

# Performing item-based recommendation for mining multi-source big data by considering various weighting parameters

Venkatesan Thillainayagam<sup>1\*</sup>, Saravanan Kunjithapatham<sup>2</sup>, Ramkumar Thirunavukarasu<sup>3</sup>

<sup>1</sup> Research Scholar, Department of Computer Science, PRIST University, Thanjavur

<sup>2</sup> Dean, Faculty of Computer Science, PRIST University, Thanjavur

<sup>3</sup> Assoc. Professor, School of Information Technology & Engineering, VIT University, Vellore

\*Corresponding author E-mail: [tvntvn09@gmail.com](mailto:tvntvn09@gmail.com)

## Abstract

In the context of big data, a recommendation system has been put forth as an efficient strategy for predicting the consumer's preferences while rating items. Organizations that are functioning with multiple branches are in the imperative need for analyzing their multi-source big data to arrive novel decisions with respect to branch level and central level. In such circumstances, a multi-state business organization would like to analyze their consumer preferences and enhance their decision-making activities based on the taste/preferences obtained from diversified data sources located in different places. One of the problems in current Item-based collaborative filtering approach is that users and their ratings have been considered uniformly while recording their preferences about target items. To improve the quality of recommendations, the paper proposes various weighting strategies for arriving effective recommendation of items especially when the sources of data are multi-source in nature. For a multi-source data environment, the proposed strategies would be effective for validating the active user rating for a target item. To validate the novelty of the proposal, a Hadoop based big data eco-system with aid of Mahout has been constructed and experimental investigations are carried out in a benchmark dataset.

**Keywords:** Big Data; Collaborative Filtering; Item-Based Filtering; Recommender System; Machine Learning; Big Data Mining.

## 1. Introduction

With the recent advancement of data generation strategies and tools, voluminous amount of data has been generated from multiple data source which are characterized by heterogeneity, velocity and deemed as "Big Data". The abundant data sources found in social networks, mobile, and IoT devices demand novel ways to handle complex data and synthesize quality patterns from those sources. The implication and outcome of such big data sources has been successfully recognized in various domains such as healthcare and biomedical (eg. Towards enhancing the efficiency in treatment methods and personalized patient health care system), Sales and marketing (eg. Targeting the sales campaign for the product promotion), Transportation (eg. Identifying optimal routes with reduced costs for speedy service), Finance (eg. Identifying low credit risk groups for offering loans), Social Networks (eg. Identification of social community forums with similar likes) and E-governance (eg. By creating societal data repositories for customized user interaction queries).

Though various kinds of pattern discovery techniques such as supervised learning, un-supervised learning, association rule mining are emerged to cater the required pattern of interest from the huge volume of multi-source big data, recommender system find its own advantage since it addresses the problem of making personalized recommendation using knowledge discovery techniques. The success behind the recommender system in the context of big data is its ability to personalize the user preferences based on the predicted rating given by them. In the context of big data, the recommender

system performs the process of information filtering where it discovers vital information from the large amount of dynamically generated data in accordance with user preferences. Also it has the ability to predict whether a particular user would prefer an item or not based on the user's profile. From the perspective of E-commerce, recommender system has been considered as a tool to help user's search based on their interest and preferences [1]. There are various approaches which have been built on the concept of recommendation system and they are (i) Collaborative Filtering – It performs recommendations by identifying other users with similar taste and uses their opinion to recommend items to the active user. (ii) Content-based Filtering – It normally performs prediction based on the user's information and ignores the contribution from other users as with the case of collaborative filtering (iii) Hybrid Filtering -which combines two or more filtering techniques in different ways in order to increase the accuracy and performance of recommender systems. The broad taxonomy of recommendation system along with various filtering techniques is shown in Figure 1. At layer 1, the system has been classified on the basis of three filtering techniques namely; content based filtering, collaborative filtering and hybrid filtering techniques. Further specialization of each filtering techniques are devised in layer 2 and layer 3 respectively.

