CLOUD BASED MOBI-CONTEXT HYBRID FRAMEWORK FOR VENUE RECOMMENDATION

Shruti. I. Timmapur* and Dr. M. M. Math**

*KLS Gogate Institute of Technology Belagavi, Vishweshwarayya Technology University, Karnataka, India
shrutitimmapur@gmail.com
**Professor, Computer Science and Engineering, KLS Gogate Institute of Technology Belagavi, Karnataka, India.
mmmath@git.edu

ABSTRACT: Recently, suggestion frameworks have seen important development in the field of information designing. A large portion of the current proposal frameworks construct their models in light of collaborative filtering approaches that make them easy to actualize. Even though, there are many filtering techniques but execution of the existing filtering based proposal framework suffers from difficulties such as cold start, data sparseness, and availability. Proposal issue is regularly characterized by the nearness of numerous incompatible goals or choice variables, for example, clients' preference and venue closeness. Mobi-Context is a hybrid cloud-based Bi-Objective Recommendation Framework is proposed for versatile informal organizations. The Mobi-Context uses multi-target optimization techniques to produce customized proposals. To give solution for the issues relating to cold start and data sparseness condition, the BORF performs information preprocessing by utilizing the Hub-Average (HA) inference model and Weighted Sum Approach (WSA) is actualized for scalar optimization and NSGA-II is connected for vector optimization to give ideal recommendations to the clients around a venue.

KEYWORDS: Bi-Objective Recommendation Framework (BORF), Collaborative Filtering, Non-Dominated Sorting Genetic Algorithm (NSGA-II).

INTRODUCTION

Recommendation frameworks are a branch of information filtering structure that look to predict "rating" that a client would provide for a thing (such as music, books or place) or social component (e.g. individuals or bunch).even though variety of approaches have been developed for Recommendation frameworks, the enthusiasm for this zone still stays high because of interest on developing applications, which can give customized suggestions and manage data over-load. Recently, these frameworks are exceptionally normal and connected in area of learning designing. This framework normally delivered the rundown of suggestions through collaborative filtering approach. However many difficulties are associated with collaborative based suggestion frameworks, for example scalability, data sparseness and cold start. In addition, there are additionally few different issues within choice variables like venue closeness with clients' preference
Recently, development of various portable long variety of social conversation administrations, for example, Face book and Google Latitude has essentially gained the interest of many users (Majid A, 2013)(Noulas A, 2012). A portable extensive choice of social conversation administration permits a client to give a check-in that is a little information of the place visited by the client. Extensive number of registration on regularly bases results in the gathering of monstrous volumes of information. Considering the information stored by such administrations, a few Venue-based Recommendation Systems (VRS) was produced (Majid A, 2013)(Ye M, 2010). Such frameworks are proposed to perform suggestion of venues to clients that mostly match with clients' preference. Despite of having exceptionally promising features, the VRS suffers with various difficulties. A significant exploration challenge for such frameworks is to process information at the real time and extract favored venues from an enormously immense and assorted dataset of clients' historical check-ins(Majid A, 2013)(Zheng Y, 2009). Further complexity to the issue is added by likewise taking into the record the continuous relevant data, for example,

(a) Venue choice by using client's preference.
(b) Venue closeness based on geographic data.

Motivation of the work

Several suggestion frameworks are proposed that depend on collaborative filtering and content based filtering, however these have some issues in providing accurate results (Bobadilla J, 2013). It is required to explore new strategies which will decrease the difficulties and give suggestions in an extensive variety of uses while considering the quality and security viewpoints. So a powerful framework is necessary to be planned which will provide recommendations.

RELATED WORK

The progressing rapid progress of Internet and easy accessibility of different e-trade and informal communities administrations, for example, Amazon, Foursquare gathered lot of data through administration provider regularly (Majid A, 2013)(Zheng Y, 2009). The continuous collection of massive amount of data has motivated center of research community from essential information recovery problem to filtering of applicable information, in this manner making it more significant and personalized to client's requirement. So, majority of research is currently synchronized towards preparation of more smart and independent data recovery frameworks, known as Recommendation Systems. Numerous analysis results are formed and subsequent section summarizes the efforts performed by research organization.

