An Effective e-Commerce Recommender System Based on Trust and Semantic Information

Qusai Y. Shambour[1], *Nidal M. Turab*[2], *Omar Y. Adwan*[3,4]

[1]*Department of Software Engineering, Faculty of Information Technology, Al-Ahliyya Amman University, Amman, Jordan*
[2]*Department of Networks and Information Security, Faculty of Information Technology, Al-Ahliyya Amman University, Amman, Jordan*
[3]*Department of Computer Science, Faculty of Information Technology, Al-Ahliyya Amman University, Amman, Jordan*
[4]*Department of Computer Information Systems, The University of Jordan, Amman, Jordan*
E-mails: q.shambour@ammanu.edu.jo   n.turab@ammanu.edu.jo   adwanoy@ammanu.edu.jo

Abstract: *Electronic commerce has been growing gradually over the last decade as a new driver of the retail industry. In fact, the growth of e-Commerce has caused a significant rise in the number of choices of products and services offered on the Internet. This is where recommender systems come into play by providing meaningful recommendations to consumers based on their needs and interests effectively. However, recommender systems are still vulnerable to the scenarios of sparse rating data and cold start users and items. To develop an effective e-Commerce recommender system that addresses these limitations, we propose a Trust-Semantic enhanced Multi-Criteria CF (TSeMCCF) approach that exploits the trust relations and multi-criteria ratings of users, and the semantic relations of items within the CF framework to achieve effective results when sufficient rating data are not available. The experimental results have shown that the proposed approach outperforms other benchmark recommendation approaches with regard to recommendation accuracy and coverage.*

Keywords: *e-Commerce, Recommender systems, Collaborative filtering, Multi-criteria, Trust network, Semantic similarity.*

1. Introduction

Electronic Commerce (e-Commerce) is a medium where the entire online process of buying and selling of products and services is done over the Internet. Through the Internet, companies can communicate with their customers in a more effective way by offering more information, choices, and personalized services. The emergence of e-Commerce is the tendency of the expansion of the market in the digital society. From the consumers' perspective, e-Commerce helps to save their time, makes shopping more advantageous and cheaper. From the sellers' perspective, e-Commerce helps them to save cost such as rent and helps them in reaching markets

that cannot otherwise be reached. However, *the popularity of e-Commerce has increased* the amount of information that consumers must process to be able to decide on which products/services can meet their needs. An effective solution to this information overload problem is the utilization of recommender systems [1-5].

Recommender systems are used to assist users in identifying the items that suit their needs and preferences in an efficient way. The deployment of recommender systems in the e-Commerce sector has been an opportunity for constant profit and business improvement. For example, big e-Commerce companies such as Amazon, Netflix, eBay, CD-Now and Alibaba use recommender systems to promote their products and services, and enhance their customer support systems [1-4]. S c h a f e r, K o n s t a n and R i e d l [4] report that recommender systems can increase sales in e-Commerce websites by: (1) Converting the browsers who often browse e-Commerce websites without any intention of buying anything into buyers by helping them in finding the things they might like or need throughout personalized recommendations; (2) Improving the cross-selling by recommending supplementary products to buyers that might complement the products that are already chosen to buy; and (3) Improving customers loyalty by using their profiles in e-Commerce websites to offer them customized recommendations that match their needs, which increases the satisfaction of them towards services provided by such websites.

Collaborative Filtering (CF) is the most widely personalized technique of choice for most recommendation algorithms that are extensively employed in several e-Commerce recommender systems and many others. CF works by predicting a number of items a user may prefer by learning the similarity patterns across users and items. CF can be classed into the user-based CF and item-based CF approaches. User-based CF works by identifying the most similar users to the active user, known as user's nearest neighbours, that have relevant preferences based on their historical data, and analyze these preferences to make personalized recommendations for the active user. Item-based CF works by identifying the most similar items to the target item, known as item's nearest neighbours, that have received similar preferences based on their historical data, and analyze these preferences to make personalized recommendations for the target item [6-8].

However, due to the huge numbers of users and items presented in major e-Commerce recommendation systems, uniquely challenging limitations of CF are emerged referred to as Sparsity and Cold Start user and item problems. The reason of the sparsity problem is that most users merely interact with a small proportion of the items in such huge systems leaving only a few ratings in these systems. As a result of sparsity, it becomes very difficult to define the similarity between users that may affect the identification of suitable neighbours, which results in making inaccurate recommendations. Cold start user problem, also known as new user problem, refers to the condition where a new user has to rate an adequate number of items before being able to get reliable and accurate recommendations. Likewise, cold-start item problem, also known as new item problem, refers to the condition where a new item cannot be recommended except it has been rated by a considerable number of users [9, 10].

An extension of overall rating based-recommender systems is the multi-criteria based-recommender systems, which are nowadays very popular and employed by modern e-Commerce services like the Yahoo! Movies, TripAdvisor, OpenTable, and so forth. In multi-criteria recommender systems, a user has the ability to provide ratings to multiple aspects, in addition to the overall rating on each item. Multi-criteria recommender systems aim to take advantage of these multi-criteria ratings to accurately model users' preferences, which accordingly leads to more accurate and effective recommendations [11, 12].

To develop an effective e-Commerce recommender system that addresses the above mentioned limitations, this study proposes a Trust-Semantic enhanced Multi-Criteria CF (TSeMCCF) approach that exploits the trust relations and multi-criteria ratings of users, and the semantic relations of items within the CF framework. The TSeMCCF approach utilizes the trust and semantic information in order to achieve effective results in alleviating the impact of sparsity, new user and new item problems when sufficient data are not available. In addition, it exploits the additional information provided by the multi-criteria ratings of users to accurately model users' preferences, which accordingly leads to more accurate recommendations. The TSeMCCF combines two recommendation approaches, the user-based MC implicit trust and the item-based semantic CF. The user-based MC implicit trust CF approach makes use of the inherent characteristics of trust and trust propagation to deal with the data sparsity and new user limitations. The item-based semantic CF approach makes use of the items' semantic relations to help lessen the impact of data sparsity and new item limitations.

