsity problems, because of a low number of common items. Therefore, these problems can reduce the quality and accuracy of the recommendations.

2.2 Asymmetric Similarity Methods

The similarity between two users may be asymmetric and the impact of the first user to the second user is different and vice versa in the real world. In order to achieve a more accurate similarity in CF methods, asymmetric weights are assigned for traditional similarities. For example, Pirasteh et al. presented a weighting scheme in which symmetric similarity became asymmetric similarity among the users by considering the fraction of common items in target user items [43,46,47]. In addition, in another study, an asymmetric similarity which used all of the users' ratings was provided [48]. The similarity measure utilized the local similarity obtained from the Pearson correlation between two users' ratings and global similarity extracted from the Bhattacharyya similarity between each pair of the items. Finally, Jaccard similarity between the two users was combined with the global and local similarities in order to provide more significance to the common items. Although these methods have a higher performance in comparison with previous subsection methods, these are still not efficient in dealing with cold start and data sparsity problems due to only use of the user-item ratings.

2.3 Contexts-Aware Methods

The Context-Aware RS considers a diversity of contextual information on the recommendation process such as time, location or social data [49]. These side information can be incorporated in the recommendation process by three approach: pre-filtering, post-filtering, and contextual modeling [50]. Several methods were recommended based on the contextual information such as time and weather [44,51]. In these methods, the pre-filtering approach was used to utilize the contextual information. Despite of the simplicity of the pre-filtering and post-filtering, these approaches eliminate much appropriate data before the recommendation process. Therefore, in the proposed method a contextual modeling approach is utilized in order to incorporate the contextual information in the recommendation process.

2.4 Hybrid Methods

As mentioned, each of the previous recommendation methods utilizes the certain information and they have several advantages and disadvantages to provide the suggestions. Therefore, two or more RS methods can be combined to use multiple resources and utilize different RSs' advantages [24]. For example, CJacMD [52] is regarded as a combination of cosine, Jaccard and mean measure divergence. In CJacMD approach, the users' ratings and the rating habits of individuals were considered to express preferences.

Table 1 displays a comparative review of advantages and disadvantages of the related works.

Table 1. Recommendation methods

Ref.	Evaluation Metric	Data-set	Advantages	Disadvantages
[15]	Precision, Mean Average Precision@50, Blended Ratio	Flickr	Considering time constraint, personalization based on user history	Ignoring context, Applying symmetric measure, Low accuracy, Explicit questions from users for time and cost
[44]	Precision, Mean Average Precision @50, Blended Ratio	Flickr	Considering context	Pre-filtering context, Symmetric similarity, Low accuracy, Ignoring time constraint
[53]	Precision	Flickr	Considering context	Low accuracy, Only evaluate precision, Pre-filtering context, Ignoring time constraint
[51]	Precision, Mean Average Precision @50, Edit Distance, n Discounted Cumulative Gain@5,@10	Flickr	Considering context	No personalization, Low accuracy, Pre-filtering context, Using probability in recommendations Ignoring time constraint
[43]	Root Mean Squared Error	MovieLens DOUBAN	Improving the traditional similarity by weighted schemes	High error value, Ignoring context, Only evaluate RMSE, Unreliable in term of Cold-Start
[47]	Root Mean Squared Error, Mean Absolute Error	MovieLens NetFlix	Improving the traditional similarity by weighted schemes, Calculating the similarity for two users with no common items	High error value, Only evaluate Errors, Ignoring context
[48]	Root Mean Squared Error, Mean Absolute Error, Precision, Recall, F1	MovieLens NetFlix Yahoo music	Calculating the similarity for two users with no common items	Having high complexity in term of time, Ignoring context
[52]	Root Mean Squared Error, Mean Absolute Error	MovieLens 100K MovieLens 1M Each movie	Considering users ratings habit	Only evaluate Errors, Is Not stable in results, Ignoring context
[45]	Root Mean Squared Error, Coverage	Movie, Food	Weighting context	Having high complexity in term of time

The aim of this study is to combine three RS methods for utilizing their advantages, and therefore, obtaining more relevant and accurate recommendations, including: a) collaborative filtering method for utilizing the interests of similar users and reducing the overspecialization problem, b) demographic based method for overcoming the cold start problem, and c) context aware method for dealing with data sparsity and cold start problems. The time and location as contexts information, and age and gender as demographic information are used to manage the users' travel and to save the time and cost for the users. To utilize a demographic RS, data from user profiles is used, and for using a context-aware system, the data from the photos shared by users is used. This information is implicitly extracted and the recommendations are also recommended to the users without their explicit request. Another advantage of the proposed method is to integrate the context and demographic information into the recommendation process in comparison with other travel recommendation methods.

3. Proposed Method

The present study seeks to propose a new personalized tourist RS called "DBCACF" based on geo-tagged photos and suggest the attractive areas for the users without any explicit requests. Therefore, the collection of community contributed geo-tagged photos, the user demographic attributes, contextual and collective information are utilized to suggest the personalized area recommendation. In demographic-based method, no ratings are necessary to create the user profile in spite of collaborative and content-based methods, and thus it is more efficient for dealing with cold-start problems. In addition, DBCACF overcomes the limitations of conventional collaborative filtering methods. As mentioned, the recognition of the user's neighbors who have similar interests with the target user is known as an important principle in the performance of collaborative filtering. Thus, a new similarity measure is proposed to overcome other similarity measure's problems such as data sparsity, relying on low amount of co-rated items and ratings, depending on the users' explicit ratings, and calculating the similarity as symmetric measure. Fig. 1 illustrates the structure of the proposed method.

