International Journal of Advance Research in Science and Engineering
Volume No.06, Issue No. 11, November 2017
www.ijarse.com
IJARSE
ISSN: 2319-8354

ENHANCED VENUE RECOMMENDATION SYSTEM BASED ON USER LOCATION AND VENUE REVIEWS

Prajakta Bothe[1], Prof.M.R.Patil[2]

[1,2]*Dept. of Computer Engineering, Smt.Kashibai Navale College of Engineering,*
Savitribai Phule Pune University, Pune, (India)

ABSTRACT

Now-a-days recommendations systems play a very crucial part in the e-commerce websites and the recommendation systems need to be enhanced and improved on a regular basis. The recommendation system is also used for recommending venues (VRS). The Venue recommendation systems majorly used the users' current location to make appropriate recommendations. However, this system does not consider the other users views for the venues while recommending. A lot of information is stored in the user's views in the form of reviews. This information was hardly ever used by previous recommendation systems; hence, we are exploiting the information generated from such reviews along with user to venue closeness to generate more precise and accurate recommendations. Thus, our system extracts the aspects (features) of the venue reviews and based on that sort venues for the recommendation. Hence our system considers important factors such as user to venue closeness, user's similarity and venue reviews while recommending a venue to a user. These enhancement in recommendations are proved by the results shown by considering the impact of information extracted from the reviews while recommending the venues to the user. Our system provides more accurate and precise recommendation as compared to existing system (i.e. mobicontext) on a large-scale data set.

Keywords: location, Recommendation System, Reviews, Similarity, VRS.

I. INTRODUCTION

As the e-commerce and the social networking sites increases along with various options available for the user it becomes very important to have recommendation systems. Along with that the growing need the recommendation systems there is also a need for us to enhance the recommendation systems. Consider an example of e-commerce website like Flipkart, which has a plethora of products in their inventory, however, recommending the product best suited for the user is a challenge. If the correct products are recommended to the user, users time of browsing through the items is saved also it boosts e-commerce websites sales. There are two types of recommendation systems (a) Content-based Recommender Systems and (b) Collaborative filtering based Recommender Systems. In

Content-Bases Recommendation systems considers the properties of the items recommended. For e.g.Consider Netflix, if a user has watched movies of comedy genre then the content bases recommendations system will recommend the users movies of comedy genre. Collaborative filtering systems recommend items based on similarity measures between users and/or items. The items recommended to a user are those preferred by similar users. Breese et al. [9] described CF algorithms as separable into two classes: memory based algorithms that require all ratings, items, and users be stored in memory and model-based algorithms that periodically create a summary of ratings patterns offline. Pure memory-based models do not scale well for real-world application.

Recommendation systems are widely used in venue recommendation systems as well. Such Recommendation systems are called as “Venue Based Recommendation Systems” or (VRS). With the advancement of technology in these days every individual is equipped with a very powerful device, smart phones. These smart phones can allow to locate the user’s location, also these devices allow to use web based application which give users an option to check-in to a particular venue [2]- [3] and also express their opinion about the venue in form of reviews. There is a lot of data collected in the form of check-ins and reviews. There needs to be an intelligent mechanism to extract the information out of the data that is collected. Collaborative Filtering is widely used in venue recommendation system. Model based and memory based CF methods are used in the “Venue Recommendation System” (VRS). However, there are certain short comings of the existing VRS that they cannot process this large amount of disorganized data. Below are the two major drawbacks of the existing VRS.

Problems faced by Existing VRS:

1. The VRS are suffer with many shortcomings and drawbacks. A major drawback for such system is to process data at a real time and extract preferred venues from a huge number of venues and their reviews.

2. Another challenge is which factors or features to consider while extracting venues from large data set in order to provide recommendations accurately or more precisely.

Hence in order to overcome drawbacks of existing VRS,in our systems we are utilizing venue reviews and extract the required information from the reviews and then exploit that information extracted from the reviews which would improve the recommendation quality.