The rest of the study is organized as follows: Section 2 briefly summarizes the recent related work on the applications of recommender systems on e-Commerce; Section 3 elaborates in details the proposed TSeMCCF approach; Section 4 demonstrates the experimental evaluation of the TSeMCCF approach; finally, the conclusions and future research of this study are presented in Section 5.

2. Related works

Over the past decade, many studies have been focusing on improving the e-Commerce sector by the utilization of various recommendation systems. In this section, we briefly summarize recent related works of the applications of recommender systems in the e-Commerce domain.

Putra et al. [13] propose an e-Commerce recommender system that utilizes user-user similarity and item-item similarity to make predictions on target items. The similarities are computed through link prediction in a graph representation. The experiment results on a dataset collected from an e-Commerce company in Indonesia show that the proposed method improves accuracy compared to the classical user-based CF and item-based CF algorithms. Whereas, Chu and Lee [14] propose an e-Commerce recommender system that utilizes user comments. The authors adopt the Word2vec technique to analyze the semantics of the comments that have been made by users, and then construct a unique vector for each word. After that, the PCA dimensionality reduction technique is used to reduce the dimensions of the vectors

into a lower dimension space. Subsequently, a clustering algorithm is applied to group the items in a few number of clusters. Finally, based on the clusters, items recommendations are produced for each user. Experimental results on a real dataset from Amazon Movie & TV Shows have shown that the proposed system performs well in terms of prediction accuracy. On the other hand, X i a o and E z e i f e [15] propose an e-Commerce recommender system, called HPCRec, that utilizes historical purchase in addition to clickstream data. The proposed system mines the consequential relationship between historical purchases and clicks to improve the quality of rating predictions. Experimental results on a real dataset from an online retailer in Europe have shown that the proposed system outperforms some of the existing recommendation methods in terms of recommendation accuracy. In the study of C h o i and K i m [16], an e-Commerce recommender system for repetitively purchasing items based on CF and association rules is proposed. Authors analyse the items' sales pattern by implicitly assigning ratings for repetitively purchasing items using a normalized repetitive purchase counts, and then these ratings are utilized by user-based and item-based CF algorithms for the purpose of making recommendations. Authors also propose an associate item-based recommendation approach that recommends top-t associate items, which are repeatedly bought by users simultaneously, using the association rule mining algorithm. Experimental results on a real dataset from Happycoop Co., an online Korean consumers' cooperative company, have shown the results of the proposed recommendation methods in terms of recommendation hit ratio, recommendation accuracy, recommendation coverage, and computation time. Authors recommend to integrate both of the item-based CF and the associate item-based algorithms to recommend the top-N items. S a s s a n i, A l a h m a d i and S h a r i f z a d e h [17] propose a cluster based CF recommendation approach, called CFGA, to enhance the performance of recommender systems in e-Commerce. The proposed approach consists of an offline and an online phases. A genetic algorithm is used in the offline phase to effectively cluster similar users, and then unrated items for cluster members of the active user are predicted in the online phase. Experimental evaluation on the Movielens dataset has shown that the proposed approach performs better than the conventional CF, genetic algorithm and k-Means methods in terms of prediction accuracy. In the study by K h o d a b a n d e h l o u [18], an e-Commerce recommender system that addresses the data sparsity problem has been proposed. Accordingly, a set of data mining methods have been utilized to provide users with reliable and accurate recommendations. Experimental results on a real dataset from an online retailer in Iran have shown that the proposed system outperforms a traditional CF-based recommendation method in terms of recommendation precision. J i a n g et al. [19] propose an e-Commerce recommender system that resolves the problem of low accuracy of the conventional slope one algorithm by fusing trusted data and user similarity. The proposed system works by selecting the trusted data, then it calculates the users' similarity and uses this similarities as weights in the rating predictions process. Experimental results on an Amazon dataset demonstrate that the proposed system outperforms the traditional slope one recommendation algorithm in terms of recommendation accuracy. I f t i k h a r et al. [20] propose an enhanced product

recommendation approach by introducing a novel triangle similarity measure between users that mitigates the limitations of existing similarity measures. The proposed similarity measure considers the non-common rating of users in addition to their common rating. Experimental evaluation on six datasets reveals that the proposed system outperforms a number of recommendation algorithms that deploy existing similarity measures in terms of recommendation accuracy. S i n g h and R i s h i [21] propose an e-Commerce recommender system using knowledge based-CF technique. The proposed system exploits a knowledge graph to be aware of the domain knowledge of users, items, and relations among them. The experimental results on the e-Shop simulated dataset show the effectiveness of the proposed system in terms of the recommendation accuracy in comparison with the baseline recommendation approaches. M o h a n a and S u r i a k a l a [3] propose an improved e-Commerce recommender system by developing an enhanced similarity measure, called as prospective Jaccard similarity measure, that alleviates the limitations of conventional similarity measures. Experimental evaluation on the MovieLens dataset shows the superiority of the proposed system in outperforming some of the recommendation algorithms that utilize conventional similarity measures in terms of recommendation accuracy.