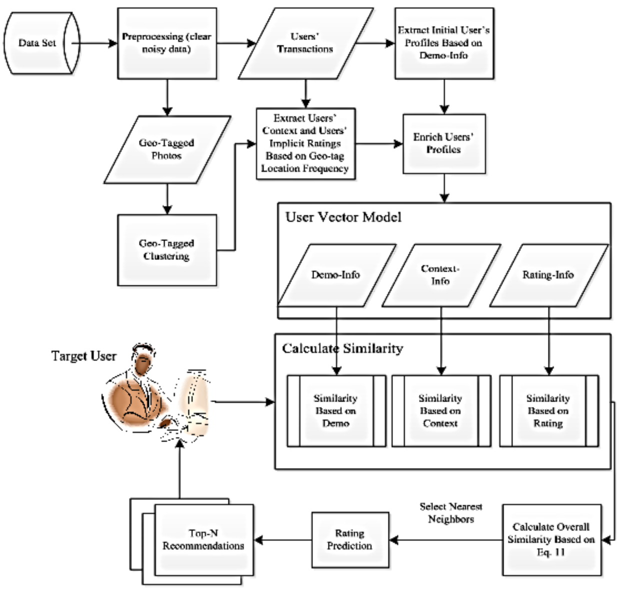

Fig. 1. Structure of proposed method

In short, after pre-processing phase, the users are identified by their ID, and their age and gender are extracted. Then, using geographic information, the photos are clustered to identify the areas of tourism. Using the geo-tagged areas, the user's interests are extracted. These phases are offline phase in the proposed method. In the online phase, the similarity between the new user and other users is calculated and the recommendations are provided to the target user. These steps are described below.

3.1 Profiling and Modeling User's Behavior

After eliminating noisy data, each user is identified based on his/her ID. The information like username, age, and gender, is used as a user vector in order to demonstrate the initial profile of each user.

Next, the geographical data were derived from geo-tagged photos' annotations and each photo is displayed as a photo vector like (photo ID, latitude, and longitude). The estimation of the geo-location of photos is regarded as a challenging task [54,55]. In this paper, the photos were clustered according to their geographical locations and using a density based clustering algorithm called DBSCAN [56], in order to find areas of interest where a photo is taken. DBSCAN clustering requires the least knowledge to determine the input parameters and can discover clusters with each arbitrary shape. In addition, this algorithm can effectively filter noisy data in large datasets.

After obtaining the cluster of each photo, the areas and their visiting time are determined based on the user's historical visiting. In fact, photos with their contextual information such as time and location trace the users who have taken the photos and record the users' temporal-spatial movement. Therefore, in order to take advantage of the user's visiting time and location information, the vector of each user is enriched by (username, age, gender, AOI cluster, and visiting time). The date taken photos are used to determine the time factor and the season visited by tourist is regarded as time context. Next, the user's implicit rating, *r*, (in this study, user's preference) for each AOI was calculated based on the frequency of visiting the AOI by the user, due to the users' disinclination to provide an explicit rating, $r_{user,AOI} = freq_{user}(AOI)$.

Next, the profile of each user is revised and a hybrid vector model is extracted based on the user's preferences and profile. The final vector model for each user is represented as (username, age, gender, user's preference for each AOI, AOI cluster, visited season). AOI cluster and season represent the contexts information.

3.2 Calculating the Similarity

Social networks allow the users to share their favorite objects. Thus, a considerable number of users are participated for producing the data. Therefore, finding the similar users among all of the users in a social network is one of the most important steps in the proposed method. To this aim, the similarity between each user with other users was calculated by adopting a new similarity measure. The proposed measure combines the demographic and context information with the rating data in the process of calculating the similarity measure. The similarity based on demographic information is to deal with the cold start condition when no preference information is available for a given user. The similarity based on context information is also used to deal with the low number of co-rated or common items in data sparsity problem. Further, the similarity based on rating data is to utilize the gain of the collective wisdom of people. Accordingly, this hybrid similarity dominates the limitations of other similarities. Eq. (1) illustrates the initial proposed similarity measure between two users, u and v, $(Sim(u,v))$.

$$Sim(u,v) = \beta(Sim_{Demo}(u,v)) + (1-\beta)\big(Sim_{CACF}(u,v)\big) \quad (1)$$

where, Sim_{Demo} represents the similarity base on demographic features between the target user (u) and other users (v), Sim_{CACF} indicates the similarity based on context and collaborative features between u and v, and β is regarded as an adjustment factor for combining two measures. In fact, the Jelinek-Mercer smoothing is used for a linear combination of two similarity measures. The smoothing was implemented to adjust two statistics and avoid the possibility of Zero.

The number of the rated items (in this study, visited AOI) by target user u, I_u, is represented as Eq. (2) in order to overcome the symmetric similarity measure problem and prevent the sensitivity to a low number of co-rated items when a user has rated less items while another user has more rated items, and transform the symmetric similarity to asymmetric similarity.

$$Sim(u,v) = \frac{1}{\gamma|I_u|}[\beta(Sim_{Demo}(u,v)) + (1-\beta)(Sim_{CACF}(u,v))] \quad (2)$$

where γ represents an adjustment factor in different conditions (see Eq. (11)).

The single attribute approach is used to obtain the similarity based on demographic information. In other words, the difference of each user's attribute is first obtained and accordingly the overall similarity of demographic feature is calculated by utilizing the weighted average of differences. The Sim_{Demo} in Eq. (2) is calculated as Eq. (3).

$$Sim_{Demo}(u,v) = \frac{1}{1+\frac{1}{N}\sum_{i=1}^{N} d_i(\alpha_u,\alpha_v)^2}. \quad (3)$$

where N represents the number of demographic feature, α_u, α_v indicates the set of the user's demographic features and $d_i(\alpha_u,\alpha_v)$ displays the difference of user u and v on the attribute i. The difference is obtained as Eq. (4).

$$d_i(\alpha_u,\alpha_v) = \begin{cases} 1 - OM(\alpha_{iu},\alpha_{iv}) & if \quad \alpha_i = nominal\ value \\ \frac{|\alpha_{iu}-\alpha_{iv}|}{max(\alpha)_i - min(\alpha)_i} & if \quad \alpha_i = Interval\ value \end{cases} . \quad (4)$$

where OM represents the overlap measure when the attribute type is nominal. The normalized Manhattan distance [57] is used when the attribute type is interval. $max(\alpha)_i and\ min(\alpha)_i$ indicate the maximum and minimum values for the attribute i, respectively.

