International Journal of Computer Science and Control Engineering
2017; 5(2): 5-10
http://www.openscienceonline.com/journal/csce

Improvised Cloud Based Venue Recommendation Framework

Roshni Vikas Misar, Balkrishna Patil, Rajesh Autil

Everest Institute of Technology, Aurangabad, Maharashtra, India

Email address

roshnitiwari2503@gmail.com (R. V. Misar), cseroyal7@gmail.com (B. Patil), rajeshauti24@gmail.com (R. Auti)

To cite this article

Roshni Vikas Misar, Balkrishna Patil, Rajesh Autil. Improvised Cloud Based Venue Recommendation Framework. *International Journal of Computer Science and Control Engineering.* Vol. 5, No. 2, 2017, pp. 5-10.

Received: May 10, 2017; **Accepted**: August 1, 2017; **Published**: September 29, 2017

Abstract

The Recommendation systems is the technological development in technology trend. Most of recommendation systems is totally based on past behavior which is termed as collaborative filtering based recommendation system which suffers some problems such as data sparseness, cold start etc. This paper is totally based bi-objective recommendation framework (BORF). Bi-objective framework is cloud based framework for mobile social network. This paper proposed a system for venue recommendation on the basis of rating given by user and nearest place to user on the basis of longitude and latitude. This paper optimized both scalar and vector optimization. Some algorithms are used for scalar and vector optimization as weighted sum approach and evolutionary NSGA-II respectively.

Keywords

BORF, Collaborative Filtering (CF), Non-dominated Sorting Genetic Algorithm (NSGA-II)

1. Introduction

There are huge number of business development social services are available in internet such as amazon, lime road, flip kart, go Walla etc. there are so many users follow this social sites for their need has been resulted in large volume of data are consumed by service on daily basis by service providers. This may cause data retrieval issue while filtering related information. So that for overcoming this issue the most enormous technique has been developed which termed as data retrieval system as recommendation system. While searching any product on any of the social service it gives result in the form of past behavior pattern means on the basis of history and similarity features to product which done by retrieval systems, recognized as Recommendation Systems.

1.1. Research Motivation

The central fundamental of popularity of e-business application is recommendation system. Recommendation systems gives results on the basis of user's interest [1]. For example, while considering social e-business site amazon has built-in recommendation system that provides users with improved recommendations for various things of interest. Recommendation systems customs a different kind of data recognition procedures on the basis of historical search record and current product similarity or best match on a user's preferences. There are various social networking sites such as Facebook, twitter has lot of attraction by various subscribers [1] [6] so it accumulates a large volume of data which faces some problems that discussed in next section. Large volume of data means multiple check-in done by or feedback given by user [2] [3]. on the basis of data accumulation by different social sites, few Venue-based Recommendation Systems (VRS) has been developed [2] [3]. On the basis of user's choices this system gives recommendation or suggestion of venues to users that closely match with user's choice. Every system have some limitations this VRS is also having some limitations and challenges. This VRS system is developed to process large amount of information and extract preferred venue from large dataset on the basis of users past check-in and similarity preferences. [2] [3] [12] [13]

1.2. Research Problem

While took survey on venue recommendation system it found that lots of work such as [1] [6] and [13] has been applied on collaborative filtering (CF) which is totally based on historical information which is difficult for recommendation system. The CF-based approaches create recommendations on the basis of similarity and routine pattern of user which suffers several limitations and challenges. [1] [2] [5] and because of that limitations the performance of recommendation system may get degrades. One of the most common problem occurs as termed as cold start problem which occurs when recommendation system has to suggest venue or place that is newer to system means if user want to search venue which is totally new, the records which is not similar to past data. The next issue is data sparseness occurs in collaborative filtering due to insufficient number of transaction and feedback from user. In this case recommendation system has to wait for sufficient entries. Scalability is the issue occurs in recommendation system. As rapid growth of e-business of social mobile sites has millions of check-in done by many users which cause difficulty to find similar user matrix among large volume of data.

2. Literature Survey

In a Context-aware Modified Travel Recommendation System based on Geo-tagged Social Media Data Mining journal it found that digital cameras and camera phones has been added for sharing photos and videos on internet groups. Large volume of data in the form of photos and videos has been accumulated from dataset. The photos which are present on internet having some recorded information such as tags, title, note and description those photos also having information related to time, at what time photo has been taken means temporal context also having information related to longitude and latitude of place where photo has taken. The objective of this journal is to predict tourist preferences in new or unknown city more precisely & generate better recommendation compared to other state of art landmark recommendation method. The main drawback of such type of approaches the system is not able to consider important factor apart from simple GPS trace which results into less optimal recommendation.