II. RELATED WORK

In previous venue recommendation systems work was majorly based on trajectorybased approaches [2]-[4]. Such approaches kept record of information about a user's visit pattern to various location, the routes taken. In [4] most popular venues are recommended by using machine learning and data mining techniques on trajectory data. These routes based systems therefore recommend locations to users based on their past routes trace or trajectories. The main drawback of such systems is that they have not considered other dominant factors other than GPS trace and hence resulted in providing poor quality of recommendations. Usually user does not visit many places often, so there are less entries exists in user-venue matrix which results in data sparseness issue. In [3] authors have proposed friend-based collaborative filtering (FCF) approach by evaluating the strong tie between social friends for recommending locations based on ratings given by social friends to commonly visited places.

In [3] authors have considered only friends when processing collaborative filtering for a targeted user, since non-friend users do not have much value for reference in recommendation. In [5] authors have proposed Geosocial DB framework which provides location-based social networking services such as location-based news feed, location-based news ranking, and location-based recommendation. In [5] they have provided recommendations for hotels or restaurants within a range distance from the user's location. Yerach Doytsher et al. [8] have implemented social-based recommendationof routes. In a social-based route recommendation, a user uprovides source and target locations and the goal is to find routes from the source to the target, that are frequently traveled by users who are related to u in the social network by using query language that is based on graph-traversal and ordering operations to speed up the formulation of queries over the network. Most of the above-mentioned approaches have used memory-based CF which provides recommendations to users on the basis of their past entries.

However, such approaches experiences common drawbacks of memory-based CF (e.g. cold start and data sparsity) which degrades their performance.In [12] authors have proposed a method to identify and extract opinion features from online reviews. It identifies candidate features that are specific to the given review domain and yet not overly generic (domain independent). The proposed method considered positive and negative opinions. In [1] authors have proposed Bi-objective recommendation framework for mobile social networks which generates recommendations by considering person's geographical location and venue closeness. But it has not considered venue reviews for recommendations. After surveying we realized Reviews are one of the dominant factors affecting the recommendation as it contains experienced user's opinions which can be greatly leveraged to improve the recommendation as reviews reflect the actual experience of the users who have visited the venues.

III. MOTIVATION

The existing Venue Recommendation Systems seldom used the large amount of data which is present in the form of reviews for venues. Reviews holds the data regarding the Venues and their features. If such information extracted properly and efficiently can be used to effectively to provide more precise and user preferred recommendations. For e.g. David goes to a location ABC. Considering the venue closeness to his location and the expert user check-ins, we get *n* venues in recommendations. However, David cannot go through the reviews of all the *n* locations manually and might not be able to select the best suited venue for him. Our system does the tedious job of going through the reviews and extracting the information (features) and rank the venues according to the expert user's opinions. Which would help the system in recommending most appropriate venues to David based on reviews.

IV. PROPOSED SYSTEM ARCHITECTURE

In the proposed system users GPS coordinates are determined and taken as an input for the VRS. The system retrieves a set of venues near to the user's location. This is done by calculating the geo-spatial distance between the user and the venue. Further we have a set of expert users who have multiple check-ins at various venues. Considering the number of check-ins at a particular venue we refine the set of venues retrieved from the closeness. The more the number of check-ins by expert user the more likely the venue will be shortlisted for recommendation. As reviews play a very vital role in judging a venue it is must to go through the reviews of the venues we obtained in 2nd step. The reviews are processed and the aspects are extracted from the reviews. Ranking is done based on the cumulative opinion of the venue from the reviews. Hence refining the recommendation and further increasing the recommendation accuracy and precision. The proposed system consist of following modules (Fig. (1)):

4.1 User Profiles

Our system maintains a database of the users in form of user profiles. The information regarding the check-ins performed by the user at various venues is stored. A user profile consists of the user identification and the Venues he has visited.

4.2 Ranking Module

Ranking module calculates a set of popular venues and expert users. We consider a venue as popular, if it is visited by many expert users, and a user as expert if he/she has visited many popular venues. The users and venues that have very low scores are not considered for further processing.

4.3 Mapping Module

The mapping module calculates similarity between the current user and the expert users. The similarity among the expert users for a given region is also calculated. The mapping module also calculates venue closeness by considering geographical distance between the current user and popular venues. The mapping module calculates the similarity between the expert users (that were identified by the ranking phase) using Pearson Correlation Coefficient (PCC). The PCC value varies from -1 to +1, where the value close to 1 signify the higher degree of similarity between two users. If this value of PCC is zero or less than zero, then this means the liking or taste of two users do not match. In order to solve data sparseness issue, we have used confidence measure. The confidence measure can be considered as a conditional probability that a venue visited by a one user is also visited by the other user in the dataset. The mapping module also calculates the geographical distance between the current user and the popular venues, the location information about current users and venues are presented as GPS coordinates. Thus, we use Haversian model [17] to compute the user-to-venue distance as follows:

$d = Ra$,

Where the parameter "a"is angular distance in radians between the current user and the venue's geospatial location. The parameter R is earth's radius.