In order to obtain the context-aware collaborative filtering, the contextual modeling approach is utilized as Eq. (5) instead of pre-filtering or post-filtering approaches.

$$Sim_{CACF} = Sim_{CA} \times Sim_{CF}. \quad (5)$$

Based on this equation, the similarity obtained from the user community, Sim_{CF}, has combined with the similarity of contexts, Sim_{CA}. In other words, the context-based similarity is multiplied with Sim_{CF} as a weight. The similarity based on contexts between two users' contexts (C_u, C_v) is calculated based on Jaccard measure (Eq. (6)).

$$Sim_{CA}(C_u,C_v) = \frac{|C_u \cap C_v|}{|C_u \cup C_v|}. \quad (6)$$

According to Eq. (6) and its combination with the similarity of CF, the similar context is not completely ignored and the users can be considered as similar by less weight in the future predictions if the conditions cannot be considered by the two users as similar based on whole matching the contexts.

The similarity-based collaborative filtering is accessed by Eq. (7) as cosine similarity and Eq. (8) as Pearson correlation similarity in order to calculate the similarity between two users' preference vectors (V_{r_u}, V_{r_v}).

$$Sim_{CF}^{COS}(u,v) = \frac{V_{r_u} \cdot V_{r_v}}{\|V_{r_u}\| \cdot \|V_{r_v}\|} = \frac{\sum_{i\in I}(r_{u,i})\cdot(r_{v,i})}{\sqrt{\sum_{i\in I}(r_{u,i})^2 \cdot \sum_{i\in I}(r_{v,i})^2}}. \quad (7)$$

$$Sim_{CF}^{PCC}(u,v) = \frac{\sum_{i\in I}(r_{u,i}-\bar{r}_u)\cdot(r_{v,i}-\bar{r}_v)}{\sqrt{\sum_{i\in I}(r_{u,i}-\bar{r}_u)^2} \cdot \sqrt{\sum_{i\in I}(r_{v,i}-\bar{r}_v)^2}}. \quad (8)$$

where, I is a set of AOIs where users, u and v, have visited.

If two users share more common contexts, their similarity collaborative weight will be more by multiplying the similarity based on collaborative and context together. In addition, the Sorensen index (SRS in Eq. (9)) is added to Eq. (2) in order to provide more significance for the users with more common items and demographic information (Eq. (10)).

$$SRS = \frac{2|I_u \cap I_v|}{|I_u| + |I_v|}. \quad (9)$$

where I_u, I_v are the number of visited AOI by target user and other users respectively.

$$Sim_{total}(u,v) = SRS + [\beta(Sim_{Demo}(u,v)) + (1-\beta)(Sim_{CACF}(u,v))]. \quad (10)$$

According to different terms and different similarities between two users, the total DBCACF similarity is defined as Eq. (11).

In this equation, the total similarity between the users is zero when they have no similarity in any respects. The overall similarity between two users is calculated by demographic information or CACF if the similarity based on CACF becomes zero, but the similarity based on demographic is nonzero (no common item or cold-start conditions), or if the similarity based on CACF becomes nonzero, but the similarity based on demographic is zero. In addition, the similarity needs to be an asymmetric measure by $\frac{1}{|I_u|}$ as an asymmetric weight. Finally, the total similarity between two users is calculated by both CACF and demographic information, if both similarities based on CACF and demographic become nonzero. Next, the existing similarity needs to be an asymmetric measure by the weight $\frac{1}{2|I_u|}$, which is added to the Sorensen index.

The obtained similarity measure can calculate the similarity in term of cold-start condition when a user has no co-rated items and suggest the recommendations to such users without relying on the co-rated items. On the other hand, the effect of similarity based on demographic is not completely delimited by increasing the co-rated and common items.

$$Sim_{DBCACF}(u,v) = \begin{cases} 0 & if \quad (Sim_{Demo}(u,v)) = 0, (Sim_{CACF}(u,v)) = 0 \\ \frac{1}{|I_u|}\left[\beta(Sim_{Demo}(u,v)) + (1-\beta)(Sim_{CACF}(u,v))\right] & if \; or \begin{matrix} (Sim_{Demo}(u,v)) = 0, (Sim_{CACF}(u,v)) \neq 0 \\ (Sim_{Demo}(u,v)) \neq 0, (Sim_{CACF}(u,v)) = 0 \end{matrix} \\ SRS + \frac{1}{2|I_u|}\left[\beta(Sim_{Demo}(u,v)) + (1-\beta)(Sim_{CACF}(u,v))\right] & if \quad (Sim_{Demo}(u,v)) \neq 0, (Sim_{CACF}(u,v)) \neq 0 \end{cases} \quad (11)$$

Next, by considering the obtained similarities, Top-*K* users with the most similarity measure with the target user as nearest neighbors are selected.

3.3 Prediction and Recommendation

In this step, the interesting areas are predicted for the current user based on his preferences, context, and personal features by considering the neighbors' interests. That is, after obtaining the nearest neighbors, the neighbor's items are extracted as candidate items. Next, the score of each candidate item is predicted for the target user by Eq. (12). Finally, the candidate items with the highest importance are suggested to the target user as a recommendation list.

$$pred(u,i) = \frac{\sum_{n \in neighbors(u)} Sim(u,n) \cdot (r_{n,i})}{\sum_{n \in neighbors(u)} Sim(u,n)} \quad (12)$$

4. Experiments and Evaluations

In this section, our proposed method is evaluated based on the Flickr dataset. Accordingly, the comparison results of the proposed method with other recommendation methods are presented.

