Enhanced Recommendation System for E-commerce Applications

S. Sumanth, Pradeep Kumar,

Abstract - ***Recommendation algorithms play a quintessential role in development of E-commerce recommendation system, Where in Collaborative filtering algorithm is a major contributor for most recommendation systems since they are a flavor of KNN algorithm specifically tailored for E-commerce Web Applications, the main advantages of using CF algorithms are they are efficient in capturing collective experiences and behavior of e-commerce customers in real time, But it is noted that , this results in the phenomenon of Mathew effect, Wherein only popular products are listed into the recommendation list and lesser popular items tend to become even more scarce. Hence this results in products which are already familiar to users being discovered redundantly, thus potential discovery of niche and new items in the e-commerce application is compromised. To address this issue , this paper throws light on user behavior on the online shopping platform , accordingly a novel selectivity based collaborative filtering algorithm is proposed with innovator products that can recommend niche items but less popular products to users by introducing the concept of collaborative filtering with consumer influencing capability. Specifically, innovator products are a special subset of products which are less popular/ have received less traction from users but are genuinely of higher quality, therefore, these aforementioned products can be captured in the recommendation list via innovator-recognition table, achieving the balance between popularity and practicability for the user***

Keywords: niche products, collaborative filtering (CF), innovator products, recommender system.

I. INTRODUCTION

In the Age of Large Data , Data is crucial in making decisions and this data analysis is also important to assist others to make better decisions . In E-commerce environment the data generated especially is important to recommend the users to make better decisions whilst keeping the interest of the user , this in a broad way paves way for a novel recommendation system that is idea of this research paper.

The recommendation system is developed to assist user in discovering niche and interested items with the initial association of collaborative filtering algorithm (CF) as it helps in zeroing-in of neighborhood of a user in accordance with collective user behavior and experiences, that is being captured by the system. But the proposed idea is different from other existing systems based on collaborative algorithms as existing systems suffer from Matthew effect , That is redundant recommendation of common items , hence making the recommendation system itself to be common in large and no user-personal interests associated with it. Therefore it is crucial element of the proposed idea to develop a system that recommends the newly added products in real time also to recommend niche products that are added recently with high quality but low visibility due to less traction from users while capturing personal interests of the user to derive at a personalized recommendation system.

Supporting the idea of the proposed system , a user survey done in this research paper shows that traditional users will hardly discover these things by themselves because of the restricted time spent on on-line looking, as newer added products will have short time to live or short time to even capture user visibility , hence this advocates that these kind of items are to be considered without which the products would not have fair chance to have user traction also these kind of products may appeal to only specific cluster of users , so to recommend such products these items may be considered as Innovative-Innovator products that is necessary to be considered with Innovative Product Index to be pushed into recommender system.

The Idea-concept proposed here supports a completely unique CF algorithmic rule, termed Selectivity-based CF is projected. particularly, Here users have a tendency to introduce the thought of influencing product listing and are capable of discovering newer-niche products into CF. A consideration is that Buyers might feel elated if the recommender system recommends a niche item having good user reviews but with economical advantage. However, once creating recommendations, in contrast to the present strategies, we have a tendency to not to force Buyers to just accept newer-innovative products as a result of different e-commerce buyers may have totally different receptiveness to product's maturity. Consequently, the projected algorithmic rule 1st calculates Innovative Product Index (IPI). The IPI is employed to classify Products into niche and traditional products. For every traditional user, the things that its nearest user have interacted with are used to construct the candidate recommendation list. Next, the neighbors' IPI and product rating of each product are integrated into the ranking, . As a result, niche products with high IPI will get high score and are candidates to be pushed into the personalized recommendation list. Algorithmic rule with this principle improves personalized classification of products for recommender system, whereas it also strikes the balance by influencing users while not forcing them to just accept innovative products in best of their interest.

Revised Manuscript Received on July 05, 2019

Mr .Pradeep Kumar, Assistant Professor in Computer Science Department of Ramaiah Institute of Technology. Bengaluru, Karnataka, India.

Mr S.Sumanth.Pai, pursuing M-Tech in CSE from MSRIT, obtained BE in CSE from VTU, Bengaluru, Karnataka, India.

