Hybrid POI recommendation system for tourism

Monishkanna Barathan
University of Houston-Clear Lake
Houston, USA.
BarathanM7346@uhcl.edu

Ershad Sharifahmadian
University of Houston-Clear Lake
Houston, USA
Sharifahmadi@uhcl.edu

***Abstract*— Due to the increase in amount of available information, finding places and planning of the activities to be done during a tour can be strenuous. Tourists are looking for information about a place in which they have not been before, which worsen the selection of places that fit better with user's preferences. Recommendation systems have been fundamentally applicable in tourism, suggest suitable places, and effectively prune large information from different locations, so tourists are directed toward those places where are matched with their needs and preferences. Several techniques have been studied for point-of-interest (POI) recommendation, including content-based which builds based on user preferences, collaborative filtering which exploits the behavior of other users, and different places, knowledge-based method, and several other techniques. These methods are vulnerable to some limitations and shortcomings related to recommendation environment such as scalability, sparsity, first-rater or gray sheep problems. This paper tries to identify the drawbacks that prevent wide spread use of these methodologies in recommendation. To improve performance of recommendation systems, these methods are combined to form hybrid recommenders. This paper proposes a novel hybrid recommender system which suggests tourism destinations to a user with minimal user interaction. Furthermore, we use sentiment analysis of user's comments to enhance the efficiency of the proposed system.**

***Index Terms*— Personalized filtering, Point of interest, Recommendation system, Tourism recommendation, Hybrid recommender.**

I. Introduction

Recommender systems are able to recommend personalized content based on past behavior. The system help users to retrieve relevant places from large databases, by automatically finding and suggesting the possible places of interest based on the user's preferences. The key point which affects the development of recommendation is from the era of data which changed from sparsity to abundance. Recommender Systems usually employ a combination of different types of recommendation techniques such as content-based, collaborative filtering, knowledge-based filtering, demographic filtering, etc.

Data plays a vital role in tourism industry. Tourists are commonly using their mobile phones to help with the interactive planning to find the places around them which makes the travel more comfortable compared to the traditional method of using maps and tourist guidebooks. Those traditional methods are not able to meet preferences of tourists which contribute to the exciting experience, and those are also designed in the manner to attract typical tourists. A tourist needs to spend more time on referencing the attractions near to their current location. Given issues can be solved by the recommendation systems which recommend the places matched with user's point of interest, and based on users preferences.

Automatic tourism place recommendation is an important topic in both university and industry. Big media, especially social media such as Facebook, Flick, Twitter, etc. offer great opportunities to address many tourism challenges, for instance, travel recommendation [2]. Travelogue websites offer rich descriptions about landmarks and traveling experiences provided by users. Recommender systems apply knowledge uncovering techniques to the problem of making personalized recommendations for information, products or services during a user's login to the system [3]. These systems, especially the collaborative are successfully bringing out the successful recommendation on the internet. In recent years, the fast grow of information and the number of online users creates big challenges for recommender systems; producing a high value recommendations, performing many recommendations per second for millions of users and places, and achieving high coverage for large amount of data. In current collaborative filtering systems, the amount of work increases with the number of participants in the system. The desired version of recommender systems are expected to produce high-quality recommendations at very fast rate in short period of time.

II. Related Works

Zaho and zhi [15], implemented Map-Reduce model for Collaborative Filtering algorithm based on User based similarity that helps to solve the scalability problem of available recommendation methods. As an advantage of this method is a personalized recommendation determining interest for similar places with high performance. However this method doesn't consider similar users in the system.

In [16], Chen and Ying focused on personalized travel recommendation and used freely accessible community-contributed photos to obtain promising applications. This method is empowered by community contribution which is a liable attribute to consider for providing suitable recommendations, but privacy issues can arise while processing community-contributed photographs, and the method is complex to execute. Jin and Yohan in [17] proposed an approach which includes product to product collaborative filtering to discover fascinating videos among a large amount of videos. This method is executed in Qizmt-a.NET Map Reduce framework-This method does not take an account of

This is an open access article under the CC BY-SA 4.0 license (https://creativecommons.org/licenses/by-sa/4.0/)

similarity interests on places, and its implementation is difficult.