4.1 Dataset Description

To demonstrate the effectiveness of our proposed method for recommendations, we conducted several experiments using Yahoo Flickr Creative Commons 100 Million Dataset (YFCC100M). The YFCC100M is the largest public multimedia collection ever released, with a total of 100 million media objects, of which approximately 99.2 million are photos and 0.8 million are videos, all uploaded to Flickr [58,59]. First, the metadata related to 16,533 geo-tagged photos were extracted. This data was captured in different cities of Iran between August 01, 1961 and March 28, 2014, and include several information such as the user ID, photo ID, geo-position, date taken and annotation tags. Then, the personal users' information such as, age and gender, was collected Crowley from Flickr social network. Table 2 represents the sample records which are used in this study. Photo ID and User ID were used anonymous for the purpose of privacy.

Table 2. Sample records of tourists and photos

Photo ID	User ID	Gender	Age	Date taken	Longitude	Latitude
1	1	Male	42	9/14/04 1:23 AM	51.460111	35.820523
2	2	-	-	10/14/04 11:38 PM	-	-
3	3	Male	40	5/1/04 4:07 PM	51.332931	35.729513
4	3	Male	40	5/2/04 4:13 PM	51.332931	35.729513

Next, the data with the following conditions was eliminated:

Those having no geographic information such as latitude and longitude,

Those without any user's profile data,

Those collected based on search result with the name of Iran in its metadata such as place, tags, description and title but its geographic information such as latitude and longitude failed to match the geographical context of the Iranian cities (Iran is located between 25-40 degrees of north latitude and between 44-64 degrees of east longitude).

Table 3 displays a descriptive statistic about the related dataset.

Table 3. Statistics on the dataset

Photos		Users	Locations
Raw	Filtered		
16533	10030	19	3873

4.2 Detecting Tourist Areas

In order to find the areas of interest, geographic information from photos was used based on the DBSCAN clustering method. For this study, the parameters of this algorithm are set as Min-Pts= 6 and ε=0.4. To this end, 58 area clusters were obtained and the vector of each photo was enriched as *(photo ID, latitude, longitude, cluster number)*.

4.3 Profiling Users

The user vector is shown as (*UID*, *Age*, *Gender*, P_1, P_2, …, P_n, C_1, C_2, …, C_m), where each $Pn \in$ Preferences represents the user preference about each AOI (implicit rating), and each $C_i \in C$ indicates the context which has been visited by a tourist, in other word if a user visits a context, the corresponding value for C_i will be 1, otherwise will be 0.

4.4 Prediction and Recommendation

The recommendation problem is defined as predicting ratings for the interesting objects that have not been seen by target user. In order to predict the tourist's AOIs, 75 % of the data was randomly selected for training and rest of the data was held for testing. The training set was utilized to predict the preferences to the target user by using each method. Accordingly, the data from the test set was used to evaluate the recommendation quality and the accuracy of prediction. Finally, a list of Top-*N* recommendation was matched with the actual list of visited areas by the target user, and the results of different methods were compared for final evaluation.

4.5 Experimental Evaluation

The performance metrics, compared methods, results and discussions to assess the proposed method are as follows.

4.5.1 Evaluation Metrics

Two statistical metrics were utilized to evaluate the prediction accuracy of the proposed DBCACFs and other methods. First, the Mean Absolute Error (MAE) is defined as Eq. (13).

$$\mathrm{MAE} = \frac{1}{|\mathrm{N}|}\sum_{\mathrm{u,i}}\left|\mathrm{r_{u,i}} - \mathrm{Pred_{u,i}}\right|. \quad (13)$$

Second, the Root Mean Squared Error (RMSE) is defined as Eq. (14).

$$\mathrm{RMSE} = \sqrt{\frac{1}{|\mathrm{N}|}\sum_{\mathrm{u,i}}(\mathrm{r_{u,i}} - \mathrm{Pred_{u,i}})^2} \quad (14)$$

where $r_{u,i}$ represents the actual implicit rating of the user u provided to item i, $Pred_{u,i}$ indicates the predicted rating of the user u imputed to item i by different methods, and $|N|$ displays the number of the test ratings. A smaller value of MAE or RMSE indicates the better prediction accuracy.

Further, three decision-support metrics including precision, recall and F-score were used to evaluate the prediction quality of the proposed methods. A larger value of precision, recall and F-score represents the better quality of prediction. Precision is regarded as the ability of RS for recommending the relevant suggestions and is interpreted as the fraction of the correct predictions in total number of predictions. Recall is considered as the ability of RS for recommending all of the suggestions which are visited by the target user. In fact, recall is interpreted as the fraction of correct predictions in total number of relevant items; however, F1-Measure can be interpreted as a weighted average of the precision and recall, where an F- Measure reaches the best and worst value at 1 and 0, respectively.

4.5.2 Compared Methods

CFcos is a Collaborative Filtering method (user-based kNN [61]). Cosine similarity is implemented to obtain similar users with target user in term of their ratings [34].

CFpcc is another Collaborative Filtering method (user-based kNN [61]) in which Pearson similarity is used to obtain similar users with target user in term of their ratings [44].

CACFpcc is regarded as Context-Aware Collaborative Filtering method. Pearson similarity is used to obtain the similar users' ratings with target user's ratings. While Jaccard similarity is utilized to access the similar contexts of two users [45].

CACFcos is described as another Context-Aware Collaborative Filtering method. Cosine similarity is used to obtain similar users' ratings with target user's ratings in this method. Like CACFpcc Jaccard similarity is implemented to access the similar contexts of two users.

ACOS is an asymmetric Collaborative Filtering method which used the Cosine, the asymmetric Jaccard and the Sorensen Index in order to obtain similar users' ratings with target user's ratings [47].

APCC is regarded as another asymmetric Collaborative Filtering method which utilized the Pearson, the asymmetric Jaccard and the Sorensen Index to obtain similar users' ratings with target user's ratings [46].

