Context-Similarity Collaborative Filtering Recommendation

HIEP XUAN HUYNH1, NGHIA QUOC PHAN2, NGHI MONG PHAM3, VAN-HUY PHAM4, LE HOANG SON5, MOHAMED ABDEL-BASSET6, AND MAHMOUD ISMAIL6

1College of Information and Communication Technology, Can Tho University, Can Tho 900000, Vietnam
2Assessment Office, Tra Vinh University, Tra Vinh 940000, Vietnam
3Thapmuoi Vocational School, Cao Lãnh 870000, Vietnam
4NLP-KD Lab, Faculty of Information Technology, Ton Duc Thang University, Ho Chi Minh City 758307, Vietnam
5VNU Information Technology Institute, Vietnam National University, Hanoi 10000, Vietnam
6Faculty of Computers and Informatics, Zagazig University, Zagazig 44519, Egypt

Corresponding author : Hiep Xuan Huynh (hxhiep@ctu.edu.vn)

ABSTRACT This article proposes a new method to overcome the sparse data problem of the collaborative filtering models (CF models) by considering the homologous relationship between users or items calculated on contextual attributes when we build the CF models. In the traditional CF models, the results are built only based on data from the user’s ratings for items. The results of the proposed models are calculated on two factors: (1) the similar factors based on rating values; (2) the similar factors based on contextual attributes. The findings from the experimentation on two datasets DePaulMovie and InCarMusic, show that the proposed models have higher accuracy than the traditional CF models.

INDEX TERMS CIBC models, contextual attributes, context-similarity matrix, CUBCF models.

I. INTRODUCTION

According to the research reviews on recommender models [8], [19], [24], [37], the CF models have been applied successfully in many applications such as e-commerce Amazon [11], Netflix [4]. This is one of the effective solutions to the information explosion problem in online systems, while the number of users increases very fast. However, these models also face issues [8], [39] that need dealing with, such as cold-start, sparsity, and scalability. The context-aware is an extensive research method nowadays [16], [17], [44]. The literature indicates that the context of a user influences user decisions, and modeling contextual information may address the sparsity [45]–[50]. Therefore, incorporating contextual information in recommender systems is an effective approach to create more accurate and relevant recommendations. Especially, some applications lack information when considering the relationship between users and items [21], [22], [40]. For example, movie recommender systems are very important to incorporate contextual information such as time, location, and companion into the process of recommending the movies to users.

In this research, we propose a new method to increase the efficiency of CF models by considering the similarity values of users or items based on contextual attributes. This method can tackle the sparse data problem of CF models. In particular, the recommender models are built based on two similarity matrices: similarity matrix based on rating values and context-similarity matrix based on the contextual attributes. The models, after being built, are experimented on two datasets of DePaulMovie and InCarMusic, and then the results are compared with the traditional CF models. This article is organized into seven parts. Part 1 introduces an overview of recommender systems and research questions. Part 2 briefly presents Context-Aware recommender models. Part 3 presents steps for calculating contextual similarity values between users or items based on the Chi-square similarity kernel. Part 4 presents the context-similarity collaborative filtering models. Part 5 presents methods for evaluating collaborative filtering models. Part 6 displays the experimental results on two datasets: DePaulMovie and InCarMusic. The final part presents the results of the proposed models.

II. CONTEXT-AWARE RECOMMENDER MODELS

Using contextual data to provide recommendations in the direction of personalization of users is a strong trend of...
development in recent research on recommender models [14], [27], [32], [42]. Usually, in different situations, users can make different decisions [7]. Comprehensive research of contextual recommendations [10] highlights the real-life that a user’s decision depends on its context, rather than it’s invariant. Similar users may prefer different items in a different context. Therefore, to accurately predict user preferences, context information needs to be integrated to support recommender models. For example, a companion person is an important context for a movie recommendation, and time and place are essential contextual attributes for holiday recommendation [5] for users. This approach has become popular in many areas, such as the proposed models applied to recommend movies according to contextual data: who to watch movies and where to watch movies [2]. A studying contextual for movie recommendation [43] proposes a context-aware recommender model that uses joint matrix factorization, combining the similarity measures with similar values that calculated from plot keywords of the movies. The models propose contextual articles [20] to users based on contextual information and adjust how to choose articles based on users’ feedback. The contextual recommendation is also prevalent in music recommendation. A study performed on the models for collecting user emotions by describing their mood based on the websites they visited [35]. This emotional context is used to recommend music for users. Beginning from the traditional recommender models as content filtering [8] collaborative filtering [8], the models often use two essential components users and items to predict the recommendation results. Therefore, these models are considered as a function with two parameters: $F$ (user, item) and return the predicted recommendation results: $F: U \times I \rightarrow R$

