Trust Based Recommendation System Using Knowledge Graph (KGTRS)

Nidhi B Channappagoudar[1], Richa Singh[2*]

{nidhi.channppagoudar2020@vitstudent.ac.in[1], richasingh.bv@gmail.com[2]}

[1,2]School of Computer Science and Engineering, Vellore Institute of Technology, Chennai, India.

Abstract. Recommender systems has proven its importance to the solution of exponentially increasing data on the internet, which provides users with more personalized services of information for better development of online services. The issue of trust has emerged into our day-to-day life since early 90s, which focuses on improvement of recommender systems. The Google Knowledge base gathers information from a variety of sources. In the presented work trust is incorporated using knowledge graph in order to improve the recommendation system that provide the opportunity to build trust with potential site visitors looking for a product/service/information. To enhance the user belief and usre acceptance trust based recommendation system is incorporated. Trust based recommender systems using knowledge graph improve the accuracy of the recommender system and user experience. They are also capable of handling some challenges of recommender systems such as cold-start problem.

Keywords: Knowledge Graph; Trust; Recommender Systems; Cold start; Sparsity.

1 Introduction

In this era of information overload, especially in the online world, Recommender Systems (RS) plays a crucial role. It find the items of user interest such as books, movies, news, music or products in general which match with their preferences.

Recommendation systems can be classified into three types, one is Collaborative Filtering (CF) and another is Content Based (CB) filtering approach. Content-based filtering filters the item on the basis of characteristics and features based on the user's preferences. The basic idea behind the CFRS is to make a decision based on the past purchases of users to find similar users [1][2][3][4][5].

But these recommendation systems, in particular, collaborative recommendation systems face some problems such as cold start (not enough information know at the start), data sparsity and attacks (fake profiles giving biased ratings).

Literature show that users tend to rely on their trustworthy friends or family unlike the strangers. Trust helps to employ the information and enhances the recommender systems that improves the accuracy as well as the users' experience. Consumer's trust is the main success behind the web based business in the context of e-commerce to perform transactions at websites. Trust is a complex social phenomenon reflecting behavioural, technological, social,

ICIDSSD 2022, March 24-25, New Delhi, India

DOI 10.4108/eai.24-3-2022.2318767

psychological and organizational interactions among human and nonhuman technological agents [6][7][8].

Trust in recommendation plays an important role in trust enhanced recommender systems. The knowledge that originates from trust networks is used by trust-enhanced recommender systems to generate more personalized recommendations: recommendations for items is used by users which are highly rated. Trust aggregation and propagation is the main strength of these systems [9][10].

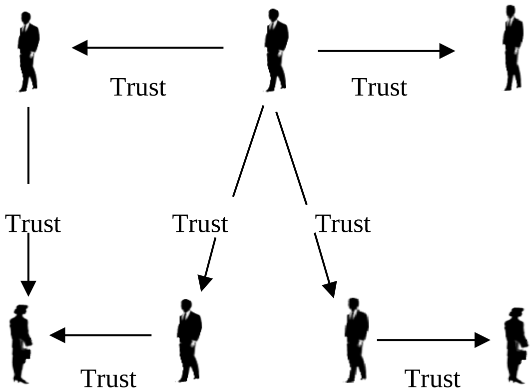


Fig. 1. Web of Trust (WOT)

The Knowledge Graph (KG) is a collection of entities i.e., objects, events or concepts [11]. The framework for data integration, analytics, unification and sharing is provided by knowledge graph when it put the data into the context. KGs are built as semantic network, which means we can calculate the semantic similarities between the entities (eg: people, place, objects). The information can be structured in the form of graph consisting of nodes that represents entities and edges representing relati006Fnship between entities. Besides, the problem of data sparsity is solved by KG and also provides recommendation explanation.

For the cold start problem involving users who have no ratings how trust information is useful is that even though we don't have any item ratings, from the data of the trust network we can get information of who the user trusts (along with weights) which lets us look at his trusted friends ratings and predict reasonable rating for the items which are far better than other models which have no basis to predict anything at the start [12].

A trust aware recommendation system using knowledge graph has been proposed. This system will use trust as weight and incorporate trust information to boost the recommendation provided by the trusted users. The proposed approach gives the solution for the sparsity and cold start problems.