4.5.3 Results and Discussions

It is a nontrivial problem to choose best values of parameters in Eq. (11) that will produce meaningful and insightful results. Therefore, after thorough experimentation of different values for parameter β (from 0 to 1 by step 0.1), it was concluded that this parameter produced interesting and informative results for $\beta = 0.3$.

Figs. 2-3 and Table 4 illustrate the performance of the proposed DBCACF and other methods in terms of Precision, Recall, F-score, MAE and RMSE.

As illustrated in Fig. 2, the increasing the number of neighbors results in increasing the precision values for all methods.

It can be seen that Collaborative Filtering methods have the lowest precision values in compared with other methods, and provide almost equal results for all neighborhood size. The main reason for these results is that CF methods cannot deal well with the sparse data.

However, by using an asymmetric similarity scheme, the recommendation results for ACOS and APCC methods improve significantly for neighborhood size from 60 to later. It should be noted that, two users with different rating vectors have a different impact on each other and this is exactly observable in the results. It is worth noting, due to the sparsity problem, these methods need more number of neighbors to provide better recommendation results.

As another observation, the CACF methods also yield better results as compared to Collaborative Filtering methods. One reason for this improvement is that, these methods incorporate additional information into the recommendation process other than the user-item data and this information can improve the recommendation results in the sparse space. However, these methods use symmetric similarity scheme to select nearest neighbors for a target user and therefore they have still a lower precision values as compared with ACOS and APCC methods.

On another hand, the precision value for the proposed method is higher than 0.34 which considerably increased in comparison with other methods. When we deal with the cold start or data sparsity problems, the users have a few ratings in their visiting history. In these cases, using users' demographic information can provide similar neighbors in term of age and gender and it is possible that they have similar interests with the target user. Based on these data, when the first user has lower ratings and the second one has more ratings, the similarity measure is different by considering our asymmetric similarity measure. That is, the second user is a valuable neighbor for the first one. Therefore, the recommendation has a higher quality in the proposed method; while the first user resembles less and is a valueless neighbor for the second one. Therefore, in the proposed method, the user is less influenced for providing recommendations based on the experimental results. In summary, the main reasons for the improvement in the proposed methods are that, incorporating additional information such as contexts and demographics in addition to user-item data into the recommendation process by using a new hybrid method and providing a new asymmetric similarity scheme to select the nearest neighbors for the target user over data sparsity and cold-start conditions.

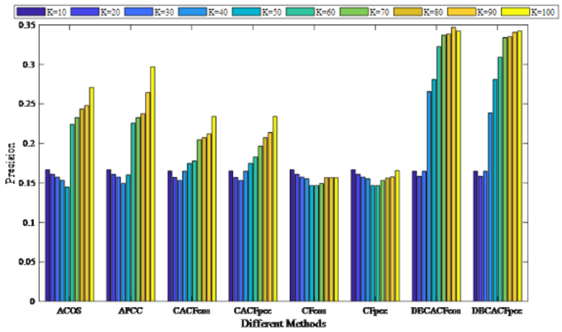

Fig. 2. Precision Measure of different methods

To further evaluate the quality of the proposed method as the recall and precision are regarded as the important metrics in this study, the effect of both metrics was evaluated in the proposed method by using the F1-measure (Fig. 3). Based on the results presented in Fig. 3 for the F-score measure, it can be concluded that the quality presented in the precision results is reliable (F-score results are almost similar to the precision values).

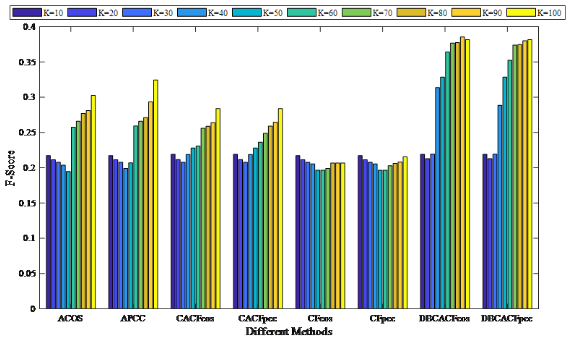

Fig. 3. F1 Measure of different methods

To better illustrate the advantages and characteristics of the proposed method, the average quality and accuracy criteria are demonstrated in Table 4.

Based on the results reported in Table 4, the proposed methods can provide better estimation of scores, and their prediction error is less than other methods. It can be observed that the proposed methods can recognize the similar neighbors for all neighborhoods, and provide the appropriate prediction error; while other methods have provided a higher error rate in prediction results. On the other hand, by considering the sparsity of this dataset, the proposed method can perform better than others on different numbers of neighbors.

Table 4. Average quality and accuracy metrics

Methods	Quality			Accuracy	
	Precision	Recall	F-Score	MAE	RMSE
CFcos	0.1556	0.4223	0.2057	0.6096	0.6966
CFpcc	0.1562	0.4235	0.2064	0.6171	0.7061
ACOS	0.1958	0.4295	0.2382	0.5586	0.6647
APCC	0.2011	0.4307	0.2424	0.5572	0.6636
CACFcos	0.1839	0.3998	0.2368	0.5403	0.6625
CACFpcc	0.1823	0.3997	0.2353	0.5381	0.6612
DBCACFcos	0.2692	0.4271	0.3152	0.5294	0.6577
DBCACFpcc	0.2644	0.4254	0.3108	0.5310	0.6586

We argue that, based on the experimental results, DBCACFcos and DBCACFpcc could yield better results as compared to ACOS, APCC, CACFs and CFs methods in term of precision, F-score, MAE and RMSE metrics; while, the recall measure for the APCC was a little higher than other methods and the ACOS was at the next rank. After that the DBCACFcos was placed at the later one. In general, better results were obtained when the contexts were considered in the recommendation process, compared to collaborative filtering methods. In fact, the results demonstrated the contexts of each user are actually regarded as useful information in tourist recommendations. In addition, the proposed methods outperformed all other recommendations methods when contextual and demographic information were utilized to provide the area recommendations. It should be note that, we deal with sparse rating data because a considerable number of users can visit awhile areas and several users are envisaged the cold start problem due to numerous areas of interest, time and cost constraints. In these conditions, utilizing demographics and contexts information integrated with the user feedbacks can provide better quality and accurate results. Furthermore, using the proposed asymmetric similarity measure could find the users which are more similar to the target user when it deals with a few numbers of co-rated and common items.