II. RELATED WORK

One of the paper proposes new business processes that are helped to be designed from scratch by proposing a system having two elements , namely Online recommendation system and offline mining , Online recommendation here observed , is similar to most user , and Offline data which is unstructured with complexity of multiple attributes , is tedious to derive at conclusions[1] .

A recommendation framework called Mobile Commerce Explorer is proposed in one of the papers which is intended to incorporate records of user behavior in terms of purchase transactions and user traction with different products in the Application .The MCE devises the framework into 2 components , a) PMCP – Personal Mobile Commerce Pattern Mine to record customer and discover customer purchase pattern , and b) similarities among stores and similar products that are offered by similar stores , considered to be two major mobile commerce elements are considered in this paper[2]. A novel framework called Behavior Predictor for Mobile commerce applications is proposed in one of the research paper, this research is the pioneering work where in to propose the idea of data mining and predicting mobile commerce customer behaviors in order to recommend stores and products which were excluded from the ambit of user purchase history. An extensive experimental and empirical evaluation is done based on simulation of user behavior But Limitations to this proposed system is it relies heavily on Empirical Experimental Methods[6].

A paper discussing about Amazon,com recommendations having item to item collaborative filtering , discusses the idea that recommendation algorithms are accurate in depicting user interests in the list , but limitations include this being traditional recommender system which lists redundant products and are extensively search based methods[3] . A research paper talking about utilizing user-product subgroups , calling to improve the collaborative recommendation , proposes to cluster product subgroups and users for that particular category of prodcuts but limitations being that ,uncertainty of status of recommendation for addition of new products and efficiency in recommending less popular products[4] . A research work done to propose an idea of recommendation for top – n user recommendation , considering a serendipitous association in the recommendation in terms of customer personalized ranking , proposes that serendipity can be relied on benefitting both customers and e-retailers , in other words pushing unexpected-useful products in tandem to product buyers , creating a win-win situation to the aforementioned parties , However there exists a thorough disparity in data generated for tail end products and popular ones , so the proposed framework limits itself to result in satisfactory serendipitous recommendations[5]. Trust is an exclusive and interesting phenomenon, especially when trust is impacting potential customer interests in a cohort of users of an e-commerce application , a research designed to incorporate the levels of trust among user activities like review and posts and rating etc is proposed , however this can be flawed as entire Onus lies in particular group of users to identify other users and products as trust worthy , which is beyond the scope of recommendation systems[8]. A ranking model , which inculcates multi-node graphs depicting diverse buyer relationship is proposed in one of the papers , here the aforementioned graph connects users with nearest neighbor of users , using an edge of the graph and level of user connectivity also can be calculated , But this results in an intricate arrangement of nodes that are difficult to analyze when the system scales . An Idea about incorporating social media is being proposed in one of the papers , providing a semantic-social fabric for recommendation , user preferences are compared with results obtained through proposed algorithm which inculcates exploration path having information about product preference in tandem with ontology of a particular domain of products. Major Limitation's in this method is system needs to analyze messages on social media platform and monitor social media activity of users resulting in breach of privacy[7]. A technique to include information-successively generated of user behavior i.e sequential information providing insights into user behavior that can be used to derive at web navigation pattern along with content information , But this system lacks when it comes to scaling of the existing applications[9].

Summary: Product popularity is sometimes the sole criterion for recommendation in traditional recommender systems, But popularity doesn't mean improving user satisfaction as it might not be the right product for him/her. An honest recommender system needn't solely rely on popularity to predict user searching behaviors to capture potential buying interest of users as well as expanding buying palate of the user. Hence the idea of a system on selectivity based CF along , wherein selectivity is defined in terms of product popularity and product quality mapped by positive user reviews makes sense , Hence this paper tackles these issues and proposes novel techniques to inculcate both the requirements.