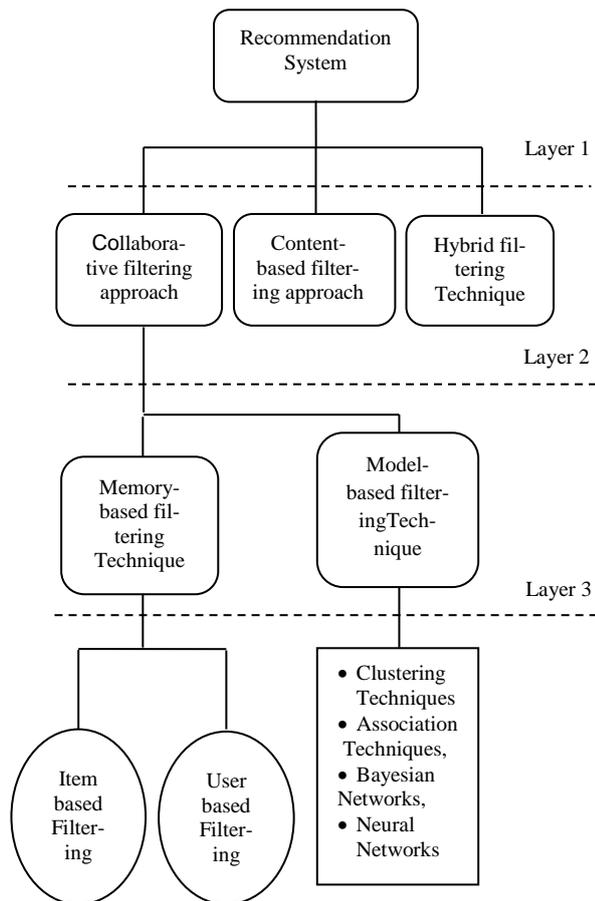


Fig. 1: General Taxonomy of Recommendation Techniques.

When the techniques listed in the above taxonomy are applied in the context of multi-source big data, novel user preferences on the basis of individual branches would be reaped to personalize the user preferences in an effective manner. In such circumstances, adaptation of weighting model for effective recommendation would emerge as an interesting research problem. In this work, we have focussed item-based collaborative filtering technique in a multi-source big data context for predicting effective preferences of an active user by considering various weighting parameters. The rest of the paper has been structured as follows: Section 2 of the paper discusses the notion of item-based collaborative filtering approach along with the state-of-the research works. Section 3 of the paper studies the experimental results and prompts the need for various weighting parameters to infer novel kinds of user preferences. The importance of considering various weighting parameters are elaborately discussed in the Section 4 and Section 5 concludes the paper.

## 2. Related work

Item-based collaborative filtering is a successful recommendation technique proposed by Sarwar et al. [2] that uses user/item matrix which includes a list of  $m$  users and a list of  $n$  items and lots of ratings. A rating  $r_{u,i}$  means how the user 'u' likes the item 'i'. To infer the rating for a target item by an active user, the techniques performs a two-step process that includes similarity computation followed by prediction. There are many measures available such as Euclidean Distance, Cosine and Pearson-Correlation for computing similarity among items. After isolating most similar items based on the similarity measures, the next step is to look into the target user's ratings and use techniques such as weighted sum and regression to obtain predictions. Figure 2 illustrates the process behind the item-based collaborative filtering.

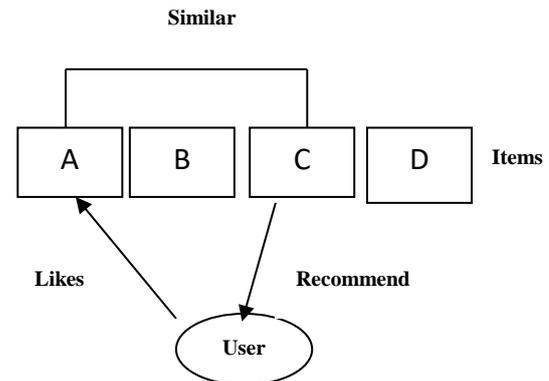


Fig. 2: Item-Based Collaborative Filtering Technique.

The above recommendation technique has been widely employed in many domains such e-commerce, tours, travels and hotel recommendations. A generic framework for employing collaborative filtering algorithms is proposed by Yang et al. [3]. The framework comprises the main procedures of typical recommendation system such as data collection, data pre-processing and recommendation. They have presented two case studies to validate the proposed framework. The first case-study explains the implementation of collaborative filtering algorithms on the basis of user behavior and the second one discusses the impact of user rating in collaborative filtering. They have performed both the case-studies with the advent of movie lens dataset. A decision supportive framework for e-commerce companies has been put forth by Geuens et al. [4]. It generates best collaborative recommendation based on the purchase history of customers. Further they have studied the efficiency of item-based collaborative filtering algorithm in terms of evaluation metrics such as accuracy, diversity and time complexity. An efficient privacy preserving item-based collaborative filtering approach has been proposed by Li et al. [5], which can protect user privacy during online recommendation process without compromising accuracy and efficiency. Further, they proposed two similarity computation algorithms for efficiently measure item similarities without compromising user privacy.