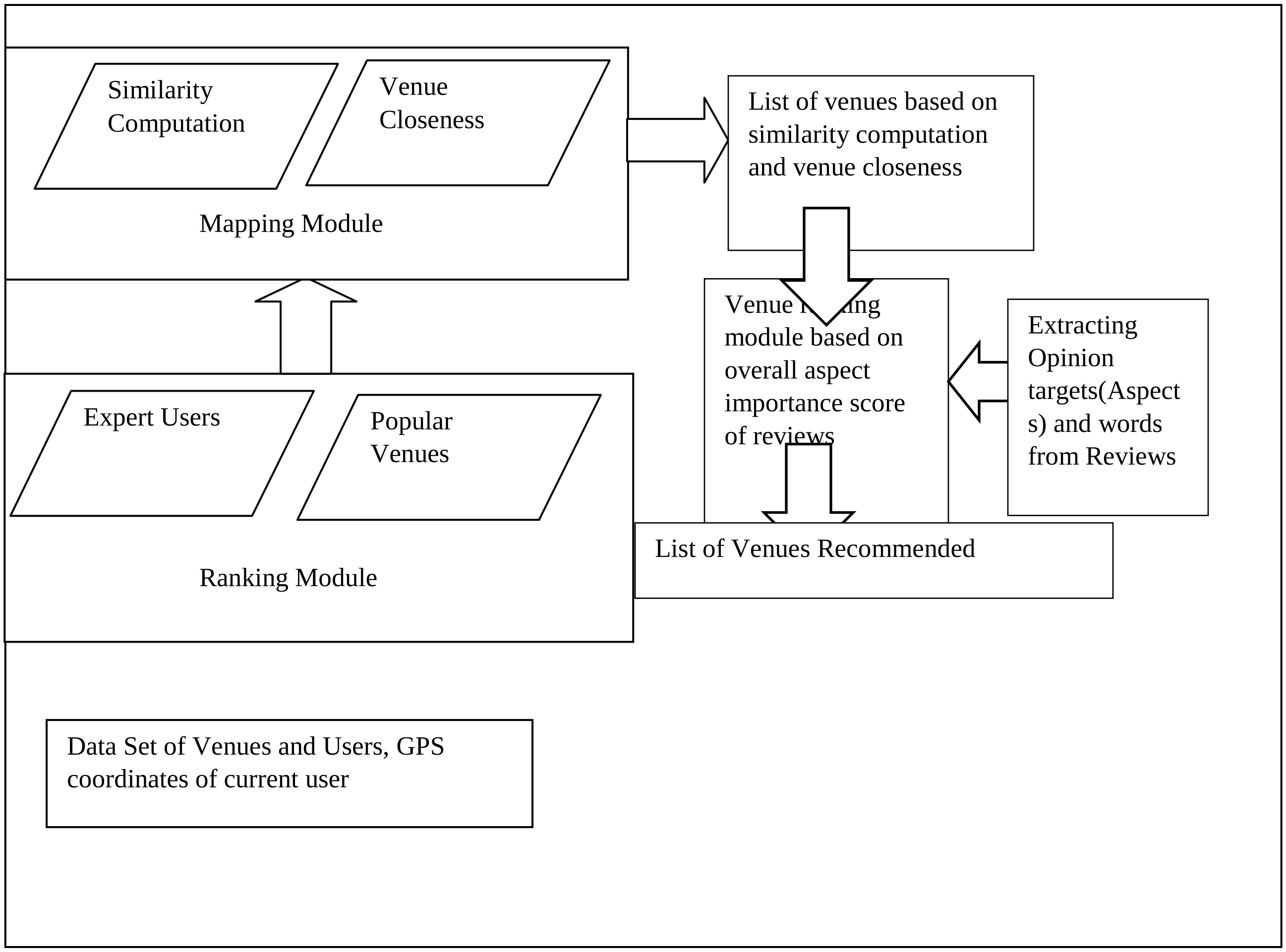

Figure 1: Proposed System Architecture

1.1 Opinion target and word extraction from venue reviews module

In this module opinion targets (aspects) and words(modifiers) from hotel reviews are extracted as a part of data pre-processing phase. This pre-processing is done offline and system admin can define the how frequently this can be done. This can be weekly, bi weekly or monthly task. In this module, we extract opinion words and target lists from venue reviews. This list further given as input to venue ranking module. An opinion target is defined as the object about which users express their opinions, typically as nouns or noun phrases. For example, Food is tasty and delicious but ambience is bad, “Food” and “ambience” are two opinion targets. Opinion words are the words that are used to express users’ opinions. In the above example, “tasty”, “delicious” and “bad” are three opinion words. Once we identify opinion targets and words, we classify opinions as good or bad.

4.4 Venue ranking module

In this phase, we rank venues based on overall aspects importance score of venue reviews. The overall opinion in a review is an accumulation of the opinions given to specific aspects in the review, and various aspects have different contributions in the accumulation. That is, the opinions on important or unimportant aspects have strong or weak influence on the generation of overall opinion. We calculate aspect importance score of each aspect and then calculate the overall aspect importance score of reviews based on that we recommend venues to user.

4.5 Venue Recommendation Module

Based on venue closeness and current user's similarity with expert users, we get optimized list of venues for recommendations. From that optimized list, we select only those venues whose review's overall aspect importance score is high. Thus, we recommend top-N venues based on venue closeness and by considering overall aspect importance score of venue reviews to the user.

V. ALGORITHM

Venue recommendation algorithm

Input: Current User: u, region: R_e

Output: V_{top}= A set of top-N venues.