Although a variety of e-Commerce have been proposed in literature, the design and development of efficient recommender systems still need to be considered by researchers and practitioners nowadays. The strength of the system proposed in this paper lies in the utilization of trust relations and multi-criteria ratings of users, and the semantic relations of items within the CF framework. This provides the system proposed with the capability to alleviate most of the known problems of the existing recommender systems including sparsity and cold start user and item problems, which hinder the effectiveness of recommender systems in playing a more significant role in the e-Commerce domain. Accordingly, by the elimination of such problem, our proposed system is able to make accurate recommendations for consumers in a wide variety of e-Commerce applications.

3. The Proposed Trust-Semantic enhanced Multi-Criteria Collaborative Filtering Approach

3.1. Preliminaries

For describing the proposed TSeMCCF recommendation approach, we define a number of concepts and entities as below.

Users: $U = \{i_1, i_2, \ldots, i_n\}$ is a set of n users *in an online community who have rated a number of items.*

Items: $I = \{x_1, x_2, \ldots, x_m\}$ is a set of m items *rated by users in U.*

Multi-Criteria Ratings: in a raw user-item MC *rating matrix, $R_{m \times n}$,* each item is articulated by means of a varied set of criteria. In particular, let $\{c_1, c_2, \ldots, c_k\}$, be a list of criteria where an item x is rated upon, each criteria represents an individual aspect of an item with a rating value. According to the principles of Multi Attribute Utility Theory [22], the total utility T of item x with respect to user i is expressed as an additive value function as follows:

$$(1) \qquad T^i(x) = \sum_{k=1}^{z} w_k^i(x) \times c_k^i(x), \ \ \text{where} \ \sum_{k=1}^{z} w_k^i(x) = 1,$$

and $w_k^i(x)$ is a weight that represents the significance of criterion c_k on item x by user i, and $c_k^i(x)$ is the rating of user i with respect to criterion c_k on item x.

Implicit Trust Network is a directed graph $G = (U, E, \omega)$, where U is set nodes that represents the set of users, E is a set of edges that represent the interactions between users, and ω is an edge weight that represents the implicit trust value between each pair of users.

Item Taxonomy is a hierarchical tree of categories given by domain experts to classify a set of items at a low level into a more general category at a higher level. The leaves of the tree denote the items and non-leaf nodes denote item classes or subclasses, each item can be associated with one or more categories. Let $T=\{G_1, G_2,..., G_g\}$ be an item taxonomy that includes a number of g categories to which items might be assigned into one or more of them. Every item is characterized as a vector of binary values [0, 1],

$$(2) \qquad \overline{V}_{x,g} = (v_{x,1}, v_{x,2}, ..., v_{x,g}),$$

where $\overline{V}_{x,g}$ represents a vector for item x and its associated categories. We define $v_{x,d}$ ($d=1, \ldots, g$) as a binary variable, as follows:

$$(3) \qquad v_{x,d} = \begin{cases} 1 & \text{if Item } x \text{ belong to category } d, \\ 0 & \text{if Item } x \text{ does not belong to category } d. \end{cases}$$

For simplicity, all the symbols and notations used in this study are described in Table 1.

Table 1. Description of *symbols and notations*

Symbol/Notation	Descriptions
U, I	The set of Users and Items, respectively
n, m	The total number of users and items, respectively
$\overline{r_i}$	The mean ratings of user i
$\overline{r_j}$	The mean ratings of user j
$\overline{r_x}$	The mean ratings of item x
$\overline{r_{i,y}}$	The mean rating of item y by user i derived from all rating criteria
$U^j(x)$	The overall rating (i.e., total utility) on item x for user j
$P_{i,x}$	The prediction rating by user i on item x
I_i	The total number of items rated by user i
$I_{i,j}$	The total number of common items among users i and j
iTrust$_{i,j}$	The direct implicit trust score among users i and j
sSem$_{x,y}$	The semantic similarity value among items x and y
NN^{iTrust}	The nearest neighbors to the active user based on their implicit trust scores
NN^{sSem}	The nearest neighbors to the target item based on their semantic similarities

3.2. The recommendation process

This section describes the recommendation process of the TSeMCCF approach, which is a hybridization of a user-based MC implicit trust CF and an item-based semantic CF techniques, in four main tasks.

3.2.1. The computation of user-based MC implicit trust

Two steps are performed to compute the user-based MC implicit trust values between every pair of users. The direct implicit trust values of every pair of users are derived based on users' ratings to construct the implicit trust network. Then, a trust propagation formula is applied on the implicit trust network to derive the indirect trust between unconnected users.

1) User-based direct implicit trust

Since trust between users in most e-Commerce applications depends on the similar views and preferences between them on products or services, their implicit trust can be explained as the similarity between their interests on shared items. Thus, the direct implicit similarity of every pair of users is computed by measuring the prediction accuracy of a given user as a reliable recommender to another user based on their historical ratings. For example, users i and j should get a high implicit trust score if user j can deliver accurate recommendations to user i based on their historical ratings. For this purpose, first the prediction method proposed by R e s n i c k et al. [23] is applied to estimate the predicted ratings for every pair of users:

$$(4) \qquad P_{i,x} = \overline{r_i} + (U^j(x) - \overline{r_j}).$$

Then, the Absolute Difference of Ratings (ADF) method [24] is employed to compute the implicit trust value between users i and j based on the absolute difference between the ratings and predicted ratings of their co-rated items:

$$(5) \qquad \mathrm{ADF}_{i,j} = \frac{\sum_{x=1}^{I_{i,j}} \exp\left(-\frac{\left|p_{i,x} - U^i(x)\right|}{\max\left(p_{i,x}, U^i(x)\right)}\right)}{I_{i,j}}.$$