In general, our method based on the Cosine similarity measurement has provided a better performance in compared with our method based on the Pearson similarity measure over data sparsity and cold-start problems.

5. Conclusions and Future Work

As mentioned, previous methods are suffering from impersonalized recommendations, low quality recommendations, low accuracy recommendations and unreliability problems due to using only user-item data, considering two different users as the same, data sparsity, cold start condition, etc. In this study, the information related to demographics, contexts, and collective wisdom of people were utilized to provide area recommendations. Based on this information a hybrid user profile was created. In addition, a new hybrid similarity measure was proposed based on an asymmetric scheme which was calculated between each pair of users in order to overcome the limitations of other methods to present the user nearest neighbors and recommend the personalized tourist's area.

Based on the experimental results, DBCACFcos and DBCACFpcc could yield better results, compared to ACOS, APCC, CACFs and CFs methods in term of Precision, F-score, MAE and RMSE values over cold-start and data sparsity conditions.

Despite the aforementioned advantages, the proposed method has several weaknesses. First, the sequence of areas was not intended to recommend the suggestions on the purpose of planning travel. Second, no time and budget constraints were considered due to the need of explicit questions from the users for each recommendation and unavailability this information in the used dataset. Finally, time and location contexts received the same values in recommendation process.

Future work can concentrate on the following directions. First, using the sequences between locations visited by the users can be emphasized as an appropriate factor for better recommendations. The sequence of visited locations by employing a sequential pattern mining algorithm can help planning travel. Second, other user's features, such as occupation, companion and the like, can be used to investigate whether these features are valuable for tourist recommendations based on community contributed geo-tagged photo collection. Third, selecting and weighting the most important contexts can be included in the recommendation model among all the available contexts. Finally, extracting hidden factors that affect the recommendation process, such as outlier, is a good topic for future research.

References

[1] D. Godoy and A. Corbellini, "Folksonomy-Based Recommender Systems: A State-of-the-Art Review," Int. J. Intell. Syst., vol. 31, no. 4, pp. 314-346, 2016.
[2] P. Yu, L. Lin, and Y. Yao, "A Novel Framework to Process the Quantity and Quality of User Behavior Data in Recommender Systems," in Proceedings of the 17th International Conference on Web-Age Information Management, WAIM 2016, Nanchang, China, Part I, 2016, pp. 231-243.
[3] Z. K. Zhang, T. Zhou, and Y. C. Zhang, "Tag-Aware Recommender Systems: A State-of-the-Art Survey," Journal of Computer Science and Technology, vol. 26, no. 5, pp. 767-777, 2011.
[4] Z. L. Zhao, C. D. Wang, Y. Y. Wan, and J. H. Lai, "Recommendation in feature space sphere," Electronic Commerce Research and Applications, vol. 26, no. Supplement C, pp. 109-118, 2017.
[5] S. Khusro, Z. Ali, and I. Ullah, "Recommender Systems: Issues, Challenges, and Research Opportunities," in Proceedings of the Information Science and Applications (ICISA), Singapore, 2016, pp. 1179-1189.
[6] M. Kolahkaj, A. Harounabadi, and M. Sadeghzade, "A Recommender System for Web Mining using Neural Network and Fuzzy Algorithm," International Journal of Computer Applications, vol. 78, no. 8, pp. 20-24, 2013.
[7] M. Kolahkaj and M. Khalilian, "A recommender system by using classification based on frequent pattern mining and J48 algorithm," in 2nd International Conference on Knowledge-Based Engineering and Innovation (KBEI), 2015, pp. 780-786.
[8] R. Safa, S. Mirroshandel, S. Javadi, and M. Azizi, "Venue Recommendation Based on Paper's Title and Co-authors Network," Journal of Information Systems and Telecommunication, vol. 1, no. 6, pp. 209-217, 2017.
[9] X. Ma, H. Lu, Z. Gan, and J. Zeng, "An explicit trust and distrust clustering based collaborative filtering recommendation approach," Electronic Commerce Research and Applications, vol. 25, no. Supplement C, pp. 29-39, 2017.
[10] R. Gao, J. Li, X. Li, C. Song, and Y. Zhou, "A personalized point-of-interest recommendation model via fusion of geo-social information," Neurocomputing, vol. 273, pp. 159-170, 2018.
[11] Y. M. Afify, I. F. Moawad, N. L. Badr, and M. F. Tolba, "A personalized recommender system for SaaS services," Concurrency and Computation: Practice and Experience, vol. 29, no. 4, p. e3877, 2017.
[12] N. M. Villegas, C. Sánchez, J. Díaz-Cely, and G. Tamura, "Characterizing context-aware recommender systems: A systematic literature review," Knowledge-Based Systems, vol. 140, pp. 173-200, 2018.
[13] I. Cenamor, T. de la Rosa, S. Núñez, and D. Borrajo, "Planning for tourism routes using social networks," Expert Systems with Applications, vol. 69, pp. 1-9, 2017.
[14] G. Cai, K. Lee, and I. Lee, "Itinerary recommender system with semantic trajectory pattern mining from geo-tagged photos," Expert Systems with Applications, vol. 94, pp. 32-40, 2018.
[15] I. Memon, L. Chen, A. Majid, M. Lv, I. Hussain, and G. Chen, "Travel Recommendation Using Geo-tagged Photos in Social Media for Tourist," Wirel. Pers. Commun., vol. 80, no. 4, pp. 1347-1362, 2015.
[16] D. Wang, S. Deng, and G. Xu, "Sequence-based context-aware music recommendation," Information Retrieval Journal, vol. 21, no. 2, pp. 230-252, 2018.
[17] D. Bachmann et al., "(CF)2 architecture: contextual collaborative filtering," Information Retrieval Journal, 2018.