2. Literature Survey

In the area of recommendation widely used approach for recommendations is collaborative filtering. Ghazanfar and Prugel-Bennett [13], proposed a solution by combining all recommender systems. This solution eliminated redundant records problem of the recommendation systems. Qian Wang, Xianhu Yuan, et al. [14] combined the item demographic information with the searching for a set of neighbouring users who have the same interests using a genetic algorithm. This method improved the system scalability. Dianping. Lakshmi et al. [15] used a spark to implement the ALS matrix factorization which has to be compared by using MLib to generate SGD matrix factorization. This generate a new list of recommended items to their user and incorporate the diversity and novelty in presented list. Sajal Halder et al. [16] proposed a movie swarm mining concept. Frequent item mining and two pruning rules were used by this algorithm. This has solved the cold start problem. Arpan V Dev [17] proposed an algorithm as extended prefix filtering based on similairty joins. Repeated overhead of computation is reduced by usign map-reduce. This work has gained high performance when compared with the other ordinal algorithms.

Rahul Katarya and Om Prakash Verma [18] has used fuzzy, a c-means technique for the users in finding the neighbourhood. This work has solved the scalability and accuracy for the movie recommender systems. Qingyu Guo [19], made a survey on illustration of KGbased recommendation which aims to use KGs is three ways: embedding-based method, the path based method and the unified method. Using KGs, provides more accurate recommendation and another benefit of KG-based recommendation is the interpretability. Swati Gupta, Sushma Nagpal [20], incorporated trust in recommendation system which was resulted in a new class of recommender system called as trust-aware recommender system(TARS). They have analysed eight different implicit trust metrics with respect to the various properties of trust proposed by research regarding to TARS. Li Xie, et al. [21], overcome problems of recommender systems of using ALS algorithms. By designing a new loss function, the issue is solved by determining the similarity joins between items and users. In different iterations by comparing the root mean square we get the algorithm output results with the other existing collaborative filtering techniques.

In the proposed work, we have combined trust in recommender systems using knowledge graph. The proposed approach gives the solution for the sparsity and cold start problems.

3. Proposed Approach

In the proposed work we have incorporated trust using knowledge graph. The trust factors are included using three factors: Co-factorization, Ensemble and Regularization Methods. Knowledge graph completion is used to handle the knowledge graph for the trust based recommendation system.

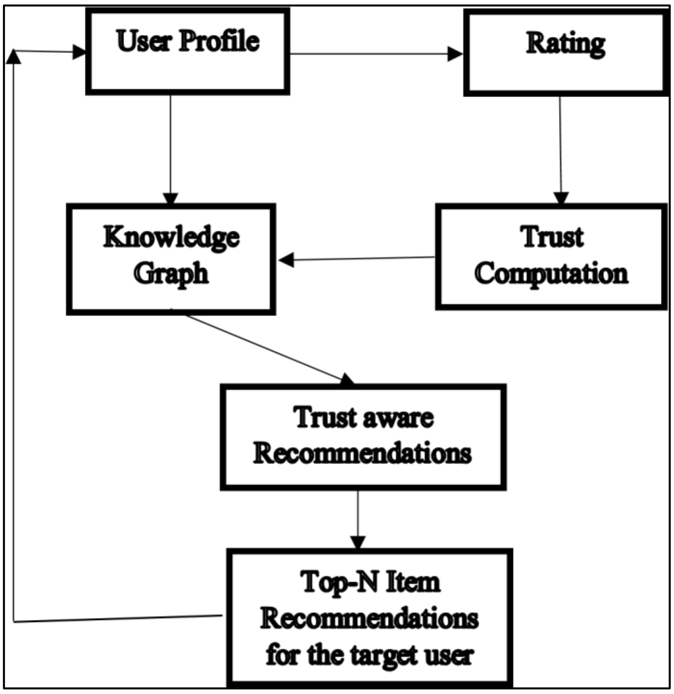


Fig. 2. Trust Based Recommendation System Using Knowledge Graph

3.1 Trust Computation Phase

The trust factors are included using three factors: Co-factorization, Ensemble and Regularization Methods. Knowledge graph completion is used to handle the knowledge graph for the trust-based recommendation system.

We further divided the trust-aware recommendation system into three groups depending upon the trust model which is introduced to capture information [22] as Co-factorization, Ensemble and Regularization methods.

These techniques are described as follows:

1. Co-factorization Methods: In this approach the prediction of unknown ratings are done using the user and the item factor matrices to determine the product of an approximate user-item matrix. Types of co-factorization methods are: SoRec (Social recommendation using probabilistic matrix factorization) and LOCABAL (exploiting global and local social context for recommendation).