In traditional models, predictive functions only consider two input parameters: users and items. Therefore, these models have some problems like cold-start, sparsity. To overcome the above problems, the contextual information of users or items is seen as an important factor to predict the recommendation results. It becomes a third input parameter for the predictive function. Meanwhile, the predictive function becomes a multidimensional function: $F: U \times I \times C \rightarrow R$

Where, $F$ is the prediction function of models; $U$ is a user set, $I$ is item set; $C$ is contexts of users or items; $R$ is user rating matrix; There are three main models used in the context-aware recommender systems. Based on how the contextual information, the models are named as follows: Pre-filtering, Post-filtering, and Contextual modeling. Pre-filtering models (Figure 1. (a)) [1], [9], [28], [37] These models use contextual properties to build an executable dataset. Then, this dataset is deployed on common 2D models such as collaborative filtering or content-based filtering.

Post-filtering models (Figure 1. (b)) [1], [9], [28], [37] These models only use contextual attributes at the last step of the recommendation process. The models take the entire 3D database to build predictable results. Then, in the final step, contextual attributes are applied to adjust the recommendation results. Contextual modeling

(Figure 1. (c)) [1], [9], [28], [37] the contextual information is used in the process of calculating the unknown ratings. The models are to deploy the context directly in the process of user rating prediction.

III. SIMILARITY BASED ON CONTEXT

A. CONTEXT INFORMATION OF USERS OR ITEMS

Consider the following examples, if users need support from the tourism recommender systems for booking tours, the contextual factors about the season and accompanied person can greatly affect to decision of users for choosing their journeys or space and time are two important contextual information if users need support to select movies for watching. The examples above show us that the contexts of users or items depend on each specific recommender systems. However, to describe the context-aware collaborative filtering models, the contexts of users or items are defined as follows:

Let $U = \{u_1, u_2, \ldots, u_n\}$ includes $n$ users; $I = \{i_1, i_2, \ldots, i_m\}$ includes $m$ items; $C = \{c_1, c_2, \ldots, c_k\}$ includes $k$ contextual attributes when users select items or when items selected by users. When the user $u_a$ select item $i_j$ or when item $i_j$ selected by user $u_a$ the value of contextual attributes is specified as follows:

$$C_{a,j} = \{c_{1a,j}, c_{2a,j}, \ldots, c_{ka,j}\}$$

B. CONTEXT SIMILARITY BASED ON CHI-SQUARE SIMILARITY KERNEL

Kernel functions have also become a popular useful tool in machine learning [30]. A kernel is a function that takes in two data objects such as images, DNA sequences, or vectors with k-dimensional and outputs a number, with the property that the function is symmetric or positive-semi definite. In this research, we use Chi-square similarity kernel [28], [30] to calculate the similarity context between two users or two items with the formula defined as follows:

Suppose user $u_a$ selects item $i_j$ and user $u_b$ selects item $i_h$ with contextual attributes defined by two vectors as follows: $C_{a,j} = \{c_{1a,j}, c_{2a,j}, \ldots, c_{ka,j}\}$ and $C_{b,h} = \{c_{1b,h}, c_{2b,h}, \ldots, c_{kb,h}\}$, then the context-similarity value of
user $u_a$ and user $u_b$ or context-similarity value of the item $i_j$ and item $i_h$ is determined by the following formula:

$$K(C_{a,j}, C_{b,h}) = \sum_{z=1}^{k} \frac{2cz_{a,j}cz_{b,h}}{cz_{a,j} + cz_{b,h}}$$

where, $K(C_{a,j}, C_{b,h})$ is the context-similarity value between the user $u_a$ and user $u_b$ or item $i_j$ and item $i_h$; $k$ is the number of contextual attributes; $cz_{a,j}$ is the value of the $z$th contextual attribute when the user $u_a$ selects item $i_j$; $cz_{b,h}$ is the value of the $z$th contextual attribute when the user $u_b$ selects item $i_h$.

C. DEFINITION OF CONTEXT-SIMILARITY MATRIX

The context-similarity matrix has a symmetric matrix structure with rows, columns as users or items, and cells as context-similarity values of two users or two items on the rows and columns respectively.