In [18], author built the tourism recommendation system based on the representative tags and images relating to GPS locations. The user is able to choose a query either as a keyword or an image, then the system compares the query with information in the database, and matches the representative images and tags with the given query. For a keyword input, a geotagged location is chosen if the representative tags comprise the keyword. For an image based input query, the geotagged locations are ordered according to the similarity between the query images and the descriptive image feature in different clusters. In [19], Author developed recommendation system based on "context-aware rating" whereas ratings and evaluations uploaded by mobile users while onsite are considered more credible and weighted higher than those uploaded by users via standard web interfaces.

III. RECOMMENDATION SYSTEM: PRINCIPLES

There are two main challenges for automatic travel recommendation with personalized filtering [1]. First, the recommended POIs should be custom-made to user interest since different users may prefer different types of POIs. As an example, some people may prefer cultural places like Museums, while others may prefer locations like the Parks. Moreover user's current interest, other attributes including preferred budget, preferable visiting season and desired visiting time may also be helpful to provide personalized tourism recommendation. Second, it is important to recommend consecutive tourist places with travel route rather than separate POIs. It is far more problematic and time consuming for users to have consecutive travel plan instead of visiting individual POIs, because the similarity between the locations and the opening time of different POIs should be deliberated. For example, it is not a noble recommendation if all the POIs recommended for one day are far from each other.

Different factors such as geographical, temporal, communal relation and contend indications affect the recommendation systems [1]. Because of the spatial and temporal properties resulted from the physical restrictions and varied information such as locations' geographical information and users' comments, the check-in activity influence recommendations.

A. Geographical Influence

Geographical influence is an important factor that differentiates the POI recommendation from outdated place recommendation because the check-in actions depend on the locations' geographical features. Evaluations of users' check-in data shows that a user acts in geographically constrained areas and prefers to visit close POIs where the user has checked-in.

B. Social Influence

This influence is inspired by the hypothesis that friends in Location Based Social Network (LBSN) share more mutual interests than non-friends. Social influence is explored to improve POI recommendation. In [4] author examined the features of the large-scale check-in data from a popular LBSNs website called Gowalla. Based on data mining, he proposed a novel Multi-center Gaussian Model to model the geographical influence of users' check-in behavior. Then he considered users' social information, and proposed a fused matrix factorization method to include the geographical influence of users' check-in locations. Results from wide experiments show that the proposed method significantly outperforms other similar approaches.

C. Temporal Influence

In [5] authors examined patterns of human mobility on three large but very diverse datasets: two sets of location-based social network check-in data spanning the whole planet and cellphone location data as well. Even though LBSN services are very dissimilar from cell phone tower location data, they found many shared patterns of human mobility in the datasets. Most surprisingly, they found that humans experience a combination of strong short-range spatially and temporally periodic movement that is not squeezed by the social network structure, while long-distance travel is more inclined by the social network ties.

D. Content Indication

In [6], authors proposed a preference model based on check-in and tip data under the statement that user's comments of a venue can help to better describe the place preference. They extended the matrix factorization approach considering the impact of both user social similarity and inter place similarity on location recommendation, named as location-based social matrix factorization (LBSMF) approach. Based on results in [6], the proposed preference model can describe user's favorites, and further the effectiveness, and efficiency of proposed LBSMF approach were validated.

IV. RECOMMENDATION SYSTEM

Debashis and Datta in [8] explained that the recommendation system usually produces many suggestions based on one of the techniques presented in Fig.1. Companies store massive data for each user and his matching likes/dislikes. Ratings, reviews, and feedback given by a customer about a place also help to collect data. Data collection, machine-learning algorithms such as collaborative filtering is applied to group a current user with known customers having similar likes and dislikes, or to group users and places together [7].

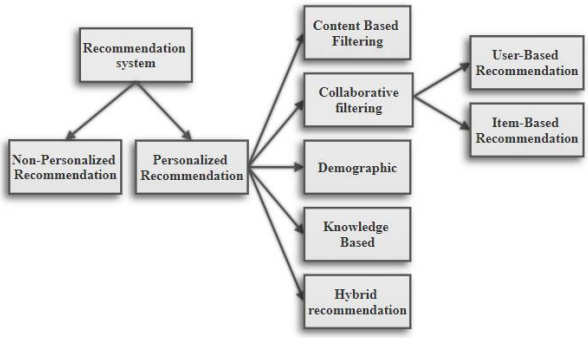

Fig. 1. Types of recommendation system

This is an open access article under the CC BY-SA 4.0 license (https://creativecommons.org/licenses/by-sa/4.0/)

A recommendation engine abstracts recommendations provided specifically to the user. For example, based on past browsing history and visiting travelogue sites, a tourist recommendation website may deduce that a customer is more aligned to like adventure travel, so adventure travel places will be suggested when he wants to visit different tourist places.