Now-a-days, almost most all e-commerce websites are having data about user rating, reviews and description of items that they are selling. Though there are many approaches exist for recommending items, they create a big list of recommendations for target users. Tewari and Barman [6] tackled the problem by combining the features of content based filtering, collaborative filtering, matrix factorization and opinion mining. Their approach proposed top-N recommendation where 'N' is small and have high precision value. A personalized recommendation framework namely item-network-based collaborative filtering has been advocated by Ha and Lee [7] by extending social network analysis techniques. Accordingly, the items and their relations are expressed as an item network and use the network indicators such as between's, closeness, and degree centralities are applied to the infer the predicted rating of user. Their result shows that the proposed approach performed better than the traditional item-based and user-based collaborative filtering approaches.

A new item oriented recommendation algorithm has been proposed by Hu et al. [8] to focus suitable users towards increasing advertisement revenues of manufacturers. The model utilizes the strategy of multi-view learning to perform recommendations. By considering the purchase records of individual users, the relationship between items and the rating differences among users have been computed as a multi-view model. Then the recommendation has been made by finding the users who are having more possibility to purchase the items. The role of user's emotion in selecting and consuming items has been tackled by Lim and Kim [9] and designates the model as tag-based recommendation model. The proposed model performed item-based recommendation based on the emotion of people and characteristics of purchased items. The model reflects the emotion of the user in terms of user's tag and the user rating for an item is the basic emotion of the tag attached to the item. Since

user's feedback about an item would be completed after consuming the item, their tag represents the detailed feelings about the purchased item. Further they represent the relationship among user, item and tag as a three-order tensor and applied tensor-factorization. Katarya and Verma [10] approached the recommendation system in the aspects of computational intelligence. Accordingly, they proposed novel recommendation system using K-means clustering and cuckoo search optimization. As a hybrid cluster and optimized recommender system it performs novel recommendations for Movie lens dataset. Further they measured the performance of their proposed approach in terms of Mean Average Error (MAE), Root Mean Square Error (RMSE) and improved accuracy is resulted. By extending the concept of knowledge based system, Aguilar et al. [11] proposed an intelligent recommendation system framework. It consists of components such as knowledge representation paradigm, learning methods and reasoning mechanisms. The framework improves the quality of recommendation due to the existence of extended user and item profile, context and domain knowledge.

Mining pattern of interest from multiple heterogeneous data sources significantly creates promising research avenues among the data science community [14], [15]. Ramkumar et al. [12] in their work discussed the importance of mining useful pattern of interest from multiple data sources by comparing the traditional process of centralized computing and local pattern analytics strategy. To define the importance of big data, Wu et al. [13] advocated HACE theorem in the context of big data analytics. The theorem defines big data as Heterogeneous Autonomous repositories that need distributed and decentralized Control towards seeking complex and Evolving relationships. Further, they put forth a three-tier big data processing framework that comprises big data mining platform, big data semantics with application knowledge, and big data mining algorithms.

### 3. Implementation scenario

This section mainly focuses on implementing item-based collaborative filtering approach in a multi-source big data context. It uses Pearson correlation co-efficient as similarity measure and implemented in mahout machine learning library of Hadoop platform. As an open source machine learning library, Mahout offers wide range of applications that are useful for the task of recommendation. The scalability nature of mahout able to support distributed processing of large data sets across cluster of nodes using Hadoop Distributed File System (HDFS).

#### 3.1. Data set

For the experiment, we have used the MovieLens100k dataset (<http://grouplens.org/datasets/movielens/>). The dataset contains 1,00,000 preferences applied to 1682 movies by 943 users of the online movie recommender service. The data are contained in three files, movies.dat, ratings.dat and tags.dat. Rating data files have at

least three columns: the user ID, the item ID, and the rating value. Each user at least rates 20 movies; the range of ratings is from 1 (lowest) to 5(highest).