III. ALGORITHM

COLLABORATIVE FILTERING ALGORITHM

The algorithm can be summarized in the following steps:
Step: 1. Select n users those who have the highest similarity.
Step: 2. Compute a prediction, Pa,u from a weighted combination. Similarity between two users is computed using the Pearson correlation coefficient

$$P_{a,u} = \frac{\sum_{i=1}^{rn}(r_{a,i} - \bar{r}_a) \times (r_{u,i} - \bar{r}_u)}{\sqrt{\sum_{i=1}^{rn}(r_{a,i} - \bar{r}_a)^3 \times \sum_{i=1}^{rn}(r_{u,i} - \bar{r}_u)^3}} \quad (1)$$

Where ra,i is the measurement given to item i by user a; and ra is the mean rating given by user a.

In step 3, predictions are computed as the weighted average of deviations from the neighbor's mean:

$$P_{a,i} = \bar{r}_a + \frac{\sum_{u=1}^{n}(r_{u,i} - \bar{r}_u) \times P_{a,u}}{\sum_{u=1}^{n} P_{a,u}} \quad (2)$$

Where Pa,i is the prediction for the user a for item i. Pa,u is the similarity between user a and u. n is the number of user in the neighborhood.

IV IMPLEMENTATION

The proposed Recommendation system is divided into two components : offline and Online.

Offline component- To train the system based on **Product Influencing Index(PII)-** an index that takes into consideration – Product rating , Purchase behavior(conformity),popularity of product, and positivity precentile in a particular category of products and comes up with influencing products called as **Innovator products** .

The offline component database is updated regularly and pushed into online component.

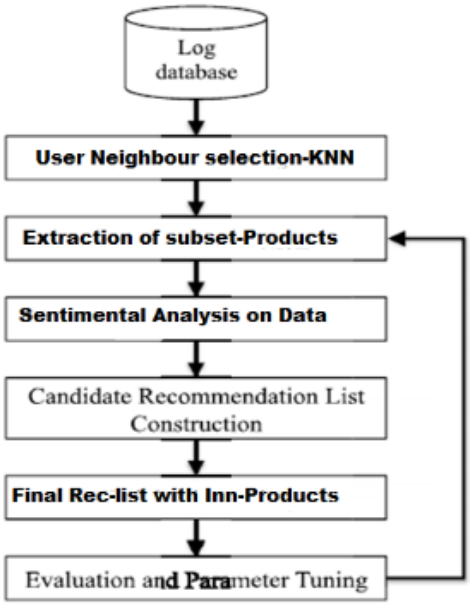

Fig 4.1 Offline Component Process

The Offline process(backend) can be summarized into User Neighbor selection- The particular user for whome a recommendation list is to be pushed is first categorized into a particular neighbourhood of users having similar purschase behavior using CF algorithm a flavor of K-Nearest Neighbour Algorithm

Extraction of subset-products: The products are filtered by rating and popularity score(t_score) in the particular neighbourhood of the user already established.

Sentimental Analysis of data : To select niche items in these products already filtered , we employ sentimental analysis on user traction i.e user reviews of less popular products , and determine positivity percentile of each potential product .

Candidate Recommendation list construction : The above steps yield us the PII-Product Influencing Index with product rating , product popularity and positivity percentile of the product , depending on the PII a recommendation list is constructed.

Final Recommendation list : Depending on the Highest Product rating to Lowest , in combination with PII of each product the final recommendation list is constructed with descending values of product rating.

Evaluation and Parameter Tuning: For consequent iterations new items may be added into the system with varied product ratings, these have to be filtered once again for potential candidate for recommendation and parameters such as popularity , rating etc should be tuned accordingly for each iteration as the system gets trained dynamically on this data.

Fig 4.2 Online Component Process

Online component – A thin client may be an e-commerce website or e-commerce mobile application, which receives recommendation list from offline component.

The above process is summarized into following sequence diagram.