Though the absolute difference of ratings plays a vital role in determining the level of trust between users, it also has limitations that confidence can resolve. The confidence signifies the reliability between two users based on the number of common items that have been rated by them. The more items that have been rated between the two users, the higher is the degree of confidence of their estimated trust [25, 26]. In order to compute the confidence, the Dice coefficient is applied as a weighted function to realize the confidence level between any pair of users, according to the next equation

$$(6) \qquad \mathrm{UDice}_{i,j} = \frac{2 \times |I_{i,j}|}{|I_i| + |I_j|}.$$

At last, the implicit trust metric for any given pair of users is defined by

$$(7) \qquad \mathrm{iTrust}_{i,j} = \mathrm{ADF}_{i,j} * \mathrm{UDice}_{i,j}.$$

2) User-based implicit trust propagation

Based on the direct implicit trust, an implicit trust network is formed where users are directly connected to each other based on the degree of trust among them. However, such network is extremely sparse due to the limited ratings given by users in most applications of recommender systems. To maximize the use of such sparse network, our approach utilizes the trust propagation feature in social networks in which trust can be propagated among various users in the implicit trust network. In the implicit trust network, any two users who have no direct link can be linked together through intermediary users, thus new indirect connections of trust can be produced in the implicit trust network. For example, assume that user i has a direct association with user j (intermediary user) and user j has a direct association with user k, it can be inferred through trust propagation that user i can trust the preferences of user k to a certain degree.

For this purpose, the following aggregation function that takes into account a confidence weight when measuring the propagated trust between users, is proposed. For any $i, j, k \in U$, the propagated trust score that denotes to what extent user i implicitly trust user k via an intermediary user j, is computed as follows:

$$(8) \qquad \mathrm{PSim}_{i,k} = \frac{\displaystyle\sum_{j \in \mathrm{intermediary}(i \, and \, k)} \left(\mathrm{UDice}_{i,j} \times \mathrm{iTrust}_{i,j} \right) + \left(\mathrm{UDice}_{j,k} \times \mathrm{iTrust}_{j,k} \right)}{\displaystyle\sum_{j \in \mathrm{intermediary}(i \, and \, k)} \mathrm{UDice}_{i,j} + \mathrm{UDice}_{j,k}}.$$

3.2.2. The computation of item-based semantic similarity

To exploit the semantic relationships existing between items, an item taxonomy has to be created in a particular domain by identifying and creating the main categories that items may belong to, and assigning each item to one or more of the main categories. The item-based semantic similarity among two items x and y is computed based on the proportion of their mutual categories using the binary Dice coefficient [27], as shown in the next equation

$$(9) \qquad \mathrm{sSim}_{x,y} = \frac{2 * C_{11}}{C_{01} + C_{10} + C_{11}},$$

C_{11} = Total number of categories where $v_{x,g}$ is 1 and $v_{y,g}$ is 1,

C_{01} = Total number of categories where $v_{x,g}$ is 0 and $v_{y,g}$ is 1,

C_{10} = Total number of categories where $v_{x,g}$ is 1 and $v_{y,g}$ is 0.

For instance, in the following movie taxonomy depicted in Fig. 1, the binary vector of Movie $X = (1, 1, 0, 0, 1, 0)$, and Movie $Y = (0, 1, 1, 0, 1, 1)$. To calculate their item-based semantic similarity, we first have to get the values of C_{11}, C_{01} and C_{10}; $C_{11} = 2$ (locations 2 and 5 have the binary value 1 in both vectors, respectively);

110

$C_{01} = 2$ (locations 3 and 6 have the binary values 0 and 1 in both vectors, respectively); $C_{10} = 1$ (location 1 has the binary value 1 in vector X and the binary value 0 in vector Y). Then in accordance with Equation (9), the item-based semantic similarity between Movies X and Y will be 0.80 ($2 \times 2/(2+1+2) = 0.80$).
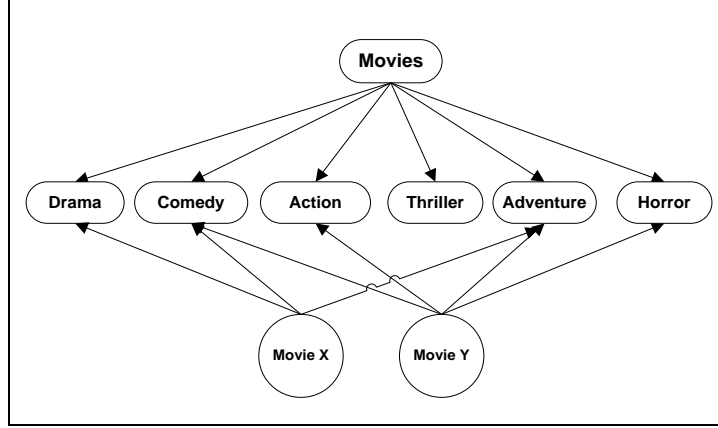

Fig. 1. The movie taxonomy

3.2.3. The selection of nearest neighbors

The most nearest neighbors to the active user and target item are selected based on the user-user implicit trust NN^{iTrust} and item-item semantic similarities NN^{sSem}, respectively. For the selection of nearest neighbors, we use the Top-k method in which a predefined number of users/items with highest trust scores/semantic similarities are selected as nearest neighbors.