[18] Z. Xiang and U. Gretzel, "Role of social media in online travel information search," Tourism Management, vol. 31, no. 2, pp. 179-188, 2010.
[19] V. Subramaniyaswamy, V. Vijayakumar, R. Logesh, and V. Indragandhi, "Intelligent Travel Recommendation System by Mining Attributes from Community Contributed Photos," Procedia Computer Science, vol. 50, pp. 447-455, 2015.
[20] R. S. Aquino, H. A. Schänzel, and K. F. Hyde, "Unearthing the geotourism experience: Geotourist perspectives at Mount Pinatubo, Philippines," Tourist Studies, vol. 18, no. 1, pp. 41-62, 2018.
[21] L. Ravi and S. Vairavasundaram, "A Collaborative Location Based Travel Recommendation System through Enhanced Rating Prediction for the Group of Users," Intell. Neuroscience, vol. 2016, p. 7, 2016.
[22] M. Pazzani and D. Billsus, "Learning and Revising User Profiles: The Identification ofInteresting Web Sites," Mach. Learn., vol. 27, no. 3, pp. 313-331, 1997.
[23] D. H. Park, H. K. Kim, I. Y. Choi, and J. K. Kim, "A literature review and classification of recommender systems research," Expert Systems with Applications, vol. 39, no. 11, pp. 10059-10072, 2012.
[24] M. Nilashi, O. bin Ibrahim, N. Ithnin, and N. H. Sarmin, "A multi-criteria collaborative filtering recommender system for the tourism domain using Expectation Maximization (EM) and PCA–ANFIS," Electronic Commerce Research and Applications, vol. 14, no. 6, pp. 542-562, 2015.
[25] G. Adomavicius and A. Tuzhilin, "Toward the next generation of recommender systems: a survey of the state-of-the-art and possible extensions," IEEE Transactions on Knowledge and Data Engineering, vol. 17, no. 6, pp. 734-749, 2005.
[26] M. Balabanovi and Y. Shoham, "Fab: content-based, collaborative recommendation," Commun. ACM, vol. 40, no. 3, pp. 66-72, 1997.
[27] W. Wu et al., "Improving performance of tensor-based context-aware recommenders using Bias Tensor Factorization with context feature auto-encoding," Knowledge-Based Systems, vol. 128, pp. 71-77, 2017.
[28] D. H. Lee and P. Brusilovsky, "Improving personalized recommendations using community membership information," Information Processing & Management, vol. 53, no. 5, pp. 1201-1214, 2017.
[29] T. Ha and S. Lee, "Item-network-based collaborative filtering: A personalized recommendation method based on a user's item network," Information Processing & Management, vol. 53, no. 5, pp. 1171-1184, 2017.
[30] A. Sesagiri Raamkumar, S. Foo, and N. Pang, "Using author-specified keywords in building an initial reading list of research papers in scientific paper retrieval and recommender systems," Information Processing & Management, vol. 53, no. 3, pp. 577-594, 2017.
[31] K. H. L. Tso-Sutter, L. B. Marinho, and L. Schmidt-Thieme, "Tag-aware recommender systems by fusion of collaborative filtering algorithms," in Proceedings of the ACM symposium on Applied computing, Fortaleza, Ceara, Brazil, 2008, pp. 1995-1999.
[32] A. M. Nagrale and A. P. Pande, "User Preferences-based Recommendation System for Services Using Map Reduce Approach for Big Data Applications," International Journal of Innovations & Advancement in Computer Science, vol. 4, pp. 528-532, 2015.
[33] M. Y. H. Al-Shamri, "Effect of Collaborative Recommender System Parameters," Adv. in Artif. Intell., vol. 2016, pp. 1-10, 2016.
[34] Z. Zhang, X. Zheng, and D. D. Zeng, "A framework for diversifying recommendation lists by user interest expansion," Knowledge-Based Systems, vol. 105, pp. 83-95, 2016.
[35] J. Bobadilla, F. Ortega, A. Hernando, and J. Alcalá, "Improving collaborative filtering recommender system results and performance using genetic algorithms," Knowledge-Based Systems, vol. 24, no. 8, pp. 1310-1316, 2011.
[36] Q. Cheng et al., "The new similarity measure based on user preference models for collaborative filtering," in Proceedings of the IEEE International Conference on Information and Automation, 2015, pp. 577-582.
[37] G. Guo, J. Zhang, and N. Yorke-Smith, "A Novel Evidence-Based Bayesian Similarity Measure for Recommender Systems," ACM Trans. Web, vol. 10, no. 2, pp. 1-30, 2016.
[38] M. Y. H. Al-Shamri, "User profiling approaches for demographic recommender systems," Knowledge-Based Systems, vol. 100, pp. 175-187, 2016.
[39] S. Zammali, K. Arour, and A. Bouzeghoub, "A Context Features Selecting and Weighting Methods for Context-Aware Recommendation," in Proceedings of the IEEE 39th Annual Computer Software and Applications Conference, 2015, vol. 2, pp. 575-584.
[40] H. Liu, Z. Hu, A. Mian, H. Tian, and X. Zhu, "A new user similarity model to improve the accuracy of collaborative filtering," Knowledge-Based Systems, vol. 56, pp. 156-166, 2014.
[41] Z. L. Zhao, C. D. Wang, and J. H. Lai, "AUI&GIV: Recommendation with Asymmetric User Influence and Global Importance Value," PLoS One, vol. 11, no. 2, pp. 1-21, 2016.
[42] S. H. Cha, "Comprehensive Survey on Distance/Similarity Measures between Probability Density Functions," International Journal of Mathematical Models and Methods in Applied Sciences, vol. 1, no. 4, pp. 300-307, 2007.
[43] P. Pirasteh, D. Hwang, and J. E. Jung, "Weighted Similarity Schemes for High Scalability in User-Based Collaborative Filtering," Mob. Netw. Appl., vol. 20, no. 4, pp. 497-507, 2015.
[44] A. Majid, L. Chen, G. Chen, H. T. Mirza, I. Hussain, and J. Woodward, "A context-aware personalized travel recommendation system based on geotagged social media data mining," Int. J. Geogr. Inf. Sci., vol. 27, no. 4, pp. 662-684, 2013.
[45] Y. Zheng, R. Burke, and B. Mobasher, "Recommendation with Differential Context Weighting," in Proceedings of the 21th International Conference on User Modeling, Adaptation, and Personalization, UMAP 2013, Rome, Italy, 2013, pp. 152-164.
[46] P. Pirasteh, J. J. Jung, and D. Hwang, "An Asymmetric Weighting Schema for Collaborative Filtering," in In: Proceedings of the New Trends in Computational Collective Intelligence, 2015: Springer, pp. 77-82.
[47] P. Pirasteh, D. Hwang, and J. J. Jung, "Exploiting matrix factorization to asymmetric user similarities in recommendation systems," Knowledge-Based Systems, vol. 83, pp. 51-57, 2015.
[48] B. K. Patra, R. Launonen, V. Ollikainen, and S. Nandi, "A new similarity measure using Bhattacharyya coefficient for collaborative filtering in sparse data," Know.-Based Syst., vol. 82, no. C, pp. 163-177, 2015.