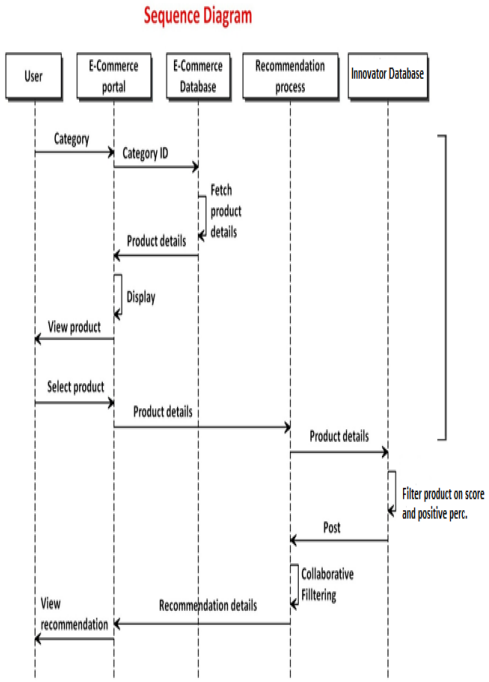

Fig 4.3 Sequence diagram of entire application.

System Architecture :

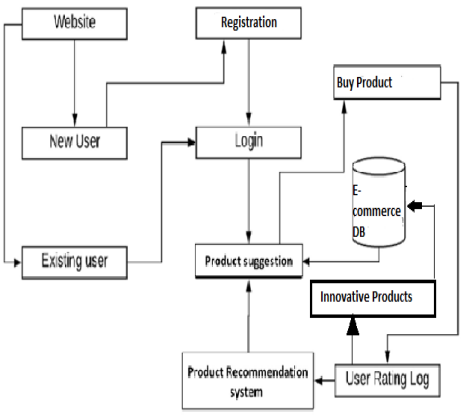

Fig 4.4 System Architecture

The System architecture follows a client-server(MVC) model wherein the client interacts with front end of the application online wherein servlets and JSP ineract with user requests and are handled at the backend , that interact with the Tomcat apache application server hosted on the local machine .

The data is collected in MySQL database application where Tomcat interacts with MySQL to service user requests. The Innovator product data base along with other relevant databases are stored and managed in MySQL.

Offline components can be a distributed databases – which has major advantage of scaling or any cloud environment like AWS or Azure to crunch user behavior and other data can be leveraged , but this paper focuses on proof of concept proposed , hence uses .xsl format files for offline component analysis.

In above architecture it can be noticeable that User rating log , Innovative products database interact with Central –Ecommerce data base to arrive at PII, that is pushed into product recommendation system .

Online component all the while capture user behavior and push the recommended list, it's a thin web client hosted on local machine .

System Performance :

COMPARISON RESULTS IN TERMS OF DIFFERENT EVALUATION MEASURES

Metrics	INVBCF	UBCF	IBCF	MostPop	RandUnpop	CBCF	PIP
Precision	0.0738	0.0744	0.0705	0.0582	0.0110	**0.0744**	0.0688
Recall	0.0746	0.0752	0.0713	0.0589	0.0111	**0.0752**	0.0695
AD	**2.8945**	1.9058	1.5077	2.5574	2.6000	2.4649	2.7621
AvgDistance	**0.7060**	0.5541	0.5495	0.2463	**0.7202**	0.4972	0.5890
AvgPopularity	**0.1509**	0.2029	0.1891	0.6350	**0.1013**	0.2087	0.1616
Coverage	**0.3694**	0.2573	0.2239	0.0040	**0.4750**	0.1950	0.2838

Fig 4.5 System Performance analysis.

The implemented System when compared to other types of recommended systems showed above results , here influencer products based collaborative filtering algorithm(first column in table) outperformed in terms of performance .

The following recommended system performances were

Snapshots of Results:

(Read Only)

	sno	item_sno	item_code	review	status
	1	1	10	good book	positive
	2	2	11	disappointed	negetive
	3	3	12	not very good , not very bad	negetive
	4	4	13	amazing	positive
	5	5	14	bad	negetive
	6	6	15	Great story , good book	positive
	7	7	16	horrible	negetive
	8	8	17	this book is very exciting,I loved reading it.	positive
	9	9	18	the most disinteresting book ever. I will not recommend this book	neutral
	10	10	19	this book is very good.I would recommend this book to everyone	positive
	11	11	20	the book is genius. It was astonishing to read something like t	positive
	12	12	21	It is absolutely marvelous	positive
	13	13	22	I really liked it. And I am not ashamed to admit it.	negetive
	14	14	23	it was annoying but I ended up liking it.	neutral
	15	15	24	worst product . dont use it ever.	negetive
	16	16	25	this product is easily accessible and amazing.	positive
	17	17	26	long staying,Highly recommended.	positive
	18	18	27	very easy to apply.	positive

Fig 4.5 Sentimental Analysis of Products review.