SoRec: It stands for Social recommendation probabilistic matrix factorization. Data sparsity has been recognised as one of the most crucial challenge that every recommender system faces. The traditional approaches of RS faces many challenges while handling very large datasets and few rating by user. Moreover, traditional recommendation systems assume that all users are identically distributed and independent and this assumption ignores the connections among users or social interactions. On this view, we proposed a factor analysis approach based on probabilistic matrix factorization to solve the data sparsity problem and also helps in enhancing trust.

Here we assumed that the i^{th} user u_i should share the same user preference vector u_i in the trust relation space and rating information (Rating space) by sharing the same user preference latent factor, the system performs a co-factorization in the useritem matrix and the user-user trust relation matrix. From rating information and the trust information, sorec learns the user preference matrix U by solving the following optimization problem:

$$\underset{U,V,Z}{min} \|W\Theta(R-U^TV)\|_F^2+\alpha\sum_{i-1}^{n}\sum_{u_k\epsilon F_i}(S_{ik}-U_i^TZ_k)^2+\lambda(\|U\|_F^2+\|V\|_F^2+\|Z\|_F^2),, \qquad (1)$$

Where, U represents the user information, V is rating information and Z is trust information, T denotes the user and item description, W handles the feedback which is implicit to encode the information related to user behaviours, quality of reviews, R show all of the ratings, α is normalization factor, K denotes the latent factors, λ helps to avoid over-fitting, controlled by parameter λ . To perform trust relation prediction, we used a reconstructed matrix T.

b. LOCABAL: With the rapid growth of data on the social media teh suggestions from their local friends' users with high reputations is considered as helpful to find relevant information. This motivates to exploit more of social relations that involve both local and global for recommendation purpose. LOCABAL is an effort that takes advantage of both the context for social relation and incorporate them in recommendation.

2. Ensemble Methods:

a. Social Ensemble Trust (STE): We used Social Trust Ensemble to represent the formulation of the social trust restrictions on the recommender systems (Learning to recommend with social trust ensemble). The basic idea of ensemble methods is that users and their trust networks should have similar ratings on items, and from the user and her trust network, missing rating for given user is predicted as linear combination of ratings. STE is a representative system in this group which models a rating from the *i*-th user u_i to the *j*-th item $v_j$as follows:

$$R_{i,j}=u_i^Tv_j+\beta\sum_{u_{k\in N_i}}T_{i,k}\,u_k^Tv_j \qquad (2)$$

Where, a weighted sum of the predicted ratings is computed for v_j from u_i's trust network, $\mathcal{N}_i$ is the number of items, and β controls the influence from trust information. STE is finally to minimize the following term:

$$||W\theta((R-U^TV)-\beta TU^TV||^2{}_F \qquad (3)$$

3. Regularization Methods:

The main idea of regularization method is to identify a solution that provides not a perfect fit to the data but instead a good data fit and enjoys a certain degree of smoothness:

a. SocialMF: For recommendation in social networks, a matrix factorization technique with trust propagation.

It focuses on a user's preferences assuming the simialrity between the user preference and trust network. A user preference u_i, using regularization methods force the preferences u_i that of users in $\boldsymbol{u_i}$'s trust network $\mathcal{N}_i$. SocialMF consists of the user preferences that is closer to with the user's trust network and provide average preference as:

$$min \sum_{i=1}^{n} (u_i - \sum_{u_{k \in N_i}} T_{i,k}\, u_k^T v_j)^2 \qquad (4)$$

Where, $\sum_{uk \in Ni} T_{ik} U_k$ is the trust network that provide the weighted average preference. Here Social MF requires each row of S to be normalised to 1. SocialMF is used to solve the optimization problem that capture trust information,

$$\underset{U,V}{min} \|W\Theta(R - U^T V)\|_F^2 + \alpha \sum_{i=1}^{n} \left(u_i - \sum_{u_k \epsilon N_i} T_{ik} U_k\right)^2 + \lambda(\|U\|_F^2 + \|V\|_F^2) \qquad (5)$$

Where, T denotes the user and item description the set of terms that appeared in item and user description, α is normalization factor.

3.2 Knowledge Graph and Item Recommendation Phase

This phase acknowledges the incorporation of knowledge graph into the recommendation phase of the items to the user.

3.2.1 Knowledge Graph: Knowledge graph is a knowledge base where to integrate data they use a graph- structured topology. They are used to store interlinked descriptions of entities like objects, events, abstract concepts or situations with free-from semantics. To one another, entity descriptions contribute to form a network and represents part of the description of the entities to provide context for their interpretation.