3.2.4. The computation of predicted rating

The process of computing the predicted rating of unrated item x by an active user i includes two main steps. First, the following prediction metric [28] is utilized twice to compute the predicted rating for each unrated item based on the user-based implicit trust, and based on the item-based semantic similarity, respectively:

$$(10) \qquad P_{i,x}^{iTrust} = \bar{r}_i + \frac{\sum_{j=1}^{NN^{iTrust}} (iTrust_{i,j} \times (\overline{r_{j,x}} - \bar{r}_j))}{\sum_{j=1}^{NN^{iTrust}} iTrust_{i,j}},$$

$$(11) \qquad P_{i,x}^{sSim} = \bar{r}_x + \frac{\sum_{y=1}^{NN^{sSim}} (sSim_{x,y} \times (\overline{r_{i,y}} - \bar{r}_y))}{\sum_{n=1}^{NN^{sSim}} sSim_{x,y}}.$$

After that, the switching hybridization strategy [29] is employed to swap the recommendation techniques according to a heuristic reflecting the recommender

capability to produce a predicted rating. Also, the harmonic mean method [30] is utilized to combine the above predictions to assure that a high predicted rating is reached when both approaches produce high predictions. Specifically, the final predicted rating for an active user i on un-rated item x, is

$$
(12) \qquad P_{i,x} = \begin{cases} 0 & \text{if} \quad P_{i,x}^{\text{iTrust}} = 0 \text{ and } P_{i,x}^{\text{sSem}} = 0, \\[2mm] P_{i,x}^{\text{iTrust}} & \text{if} \quad P_{i,x}^{\text{iTrust}} \neq 0 \text{ and } P_{i,x}^{\text{sSem}} = 0, \\[2mm] P_{i,x}^{\text{sSem}} & \text{if} \quad P_{i,x}^{\text{iTrust}} = 0 \text{ and } P_{i,x}^{\text{sSem}} \neq 0, \\[2mm] \dfrac{2 \times P_{i,x}^{\text{iTrust}} \times P_{i,x}^{I}}{P_{i,x}^{\text{iTrust}} + P_{i,x}^{I}} & \text{if} \quad P_{i,x}^{\text{iTrust}} \neq 0 \text{ and } P_{i,x}^{\text{sSem}} \neq 0. \end{cases}
$$

4. Experimental evaluation

A number of experiments have been carried out to evaluate the performance of the proposed TSeMCCF recommendation approach on a real-world MC dataset. Then, three commonly used evaluation measures have been adopted to compare the recommendation results of the proposed approach with other recommendation methods.

4.1. Dataset description and evaluation methods

The Yahoo! Movies MC dataset [31], which was collected from the Yahoo! movies website (**http://movies.yahoo.com**), is used to evaluate the performance of the proposed TSeMCCF recommendation approach. The dataset consists of 1716 users, 965 movies and 34,800 ratings on the scale from 1 to 5 on four criteria: story, acting, direction and visuals, plus an overall rating. We built a two-level hierarchical taxonomy based on the movie genres. The first level contains 32 main categories that represent the movie genres including: Sci-fi, adventure, comedy, action, etc. The second level includes the movies as leaf nodes. Each movie is assigned to one or more of the categories in the first level.

The accuracy of prediction of the proposed approach is measured in terms of Mean Absolute Error (MAE) and Root Mean Square Error (RMSE) metrics, whereas the prediction coverage is measured using the Rating Coverage metric. MAE is calculated as an average of the absolute difference of the predicted ratings from the respective actual ratings. RMSE is similar but more stringent than the MAE as it measures the large errors in extreme cases. In both metrics, the smaller the values a recommendation approach obtains, the more accurate the predictions are. In addition, the Rating Coverage calculates the percentage of ratings for which a given recommendation approach is able to produce predictions of some value (refer to [32] for more information on the metrics).

4.2. Comparison methods

For comparison purposes, we have considered the following traditional and stat-of-the-art user-based and item-based CF recommendation approaches to compare the results of the proposed TSeMCCF approach with their results:

- The Multiple-Criteria User Based CF (MC-UBCF) [33]: the Pearson Correlation Coefficient has been used as similarity metric between users to make personalized recommendations.
- The Multiple-Criteria Item Based CF (MC-UBCF) [33]: the Pearson Correlation Coefficient has been used as similarity metric between items to make personalized recommendations.
- The Multi-Criteria User-based Trust-enhanced CF (MC-TeCF) [11]: the multi-criteria ratings and implicit trust relations among users have been utilized to make personalized recommendations.
- The Multi-Criteria Item-based Semantic-enhanced CF (MC-SeCF) [12]: the multi-criteria ratings and semantic relations among items have been utilized to make personalized recommendations.

4.3. Experimental results

In this section a set of experiments have been conducted to validate the effectiveness of the TSeMCCF approach over a number of traditional and state-of-the-art user-based and item-based CF recommendation approaches. In experiments the effect of the TSeMCCF approach over the prediction accuracy, data sparsity, new user and new item problems in comparison with benchmark approaches are presented.

4.3.1. Performance comparison of different number of nearest neighbors

Table 2 shows the prediction accuracy in terms of MAE and RMSE of the five approaches on the Yahoo Movies! dataset. It can be seen that with the increasing number of nearest neighbors, the prediction accuracy of all algorithms steadily increases. The best performance of the proposed approach is obtained when the number of k-NN is around 30. Then, the prediction accuracy decreases to some extent with some inadequate ratings being exploited. Furthermore, it can also be seen, the proposed TSeMCCF approach achieves better accuracy compared with other approaches at all varied sizes of nearest neighbors, and the MC-UBCF approach performs the worst. The percentages of MAE improvement of the TSeMCCF over the MC user-based CF, MC item-based CF, MC trust-enhanced CF, and MC semantic-enhanced CF benchmark approaches are 19.73%, 13.13%, 6.15%, and 3.10%, respectively. The percentages of RMSE improvement of the TSeMCCF over the MC user-based CF, MC item-based CF, MC user-based trust-enhanced CF, and MC item-based semantic-enhanced CF benchmark approaches are 24.65%, 19%, 6.17%, and 1.27%, respectively.