1 Result | 2 Messages | 3 Table Data | 4 Objects | 5 History

(Read Only)

	sentiment_code	item_sno	item_id	positive	negative	neutral	total_score	percentage	status
	1	1	10	1	0	0	1	100	(NULL)
	2	2	11	0	1	0	1	0	(NULL)
	3	3	12	1	3	0	4	0	(NULL)
	4	4	13	1	0	0	1	100	(NULL)
	5	5	14	0	1	0	1	0	(NULL)
	6	6	15	2	0	0	2	100	(NULL)
	7	7	16	0	1	0	1	0	(NULL)
	8	8	17	1	0	0	1	100	(NULL)
	9	9	18	1	1	0	2	0	(NULL)
	10	10	19	1	0	0	1	100	(NULL)
	11	11	20	3	0	0	3	100	(NULL)
	12	12	21	1	0	0	1	100	(NULL)
	13	13	22	1	2	0	3	0	(NULL)
	14	14	23	1	1	0	2	0	(NULL)
	15	15	24	0	2	0	2	0	(NULL)
	16	16	25	2	0	0	2	100	(NULL)
	17	17	26	1	0	0	1	100	(NULL)

Fig 4.6 Innovative Product repository with IPI.

```
Console  Problems  Debug Shell
<terminated> TestRecommender [Java Application] C:\Program Files\Java\jre1.8.0_211\bin\javaw.exe (May 7, 2019, 7:52:19 PM)
Please rate some item to get the recommendations :
Min Rating: 1   Max Rating : 5
Please rate the item : Item1560
4
Min Rating: 1   Max Rating : 5
Please rate the item : Item51
3
Min Rating: 1   Max Rating : 5
Please rate the item : Item41
4
Min Rating: 1   Max Rating : 5
Please rate the item : Item1665
2
Min Rating: 1   Max Rating : 5
Please rate the item : Item1041
4
Recommended for you :
Item: Item770, Rating: 5.0
Item: Item489, Rating: 5.0
Item: Item478, Rating: 5.0
Item: Item187, Rating: 5.0
Item: Item136, Rating: 5.0
Item: Item81, Rating: 5.0
Item: Item127, Rating: 4.943328417969477
```

Fig 4.7 CF algorithm applied to find neighbor hood of a user and finally recommending a list of items referring to Innovative Product Repository.