Notations: U_{loc}=location of current user u, $Venue_u$= set of venues or hotels visited by current user, $Venue_e$= set of venues or hotels visited by expert user e, R_v= set of recommended venues after similarity computation and venue closeness, E_{sim}=set of expert users similar to the current user u, V_{ue}= set of venues visited by expert users but not visited by current user. clo_{eu} =closeness measure of the expert user e with the location of current user u, $S_{ue}$is similarity of the user u with the expert user e. Sv is the set of venues after extracting aspects from reviews and calculating overall aspect importance score of venue reviews.

1: $R_v \leftarrow \emptyset$; $z_{agg} \leftarrow \emptyset$;

2: $E_{sim} \leftarrow$ calculate$simset$ (u, E)

3: for $each e \in E_{sim}$do

4: $V_{ue} \leftarrow \{v: Venue_e | v \notin Venue_u\}$

5: $clo_{eu} \leftarrow max$ (calculatelocclosenes$s$ (U_{loc}, V_{ue}))

6: $zag[e] \leftarrow$ calculateagg (S_{ue}, clo_{eu})

7: end for

8: $R_v \leftarrow$ calculateRec $(u, zagg)$

9: $S_v \leftarrow$ calculateoverallReviewAspectsImportancescore (R_v)

10. $V_{top} \leftarrow$ sort (S_v)

The algorithm considers parameters as: 1. current user identification that generates a recommendation query and 2. Geographical region where the current user is located. The aggregated utility function (Line 2–Line 7) calculates the users' similarity in terms of venue preferences and calculates the user-to-venue closeness. The *c*alculatesimset () function computes similarity among current user u with the expert users by using Pearson Correlation Coefficient (PCC). In Line 5, the function calculatelocclosenes () calculates location closeness between current user and venues of the expert user and then collects those venues of the expert user that are nearest to current user. In line 6 calculateagg () function calculates the overall aggregate similarity with respect to preferred venues and user-to-venue closeness. In Line 8 the algorithm provides the top-N venues for the user by considering both factors such as similarity computations and venue closeness. In line 9 aspects are extracted from reviews and aspect importance score is calculated. In line 10 venues are ranked based on overall aspect importance score of reviews and finally venues with higher overall aspect importance score are recommended to the user.

VI. PERFORMANCE EVALUATION

We compared our approach with the following existing approach (a) Mobicontext(CF-BORF) proposed by R.Irfan, Khalid etc. in [1] in which venue recommendation to the user is based upon user to venue closeness and users similarity.