Check-in based methodologies are used where the clients give little inputs as check-in about the place went by them (Chow C, 2010). Apart from rating based methodologies, few of the strategies have their models depend on registration based methodologies where the clients give little information as check-in about the places they have visited. Check in based methodologies works based on memory-based CF that empowers such ways to react with give proposals to clients based on their history sections (Ye M, 2010). However, such methodologies experience the outcomes of normal disadvantages of memory-based CF which decrease their execution. The idea of suggestion system was familiar with manages the difficulties of data over-burden, to investigate the substantial data sets and to improve the most significant data (Limin Zhang). Through customized proposals, clients are recommended about the things or administrations on the based on buys of comparable things or administrations by alternate clients. To finish the task of suggestions, the recommender frameworks as rule utilize each of the accompanying proposal approaches such as Collaborative Filtering, Content Based filters, Hybrid filtering etc.

The CF-based methodologies in VRS have a tendency to create proposals in light of the closeness in activities and schedules of clients (Limin Zhang). The CF approach will work using the premise of observed behavior of users while collaborating with the frameworks and extract out those with the relative behavior. The CF based methodologies are further sorted into memory based and model based calculations. Memory construct calculations work with respect to the user rating matrix and make suggestions in sight of the things evaluated by the clients before. Then again, the model based calculations utilize the client ratings to take in the model that in this way are utilized to perform the assignment of prediction.

Content based filtering method gives suggestions in sight of substance things that were focused by the clients in the past searches (Majid A, 2013)(Limin Zhang). By contrasting different applicant things and the things evaluated in past by various clients, the best coordinating things are prescribed. Recommender frameworks that utilization a combine of two or additionally separating procedures are called hybrid frameworks. Such frameworks are asserted to have enhanced suggestion exactness by successful the disadvantages of individual methodologies. Despite of providing accurate services the existing framework suffers from many problems. The following are the most prominent factors which affects performance of numerous current CF-based proposal frameworks:

- Cold Start: This issue happens when a proposal framework needs to give recommendations to user who is more recent to the framework. Lacking check-ins for the new client results in zero similarity value that corrupts the execution of the proposal framework (Zheng Y, 2009)(Bao J, 2012)(Noulas A, 2012). The main path for the framework to give suggestion in such situation is to wait for sufficient check-ins by the client at various venues.
- Data sparseness: Many existing suggestion frameworks experiences data sparseness issue, it happens when clients have visited to just some number of venues (Alex, 2010). This results into a sparsely filled client to venue registration matrix. The sparseness condition of such matrix makes trouble in finding sufficient reliable comparable clients to produce great quality suggestion.
- Scalability: Majority of conventional proposal frameworks experiences scalability issues (Zheng Y, 2009)(Alex, 2010). The quick and dynamic progress of number of clients causes recommender framework to parse a big amount of check-in records to situate the arrangement of comparative clients.

The trajectory based methodologies record data around a client’s visit design (as GPS directions) to different areas, the routes taken. Despite the fact that, trajectory based methodologies prescribe areas to clients in light of their past trajectory, important disadvantage of such methodologies is that same time they can't consider other variables separated from basic GPS follow that makes them deliver less ideal proposals (Bao J, 2012). Another issue is that the trajectory based methodologies experience data sparseness issue as normally a client does not visits numerous places, which results in insufficient client venue matrix. In addition, the trajectory based methodologies experience scalability issues as very big amount of trajectory information should be processed which will cause impressive overhead.

To address all the issues discussed above a cloud-based framework is proposed which make use of bi-objective development strategies named as CF-BORF and greedy BORF. The GA-BORF is based Genetic Algorithm uses NSGA-II (D, 2012) to improve the venue proposal issue. A pre-preparing stage is presented which performs data refinement utilizing Hub-Average (HA) interface.

PROPOSED SYSTEM

Proposed cloud-based system that produces improved suggestions by all the while considering the exchange offs among genuine physical variables for example, individual's land area and area closeness likewise makes utilization of more logical data as target capacities, for example, the check-in time, clients' profiles, and interests. The fundamental issue is partitioned into sub problems and it provides solutions for adaptability, data sparseness, and cold start issues. The proposed framework comprising of bi-objective advanced techniques named as CF-BORF and Greedy BORF. The Genetic Algorithm based BORF (GA-BORF) uses Non-dominated Sorting Genetic Algorithm (NSGA-II) (D, 2012) to advance the venue proposal issue. A pre-preparing stage is presented which performs information refinement utilizing Hub-Average (HA) interface. The unified design for venue proposals should at the same time consider clients' preference, check-in history, and social setting to create ideal venue suggestions.