3.2.2 Item Recommendation: The Y={(u,i)} represntes user-item interactions and implicit feedback is used for the pair of user and item that shows that user 'u' consumes the item i ∈ I. We have used knowledge graph to provide top-N items that enhances trust in recommendation system.

3.2.3 KG Completion: A directed graph that is composed of the triple facts having subject-property-object and is known as Knowledge Graph KG. The purpose of triplate is to show the relationship of e_h to e_t that means from head to tail entity. The triplet is denoted as (e_h,e_t, r), where e_h,e_t are the entities and r is the relations. In order to predict the missing entity KG completion is used which is considered as top-N entities for a target (e_t ,r) or (e_h,r). Here we have used KG completion to repair the incomplete knowledge graph to avoid data sparsity and cold start problem.
The energy score function [6] is defined as follows for a triplet:

$$f(e_h, e_t, r) = \left|\left| e \perp \mathrm{h} + \mathrm{r} - \mathrm{e} \perp \mathrm{t} \right|\right| \qquad (6)$$

Where, f (e_h, e_t ,r) denotes the possibility of true of the triplet, else no. e ⊥ h and e ⊥ t are projected entity vectors:

$$e \perp \mathrm{h} = e_h - w_r^T e_h w_r \qquad (7)$$

$$e \perp \mathrm{h} = e_h - w_r^T e_t w_r \qquad (8)$$

where, w_r and r are two learned vectors of relation r, w_r denotes the projection vector of the corresponding hyperplane, and r is the translation vector.

3.2.4 Top-N Item Recommendation

The knowledge graph completion provides the final list of items to be recommended. The top-N items are selected and further pushed to the target user as list of recommendation.

4. Experimental and Evaluation Of Proposed Approach

The proposed work has evaluated using two factors: Time and Efficiency. In order to include the factors, we have used the standard metrics: MAE and RMSE.

4.1 Dataset

In the proposed work we have incorporated trust with knowledge graph using the toy dataset. We have combined the trust factors SoRec and SocialMF in order to get the trust factors between the users. We have used the epinion dataset for the evaluation of the two methods [23]. Epinions was a general consumer review site. The dataset contains ratings from users on items, item_id and user_id. Another one personally by scraping the site https://www.zomato.com/. We made it in the same structure as the Epinions dataset. The difference between zomato and epinions is that this has way more trust relations for a group of users. The toy dataset used for the evaluation purpose is as follows:

Table 4.1.1: User-Item Matrix

	i1	i2	i3	i4	i5	i6	i7	i8
u1	5	2		3		4		
u2	4	3			5			
u3	4		2				2	4
u4								
u5	5	1	2		4	3		
u6	4	3		2	4		3	5

The above is the user item rating matrix and represent a toy problem. The predicted user item rating matrix filled is as follows using the proposed method:

Table 4.1.2: Predicted User-Item Matrix

	i1	i2	i3	i4	i5	i6	i7	i8
u1	5	2	**2.5**	3	**4.8**	4	**2.2**	**4.8**
u2	4	3	**2.4**	**2.9**	5	**4.1**	**2.6**	**4.7**
u3	4	**1.7**	2	**3.2**	**3.9**	**3.0**	2	4
u4	**4.8**	**2.1**	**2.7**	**2.6**	**4.7**	**3.8**	**2.4**	**4.9**
u5	5	1	2	**3.4**	4	3	**1.5**	**4.6**
u6	4	3	**2.9**	2	4	**3.4**	3	5

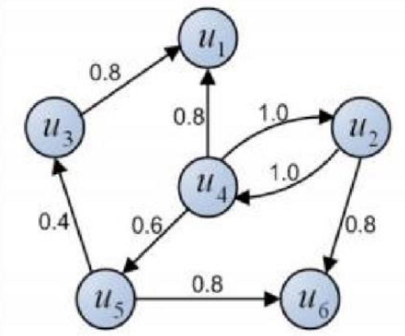


Fig. 3. Social Network Graph

4.2 Evaluation

In order to evaluate time and efficiency of our proposed approach, we have used the standard metrics: MAE and RMSE.

MAE (Mean Absolute Error) targets the average derivation between the ratings given by proposed system and actual rating in recommender system.

$$MAE = \sum_{i,j} |\widehat{r_{u,i}} - r_{u,i}|/N \qquad (9)$$

Where, is $r_{u,i}$ is rating of item and $\hat{r}_{u,i}$ represents the predicting rating, and N is the total number of items.