6.1 Results and analysis

We used "Gowalla" dataset which consist of 1, 82,175 records which consist of attibutes like User_id, Checkin, Latitude and Longitude, Loc_id, Review_file_path ,Hotel_name.

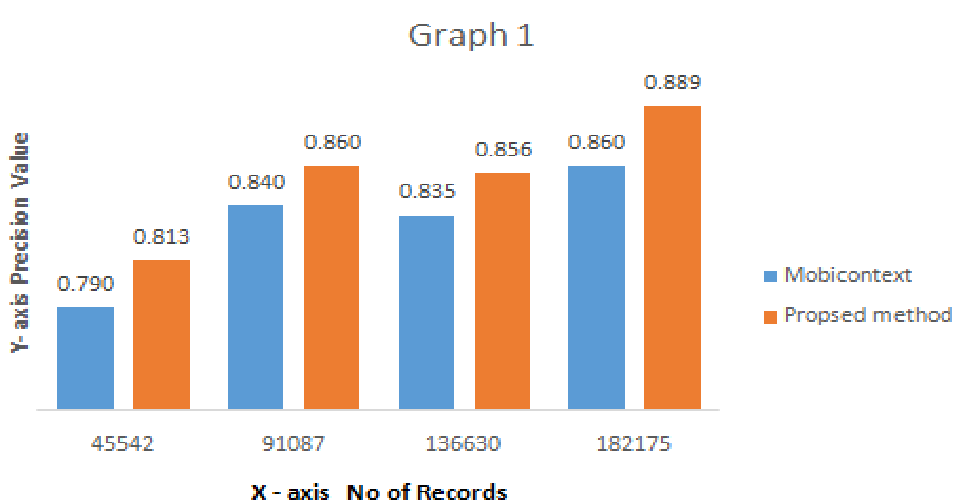

Figure 2: Precision versus number of records in dataset

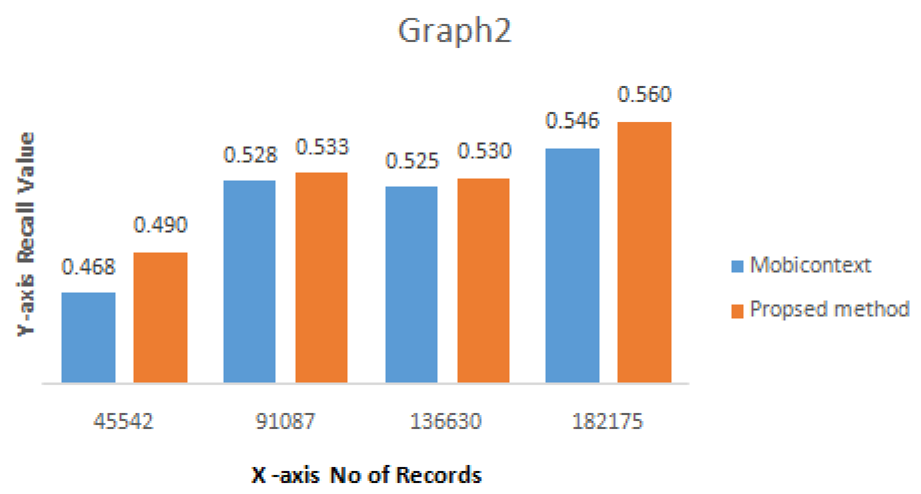

Figure 3: Recall versus number of records in dataset

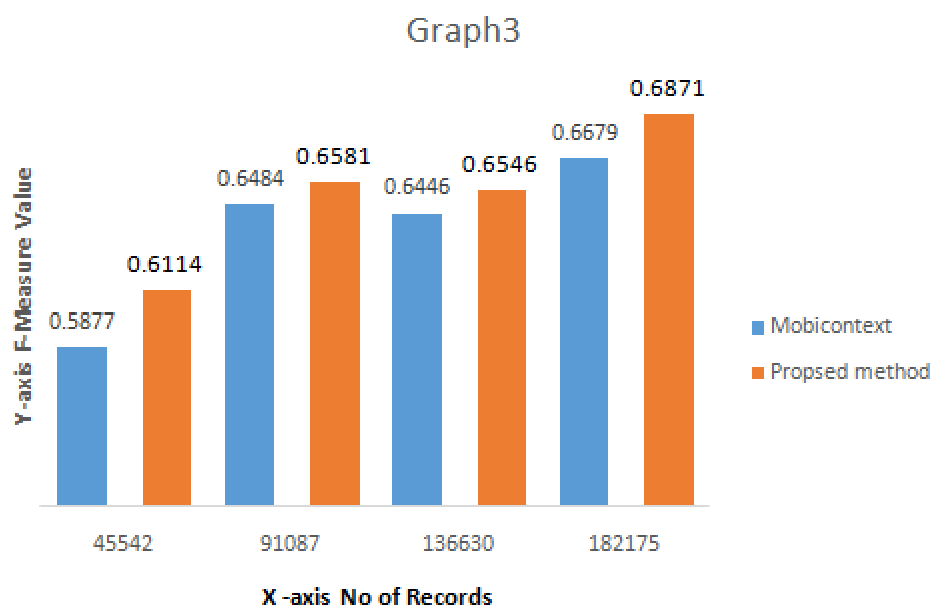

Figure 4: F-measure versus number of records in dataset

We have used the standard performance evaluation matric to assess the performance of our approach with existing method such as (1) precision, (2) recall, (3) F-measure. The precision presents a ratio of the relevant recommendations among the total number of recommendations (true positive (tp)) to the total number of recommendations (tp+ false positive (fp)).

$$Precision = \frac{tp}{tp + fp}$$

$$Precision = \frac{|\{relevant\ recommendations\} \cap \{total\ no.of\ recommendations\}|}{|\{total\ no.of\ recommendations\}|}$$

The recall presents a ratio of the relevant recommendations among the total number of recommendations (true positive (tp)) to the number of relevant recommendations (tp+ false negative (fn)).

$$recall = \frac{tp}{tp + fn}$$

$$recall = \frac{|\{relevant\ recommendations\} \cap \{total\ no.of\ recommendations\}|}{|\{relevant\ recommendations\}|}$$

The F-measure is harmonic mean of precision and recall and is given as follows:

$$Fmeasure = 2 \times \frac{precision \times recall}{precision + recall}$$

Fig. (2),(3),(4) shows the precision, recall, f-measure results w.r.t number of records in the dataset for our proposed approach and existing method (i.e. Mobicontext).As observed in fig. (2), (3), (4) our proposed approach shows better performance in terms of precision, recall, f-measure as compared to existing system (i.e. Mobicontext).This improvements in results is due to consideration of venue(hotel)reviews along with the venue closeness and users similarity. As observed in fig. (2), (3), (4), we verified performance of our proposed system and existing system with the increasing number of records in the dataset and observed that performance gradually increased. Thus our proposed approach provides more precise and accurate recommendations due to incorporation of venue reviews.