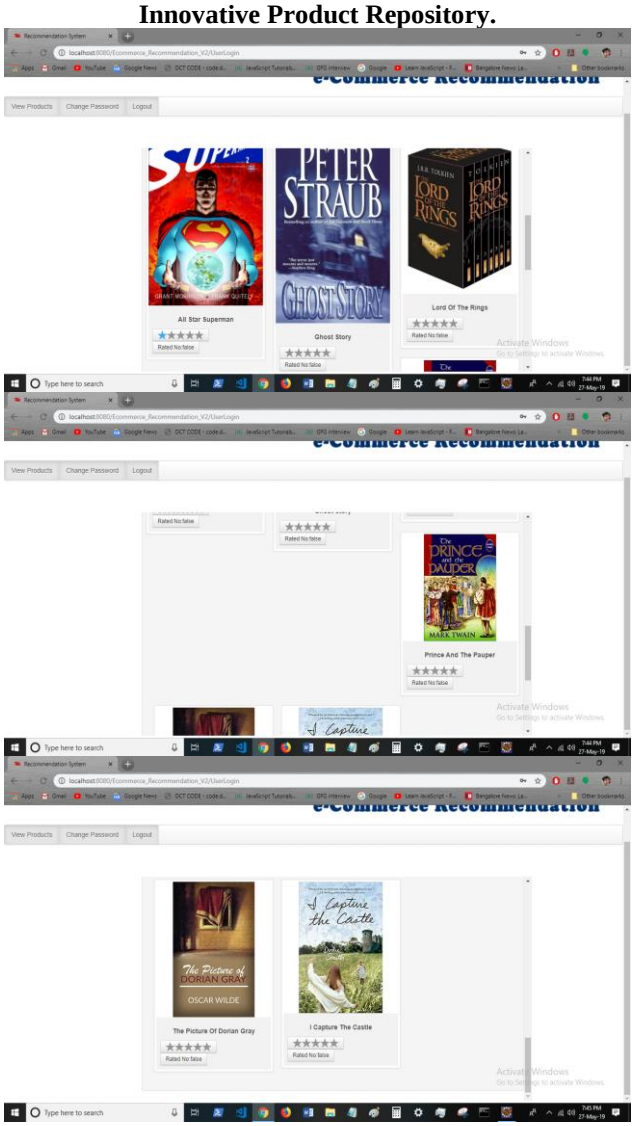

Fig 4.8 Recommendation list for USER A.

Retrieval Number :H6914068819/19©BEIESP
DOI: 10.35940/ijitee.H6914.078919

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

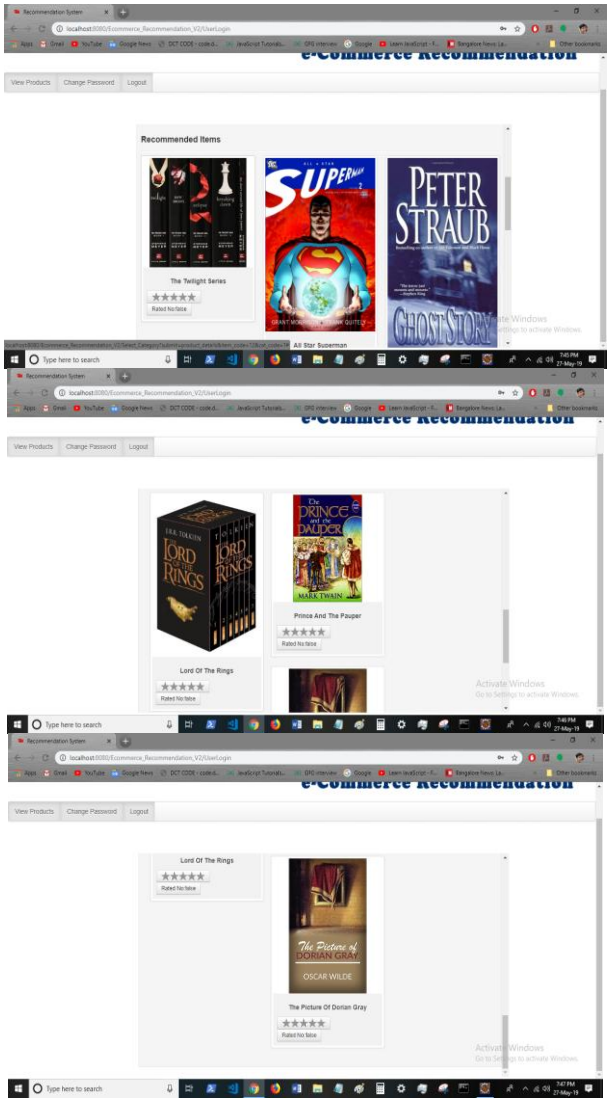

Fig 4.9 Recommendation list for USER D.

PS: THE PURCHASE BEHAVIOUR OF USER A AND USER D ARE PURPOSEFULLY MADE DIFFERENT FOR DEMONSTRATION PURPOSES.

The Following Observations from the results can be noted :

In snapshots above we can see final results when user A and user D both with different user behavior interact with the Application choosing to view a product in Books category following is clearly visible .

a) the most recommended book for user A is All star superman , like wise for user D is twilight series using CF algorithm as well as sentimental analysis .

b) least recommended in the list according to product rating by CF algorithm i.e product that the user may buy with lesser probability , for user A is the book "I capture the castle" while for user D is "The picture of dorian gray".

c) most importantly twilight series is not at all recomended to user A and “I capture the castle” is not at all recommended to user D. this proves the proof of my concept of project . It is also noticeable that recommended list may have same products in both users but the preference of recommendation i.e their position in recommendation list is different thereby proving that products most recommended appear earlier in the list.