#### 3.2. Similarity computation

Similarity computation among items or users plays a vital role in memory-based collaborative filtering algorithms. For item-based based algorithms, the user/item matrix has been used to extract the similarities among the items rated by the active user. To find the similarity among items, we have used Pearson correlation coefficient since it performs fast recommendation in Mahout Environment.

$$\text{similarity}(i, j) = \frac{\sum_{u \in U} (R_{u,i} - \bar{R}_i)(R_{u,j} - \bar{R}_j)}{\sqrt{\sum_{u \in U} (R_{u,i} - \bar{R}_i)^2} \sqrt{\sum_{u \in U} (R_{u,j} - \bar{R}_j)^2}} \quad (1)$$

Where,

U - Domain of all users

u - A user

i, j - Some items

$R_{u,i}$  - Rating of user, 'u' on item, 'i'

$\bar{R}_i$  - Mean rating value for item, 'i'

$R_{u,j}$  - Rating of user, 'u' on item, 'j'

$\bar{R}_j$  - Mean rating value for item, 'j'

Then the prediction for user 'u' on item 'i' denoted by  $p_{u,i}$  is calculated as follows:

$$P_{u,i} = \frac{\sum_{\text{all similar items } N(S_{i,N} * R_{u,N})}{\sum_{\text{all similar items } N(|S_{i,N}|)} \quad (2)$$

#### 3.3. Experimental study

The experimental set up has been created in a Hadoop Version of 1.2.1 with the JDK version of 1.8 and the mahout library version of 11.00 has been used. To create the multi-sourced big data scenario the 1,00,000 preferences of movielenz data set has been divided into three data sources namely 'Site 1', 'Site 2', 'Site 3' and each contains the preferences of 40,000, 50,000 and 10,000 respectively. We have created three Hadoop Cluster for processing these three data sources and execute item based collaborative filtering in all three nodes for inferring the taste of the user. We have tabulated the predicted rating for top 10 recommended items for first 10 users in each node [Table 1]. Though rating can be inferred by performing similarity computation in each site, assigning weightages for the predicted rating would be an interesting solution while synthesize the rating values from multiple source of data. For example, item 7 got an inferred rating value of 5 [which is maximum] in all participating sites. How to synthesize the final global rating value since it found in all participating sites.

**Table 1:** Predicted Recommendation of 10 Items for First 10 Users

Hadoop Cluster (Sites)	User id	Predicated Rating {Item id :rating}
Site1 with a preference population of 40000	1	[[{5:5},{9:5},{2:5},{4:5},{7:5},{8:5}, {1039:5},{1:5},{1042:5},{12:5}], {909:5},{360:5},{1463:5},{651:5}, {889:5},{882:5},{1386:5},{86:5}, {15:5},{387:5}]]
	2	[[{1160:5},{1463:5},{882:5},{303:5}, {313:5},{305:5},{1423:5},{327:5}, {293:5},{361:5}]]
	3	[[{1405:5},{305:5},{1127:5},{811:5}, {292:5},{303:5},{270:5},{1388:5}, {286:5},{1137:5}]]
	4	[[{611:5},{517:5},{490:5},{267:5}, {70:5},{467:5},{166:5},{479:5}, {114:5},{1451:5}]]
	5	[[{549:5},{419:5},{517:5},{173:5}, {72:5},{614:5},{188:5},{1168:4},
	6	