VII. CONCLUSION

The present VRS system considered the venue closeness as a major factor while recommending, however with the rise in number of venues it's essential to recommend the best suited venue to a user. Information regarding a venue can be accurately extracted from the reviews of the users as they convey the experience of a user had at a venue. This piece of information is very useful while recommending the venues. Thus, we are exploiting the information in venue reviews along with venue closeness in our system to improve recommendation quality. Ranking the venues based on review's aspects importance score along with the venue closeness gives us more refined recommendation suited for the user. Thus, improving the recommendation precision and accuracy of the existing venue recommendation system. This system can be used for products on the e-commerce websites. In proposed approach, we have incorporated the reviews of the venues in order to recommend the best venues to a user. However, we can improve the recommendation accuracy further by considering the factors such as recommending personalized venues by considering the pages the user has liked on social networking sites like Facebook etc. At times a user prefers not to visit a crowded venue depending on the company and situation, hence we can add a module which will keep a track on of number of check-ins at a particular time of the day.

VIII. ACKNOWLEDGEMENTS

First and foremost, I would like to thank my guide, Prof. M. R. Patil for her guidance and support. I am very much indebted to head of the department Dr. P. N. Mahalle and Principal Dr. A. V. Deshpande, for providing all the facilities needed for successful completion of this paper and providing necessary assistance while preparing for this paper work.

REFERENCES

[1] Rizwana I., O. Khalid, M. Khan, C.Chira, R.Ranjan, F.Zhang, "MobiContext: A Context-aware Cloud-Based Venue Recommendation Framework," *IEEE TRANSACTIONS ON CLOUD COMPUTING, vol. pp, issue 99*, pp. 1-14, 2015.