A Non-Personalized recommendation system

This system is more popular, and would recommend the places liked by the most number of the users. The main disadvantage of this approach is the lack of user personalization.

B Personalized recommendation system

This system recommends desired places based on past behavior of user, and interpersonal relationship of LBSN by considering three standpoints: Interpersonal influence, Interest circle derivation, User individual interest which affects which places would be of interest [9]. Personalized recommendation systems are categorized into five different categories [10].

i .Demographic recommendation technique

This technique categorizes a user based on his own preferable features and recommends them based on the demographic information about the places. The main advantage of this method is that there is no need for past rating from user, whereas content based recommender and the collaborative methods depend on user rating history.

ii. Knowledge based recommendation technique

This technique suggests a POI based on the speculations about user's demands and preferences. This technique knows how well the POI meets the user demands. Hybrid recommendation technique uses two or more recommendation techniques discussed before to gain high efficiency, and is able to solve challenges that individual recommendation techniques faced with.

iii .Content-based filtering technique

Content based filtering technique builds the user profile based on the features presented in the POI where is already rated by the user. The algorithm starts to mine the description of POI from internet. For each user, the algorithm recommends places where are similar to the places were visited by users. The algorithm does the following steps:

- Extracting features of the place by mining descriptions from internet.
- Defining user profiles based on previous visits or user's preferable features.
- Recommending each user the places matched with his preferable features or previous visited places.

User profiles are based on both the user preference and previous visited places. For new user the places matched with user's preference is recommended.

iv. Collaborative filtering technique

Collaborative recommendation is widely used, and implemented for many applications. This method recognizes the similarity between users and POIs based on available ratings and generates recommendations. This technique aims at looking for patterns of agreement among user's ratings for different places [3]. The main instinct behind this technique is that if a user agrees with his neighbors in the past, he will continue to do so in the future. There are two types of collaborative filtering: place-based collaborative filtering and User-based collaborative filtering; place-based collaborative filtering recommends a place to the user where is similar to the places already visited by the user. For each two places, place-based collaborative filtering measures how similar those places are in terms of having received similar ratings by similar users. Then, for each place, the k-most similar places (using cosine or Pearson Correlation) are identified. Next, for each user, the places where are most similar to the user's previous visiting are identified. User-based collaborative filtering recommends the place to the user where are most preferred by similar users. For every new user, user-based collaborative filtering measures how similar each user is to the new one (using cosine or Pearson Correlation). Then, the most similar users is identified by taking account of top k users by k-nearest neighbors or the users for which their similarity is above a defined threshold.
Next, the places visited by most similar users are rated using average rating or weighted average rating. Eventually, the top-rated places will be selected and recommended to the user.
Figure 2 explains the place-based collaborative filtering process; rows represent the users, and columns represent the places. Various mathematical formulations can be used to calculate the similarity between two POIs. Cosine Similarity method treats the absent ratings as negative i.e. 0. To overcome the disadvantages of the cosine method, we need to normalize the rating by subtracting mean of rows.

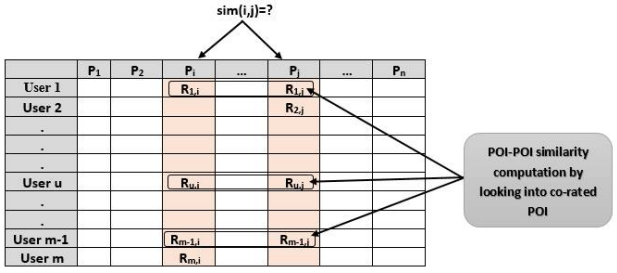

Fig 2. Selection of POIs; POIs have been similarly rated by different users.