Table 2. A performance comparison of prediction accuracy on different number of nearest neighbors

k-NN	MC-UBCF		MC-IBCF		MC-TeCF		MC-SeCF		TSeMCCF	
	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE
5	0.872	1.393	0.810	1.309	0.724	1.093	0.698	1.038	0.693	1.035
10	0.847	1.369	0.781	1.264	0.714	1.082	0.695	1.029	0.676	1.020
15	0.838	1.356	0.772	1.250	0.713	1.081	0.697	1.030	0.675	1.018
20	0.837	1.350	0.771	1.249	0.715	1.082	0.697	1.030	0.675	1.012
30	0.830	1.335	0.768	1.244	0.718	1.084	0.696	1.030	0.669	1.014
50	0.832	1.330	0.770	1.241	0.723	1.084	0.696	1.030	0.671	1.018
70	0.832	1.324	0.771	1.241	0.727	1.089	0.696	1.030	0.671	1.018
Average	0.841	1.351	0.777	1.257	0.719	1.085	0.697	1.031	0.676	1.019

4.3.2. Performance comparison for various levels of sparsity

Table 3 demonstrates the prediction accuracy and coverage in terms of MAE and Coverage measures of the five approaches on different levels of sparsity to validate whether the proposed TSeMCCF approach is positively better than the other benchmark approaches in alleviating the data sparsity problem. It can be seen that with the decreasing level of sparsity, the prediction accuracy and coverage of all approaches steadily increases. This is because the less the sparsity of the rating data the more rating data is available, which leads to better formulation of proper neighbours resulting in producing accurate predictions. Additionally, it can also be observed that the proposed TSeMCCF approach achieves better accuracy and coverage compared with other approaches at all varied levels of sparsity, and the MC-UBCF approach performs the worst. The percentages of MAE improvement of the TSeMCCF over the MC user-based CF, MC item-based CF, MC user-based trust-enhanced CF, and MC item-based semantic-enhanced CF benchmark approaches are 58.35%, 55.19%, 17.46%, and 19.74%, respectively. The percentages of *Coverage* improvement of the TSeMCCF over the MC user-based CF, MC item-based CF, MC user-based trust-enhanced CF, and MC item-based semantic-enhanced CF benchmark approaches are 46.67%, 42.44%, 6.12%, and 9.88%, respectively.

Table 3. A performance comparison of prediction accuracy and coverage at various Sparsity levels

Sparsity	MC-UBCF		MC-IBCF		MC-TeCF		MC-SeCF		TSeMCCF	
	MAE	Coverage	MAE	Coverage	MAE	Coverage	MAE	Coverage	MAE	Coverage
99.8	4.043	0.30	4.037	0.30	2.731	40.550	2.601	44.510	2.070	64.940
99.5	3.734	9.80	3.650	12.62	1.417	90.020	1.658	72.080	0.961	94.770
99.0	2.490	49.04	2.106	58.26	0.851	97.410	0.956	93.140	0.880	98.860
98.8	1.936	65.66	1.599	72.67	0.823	98.070	0.839	95.440	0.764	99.070
98.5	1.427	79.55	1.262	83.73	0.804	98.100	0.813	97.570	0.723	99.420
98.0	0.979	92.53	0.924	92.85	0.746	98.460	0.713	98.960	0.688	99.640
Average	2.435	49.48	2.263	53.41	1.229	87.102	1.263	83.617	1.014	92.783

4.3.3. Performance comparison for cold start users

In Table 4 the prediction accuracy and coverage in terms of MAE and Coverage measures of the three approaches on different number of ratings for new users are given to verify the impact of the proposed TSeMCCF approach in alleviating the new user problem compared with two benchmark user-based CF approaches. It can be

shown that with the increasing number of ratings in the new user profile, the prediction accuracy and coverage of all approaches gradually increases. This is because the more rating data is available, the better is the formulation of suitable neighbors of users resulting in producing more accurate predictions. Furthermore, the proposed TSeMCCF approach achieves better accuracy and coverage compared with other approaches at all varied number of ratings, and the MC-UBCF approach performs the worst. The percentages of MAE improvement of the TSeMCCF over the MC user-based CF and the MC user-based trust-enhanced CF benchmark approaches are 68% and 9%, respectively. The percentages of Coverage improvement of the TSeMCCF over the MC user-based CF and the MC user-based trust-enhanced CF benchmark approaches are 51.47% and 6.19%, respectively.

Table 4. A performance comparison of prediction accuracy and coverage on different numbers of ratings for new users

Number of users	MC-UBCF		MC-TeCF		TSeMCCF	
	MAE	Coverage	MAE	Coverage	MAE	Coverage
10	3.234	22.940	1.207	80.000	1.076	95.180
12	2.877	36.510	0.862	91.000	0.788	98.320
14	2.557	46.750	0.815	93.000	0.760	98.730
16	2.300	54.110	0.811	95.860	0.718	99.000
18	2.045	61.740	0.802	96.010	0.728	99.540
20	1.998	62.710	0.783	97.650	0.737	99.260
Average	2.502	47.460	0.880	92.253	0.801	98.338

4.3.4. Performance comparison for cold start items

In Table 5 the prediction accuracy and coverage are given in terms of MAE and Coverage measures of the three approaches on different number of ratings for new items to investigate the impact of the proposed TSeMCCF approach in alleviating the new item problem compared with two benchmark item-based CF approaches. It can be demonstrated that with the increasing number of ratings in the new item profile, the prediction accuracy and coverage of all approaches progressively increases.