In Location based & Preferences aware recommendation using sparse geo-social Networking data focus is on location recommendations in the perspective of social networking, it is on user's geo-social accomplishments. This system is grouped into two parts as standalone which gives exact location that match with user's preferences and another is sequential location which provides series of locations such as popular travel route in city based on user's preferences with constraints as time and cost. But this type of Recommendations in LBSNs can be costlier. As LBSNs is used in this system that is why it can be very active. They visit so many locations within very short time spans, which adds data related to their preferences at a high degree.

In interesting locations and travel sequences from GPS trajectories mining of interested locations and classical travel sequences by using the GPS paths created by many users. This information used to understand the correlation between users and locations, and it will enable recommendation for travel as well as mobile tourist guidance. So that concern a person's visit to a place that is link from the person to the location and weight these links in form of users' travel experiences in various sections. This work will have extended to improve the efficiency of sequence mining and grouping users on the basis of their location histories or clustering locations in terms of people's visits.

A recommendation method based on personalized random walks is system over a graph that used to encode links between users, objects and contextual attributes. The general theme of the research in recommendation systems has been totally focused in improving accuracy, variety and originality. This system is best for in term of security with multidimensional information. In this system the performance of recommendation may accurate.

In Route sin Socio-spatial Networks and Supporting Social-based Route Recommendation system the co-ordination of cellular phone and GPS trajectory of user with traveled route for location based has been developed. The graph model has been developed for storing socio-spatial information. This model stores frequently traveled route with user information. Query language has been used for this type of recommendation system.

In Non-Dominated Sorting Genetic Algorithm for Shortest Path Routing Problem in Computer Networks multiobjective evolutionary algorithm is used based on the Non-Dominated Sorting Genetic Algorithm (NSGA) which is used for solving shortest distance routing problem in computer networks. For population initialization priority based encoding algorithm has been used. The outcome obtained from NSGA and single objective weighting factor for genetic algorithm has been compared.

3. Proposed System

3.1. Training Module

A self-learning system is proposed that will also take into account, past user searches and consider scalar as well as vector queries to recommend appropriate results.

3.2. Quality of Predictions

There will be a definite prediction is that system want better recommender to make noble recommendations. So that this system performs better than any "dumb" guess algorithm which just uses universal data which is as an average rating for items.

3.3. Speed/Scalability

Maximum recommender systems work in a marketable and/or accessible setting and it is important that they can start building recommendations for a user nearly instantly. It means that the algorithm will not take too long to make any predictions - it has to work, and work fast. speed is directly related to scalability of the algorithm. Again, systems in a profitable and/or online setting can have a massive dataset. The algorithm must keep its speed even if there are many billions of ratings.