IV. CONCLUSION

Our research here concludes by obtaining the Final Result where demonstrate personalized recommendation in an E-commerce environment for a Buyer. The paper completes the objective proposed by implementing collaborative filtering algorithm , so that the user is grouped among other users of same buying behavior, then the potential products that are intended to be bought by the user are pushed in the recommendation list of the user by a)calculating the rating during recommendation phase of CF algorithm , b) determining the popularity of the product(t_score) and positive user reviews of that product(Positivity percentage) , these values retrieved from innovator product data collected by the system .Hence these process ultimately results in personalized(tailor made) recommendation for the user and also pushing lesser popular products with high positive user ratings(niche)

REFERENCES

1. [1]:A Recommendation System to Facilitate Business Process Modeling, Shuiguang Deng, Member, IEEE, Dongjing Wang, Ying Li, Bin Cao, Jianwei Yin, Zhaohui Wu, Senior Member, IEEE, and Mengchu Zhou, Fellow, IEEE,2017
2. [2]A Framework for Personal Mobile Commerce Pattern Mining and Prediction, QusaiAbuein, Ali Shatnawi Faculty of Computer and Information Technology Jordan University of Science and Technology Irbid, Jordan qabuein@just.edu.jo, ali@just.edu.jo Huda Al-Sheyab Computer Engineering Department Jordan University of Science and Technology Irbid, Jordan hyalsheyab12@cit.just.edu.jo, IEEE 2016
3. [3]Amazon.com recommendations: item-to-item collaborative filtering, Rajhans Mishra a,∗, Pradeep Kumar b, Bharat Bhasker b a Indian Institute of Management, Indore, Information Systems Department, Indore, India b Indian Institute of Management, Lucknow, IT and Systems Department, Lucknow, India,IEEE 2016
4. [4]Improving Collaborative Recommendation via User-Item Subgroups, Eric Hsueh-Chan Lu, Wang-Chien Lee, Member, IEEE, and Vincent S. Tseng, Member, IEEE 2017
5. [5]Serendipitous Personalized Ranking for Top-N Recommendation, S. Neeraj, C. Oswald, B. Sivaselvan Department of Computer Engineering Indian Institute of Information Technology Design & Manufacturing Kancheepuram CHENNAI, INDIA {coe13b025, coe13d003, sivaselvanb}@iiitdm.ac.in,IEEE 2017
6. [6] Serendipitous Recommendation for Mobile Apps Using Item-Item Similarity Graph Upasna Bhandari1, Kazunari Sugiyama1, Anindya Datta1, and Rajni Jindal21 School of Computing, National University of Singapore,Computing 1, 13 Computing Drive, Singapore 1174172 Department of Computer Science, Delhi Technological University,ShahbadDaulatpur, Main Bawana Road, Delhi-110042, India{a0106246,sugiyama,datta}@comp.nus.edu.sg, rajnijindal@dce.ac.in
7. [7]Multirelational Social Recommendations via Multigraph RankingMingsong Mao, Jie Lu, Senior Member, IEEE, Guangquan Zhang, Member, IEEE, and Jinlong Zhang,2016
8. Student Member, IEEE, Chun Chen, Member, IEEE,Xiaofei He, Senior Member, IEEE, and Deng Cai, Member, IEEE, 2017
9. [8] Yi-Tao W U, Zhang X M, Wang X M, et al. User fuzzy similaritybased collaborative filtering recommendation algorithm[J]. Journal on Communications, 2016.
10. [9] Zhang X, Li Y. Use of collaborative recommendations for web search: an exploratory user study[M]. Sage Publications, Inc. 2018

AUTHORS PROFILE

Mr S.Sumanth.Pai pursuing M-Tech in CSE from MSRIT, obtained BE in CSE from VTU , Having experience in working in cutting edge technologies in companies like BOSCH india, and Intel India.

Mr .Pradeep Kumar, Mr. Pradeep Kumar D. is working as an Assistant Professor in Computer Science Department of Ramaiah Institute of Technology. His areas of interest include data mining, data sciences, big data, machine learning and internet of things.