[2] A. Majid, L. Chen, G. Chen, H. Turab, I. Hussain, and J.Woodward, "A Context-aware Personalized Travel Recommendation System based on Geo-tagged Social Media Data Mining," *International Journal of Geographical Information Science*, pp. 662-684, 2013.

[3] M. Ye, P. Yin, and W. Lee, "Location recommendation for location-based social networks," *In Proceedings of the 18t SIGSPATIAL International Conference on Advances in Geographic Information Systems, ACM*, pp. 458-461, 2010.

[4] Y. Zheng, L. Zhang, X. Xie, and W.Y. Ma, "Mining interesting locations and travel sequences from gps trajectories," *In Proceedings of the 18th international conference on World wide web, ACM*, pp. 791-800, 2009.

[5] C. Chow, J. Bao, and M. Mokbel, "Towards Location-Based Social Networking Services," *In Proceedings of the 2nd ACM SIGSPATIAL International Workshop on Location Based Social Networks, ACM*, pp. 31-38, 2010.

[6] P. G. Campos, F. Díez, I. Cantador, "Time-aware Recommender Systems: A Comprehensive Survey and Analysis of Existing Evaluation Protocols," *User Modeling and User-Adapted Interaction, vol. 24, no.1-2,* pp. 67-119, 2014.

[7] A. Noulas, S. Scellato, N. Lathia, and C. Mascolo, "A Random Walk around the City: New Venue Recommendation in Location-Based Social Networks," *In Proceedings of International Conference on Social Computing (SocialCom)*, pp.144-153, 2012.

[8] Y. Doytsher, B. Galon, and Y. Kanza, "Storing Routes in Sociospatial Networks and Supporting Social-based Route Recommendation," *In Proceedings of 3rd ACM SIGSPATIAL International Workshop on Location-Based Social Networks, ACM*,pp. 49-56, 2011.

[9] J. S. Breese, D. Heckerman, and C. Kadie, "Empirical analysis of predictive algorithms for collaborative filtering," *in Proc.14th Annu. Conf.Uncertainty Artif. Intell*. pp. 43–52, 1998.

[10] O. Khalid; M. Khan; Samee U. Khan; Albert Y. Zomaya, "OmniSuggest: A Ubiquitous Cloud-Based Context-Aware Recommendation System for Mobile Social Networks*," IEEE Trans. Services Computing. vol. 7, no. 3*, pp. 401 - 414, 2014.

[11] M. Hu and B. Liu, "Mining and summarizing customer reviews," *in Proc. 10th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining, Seattle, WA, USA*, pp. 168–177, 2004.

[12] Z. Hai, K. Chang, J.-J. Kim, and C. C. Yang, "Identifying features in opinion mining via intrinsic and extrinsic domain relevance," *IEEE Trans. Knowledge Data Eng., vol. 26, no. 3,* p. 623–634, 2014.

[13] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, "A Fast and Elitist Multi-Objective Genetic Algorithm: NSGA-II," *IEEE Transaction on Evolutionary Computations, vol. 6, no. 2*, pp. 182-197, 2002.

[14] M. Hu and B. Liu, "Mining opinion features in customer reviews," *in Proc. 19th Nat. Conf. Artif. Intell., San Jose, CA, USA*, pp. 755–760, 2004.

[15] L. Zhang, B. Liu, S. H. Lim, and E. O'Brien-Strain, "Extracting and ranking product features in opinion documents," *in Proc. 23th Int.Conf.Comput. Linguistics, Beijing, China*, pp. 1462–1470, 2010.

[16] H. Nasiri, M. Maghfoori, "Multiobjective Weighted Sum Approach Model reduction by Routh-Pade approximation using Harmony Search," *Turkish Journalof Electrical Engineering and Computer Science, vol. 21, no. 2,* pp. 2283-2293, 2013.

[17] S. Seema, and S. Alex, "Dynamic Bus Arrival Time Prediction, using GPS Data," *In Proceedings of the Nat. ConferenceTechnological Trends (NCTT),* pp. 193-197, 2010.

[18] Prajakta Bothe, "A Survey on Venue Recommendation System", *International Journal of Innovative Research in Computer and Communication Engineering, vol.5, issue 4,* pp.7413-7417, April 2017.