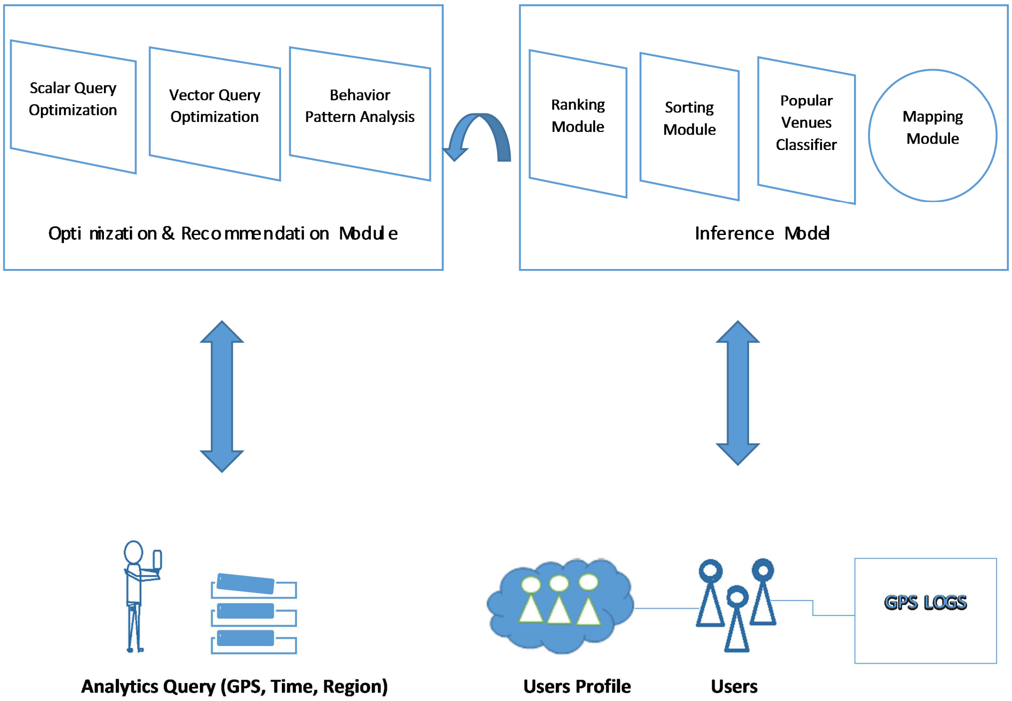

Figure 1. Architectural diagram of VRS.

3.4. Easily Updated

The datasets which is used is behind recommender systems are constantly being updated with updated ratings from users. The algorithm which is used must handle this updated information rapidly and efficiently. So that if the algorithm for required taking too much time for debug then it may definitely cause some issues and user may get disturbed so it might miss chance to make recommendation.

3.5. User Profiles

As shown in Figure 1 the module has made on VRS keeps records of user's profiles for each geographical region. users to venues arrow indication at lower right of Figure 1 shows the entries performed by each user at many venues. A user's profile holds user's identification, venues with longitude, latitude date with check-in time at a venue.

3.6. Ranking Module

This module having important aspect because when user visits venue, user will give ranking in numbers from 1 to 5 and based on this venue it will goes into recommendation list. It is present on the top of user profile. The ranking module completes functionality during the pre-processing phase of data improvement. The pre-processing can be done in the form of periodic batch jobs running at monthly or weekly basis i.e. handled by system administrator. HA interface method applies on ranking module with user profile to give optimal ranking.

3.7. Mapping Module

In mapping module similarity graph generated in between expert users and popular venues for selected region and this work is done during preprocessing phase. The result of similarity graph calculation is generation of network of likeminded people who shares similar liking for various venues they visit in geographical region. It also maps closest venue to current user who are searching for best recommendation into current context with longitude and latitude.

3.8. Recommendation Module

This module developed with the help of apache mahout recommender. In this module user may visit place and if user likes that particular place then there is option for user to rate the place in the form of number range from 1 to 5. this data will save into local host. In this module the current location of user gets automatically fetched with longitude, latitude, date and check-in time. Recommendation is done on the basis of mapping and ranking which is done in preprocessing phase. It is categorized into two form as nearby venue and rating based. If user wants to search a place with nearby distance, then recommendation module suggests a top N venue which is nearby to current location. And if user wants a location which is based on rating then this module recommends top N venues on the basis of best rating. In recommendation framework both scalar and vector optimization utilized so that it is termed as bi-objective framework.

4. Performance Analysis

UCF generates like-minded users who visited the similar venues in the past are most likely visit venues in the future [1], [3].

Existing dataset "GOWALLA" consists of 6,442,890 check-ins performed by 150,734 users in total number of 1,280,969 venues [6]. That's why this system has performed general experiments on internal Open Nebula cloud setup which is execution on 96 core Super Micro Super Server SYS- 7047GR-TRF systems. From selected dataset from the total records, 80% of the record is used as the training set and 20% constitute test set for the evaluation. While considering every data point, system did 25 self-determining runs so that system used a standard 5-fold cross validation method for evaluating accuracy rate of the framework. By using three standard evaluation techniques to calculate proposed recommendation framework as: (a) precision, (b) recall, and (c) F-measure.

The following figures present precision, recall, and f-measure results for showing better performance in terms these three techniques. System get Such improvement in results because of pre-processing phase decreases the negative effect of data sparseness over recommendation quality. Data sparseness results in unique similarity values in collaborative filtering, and with large number of zero entries in user-to-User.