Let $U = \{u_1, u_2, \ldots, u_n\}$ includes $n$ users and $I = \{i_1, i_2, \ldots, i_m\}$ includes $m$ items:

Context-similarity matrix based on users defined as follows:

$$Matrix_{sim}(C_{u}) = \begin{pmatrix} su_{1,1} & su_{1,2} & \cdots & su_{1,m} \\ su_{2,1} & su_{2,2} & \cdots & su_{2,m} \\ \vdots & \vdots & \ddots & \vdots \\ su_{n,1} & su_{n,2} & \cdots & su_{n,n} \end{pmatrix}$$

where $su_{a,b}$ is a context-similarity value between the user $u_a$ and user $u_b$. This value is calculated by the Eq. (1).

Context-similarity matrix based on items defined as follows:

$$Matrix_{sim}(C_{i}) = \begin{pmatrix} si_{1,1} & si_{1,2} & \cdots & si_{1,m} \\ si_{2,1} & si_{2,2} & \cdots & si_{2,m} \\ \vdots & \vdots & \ddots & \vdots \\ si_{m,1} & si_{m,2} & \cdots & si_{m,m} \end{pmatrix}$$

where $si_{j,h}$ is a context-similarity value between item $i_j$ and item $i_h$. This value is calculated by Eq. (1).

IV. CONTEXT-SIMILARITY COLLABORATIVE FILTERING RECOMMENDER MODELS

In this section, we propose two models: Context-similarity, user-based collaborative filtering recommender model (CUBCF), and context-similarity item-based collaborative filtering recommender model (CIBCF). The CUBCF model is developed based on the user-based collaborative filtering recommender model and the contextual attributes of users. The CIBCF model is developed based on the Item-based collaborative filtering recommender model and the contextual attributes of items. Both models are based on the rating data and contextual information. However, the CUBCF model is based on the similarity between users, while the CIBCF model is based on the similarity between items.

A. CONTEXT-SIMILARITY USER-BASED COLLABORATIVE FILTERING RECOMMENDER MODEL

The CUBCF model is defined as follows:

Let $U = \{u_1, u_2, \ldots, u_n\}$ include $n$ users; $I = \{i_1, i_2, \ldots, i_m\}$ include $m$ items; $C = \{c_1, c_2, \ldots, c_k\}$ include $k$ contextual attributes; $R = \{r_{c,j}\}$ be the rating matrix, each row representing a user $u_c$ and each column representing an item $i_j$. The CUBCF model is defined as follows:

$$Output: \text{List of N items recommends to the user } u_a; \text{ Begin}$$

Step 1: Construct a similarity matrix based on users ($SU_{R}$)
- Use user splitting techniques;
- Construct a similarity matrix based on users;

Step 2: Construct a context-similarity matrix based on users ($SU_{C}$)
- Calculate the context-similarity value between two users;

$$K(C_{a,j}, C_{b,h}) = \sum_{z=1}^{k} \frac{2cz_{a,j}cz_{b,h}}{cz_{a,j} + cz_{b,h}}$$

- Construct a context-similarity matrix based on users;

Step 3: Construct a similarity integrated matrix ($SU_{I} = SU_{R} + SU_{C}$);

Step 4: Build CF models based on similarity integrated matrix;

Step 5: Identify a list of similar items for user $u_a$;

Step 6: Recommend user $u_a$ N items the highest similarity value;

C. CONTEXT-SIMILARITY ITEM-BASED COLLABORATIVE FILTERING RECOMMENDER MODEL

The CIBCF model is defined as follows:

The CIBCF model is defined as follows:

Let $U = \{u_1, u_2, \ldots, u_n\}$ include $n$ users; $I = \{i_1, i_2, \ldots, i_m\}$ include $m$ items; $C = \{c_1, c_2, \ldots, c_k\}$ include
Context-Similarity Collaborative Filtering Recommendation

$k$ contextual attributes; \( R = \{ r_{c,j} \} \) be the rating matrix, each row representing a user \( u_c (1 \leq c \leq n) \), each column representing an item \( i_j (1 \leq j \leq m) \), \( r_{c,j} \) be the rating value based on item contexts; \( N \) be the item list that has the high rating value, and \( u_a \) be the current user with contextual attributes:

\[
C_{a,j} = \{ c_{1,a,j}, c_{2,a,j}, \ldots, c_{k,a,j} \}.
\]

The CIBCF model is represented as follows:

Figure 3 shows the context-similarity item-based collaborative filtering recommender model. In the first step, item splitting techniques are used to build a similarity matrix based on contextual attributes of items. In the second step, the context-similarity matrix is built based on the contextual attributes of items. In the third step, the similarity integrated matrix between the similarity matrix and the context-similarity matrix is built. In the final step, the similarity integrated matrix is used to build the item-based FC models.