Table 5. A performance comparison of prediction accuracy and coverage on different numbers of ratings for new items

Number of items	MC-IBCF		MC-SeCF		TSeMCCF	
	MAE	Coverage	MAE	Coverage	MAE	Coverage
4	2.778	0.000	2.611	37.040	2.852	46.300
6	2.809	3.610	2.170	47.420	1.979	59.280
10	2.599	14.690	1.836	58.990	1.485	81.580
15	2.255	33.330	1.461	73.580	1.156	90.790
20	1.889	51.600	1.220	83.030	1.061	93.600
25	1.669	63.600	1.106	87.190	1.009	94.880
Average	2.333	27.805	1.734	64.542	1.590	77.738

The reason is that as more rating data is available, more suitable neighbours of items can be identified that resulting in producing more accurate predictions. Moreover, the proposed TSeMCCF approach attains improved accuracy and coverage compared with other item-based CF approaches at all varied number of

ratings, whereas the MC-IBCF approach performs the worst. The percentages of MAE improvement of the TSeMCCF over the MC item-based CF and the MC item-based semantic-enhanced CF benchmark approaches are 31.85% and 8.3%, respectively. The percentages of Coverage improvement of the TSeMCCF over the MC item-based CF and the MC item-based semantic-enhanced CF benchmark approaches are 64.23 and 17%, respectively.

5. Conclusion and future work

This paper explores the use of an effective recommender system to alleviate the influence of the information overload problem in the e-Commerce industry, mainly, assisting consumers in seeking the best available products and services according to their needs and preferences without any difficulties. Consequently, a Trust-Semantic enhanced Multi-Criteria CF approach that exploits the trust relations and multi-criteria ratings of users, and the semantic relations of items within the CF framework has been proposed. The TSeMCCF approach utilizes the trust and semantic information to achieve effective results in alleviating the impact of sparsity, new user and new item problems when sufficient rating data are not available. In addition, it exploits the additional information provided by the multi-criteria ratings of users to accurately model users' preferences, which accordingly leads to more accurate recommendations. The experimental results confirm the effectiveness of the proposed approach in regards to its applicability in dealing with high levels of sparse datasets, in addition to the new user and new items scenarios when compared with other *traditional and stat-of-the-art user-based and item-based CF* benchmark *recommendation* approaches. Future research will focus on examining the impact of incorporating more additional information in the recommendation process, such as product reviews, in further enhancing the performance of the proposed approach.

R e f e r e n c e s

1. P a t r o, S. G. K., B. K. M i s h r a, S. K. P a n d a, R. K u m a r, A. A p o o r v a. Hybrid Social Recommender Systems for Electronic Commerce: A Review. – In: 2020 International Conference on Computer Science, Engineering and Applications (ICCSEA'20), IEEE, 2020, pp. 1-6.
2. L i, Y. M., C. T. W u, C. Y. L a i. A Social Recommender Mechanism for e-Commerce: Combining Similarity, Trust, and Relationship. – Decision Support Systems, Vol. **55**, 2013, No 3, pp. 740-752.
3. M o h a n a, H., M. S u r i a k a l a. An Enhanced Prospective Jaccard Similarity Measure (PJSM) to Calculate the User Similarity Score Set for e-Commerce Recommender System. – In: S. Satapathy, V. Bhateja, B. Janakiramaiah, Y. W. Chen, Eds. Advances in Intelligent Systems and Computing. Singapore, Springer, 2021, pp. 129-142.
4. S c h a f e r, J. B., J. K o n s t a n, J. R i e d l. Recommender Systems in e-Commerce. – In: Proc. of 1st ACM Conference on Electronic Commerce, ACM, 1999, pp. 158-166.
5. S h a m b o u r, Q., J. L u. An Effective Recommender System by Unifying User and Item Trust Information for B2B Applications. – Journal of Computer and System Sciences, Vol. **81**, 2015, No 7, pp. 1110-1126.
6. A g g a r w a l, C. C. Neighborhood-Based Collaborative Filtering. – In: Recommender Systems: The Textbook, Springer International Publishing, Cham, 2016, pp. 29-70.

7. A m b u l g e k a r, H. P., M. K. P a t h a k, M. B. K o k a r e. A Survey on Collaborative Filtering: Tasks, Approaches and Applications. – In: M. Chakraborty, S. Chakrabarti, V. Balas, J. Mandal, Ed. Advances in Intelligent Systems and Computing, Singapore, Springer, 2019, pp. 289-300.

8. S h a m b o u r, Q. Y., M. M. A b u-A l h a j, M. M. A l-T a h r a w i. A Hybrid Collaborative Filtering Recommendation Algorithm for Requirements Elicitation. – International Journal of Computer Applications in Technology, Vol. **63**, 2020, No 1-2, pp. 135-146.

9. S c h a f e r, J. B., D. F r a n k o w s k i, J. H e r l o c k e r, S. S e n. Collaborative Filtering Recommender Systems. – In: P. Brusilovsky, A. Kobsa W. Nejdl, Ed. The Adaptive Web: Methods and Strategies of Web Personalization, Berlin, Heidelberg, Springer, 2007, pp. 291-324.

10. A d o m a v i c i u s, G., A. T u z h i l i n. Toward the Next Generation of Recommender Systems: A Survey of the State-of-the-Art and Possible Extensions. – IEEE Transactions on Knowledge and Data Engineering, Vol. **17**, 2005, No 6, pp.734-749.

11. S h a m b o u r, Q. A User-Based Multi-Criteria Recommendation Approach for Personalized Recommendations. – International Journal of Computer Science and Information Security, Vol. **14**, 2016, No 12, pp. 657-663.