Pearson Correlation is the cosine similarity bounded between -1 and 1. The Pearson correlation captures the intuition better as it considers the missing ratings as average and handles the 'tough raters' and 'easy raters'.

This is an open access article under the CC BY-SA 4.0 license (https://creativecommons.org/licenses/by-sa/4.0/)

$$sim(i,j) = \frac{\sum_{u \in U}(R_{u,i} - \overline{R_i})(R_{u,j} - \overline{R_j})}{\sqrt{\sum_{u \in U}(R_{u,i} - \overline{R_i})^2} \cdot \sqrt{\sum_{u \in U}(R_{u,j} - \overline{R_j})^2}} \quad (1)$$

After the similar users are identified by the Pearson similarity, users are prioritized in the list of the top N similar users, and their ratings are used to predict the rating of the user using equation (2).

$$r_{u,i} = \frac{\sum_{y \in N} sim(i,y).r_{u,y}}{\sum_{y \in N} sim(i,y)} \quad (2)$$

Where $r_{u,i}$ is the predicted rating of the user u for the place i using the similar N users. The rating prediction for the user u for the place i is shown in figure3 (here, 5 neighbors are considered).

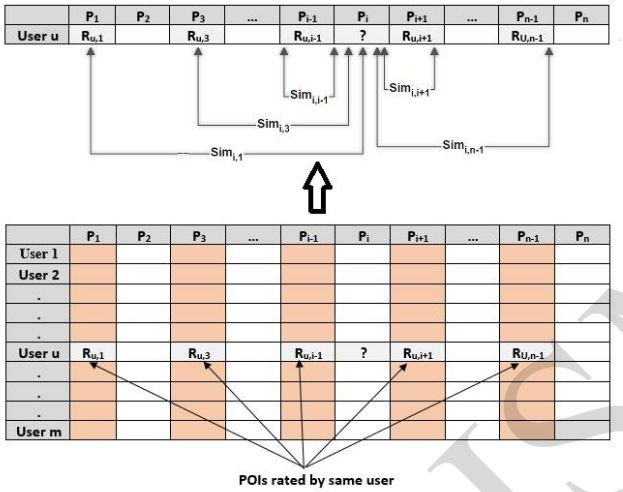

Fig 3. Rating prediction algorithm for POI based on collaborative filtering; finding possible rating by user u for place i.

v. Latent factor model

In this method, the primary goal is to infer the latent factors from the user's rating for different places using mathematical techniques such as the Singular Value Decomposition (SVD) –SVD is used to find the low rank optimization of the User-Rating matrix–.

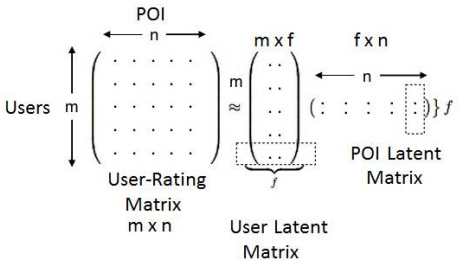

Fig 4. Latent factor model; Splitting User-Rating matrix into user and POI latent matrix

In this approach User rating matrix is decomposed based on SVD, and reconstructed based on latent features (figure 4). This gives an ability to predict missing values in the User-Rating matrix. The SVD approximation method is similar to the matrix factorization, but in this latent factor model, latent factors of users and places are retrieved in another way. Latent Factor models can be described in terms of matrix factorization. Matrix factorization aims to decompose a matrix into two low-rank terms for which their product reconstructs or approximates the original matrix as close as possible. Typically, a matrix R is considered to represent the observed feedback from m users and n places. The matrix R can then be factorized into one user-specific matrix U ($m \times f$) and one place-specific matrix V ($f \times n$). In other words, $R \approx UV^T$, where f is the dimensionality of the latent factors which characterizes the features of a user or a place. Accordingly, $U_u V_i^T$ captures user u's preference toward place i.

V. Hybrid Recommendation Techniques

By combining recommender methods explained before hybrid recommendation techniques can be classified into five different categories as follows [21].

i. *Weighted technique:* The voting of several recommender techniques is used to form a single recommendation.
ii. *Switching technique:* Depending on the user's available rating, it switches between recommendation techniques.
iii. *Mixed recommendation:* Each method produces its own recommendation and all recommendations are presented to the user at the same time.
iv. *Feature Combination technique:* Features collected from different recommendation methods are fed as input to the single recommendation algorithm.
v. *Cascading technique:* Recommendation given by one technique is refined by another recommendation technique.