D. CONTEXT-SIMILARITY ITEM-BASED COLLABORATIVE FILTERING RECOMMENDER ALGORITHM

Context-Similarity Item-Based Collaborative Filtering Recommender Algorithm

| Input: Rating data of users; Context-data of items; User need recommendation (\( u_a \)); |
| Output: List of \( N \) items recommends to user \( u_a \); |

Begin

**Step 1:** Construct a similarity matrix based on items (\( SI_R \))
- Use item splitting techniques;
- Construct a similarity matrix based on items;

**Step 2:** Construct a context-similarity matrix based on items (\( SI_C \)).
- Calculate the context-similarity value between two items;
\[
K(C_{a,j}, C_{b,h}) = \sum_{z=1}^{k} \frac{2c_{a,j} z c_{b,h}}{c_{a,j} z + c_{b,h} z}
\]
- Construct a context-similarity matrix based on items;

**Step 3:** Construct a similarity integrated matrix (\( SI_I = SI_R + SI_C \));

**Step 4:** Build CF models based on similarity integrated matrix;

**Step 5:** Identify a list of similar items for user \( u_a \);

**Step 6:** Recommend user \( u_a \) \( N \) items the highest similarity value;

End

V. EVALUATE THE ACCURACY OF COLLABORATIVE FILTERING RECOMMENDER MODELS

Evaluation based on predicted rating value of models is a method based on the accuracy of the models by comparing the predicted rating value with the actual rating value [31]. This method often uses the five measures used to calculate errors in statistics: Mean Squared Error (MSE), Root Mean Square Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE) and Symmetric Mean Absolute Percentage Error (SMAPE). It is considered good when quantities of these errors are of low values [12], [15].
Mean Squared Error (MSE) [25], [31]: The average square error between user rating value and model’s predicted rating value.

\[ \text{MSE} = \frac{\sum_{(i,j) \in \kappa} (r_{ij} - \hat{r}_{ij})^2}{|\kappa|} \]  

(2)

Root Mean Square Error (RMSE) [25],[31]: The square root of the mean square error between user rating value and the model’s predicted rating value.

\[ \text{RMSE} = \sqrt{\frac{\sum_{(i,j) \in \kappa} (r_{ij} - \hat{r}_{ij})^2}{|\kappa|}} \]  

(3)

Mean Absolute Error (MAE) [25],[31]: The average absolute error between user rating value and model’s predicted rating value.

\[ \text{MAE} = \frac{1}{|\kappa|} \sum_{(i,j) \in \kappa} |r_{ij} - \hat{r}_{ij}| \]  

(4)

Mean Absolute Percentage Error (MAPE) [25], [31]: The average absolute percentage error between user rating value and model’s predicted rating value.

\[ \text{MAPE} = \frac{1}{|\kappa|} \sum_{(i,j) \in \kappa} \left| \frac{r_{ij} - \hat{r}_{ij}}{r_{ij}} \right| \]  

(5)

Symmetric Mean Absolute Percentage Error (SMAPE) [25], [31]: The symmetric average absolute percentage error between user rating value and model’s predicted rating value.

\[ \text{SMAPE} = \frac{1}{|\kappa|} \sum_{(i,j) \in \kappa} \frac{2 |r_{ij} - \hat{r}_{ij}|}{r_{ij} + \hat{r}_{ij}} \]  

(6)

where \( \kappa \) is the number rating values; \( r_{ij} \) is the actual rating value for item i by user i; \( \hat{r}_{ij} \) is the predicted value of the models.

B. EVALUATION BASED ON RECOMMENDATION RESULTS

Evaluation based on recommendation results of models is a method used to measure the effectiveness of models by comparing the proposed results of the models with the actual choice of users [31]. To compare these two values, we use three measures: Precision, Recall, and F-measure [25]. In many studies, Precision and Recall are inversely proportional, the F-measure indicator is used to evaluate the overall efficiency of the models by harmonizing the Precision indicator and the Recall indicator. If three indicators are high value, the models are rated well [12], [15], [25].

The formulas of three evaluation indicators are presented as follows [25], [31]:

\[ \text{Precision} = \frac{TP}{TP + FP} \]  

(7)

\[ \text{Recall} = \frac{TP}{TP + FN} \]  

(8)

\[ F - \text{measure} = \frac{2 * \text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}} \]  

(9)

where:
- \( \text{TP (True Positives)} \): The items recommended by the model that has been selected.
- \( \text{FP (False Positives)} \): The items recommended by the model that has not been selected.
- \( \text{TN (True Negatives)} \): The items not recommended by the model that has been selected.
- \( \text{FN (False Negatives)} \): The items not recommended by the model that has not been selected.