		{52:4},{479:4}]
	7	[[{7:5},{13:5},{4:5},{6:5},{8:5}, {10:5},{1:5},{2:5},{5:5},{17:5}]
	8	[[{220:5},{176:5},{1228:4},{686:4}, {197:4},{449:4},{642:4},{797:4}, {42:4},{459:4}]
	9	[[{447:5},{576:5},{578:5},{423:5}, {566:5},{568:5},{411:5},{417:5}, {659:5},{233:5}]
	10	[[{566:5},{14:5},{739:5},{1103:5}, {10:5},{570:5},{1115:5},{2:5}, {561:5},{572:5}]
Site2 with a preference pop- ulation of 5000	1	[[{16:5},{19:5},{7:5},{13:5},{17:5}, {18:5},{4:5},{5:5},{11:5},{20:5}]
	2	[[{1089:5},{566:5},{327:5},{900:5}, {363:5},{496:5},{168:5},{144:5},{176:5},{1099:5}]
	3	[[{267:5},{763:5},{850:5},{410:5}, {1123:5},{128:5},{619:5},{669:5}, {551:4},{800:4}]
	4	[[{31:5},{20:5},{1062:5},{27:5},1056:5},{19:5},{16:5},{1:5},{1045:5},{1098:5}]
	5	[[{12:5},{15:5},{6:5},{11:5},{13:5}, {14:5},{3:5},{4:5},{7:5},{16:5}]
	6	[[{8:5},{11:5},{4:5},{7:5},{9:5}, {10:5},{1:5},{2:5},{6:5},{12:5}]
	7	[[{6:5},{11:5},{3:5},{5:5},{9:5}, {10:5},{1:5},{2:5},{4:5},{12:5}]
	8	[[{8:5},{13:5},{4:5},{7:5},{11:5}, {12:5},{1039:5},{1:5},{5:5},{1056:5}]
	9	[[{316:5},{304:5},{309:5},{300:5},{24:5},{1146:5},{1267:5},{1258:5}, {173:5},{472:5}]
	10	[[{6:5},{9:5},{3:5},{5:5},{7:5},{8:5},{1:5},{2:5},{4:5},{11:5}]
Site3 with a preference pop- ulation of 10000	1	[[{32:5},{7:5},{17:5},{22:5},{9:5},{568:5},{159:5},{15:5},{561:5},{514:5}]
	2	[[{412:5},{134:5},{99:5},{946:5}, {1161:5},{567:5},{332:5},{324:5}, {1199:5},{147:5}]
	3	[[{755:2},{194:2},{357:2},{1009:1}, {527:1},{196:1},{216:1},{875:1}, {288:1},{654:1}]
	4	[[{285:5},{568:4},{432:4},{301:4}, {1063:4},{502:4},{98:4},{26:4}, {217:4},{111:4}]
	5	[[{151:5},{274:5},{45:5},{1110:5},{141:5},{96:5},{15:5},{9:5}, {61:5},{340:5}]
	6	[[{12:5},{23:5},{7:5},{11:5},{1060:5},{22:5},{189:5},{1:5},{10:5},{24:5}]
	7	[[{451:4},{449:4},{694:4},{834:4},{311:4},{576:4},{1099:4},{1375:4}, {559:4},{195:4}]
	8	[[{492:5},{89:5},{479:5},{1091:5}, {419:5},{485:5},{507:4},{436:4}, {300:4},{951:4}]
	9	[[{168:4},{480:4},{402:4},{163:4}, {689:4},{26:4},{330:4},{603:4}, {285:4},{544:4}]
	10	[[{177:5},{298:5},{96:5},{66:5},{53:5},{281:5},{47:5},{4:5},{1048:5},{323:5}]

A common problem of current item based collaborative filtering is that all users have the same weights when item similarities are computed. It is common knowledge that some user's recommendation are more important than those of others. Thus for item based collaborative recommendations, some users and their rating should be weighted higher than others. Further inclusion of those weighting parameters is essential when the data sources are spread in multiple locations. The importance of those weighting parameter is discussed elaborately in the following section.

#### 4. Consideration of various weighting parameters

This section presents certain novel ideas in adding various weighting parameters over the predicted rating in each site of a multi-sourced big data.

##### 4.1. Weighted rating based on user rank

In any social group, some persons have higher prestige than others because they have been in the group for a long time or they have made greater contribution to the group. Hence if user weights are

taken into consideration in item-based collaborative filtering, the similarities between items would be more realistic. Hence strategies such as ranking the importance of users while performing recommendation would be an efficient strategy. The underlying concept is similar to the page rank algorithm, which computes the rank vector to rate the importance of all web pages by analyzing their hyperlinks. Similarly there may be different degree of correlation between users in a recommendation system, Eg. if more items have been rated by two users, 'U<sub>i</sub>' and 'U<sub>j</sub>', they closer with each other and rank with high values of similarity. Then other users who are linked with these high ranked users should also enjoy significant importance in computing of rating.

##### 4.2. Weighting user based on both rating and interesting similarities

There are various relationships between users that can be found in a recommendations system such information can also be exploited for the calculations of user weights in rating prediction. Traditionally, the rating similarity has been perceived as one aspect of user similarity. Besides there are certain other relationship could also be found behind the ratings. Assume that there are two users 'u' and 'v' and there are many items rated by them. Though their ratings

are different and similarity is low, their degree of relationship is high since the set of items they have rated are similar. Such interesting relationship is called as interesting similarity. Hence when constructing data model for user-relationship matrix, due weightage has to be introduced for both rating and interesting similarities. As a result, when preferences are emerged from users belong to multiple locations, the estimated weighting value would identify certain novel information.