VI. Proposed Hybrid POI Recommendation System for Tourism

Hybrid recommendation technique can alleviate disadvantages associated with collaborative, content-based and several other recommendation techniques. Our proposed system is a Hybrid recommendation system which utilizes content-based filtering, collaborative filtering and latent factor model to reduce the effect of the "ramp-up" problem. Ramp-up commonly refers to new user and new POI problem. In fact, the new users with few ratings make trouble to find similar users accurately. Similarly, a new POI without high number of ratings cannot be considered in process of recommendations. Furthermore, the sparsity issue happens while collaborative filtering is just used, in other words few users rate the same POI. This sparsity is reduced in the latent factor model based filtering technique. Therefore, we utilize

IJSMS

This is an open access article under the CC BY-SA 4.0 license (https://creativecommons.org/licenses/by-sa/4.0/)

the Singular value decomposition along with the collaborative filtering to overcome sparse data. New places and new users can be alleviated using the content based filtering technique which analyzes POI features by considering descriptions of the place. The proposed system suggests a POI with highest similarity based on its features like categories, timing, the type of the place, etc. If users have preferences for a certain type of POI, then the proposed system recommends a similar POI to the user who likes those features. Fig 5 shows our proposed recommendation system. In the proposed system, users need to create their own profile with their own preferable preferences. Users are also able to restrict POI recommendations constrained by distance.

To reach high efficiency, defining certain categories based on the user's inputs is important. Then, the entire data is processed by content, collaborative and latent factor methods and individual recommendations are obtained. In next step, the results are fused to give optimum recommendations to the user (figure 6).