Mathematical model:

Table 1. *Experimental Result for NSGA Algorithm.*

f(1)	f(2)	Precision	Recall
0.1	0.5	0.8534	0.9365
0.2	0.3	0.8544	0.9399
0.4	0.1	0.8664	0.9341
0.5	0.6	0.8794	0.9374
0.6	0.1	0.8914	0.920
0.5	0.9	0.8934	0.929
0.9	0.7	0.8993	0.960

Notations: tp -True positive, tp –False positive, fn –false negative

Precision is ratio of accurate recommendations (tp) to total no of expected recommendations. (tp+fp). And accurate recommendations can be defined as recommendations that has been predicted accurately in top N recommendations.

$$\text{Precision} = \frac{\text{tp}}{\text{tp+fp}} \quad (1)$$

Recall is the ratio of correct recommendations to the total no of recommendations.

$$\text{Recall} = \frac{\text{tp}}{\text{tp+fn}} \quad (2)$$

F-Measure is the harmonic mean of precision and recall.

$$\text{F} - \text{measure} = \frac{2\times\text{Precision}\times\text{Recall}}{\text{Precision+Recall}} \quad (3)$$

Graph gives Improvised Recommendation Module using Apache Mahout demonstrates better performance in terms of precision and recall as compared to t rest of the schemes (NSGA-II and G-BORF). While the performance of the NSGA-II and greedy-BORF present slightly lower because of the aggregation method that maps the users' preferences and location closeness into single objective function. This aggregation may not provide accurate results especially when there is tradeoff between the user's preferences and location closeness. For instance, in the case of G-BORF, when there is no similarity between two users' preferred locations, the venue will be suggested to the current user on the bases of user-to-venue closeness this result does not give optimal recommendation which indicates lower performance in terms of precision and recall presented graph.

Multi-objective performance measure for Improvised Recommendation Module using Apache Mahout:

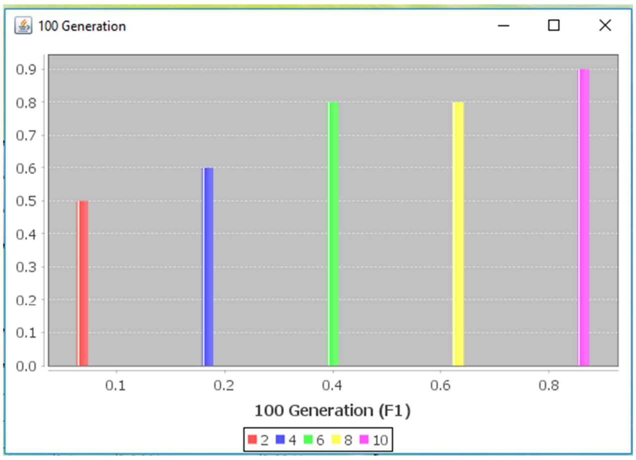

Figure 2. *first Generation graph.*

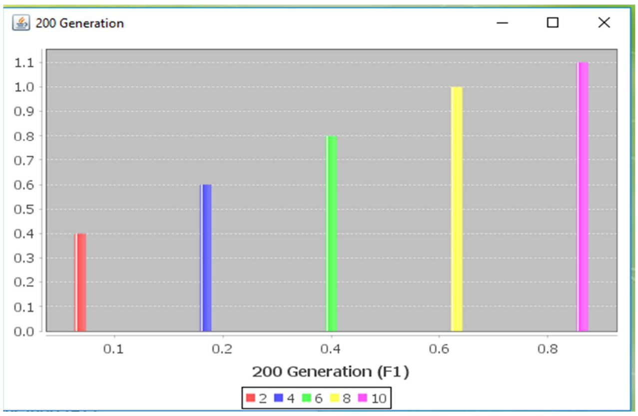

Figure 3. *second Generation graph.*

Graph shows the impact of user's preferences (vertical axis) and venue closeness (horizontal axis) on the recommendations. Numerous simulations are performed to show the effect of generation size on Improvised Recommendation Module using Apache Mahout's performance. It can be observed from Figure 2 that with the generation size of 5, the solutions are not converging, which indicates small generation size does not yield good recommendations. For generation size of 100 from Figure. 2, the solutions appear to converge slightly, thereby improving recommendation quality. Figure 3, Presents the performance of Improvised Recommendation Module using Apache Mahout by increasing the number of generations to

200, which shows the maximum convergence with improved solution quality in terms of recommendations. The convergence Bar depicted by Figure 3, Contains better spread and solutions by considering the two objectives (users' preferences and location closeness).