RMSE is given as follows:

$$RMSE = \sqrt{\frac{\sum_{(u,i)\in TestSet}(r_{u,i} - r_{u,i})^2}{n}} \qquad (10)$$

where is $r_{u,i}$ is rating of item and $\hat{r}_{u,i}$ represents the predicting rating, N is the total number of items.

As per the convention we have used the 80% of data as Train set and rest 20% data as Test set. We have also compared the MAE in the varying user and item numbers. In that variation the MAE has been recorded between 0.82 to 0.87.

Table 4.1.3: Time Comparison between KGTRS and TARS

User/Items	Run Time	MAE in Proposed System	MAE in CF	Training Data
49290/139738	29 min	0.82	0.90	99%
49290/139738	26min	0.83	0.932	80%
7000/21000	20 sec	0.93	0.932	80%
3000/9000	5 sec	0.90	0.932	80%
7000/21000	20 sec	0.88	0.90	99%
3000/9000	4 sec	0.87	0.90	99%

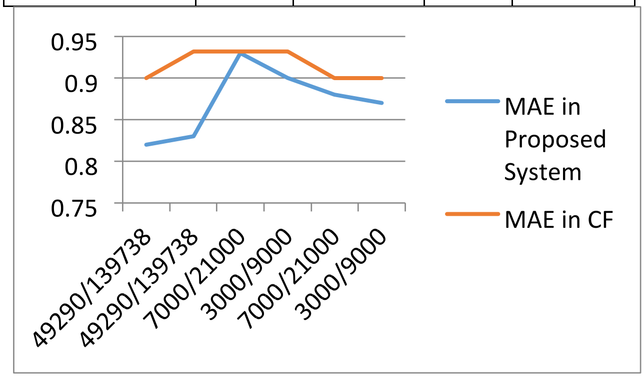


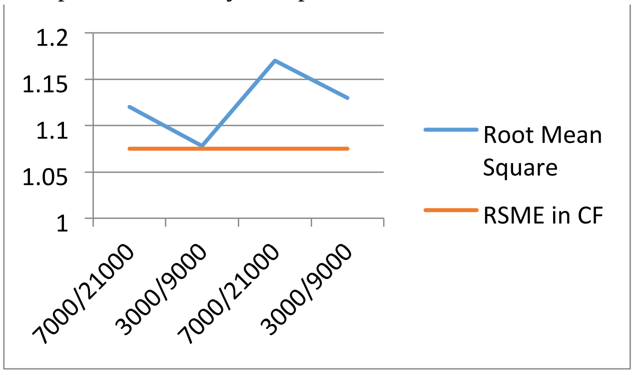
Fig 4. Graph for Time Comparison between KGTRS and CF

Table 4.1.5: Efficiency Comparison between KGTRS and *CF*

User/Items	RMSE in proposed KGTRS	RMSE in CF	Training Data
7000/21000	1.12	1.075	90%
3000/9000	1.078	1.075	90%

7000/21000	1.17	1.075	80%
3000/9000	1.13	1.075	80%

4.1.6: Graph for Efficiency Comparison between KGTRS and TARS



4.2.1 Discussion

Comparison of Hybrid Sorec and SocialMF

The i^{th} user u_i share the user preference vector v_i in the rating information and the trust relation space.

As per the convention we have used 90% of data as Train set and rest 10% data as Test set. We have compared RMSE between the two trust factors SocialMf and Hybrid which recorded 1.078 RMSE and and 1.054 RMSE respectively.

Below is the comparison between epinions and zomato

Table 6.1.7: Comparison between Epinions and Zomato

Parameters/Dataset	Epinions	Zomato
Users	49,290	8028
Items	1,39,738	92,324
Item Ratings	6,64,823	2,33,974
RMSE	1.078	1.054

As per the convention we have used 90% of data as train set and rest 10% data as Test set with 3000 users and 9000 items. We have compared MAE between the two datasets which recorded 0.87 and 0.53 respectively.

5. Conclusion

Recommender system is an information filtering approach that assist overburden users in decision making process. The proposed work includes the trust and knowledge graph in recommendation where trust help to increase the user acceptance towards the presented recommendation. While handling trustworthiness in recommendation it is important to maintain the relation between various entities. The incorporation of knowledge graph has helped to deal with not only the above-mentioned aspects but also to deal with the sparsity and cold start problem. The results shows that the proposed system outperforms as compared to the traditional approach of recommendation.

References

[1] Ponnam, Lakshmi Tharun, et al. "Movie recommender system using item-based collaborative filtering technique." Emerging Trends in Engineering, Technology and Science (ICETETS), International Conference on. IEEE, 2016.