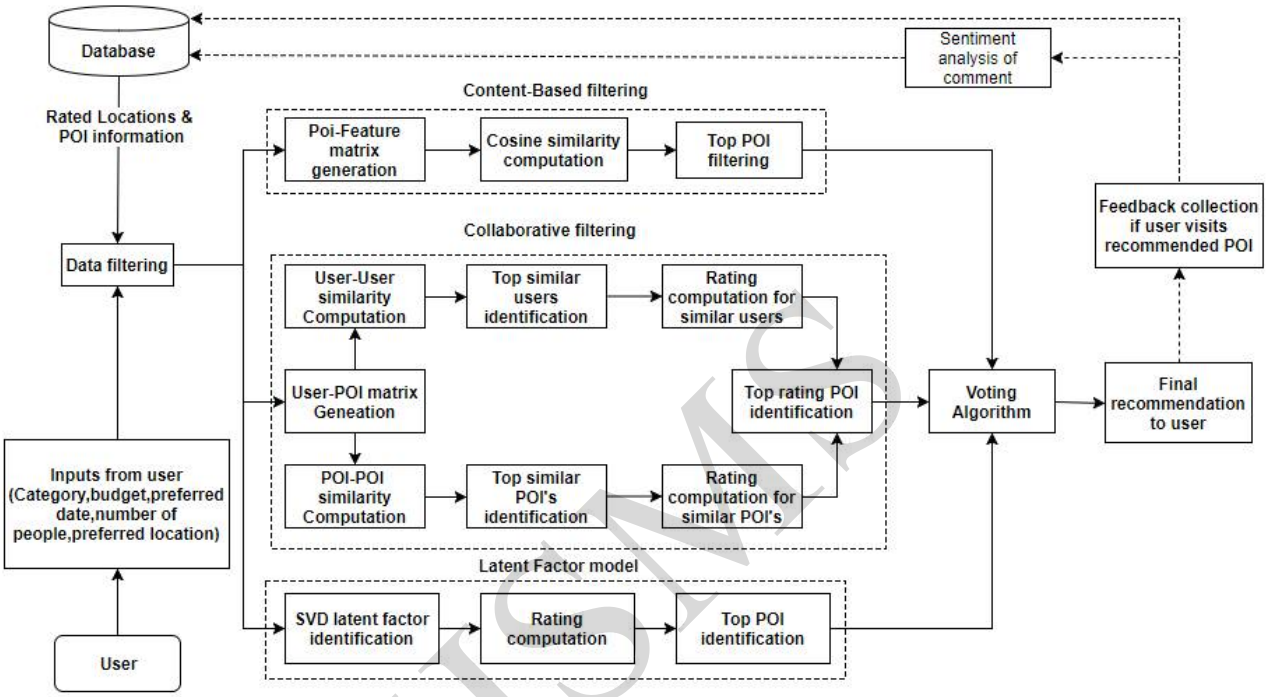

Fig. 5. Proposed POI recommendation system

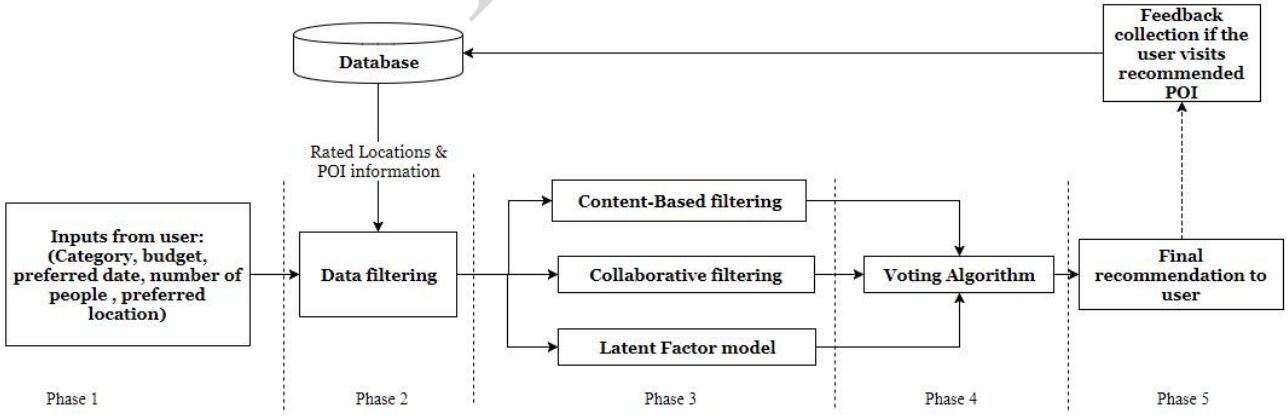

Fig 6. Phase diagram of proposed POI recommendation system

This is an open access article under the CC BY-SA 4.0 license (https://creativecommons.org/licenses/by-sa/4.0/)

VII. Discussion and Conclusion

Content based filtering is based on the customer preferences, and has the merits such as lots of input data is not required, and no data sparsity. The point of content-based filtering process is that the features of both user and places should be known. Typically, user-profile and place-profile are created by the features of shared attributes. The main disadvantage of this technique is the affinity towards the place is not estimated.

Collaborative filtering algorithm uses user's behavior for recommending places to the user. It exploits behavior of other users in terms of user's visit, ratings, comment, and visiting information. In addition, behavior and preferences from other users for different places are used to recommend places to new users. In this case, features of the places are not known. Collaborative systems consider both the users past ratings and similarity among users and places. The main advantage of this method is its ability to estimate the best place based on all other users' rating, however, collaborative filtering suffers from the cold start problem.

All available recommender systems use one of basic recommendation techniques; content-based, collaborative, demographic, utility-based, and knowledge-based techniques. Studying these techniques showed those have pros and cons. Therefore, the hybrid recommender system which combines some of given techniques is used to improve performance of the recommendation system.

Our recommender system combines the content-based filtering, collaborative filtering, and latent factor model to provide more accurate recommendations to users. Furthermore, the proposed recommendation system uses semantic analysis of user's comments to provide better recommendations.

References

[1] Shenglin Zhao, Irwin King, Michael R. Lyu, "A Survey of Point-of-interest Recommendation in Location-based Social Networks"arXiv: 1607.00647 [cs.IR], Jul.2016.

[2] S. Jiang, X. Qian, J. Shen, Y. Fu, and T. Mei, "Author topic model based collaborative filtering for personalized POI recommendation," IEEE Trans. Multimedia, vol. 17, no. 6, pp. 907–918, Jun. 2015.