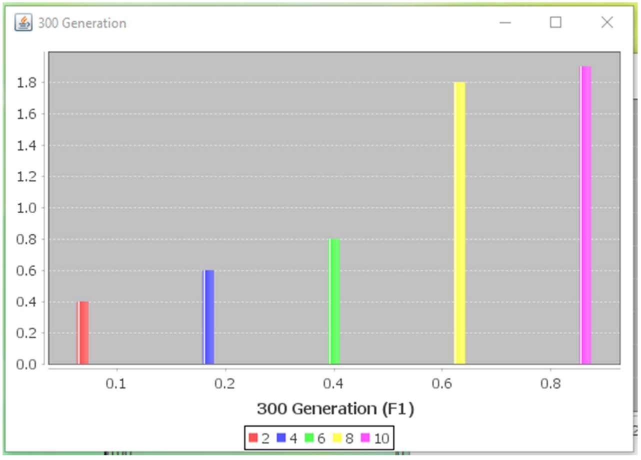

Figure 4. *third Generation graph.*

Table 2. *P-SCORE of NSGA-II with Other Algorithms.*

Algorithm	P-Score
CF-BORF	0.96
G-BORF	0.97
MF	0.99
RWR	0.97
UCF	0.97

The Experimental Results of Improvised Recommendation Module using Apache Mahout

Table 1 presents performance of both objective functions in terms of precision and recall. The Improvised Recommendation Module using Apache Mahout demonstrates better performance by increasing the number of generation size. For instance, in case of generation, the precision of recommendations is greater than 0.02. Similarly, in case of increasing generation size the precision is greater than 0.08. The improved performance is because of the fact that the Improvised Recommendation Module using Apache Mahout generates the optimal recommendations by maximizing both the objective functions simultaneously. So that it will produce result as those venues are suggested to the current users that are not only similar to user's preferences but also located in the closest proximity of user's current location. While the number of generation size increases all dominated solution may get decreases.

5. Conclusion

This system offered a cloud-based venue recommendation framework which produces optimized recommendations by considering the trade-offs between real-world physical factors, such as person's geographical location and location closeness. The importance and originality of the venue recommendation framework which is the reworking of collaborative filtering and bi-objective optimization approaches, such as scalar and vector. In venue recommendation framework data sparseness problem is addressed by integrating the user-to-user likeness computation with confidence measure that computes the size of similar interest indicated by the two users in the venues commonly visited by both of them. Also, a solution to cold start issue is deliberated by introducing the inference model which is combination of different modules that assigns ranking to the users and has a precompiled set of popular unvisited venues that can be recommended to the new user.

References

[1] Internal journal of Advance scientific research and engineering trends || Volume 1 ||Issue 5 ||Sept 2016||ISSN (Online) 2456-0774 improvised cloud based venue recommendation content aware using KNN method.

[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 18th 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 Routesin 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] S. Seema, and S. Alex, "Dynamic Bus Arrival Time Prediction, using GPS Data," In Proceedings of the Nat. Conference Technological Trends (NCTT), pp. 193-197, 2010.

[10] B. Chandra, S. Bhaskar, "Patterned Growth Algorithm using Hub-Averaging without Pre-Assigned Weights," In Proceeding of IEEE International Conference on Systems, man, and Cybernetics (SMC), pp.3518-3523, 2010.

[11] B. Hidasi, and D. Tikk, "Initializing Matrix Factorization Methods on Implicit Feedback Database," Journal of Universal Computer Science, vol. 19, no. 12, pp. 1835-1853, 2013.

[12] C. Chitra and P. Subbaraj, "A Non-Dominated Sorting Genetic Algorithm for Shortest Path Routing Problem in Computer Networks," Expert Systems with Applications, vol. 39, no. 1, pp. 1518-1525, 2012.

[13] Y. Wang, S. Wang, N. Stash, L. Aroyo, and G. Schreiber, "Enhancing Content-Based Recommendation with the Task Model of Classification," In Proceedings of the Knowledge and Management, pp. 431-440, 2010.

[14] J. Bobadilla, F. Ortega, A. Hernando, A. Gutiérrez, "Recommender Systems Survey," Knowledge-Based Systems, vol. 46, pp. 109-132, 2013.

[15] J. Bao, Y. Zheng, M. F. Mokbel, "Location-based and Preference Aware Recommendation using Sparse Geo-Social Networking Data," In Proceeding of 20th International Conference on Advances in Geographic Information Systems, ACM New York, pp.199-208, 2012.