After careful analysis, the system has identified to have the following modules

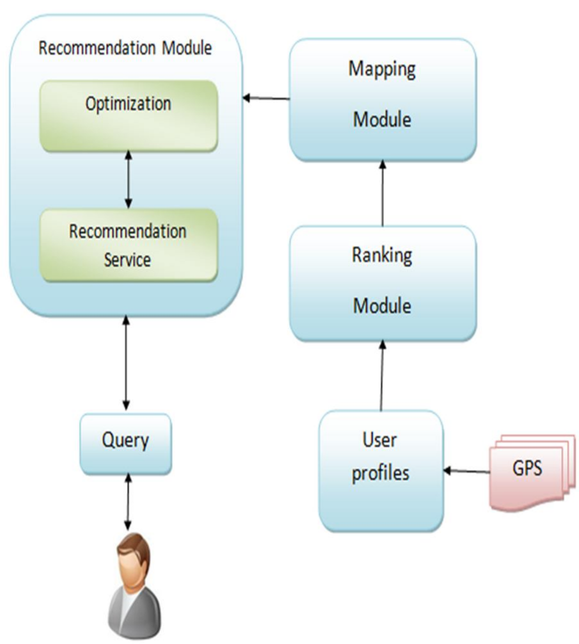

Fig. 1. Venue Recommendation Framework Architecture

User Profiles

The planned framework keeps up records of clients' profiles for each land district. A client's profile comprises of the client's recognizable proof, venues went by the client, and check-in time at a venue.

Ranking Module

Beyond clients' profiles, the mapping module performs usefulness among the pre-processing period of information refinement. The pre-handling can be performed as irregular cluster occupations running at month to month or week by week premise as arranged by framework manager.

The ranking module utilizes model-construct HA inference technique with respect to clients' profiles to set out positioning to the arrangement of clients and venues including shared support relationship. The thought is to take out an arrangement of well-known venues and expert clients. A venue is called as popular, which is visited by numerous expert clients and a client as expert in the event that when he has gone to various popular venues. The clients and venues that have low scores are pruned from database along with logged off pre-handling stage to decrease the online calculation time.

Mapping Module

The mapping module registers similarity charts among master clients for a given conditions among pre-processing stage. The motivation behind comparability diagram allocates comparative preferences for different venues which calculation is to create a system of similar individuals who are visited by them in a geographical location. Mapping module likewise registers venue closeness in view of geological separation between the present client and famous venues.

Recommendation Module

Online suggestion module has a service which is used to get suggestion from clients. A client's queries comprises of: current context, (for example, GPS area of client, time, and region) and a limited region encompassing the client from where the top N venues will be chosen for the present client (N is number of venues).The suggestion administration passes the client's query to advancement module that uses scalar and vector optimization strategies to create an ideal arrangement of venues. In proposed structure, the scalar optimization strategy makes uses of many techniques such as CF-based methodology, greedy heuristics to create client favored suggestions. The vector optimization method, in particular GA-BORF, uses transformative calculations, for example, NSGA-II to deliver upgraded suggestions.

RECOMMENDATION MECHANISM

The online suggestion module uses bi-objective advancement to produce an upgraded rundown of venues. Assume a present client is interested on some venue type that must be found nearest to the present area of the current client inside a particular location. In such a situation, the present client requires the best favoured venues too as the nearest venues from the client's present area. To meet both the previously stated destinations, bi objective enhancement is used in the proposed Mobi-Context proposal structure. The advancement module all the while boosts the accompanying two destinations: popular venues and venues' closeness.

Collaborative Filtering Technique

The CF-based methodologies in VRS have a tendency to produce proposals taking into account the likeness in activities and schedules of clients (Tikk, 2013). The CF approach works on the premise of watched practices of the clients while cooperating with the frameworks and channels out those with the comparative conduct. The CF method has been utilized as a part of various sorts of recommender frameworks, for example, e-business, e-learning and tourism. The CF based methodologies are further arranged into memory based and model based calculations. Memory construct calculations work in light of the client rating network and make suggestions taking into account the things appraised by the clients before. Then again, the model based calculations utilize the client evaluations to take in the model that consequently are utilized to perform the task of expectation.

CF-BORF Based Venue Selection
Input- Current User C, Region R
Output- Toprec = A set of top N venues
Definition, - Ve = set of venues visited by expert user e, Nc = set of recommended venues, Lc = location of current user c, Vc = ser of venues visited by current user, Sr = set of expert user similar to current user c, Cce=closeness measure of expert user e with location of current user c, Sce is similarity of user c with the expert user e.