[2] Catherine, Rose, and William Cohen. "Personalized recommendations using knowledge graphs: A probabilistic logic programming approach." Proceedings of the 10th ACM Conference on Recommender Systems. 2016.

[3] Massa, Paolo, and Paolo Avesani. "Trust-aware recommender systems." Proceedings of the 2007 ACM conference on Recommender systems. 2007.

[4] Avesani, Paolo, Paolo Massa, and Roberto Tiella. "A trust-enhanced recommender system application: Moleskiing." Proceedings of the 2005 ACM symposium on Applied computing. 2005.

[5] Sohail, Shahab Saquib, Jamshed Siddiqui, and Rashid Ali. "Classifications of recommender systems: A review." Journal of Engineering Science & Technology Review 10, no. 4 (2017).

[6] Pazzani, Michael J., and Daniel Billsus. "Content-based recommendation systems." The adaptive web. Springer, Berlin, Heidelberg, 2007. 325-341.

[7] Bedi, Punam, Sumit Kumar Agarwal, and Richa. "Trust and reputation-based multi-agent recommender system." International Journal of Computational Science and Engineering 16, no. 4 (2018): 350-362.

[8] Richa and Punam Bedi. "Combining trust and reputation as user influence in cross domain group recommender system (CDGRS)." Journal of Intelligent & Fuzzy Systems 38, no. 5 (2020): 6235-6246.

[9] Sharma, Lalita, and Anju Gera. "A survey of recommendation system: Research challenges." *International Journal of Engineering Trends and Technology (IJETT)* 4.5 (2013): 1989-1992.

[9] P. Bedi, S. K. Agarwal, V. Jindal, and Richa, "MARST: Multi-Agent Recommender System for e-Tourism Using Reputation Based Collaborative Filtering," in International Workshop on Databases in Networked Information Systems, Aizu-Wakamatsu City, Japan, 2014, pp. 189-201.

[11] Victor, Patricia, Martine De Cock, and Chris Cornelis. "Trust and proposals." *Recommender systems handbook*. Springer, Boston, MA, 2011. 645-675.

[12] Richa, and Punam Bedi. "Trust and distrust based cross-domain recommender system." Applied Artificial Intelligence 35, no. 4 (2021): 326-351.

[13] Mustansar Ali Ghazanfar and Adam Prugel-Bennett," A Scalable, Accurate Hybrid Recommender System", 2010 Third International Conference on Knowledge Discovery and Data Mining.

[14] Qian Wang, Xianhu Yuan, Min Sun "Collaborative Filtering Recommendation Algorithm based on Hybrid User Model", FSKD, 2010.

[15] Cao, Yixin, et al. "Unifying knowledge graph learning and recommendation: Towards a more robust understanding of user preferences." The planet wide web conference. 2019.

[16] Golbeck, Jennifer, Bijan Parsia, and James Hendler. "Trust networks on the semantic web." International workshop on cooperati4e information agents. Springer, Berlin, Heidelberg.

[17] Dev, Arpan V., and Anuraj Mohan. "Recommendation system for big data applications based on set similarity of user preferences." Next Generation Intelligent Systems (ICNGIS), International Conference on. IEEE, 2016.

[18] Katarya, Rahul, and Om Prakash Verma. "A collaborative recommender system enhanced with particle swarm optimization technique." Multimedia Tools and Applications 75.15 (2016): 9225-9239.

[19] Guo Q, Zhuang F, Qin C, Zhu H, Xie X, Xiong H, He Q. A survey on knowledge graph-based recommender systems. IEEE Transactions on Knowledge and Data Engineering. 2020 Oct 7.

[20] Gupta S, Nagpal S. Trust aware recommender systems: a survey on implicit trust generation techniques. International Journal of Computer Science and Info rmation Technologies. 2015;6(4):3594-9.
[21] Xie, Li, Wenbo Zhou, and Yaosen Li. "Application of Improved Recommendation System Based on Spark Platform in Big Data Analysis." Cybernetics and Information Technologies 16.6 (2016): 245- 255
[22] Schwabe, Daniel. "Trust and Privacy in Knowledge Graphs." Companion Proceedings of The 2019 World Wide Web Conference. 2019.
[23] Linden, Greg, Brent Smith, and Jeremy York Amazon.Com. "Industry report: Amazon. com recommendations: Item-to-item collaborative filtering." IEEE Distributed Systems Online. 2003.