#### 4.3. Inclusion of semantic information

In item based collaborative filtering systems, the goal is to find similar items and use them to make predictions. Such kind of similarities can capture user's likes and dislikes of items, which is a subjective indicator. Hence it is highly essential to capture intrinsic relationship between items, which would serve as an objective indicator of similarities. Such kind of objective indicator of similarities would be established by studying the semantic information about items. It may include attribute of the item, the item's relationships to other items, its role in the relationship to other items, and other meta information. For example, in a movie scenario, it contains many attribute such as title, director, cast, release date and has relationship with other movies. Such semantic information and relationship among items extracts interesting feature about items. When semantic similarities are included as weighting parameters along with the user's rating, intelligent computation of item similarities can be performed that leads to better recommendations. But computation of semantic similarity is a domain dependent and requires knowledge of the underlying structure and relationship among items.

#### 4.4. Incorporation of 'trust – relation' between user as weighting parameter

The increasing trends of social media greatly enrich people's social relations and allow users to share information and opinions with connected users. In such circumstances, a trust network can be generated by asking users to rate other users directly. Trust is an important relation in a social network and a basis for forming qualified and credible communities. A trust-based recommendation system assumes that user have similar taste at par with the trusted users. The weight of each user is determined based on the similarity of the target user in a trust network. Hence it is appropriate to develop recommendation model by including user-item rating history and a trust network to recommend potential items of interest to a user by duly considering his/her trust relation.

#### 4.5. Incorporation of customer's RFM

While considering the predicted rating for an item by a customer, their RFM (Recency, Frequency, Monetary) characteristics are to be taken as weighting parameter. The attribute 'Recency' represents the time since the last purchase, a lower value indicating a higher probability of customer making repeat purchase. The attribute 'Frequency' denotes the number of purchase made by the customer with in a specific time period. The higher value with respect to frequency indicates higher loyalty towards customer. The final attribute 'Monetary' means the amount of money spent in the specified time period, a higher value indicates a profitable customer that organization must focuses. By incorporating these RFM features along with customer life time value, the multi-source big data organization would reap meaningful pattern in the knowledge discovery process.

#### 4.6. Impact of geographical information

In a multi-source big data scenario, predicted ratings for attributes can be weighted in terms of geographical information of users. With the advent of mobile devices and ambiguities sensors, it is possible to incorporate location based information in rating a item. Then the rating over an item can be modeled as 4-tuple information such as (userid, user location, user rating, movieid). From that one can draw

some conclusion that users nearby can share similar preferences to a given topic of interest item or not. This kind of weighted parameter would be highly useful for location based services.

#### 4.7. Inclusion of data source weight in prediction of rating

Organization that function across multiple locations or regions can significantly be influenced by their data source nature and volume of operation performed on it. In such circumstances, prediction of rating for an item can be weighted in terms of data source it belongs. Hence inclusion of data source weight in predicting rating for an item would be an efficient way for assigning the weighted rate of an item especially when the data source are spread over multiple location. The weight of the data source can be estimated in terms of various features such as transaction population, turnover or cost items sold in the branches and also assigning weights for transactions made in the branches. Consider two sites 'S1' and 'S2' with the respective population of 10,000 and 1,00,000 transaction. By assigning weights for the data sources, weight of the data source, 'S2' would be 10 times higher than 'S1'. The estimate would not be fair one, if the turnover of 'S1' is higher than of 'S2'. To improve decision, predicted rating of an item can be weighted on the basis of turnover quantity or cost of items sold may be carried out. When the quantitative information is included in the allocation of site weight, synthesizing model would be an efficient choice for finalizing the predicted rating of an item.