1. Nc ← 0; Zagg <- 0;
2. Sr ← computesimset (c, E)
3. for each e ε Sr do
4. S ← {v: Ve|v ε $\cancel{Vc}$}
5. Sce ← max (computsimD (Lc, S)
6. Zagg[e] ← Computeagg (Sce, Cce)
7. End for
8. Nc ← ComputeRec(C, Zagg)
9. Toprec ← sort (Nc)

Greedy Filtering Technique

Greedy approach produces an arrangement of top- N venue proposals by navigating a graph of the expert clients (Tikk, 2013).. The proposed approach appoints a starting weight on the connections among hubs in the chart of master clients. In this way, the venues are suggested by those clients that are not just the most like the current client; additionally give top level input of the venues that should be prescribed to the current client. Thus, the Greedy-BORF approach finds an ideal way on the chart that conveys an aggregate assessment about venues by a gathering of expert clients.

Greedy-BORF approach for Venue Recommendation
Input- current user=s, Type=C, Region =R
Output –A set V' top- N venues visited by expert user similar to current user.
Definitions -Kj= neighbor set of node j, δij= edge count between I and j, η (i, j) = 1/δij and Zj= number of required venues found at a node j, visited list=0.

1. a←c; δ←1;
2. Gc← getSimGraph(C,R)
3. Ka←{x:Gc|sim(a,x)>0)
4. visitedlist←a
5. Sort Ka in terms of [Sim (a,x) × η(I,j)], j ε Ka (descending)
6. for each e ε ka do
7. S← {v:Ve|v ε Ɐa}
8. M←M.append(e,S)
9. visitedlist←visitedlist u {e}
10. end for
11. if venueCount(M) >= N then
12. go to line 23
13. else
14. ¥ j ε Ka, set a ←j, such that we have arg max[Sim(a,j) × η(i,j) ×$\frac{Zi}{N}$] ^ kj ≠ Ø ^ ¥g ε Kj | g ε visitedlist
15. if No any such node found in step 15 then
16. go to line 22
17. else
18. δ←δ+1;
19. go to Line 6
20. end if
21. end if
22. D'←computelist(Lc,M)
23. V'= aggregateranking(M,D')
24. return v'

NSGA II Based Selection Technique

NSGA-II in view of its across the board fame in solving multi objective enhancement issues. It has been appeared previously that NSGA-II (D, 2012) can discover better spread of arrangements, and better union close ideal arrangement, with low multifaceted nature contrasted with numerous other partner calculations. The NSGA-II calculation proposes ideal top-N proposals and is separated into two stages: suggestion era and suggestion enhancement.

NSGA II based venue selection
Input – R=set of recommendations.
Output – Top N recommendations based on bi-objective optimizations.
Definitions – Pop=set of populations, Epop=set of population after evaluation, gen=number of generations, Qt=Set of top –N optimized recommended venues, Psize=total size of population.

1. Parents← 0; fl←0;
2. pop←randopo(psize,R)
3. Epop←evaluate(pop)
4. PP←nondominsort(Epop)
5. S←selectParent(PP,Psize)
6. Qt←crossoverMut(S,Pcros,Pmut)
7. While(gen<=maz_gen)
8. CC←evauate(Qt)
9. Rc←PP U Qt
10. F←nondominsort(Rt)

```
for each fi ε F do
      CDA ←cda(fi)
      if size(parent) > Psize
          fl←i
      else
          parent←parents U fl
      end if
end for
if size(parents)<psize then
      fl←ccf(fl)
parents←parents U fl
end if
   S ←selectParent(parent,Psize)
   pop←Qt
    Qt← crossoverMut(S,Pcros,Pmut)
end while
return Qt
```

ANALYSIS RESULTS

The outcomes are compared with related schemes: User-based Collaborative Filtering (UCF), Matrix Factorization (MF) and Random Walk with Restart (RWR). The proposed framework for venue suggestion with current UCF, MF, and RWR methods CF-BORF, greedy BORF and GA-BORF present the better execution in terms of accuracy and recall when contrasted with current plans. The enhanced execution is on the grounds that proposed methods simplify the proposal by taking into account the client preference in light of similarity calculation and client venue closeness. The venue recommendations are provided using user to venue closeness information. The standard execution assessments metric are used to assess the proposed suggestion systems are precision and recall. The precision shows a proportion of the precise suggestions (genuine positive (tp)) to the aggregate number of expected proposals (tp+ false positive (fp)). An exact suggestion is the proposal that has been anticipated effectively in the top-N prescribed venues. Precision is given as:

$$Precision = tp/\,(tp + fp) \qquad (1)$$

The recall measures the single client suggestion viability by registering the normal nature of the individual proposals. Recall is characterized as the proportion of right proposals (tp) to the aggregate number of suggestions (tp + fn). The recall introduces the extent of all the precise proposals in the top-N prescribed venues and can be given to as:

$$Recall = tp/\,(tp + fn \qquad (2)$$

Fig. 2 and 3 display the exactness and recall results without the pre-processing stage. The outcomes in Fig 4 and 5 indicate better execution as far as accuracy, recall when contrasted with the outcomes in Fig.2 and 3. Such change in results is because of the way that the pre-processing stage reduces the negative impact of data sparseness condition over proposal quality. Data sparseness condition results in zero likeness values in collaborative filtering and with huge number of zero sections in client to-client comparability index the suggestion quality decreases. To reduce the quantity of zero passages in client to-client weighted lattice in the previously stated situation, closeness values is enlarged with certainty. In this way, if similarity of two persons is zero however they have gone by almost same set of venues then they won't be relegated a zero weight in the client to-client grid, which general enhances the proposal quality. As reflected in Fig. 4 and 5, NSGA-II exhibits the better execution as far as precision and recall when contrasted with other existing approaches.

Interestingly, the CF-BORF and greedy BORF approaches introduces the accumulation strategy that maps the clients' preference and area closeness into single target capacity. Such accumulation will not provide exact results particularly when there is tradeoff between the client's preference and area closeness. The proposed procedures for venue suggestion are compared with the current UCF, MF, and RWR systems. CF-BORF, greedy BORF and GA-BORF present the better execution as far as precision and recall when compared with the current plans. The Enhanced execution is on the grounds that the proposed methods upgrade the suggestion by considering the client inclinations in view of comparability calculation and client venue closeness. The venue recommendations in light of such advancement are not just the most best for a given client, additionally situated in the nearest area of a client's present area.