[3] Badrul Sarwar, George Karypis, Joseph Konstan, and John Riedl," Item-based collaborative filtering recommendation algorithms," Hong Kong.ACM 1-58113-348-0/01/0005, May.2005.

[4] Chen Cheng, Haiqin Yang, Irwin King, and Michael R Lyu,"Fused matrix factorization with geographical and social influence in location-based social networks", Proceedings of the Twenty-Sixth AAAI Conference on Artificial Intelligence, Jan.2012.

[5] Eunjoon Cho, Seth A Myers, and Jure Leskovec,"Friendship and mobility: user movement in location-based social networks." Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining, Mar.2011.

[6] Dingqi Yang, Daqing Zhang, Zhiyong Yu, and Zhu Wang,"A sentiment-enhanced personalized location recommendation system" Proceedings of the 24th ACM Conference on Hypertext and Social Media, pages 119–128. ACM, 2013.

[7] Charlotte Castelino , Reena Shaw Muthalaly , Aloma Lopes ,Sweedal Lopes, "Recommender Systems: Increasing Profits and Efficiency in Business." International Journal of Engineering Trends and Technology (IJETT), V16 (3), 119-123, ISSN: 2231-5381, Oct.2014.

[8] Debashis Das,Laxman Sahoo,Sujoy Datta "A Survey on Recommendation System" International Journal of Computer Applications, Vol.160, No.7, Feb.2017

[9] GK, Kiran Kumar Bathi, and N. Sheela, "Personalized Recommender System." PARIPEX-Indian Journal of Research 4.8,2016.

[10] Asabere, Nana Yaw. "Review of Recommender Systems for Learners in Mobile Social /Collaborative Learning." International Journal of Information 2.5,2012.

[11] Priyanka,"A Survey Paper on Various Algorithm's based Recommender System" IOSR Journal of Computer Engineering (IOSR-JCE) e-ISSN: 2278-0661, p-ISSN: 2278-8727, Volume 19, Issue 3, Ver. IV, PP 27-32, May - June 2017.

[12] Shuhui Jiang, Xueming Qian,Tao Mei and Yun Fu, Senior Member, IEEE" Personalized Travel Sequence Recommendation on Multi-Source Big Social Media" ,IEEE TRANSACTIONS ON BIG DATA, VOL. 2, NO. 1, Jan-Mar 2016.

[13] Zhao, Zhi-Dan, and Ming-Sheng Shang. "User-based collaborative-filtering recommendation algorithms on hadoop." Knowledge Discovery and Data Mining, Third International Conference on. IEEE, 2010.

[14] Chen, Yan-Ying, An-Jung Cheng, and Winston H. Hsu. "Travel recommendation by mining people attributes and travel group types from community-contributed photos." IEEE Transactions on Multimedia 15.6: 1283- 1295, 2013.

[15] Jin, Yohan, et al. "MySpace video recommendation with map-reduce on qizmt." Semantic Computing (ICSC), IEEE Fourth International Conference, 2010.

[16] Liangliang Cao, Jiebo Luo, Andrew Gallagher, Xin Jin, Jiawei Han and Thomas S. Huang, "A worldwide tourism recommendation system based on geo tagged web photos",IEEE International Conference on Acoustics, Speech, and Signal Processing,978-1-4244-4296-6/10, 2010.

[17] Damianos Gavalas,Michael Kenteris,"A web-based pervasive recommendation system for mobile tourist guides" Springer-Verlag London Limited,2011.

[18] Luo, Zhiling, Ying Li, and Jianwei Yin. "Location: a feature for service selection in the era of big data." ICWS, 20th International Conf. on. IEEE, 2013.

[19] Liu, Hongyan, et al. "Combining user preferences and user opinions for accurate recommendation." Electronic Commerce Research and Applications 12.1, 14-23, 2013.

[20] Esparza, Sandra Garcia, Michael P. O'Mahony, andBarry Smyth. "Mining the real-time web: a novel approach to product recommendation." Knowledge Based Systems 29: 3-11, 2012.

[21] R Burke," Hybrid recommender systems: Survey and experiments", User modeling and user-adapted interaction, Springer, 2002.

This is an open access article under the CC BY-SA 4.0 license (https://creativecommons.org/licenses/by-sa/4.0/)