12. S h a m b o u r, Q., M. H o u r a n i, S. F r a i h a t. An Item-Based Multi-Criteria Collaborative Filtering Algorithm for Personalized Recommender Systems. – International Journal of Advanced Computer Science and Applications, Vol. **7**, 2016, No 8, pp. 274-279.

13. P u t r a, A. A., R. M a h e n d r a, I. B u d i, Q. M u n a j a t. Two-Steps Graph-Based Collaborative Filtering Using User and Item Similarities: Case Study of e-Commerce Recommender Systems. – In: Proc. of 2017 International Conference on Data and Software Engineering (ICoDSE'17), IEEE, 2017, pp. 1-6.

14. C h u, P., S. L e e. A Novel Recommender System for e-Commerce. – In: Proc. of 10th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI'17), IEEE, 2017, pp. 1-5.

15. X i a o, Y., C. I. E z e i f e. e-Commerce Product Recommendation Using Historical Purchases and Clickstream Data. – In: C. Ordonez, L. Bellatreche, Eds. Big Data Analytics and Knowledge Discovery. Cham, Springer International Publishing, 2018, pp. 70-82.

16. C h o i, Y. K., S. K. K i m. A Recommendation System for Repetitively Purchasing Items in e-Commerce Based on Collaborative Filtering and Association Rules. – Journal of Internet Technology, Vol. **19**, 2018, No 6, pp.1691-1698.

17. S a s s a n i, B., A. A l a h m a d i, H. S h a r i f z a d e h. A Cluster Based Collaborative Filtering Method for Improving the Performance of Recommender Systems in e-Commerce. – In: K. Arai, R. Bhatia, S. Kapoor, Eds. Advances in Intelligent Systems and Computing. Cham, Springer International Publishing, 2019, pp. 990-1001.

18. K h o d a b a n d e h l o u, S. Designing an e-Commerce Recommender System Based on Collaborative Filtering Using a Data Mining Approach. – International Journal of Business Information Systems, Vol. **31**, 2019, No 4, pp. 455-478.

19. J i a n g, L., Y. C h e n g, L. Y a n g, J. L i, H. Y a n, X. W a n g. A Trust-Based Collaborative Filtering Algorithm for e-Commerce Recommendation System. – Journal of Ambient Intelligence and Humanized Computing, Vol. **10**, 2019, No 8, pp. 3023-3034.

20. I f t i k h a r, A., M. A. G h a z a n f a r, M. A y u b, Z. M e h m o o d, M. M a q s o o d. An Improved Product Recommendation Method for Collaborative Filtering. – IEEE Access, Vol. **8**, 2020, pp. 123841-123857.

21. S i n g h, M. K., O. P. R i s h i. Event Driven Recommendation System for e-Commerce Using Knowledge Based Collaborative Filtering Technique. – Scalable Computing: Practice and Experience, Vol. **21**, 2020, No 3, pp. 369-378.

22. D y e r, J. S. MAUT – Multiattribute Utility Theory. – In: Multiple Criteria Decision Analysis: State of the Art Surveys. New York, Springer, 2005, pp. 265-292.

23. R e s n i c k, P., N. I a c o v o u, M. S u c h a k, P. B e r g s t r o m, J. R i e d l. GroupLens: An Open Architecture for Collaborative Filtering of Netnews. – In: Proc. of 1994 ACM Conference on Computer Supported Cooperative Work, ACM, 1994, pp. 175-186.

24. G a z d a r, A., L. H i d r i. A New Similarity Measure for Collaborative Filtering Based Recommender Systems. – Knowledge-Based Systems, Vol. **188**, 2020, pp.58-105.

25. B a r z e g a r N o z a r i, R., H. K o o h i, E. M a h m o d i. A Novel Trust Computation Method Based on User Ratings to Improve the Recommendation. – International Journal of Engineering, Vol. **33**, 2020, No 3, pp.377-386.

26. P a p a g e l i s, M., D. P l e x o u s a k i s, T. K u t s u r a s. Alleviating the Sparsity Problem of Collaborative Filtering Using Trust Inferences. – In: P. Herrmann, V. Issarny, S. Shiu, Ed. iTrust 2005: Trust Management. Berlin, Heidelberg, Springer, 2005, pp. 224-239.

27. F r a k e s, W. B., R. B a e z a-Y a t e s. Information Retrieval: Data Structures and Algorithms. Prentice Hall, 1992.

28. H e r l o c k e r, J., J. A. K o n s t a n, J. R i e d l. An Empirical Analysis of Design Choices in Neighborhood-Based Collaborative Filtering Algorithms. – Information Retrieval, Vol. **5**, 2002, No 4, pp. 287-310.

29. B u r k e, R. Hybrid Web Recommender Systems. – In: P. Brusilovsky, A. Kobsa W. Nejdl, Ed. The Adaptive Web: Methods and Strategies of Web Personalization. Berlin, Heidelberg, Springer, 2007, pp. 377-408.

30. O'D o n o v a n, J., B. S m y t h. Trust in Recommender Systems. – In: Proc. of 10th International Conference on Intelligent User Interfaces, ACM, 2005, pp. 167-174.

31. A l o d h a i b i, K. Decision-Guided Recommenders with Composite Alternatives. Information Technology, George Mason University, Virginia, 2011.

32. A g g a r w a l, C. C. Evaluating Recommender Systems. – In: Recommender Systems: The Textbook, Springer International Publishing, Cham, 2016, pp. 225-254.

33. A d o m a v i c i u s, G., Y. O. K w o n. New Recommendation Techniques for Multicriteria Rating Systems. – IEEE Intelligent Systems, Vol. **22**, 2007, No 3, pp. 48-55.