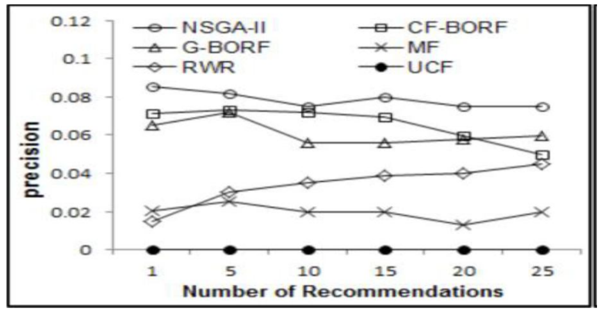

Fig.2. Performance Evaluation without preprocessing for Precision

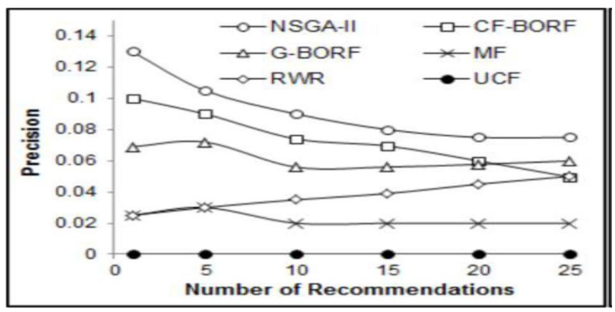

Fig.3. Performance Evaluation with preprocessing for Precision

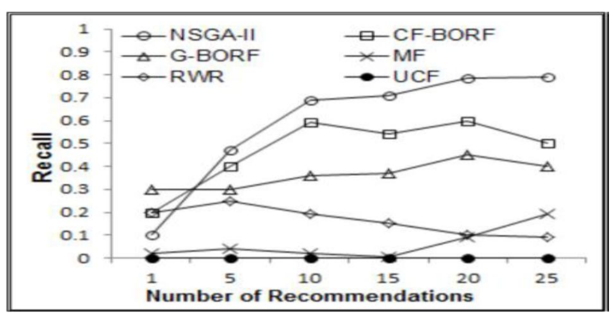

Fig.4. Performance Evaluation without preprocessing for Recall

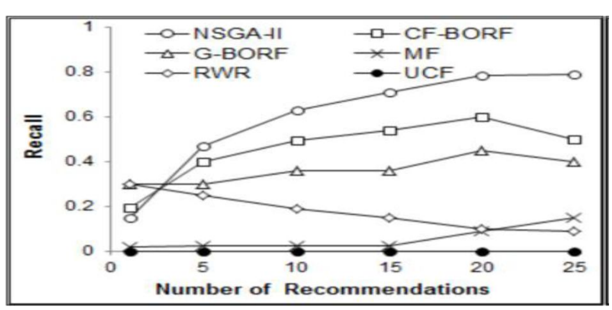

Fig.6. Performance Evaluation with preprocessing for Recall

Paired t-test is used to assess measurable essentialness among the calculations (Wiley, 2008). Null hypothesis is utilized for algorithmic comparison, which represents that the genuine mean difference between the algorithms is zero. The zero mean difference demonstrates that the algorithms are fundamentally related. The p-score in combined t-test ranges from 0 to 1. The p-score esteem more like 0 implies the calculations are essentially related. Then again, the p-score esteem more like 1 implies that the algorithms are essentially different. The paired test between NSGA-II (5 generations and 100 generations) for precision values yielded the normal distinction: 0.0913, standard deviation: 0.2007, standard blunder of mean contrast: 0.0186, t-score: 0.4901, and p-score: 0.6875. In addition, combined t-test between NSGA-II (100 eras and 200 eras) for precision presents the normal distinction: 0.0023, standard deviation: 0.1165, standard mistake of mean distinction: 0.0108, t-score: 0.2151, what's more, p-score: 0.5849. The Table 1 shows the p-score of NSGA-II contrasted and different calculations. It can be seen from the Table 1 that NSGA-II demonstrates high accurate distinction and change as far as accuracy contrasted with alternate calculations.

Table .1. P-score of algorithms with NSGA II

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

CONCLUSION

In proposed cloud based structure Mobi-context addresses hybrid cloud bi-objective recommendation framework for versatile informal organization. It also makes use of multi target streamlining systems to produce customized proposals to provide solutions for the issues relating to cold start and data sparseness conditions. Importance of proposed method is the adjustment of collaborative filtering and bi-objective optimization methodologies, for example, scalar and vector. Proposed approach addresses data sparseness issue by incorporating user to user similarity estimation with configuration measure which computes measure of comparative interests verified by two clients in venues regularly visited by them. In addition, solution for cold start issue is addressed by using the Hub-Average model which allocates rankings to clients and has precompiled set of popular unvisited venues which will be prescribed to new client.
In future the work can be extended by incorporating other approaches such as machine learning, data mining to refine the existing system.

REFERENCES

[1] Alex, S. S. (2010). Dynamic Bus Arrival Time Prediction using GPS Data. Nat. Conference Technological Trends (NCTT), 193-197.
[2] Bao J, Z. Y. (2012). Location-based and Preference Aware Recommendation using Sparse Geo-Social Networking Data. International Conference on Advances in Geographic Information Systems, 199-208.
[3] Bobadilla J, O. F. (2013). Recommender Systems Survey. Knowledge-Based Systems , 109-132.
[4] Chow C, B. J. (2010). Towards Location-Based Social Networking Services. International Workshop on Location Based Social, 31-38.
[5] D, A. (2012). A Non-linear Weights Selection in Weighted Sum for Convex Multi-objective Optimization. Mathematics and Information .
[6] Limin Zhang, A. A. (n.d.). A Survey on Context-aware Recommender Systems Based on Computational Intelligence Techniques.
[7] Majid A, C. L. (2013). A Context-aware Personalized Travel Recommendation System based on Geo-tagged Social Media. International Journal of Geographical Information Science , 662-684.
[8] Noulas A, S. S. (2012). A Random Walk around the City: New Venue Recommendation in Location-Based Social Networks. International Conference on Social Computing (SocialCom), 144-153.
[9] Tikk, H. B. (2013). Initializing Matrix Factorization Methods on Implicit Feedback Database. Journal of Universal Computer Science , 1835-1853.
[10] Wiley. (2008). Paired t test. Wiley Encyclopedia of Clinical Trials.
[11] Ye M, Y. P. (2010). Location recommendation for location-based social networks. International Conference on Advances in Geographic Information Systems, 458-461.
[12] Zheng Y, Z. L. (2009). Mining interesting locations and travel sequences from gps trajectories. 18th international conference on World wide web, 791-800.

AUTHORS' PROFILE

Miss Shruti I Timmapur (Mtech in CSE) from Vishweshwarayya Technology University Student of KLS Gogate Institute Of Technology, Belagavi, Karnataka, India.
Dr. M.M.Math, Professor, Computer science and engineering, KLS Gogate Institute of Technology, Belagavi, Karnataka, India.