[49] S. Vairavasundaram, V. Varadharajan, I. Vairavasundaram, and L. Ravi, "Data mining-based tag recommendation system: an overview," Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, vol. 5, no. 3, pp. 87-112, 2015.
[50] G. Adomavicius and A. Tuzhilin, "Context-Aware Recommender Systems," in Recommender Systems Handbook1st ed.: Springer US, 2011, pp. 217-253.
[51] A. Majid, L. Chen, H. T. Mirza, I. Hussain, and G. Chen, "A system for mining interesting tourist locations and travel sequences from public geo-tagged photos," Data & Knowledge Engineering, vol. 95, pp. 66-86, 2015.
[52] Suryakant and T. Mahara, "A New Similarity Measure Based on Mean Measure of Divergence for Collaborative Filtering in Sparse Environment," Procedia Computer Science, vol. 89, pp. 450-456, 2016.
[53] Z. Xu, "Trip similarity computation for context-aware travel recommendation exploiting geotagged photos," in Proceedings of the 30th IEEE International Conference on Data Engineering Workshops, 2014, pp. 330-334.
[54] S. Kisilevich, F. Mansmann, and D. Keim, "P-DBSCAN: a density based clustering algorithm for exploration and analysis of attractive areas using collections of geo-tagged photos," in Proceedings of the 1st ACM International Conference and Exhibition on Computing for Geospatial Research & Application, Washington, D.C., USA, 2010, pp. 1-4.
[55] L. J. Li, R. K. Jha, B. Thomee, D. A. Shamma, L. Cao, and Y. Wang, "Where the Photos Were Taken: Location Prediction by Learning from Flickr Photos," in Proceeding of the Large-Scale Visual Geo-Localization, 2016, pp. 41-58.
[56] M. Ester, H. P. Kriegel, r. Sander, and X. Xu, "A density-based algorithm for discovering clusters in large spatial databases with noise," in Proceedings of the Second International Conference on Knowledge Discovery and Data Mining, Portland, Oregon, 1996, pp. 226-231.
[57] J. Han, M. Kamber, and J. Pei, "Data Mining: Concepts and Techniques," Third ed. Boston: Morgan Kaufmann, 2012, pp. 39-82.
[58] B. Thomee et al., "YFCC100M: the new data in multimedia research," Commun. ACM, vol. 59, no. 2, pp. 64-73, 2016.
[59] Flickr. (2018, 6 jan. 2017). [Online]. Available: http://www.Flickr.com.
[60] WebscopeYahooLabs. (15 Feb 2017). [Online]. Available: https://webscope.sandbox.yahoo.com/catalog.php?datatype=i&did=67.
[61] A. Bellogín and P. Sánchez, "Collaborative filtering based on subsequence matching: A new approach," Information Sciences, vol. 418, pp. 432-446, 2017.

Maral Kolahkaj received the B.S. degree in Computer Engineering from Azad University, Dezfoul Branch, Iran in 2009, and M.S. degree in Software Systems from Azad University, Science & Research Branch, Iran, in 2013. Currently she is Ph.D. Candidate in Azad University, Karaj Branch, Iran. Her research interests include Information retrieval, Recommendation systems, Human-computer interaction, Context awareness, Web mining and Data mining.

Ali Harounabadi is an assistant professor of Computer Engineering at Azad University, Central Tehran Branch, Iran. His research is focused on Recommender Systems, Web mining and Methodologies in Software Engineering.

Alireza Nikravanshalmani is a full-time Assistant-Professor of Computer Engineering in the Faculty of engineering at Azad University, Karaj Branch, Iran. He received his Ph.D. degree in Computer Science from Kingston University-London, England in 2011. He is conducting research activities in the areas of Image and Data Analysis Software Systems.

Rahim Chinipardaz received the B.S. degree in Statistics and Computer Science from Shahid Chamran University, Ahwaz, Iran in 1984, and M.S. degree in Statistics from Tarbiat Modares University, Tehran, Iran, in 1989. He received his Ph.D. degree from Newcastle University of England in 1996. Now, he works as full-professor in the Faculty of Mathematics, Statistics and Computer Science at Shahid Chamran University, Ahwaz, Iran. His area research interests include Statistics, Probability, Time series analysis and Discriminant Analysis.