### 5. Conclusion

With the advent of emerging web repositories and big data of multi-sourced nature, organizations are in imperative need for analyzing their user preferences/taste in an effective manner. Though various recommends system have been emerged due to the explosive growth of web, Item-based collaborative filtering approach emerged as an effective way for predicting the preference due to advantages such as avoidance of cold-start problem, ability in handling scalable data item and user. In this paper, an item based collaborative filtering technique has been put forth for an inter-state business organization to infer customer preferences from voluminous of transaction performed in multiple locations. In traditional item-based collaborative filtering approach, all users have been designated with uniform weights. But in real sense, weighting attributes are to be considered for predicting the preferences of user located in multiple location of an inter-state business organization. In the context of multi-source big data, this paper recommends seven prominent weighted parameters that could be considered while synthesizing predicted rating of an item from multiple data sources. In our opinion, allocation of site weights would be an efficient strategy for synthesizing the predicated rating /preferences of an item. If the allocation is based on turn-over or quantity of item sold in the branch, novel preferences of user would be emerged

### References

- [1] F.Q. Isinkaye, Y.O. Folajimi, and B.A. Ojokoh (2015), Recommendation systems: Principles, methods and evaluation. *Egyptian Informatics Journal*, 16(3), 261-273. <https://doi.org/10.1016/j.eij.2015.06.005>.
- [2] B. Sarwar, G. Karypis, J. Konstan and J. Riedl (2001), Item-based collaborative filtering recommendation algorithms. In *Proceedings of the 10th international conference on World Wide Web*, pp. 285-295. ACM. <https://doi.org/10.1145/371920.372071>.
- [3] Z. Yang, B. Wu, K. Zheng, X. Wang and L. Lei (2016), A survey of collaborative filtering-based recommender systems for mobile Internet applications. *IEEE Access*, 4, 3273-3287. <https://doi.org/10.1109/ACCESS.2016.2573314>.
- [4] S. Geuens, K. Coussement and K.W. De Bock (2018), A framework for configuring collaborative filtering-based recommendations derived from purchase data. *European Journal of Operational Research*, 265(1), 208-218. <https://doi.org/10.1016/j.ejor.2017.07.005>.

- [5] D. Li, C. Chen, Q. Lv, L. Shang, Y. Zhao, T. Lu and N. Gu (2016), An algorithm for efficient privacy-preserving item-based collaborative filtering. *Future Generation Computer Systems*, 55, 311-320. <https://doi.org/10.1016/j.future.2014.11.003>.
- [6] A. S. Tewari and A.G. Barman (2018), Sequencing of items in personalized recommendations using multiple recommendation techniques. *Expert Systems with Applications*, 97, 70-82. <https://doi.org/10.1016/j.eswa.2017.12.019>.
- [7] T. Ha and S. Lee (2017), Item-network-based collaborative filtering: A personalized recommendation method based on a user's item network. *Information Processing & Management*, 53(5), 1171-1184. <https://doi.org/10.1016/j.ipm.2017.05.003>.
- [8] Q.Y. Hu, Z.L. Zhao, C. D. Wang and J. H. Lai (2017), an item oriented recommendation algorithm from the multi-view perspective. *Neurocomputing*, 269, 261-272. <https://doi.org/10.1016/j.neucom.2016.12.102>.
- [9] H. Lim and H. J. Kim (2017), Item recommendation using tag emotion in social cataloguing services. *Expert Systems with Applications*, 89, 179-187. <https://doi.org/10.1016/j.eswa.2017.07.046>.
- [10] R. Kataraya and O.P. Verma (2017), an effective collaborative movie recommender system with cuckoo search. *Egyptian Informatics Journal*, 18(2), 105-112. <https://doi.org/10.1016/j.eij.2016.10.002>.
- [11] J. Aguilar, P. Valdiviezo-Díaz and G. Riofrio (2017), A general framework for intelligent recommender systems. *Applied computing and informatics*, 13(2), 147-160. <https://doi.org/10.1016/j.aci.2016.08.002>.
- [12] T. Ramkumar, S. Hariharan and S. Selvamuthukumaran (2013), A survey on mining multiple data sources. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 3(1), 1-11. <https://doi.org/10.1002/widm.1077>.
- [13] X. Wu, X. Zhu, G.Q. Wu and W. Ding (2014), Data mining with big data. *IEEE transactions on knowledge and data engineering*, 26(1), 97-107. <https://doi.org/10.1109/TKDE.2013.109>.
- [14] B. B. Ahamed, T. Ramkumar(2016), An intelligent web search framework for performing efficient retrieval of data, *Computers and Electrical Engineering*, 56, 289-299. <https://doi.org/10.1016/j.compeleceng.2016.09.033>.
- [15] T. Ramkumar., R. Srinivasan,(2008), Modified Algorithms for Synthesizing High-Frequency Rules from Different Data Sources, *Knowledge and information systems*, 17(3), 313-334. <https://doi.org/10.1007/s10115-008-0126-6>.