VI. EXPERIMENT

A. EXPERIMENTAL DATASETS

Experimental parts of the models are implemented in two different scenarios: Scenario 1 is performed on the DePaulMovie dataset [50], and Scenario 2 is performed on the InCarMusic dataset [3]. DePaulMovie is dataset collected from students’ survey results that rate films according to different contextual attributes: time, location, and companion. This dataset contains 5043 rating values of 97 students for 319 films based on three contextual attributes mentioned above. In particular, the rating attribute has a continuous value between 1 and 5 with 16.44% rated 1; 12.39% rated 2; 19.97% rated 3; 24.03% rated 4; 25.92% rated 5. InCarMusic is the dataset that collected suggestions of contextual music of automobile drivers on web services. This dataset includes 4012 rating values of 43 users for 139 songs based on eight contextual attributes. In particular, the rating attribute has a continuous value between 0 and 5 with 16.44% rated 1; 12.39% rated 2; 19.97% rated 3; 24.03% rated 4; 25.92% rated 5.

B. TOOL USED FOR EXPERIMENT

The proposed models have experimented on the ARQAT tool [33]. This tool has been developed by our team in the language R. It includes functions: process data; calculate context-similarity matrix based on Chi-square similarity kernel; build and evaluate the CF models [13].

C. SCENARIO 1

1) PROCESS DEPAULMOVIE DATASET ACCORDING TO CONTEXTUAL ATTRIBUTES

From students’ rating data for movies, we build a real numerical rating matrix of DePaulMovie dataset with 97 rows, 319 columns, and 5043 rating values. To build a context-similarity matrix of users, the contextual attributes are used to split users according to specific contexts when students choose to watch movies. This leads to the set of original users being split into 488 different contextual users. Therefore, the context-similarity matrix of users is a square matrix with a structure of \( 488 \times 488 \). To build a context-similarity matrix of items, the contextual attributes are used to split items according to specific contexts when the movies are selected by the students. This leads to the set of original items being split
into 1610 different contextual items. Therefore, the context-similarity matrix of items is a square matrix with the structure \(1610 \times 1610\).

2) EVALUATE BASED ON PREDICTED RATING VALUES
To compare the accuracy of the two proposed models with two traditional CF models (UBCF and IBCF), we developed evaluation data for models using the k-fold method (\(k = 5\)) on the rating matrix. After that, the experimental process is implemented on four models: CUBCF, CIBCF, UBCF, and IBCF. From the empirical results, we measured error indicators: Mean Squared Error (MSE), Root Mean Square Error (RMSE), Mean Absolute Error (MAE) of each model. Table 1 presents the value of the error indicators of the models. This result shows that context-similarity collaborative filtering recommender models have lower error indicators than traditional CF models on the DePaulMovie dataset. Specifically, the RMSE, MSE, MAE indicators of CIBCF models are all lower than those on the IBCF models; Similar to CIBCF models, the RMSE, MSE, MAE indicators of CUBCF models are all lower than those on the UBCF models.

3) EVALUATE BASED ON RECOMMENDATION RESULTS
This evaluation method runs experiments on four models: CUBCF, CIBCF, UBCF, IBCF and calculate the Precision, Recall, and F-measure indicators on the empirical data divided by the k-fold method (\(k = 5\)). Figure 4 presents the results of comparing the three above indicators of the models. This comparing result shows that the indicators for measuring the accuracy of both models (CUBCF and CIBCF) are higher than those in the traditional CF models (UBCF and IBCF). In particular, CIBCF models have the highest accuracy compared to the remaining 3 experimental models.

4) COMPARE ACCURACY OF MODELS BASED ON THE PREC/REC RATIO
To evaluate the effectiveness of the two proposed models on the DePaulMovie dataset, we chart the Prec/Rec ratio of these two models and compare them with the two traditional CF models UBCF and IBCF. In particular, the number of movies recommends to users of these models is gradually increased from 1 to 10. The results of comparing the accuracy of each pair of models are presented in Figure 5. The chart comparing the results of the models shows that the Prec/Rec ratios of CUBCF and CIBCF models are always higher than Prec/Rec ratios of UBCF and IBCF models when the number of recommendation movies per user varies from 1 to 10.

D. SCENARIO 2
1) PROCESS INCARMUSIC DATASET ACCORDING TO CONTEXTUAL ATTRIBUTES
From the rating data of automobile drivers for songs, we build a real numerical rating matrix of InCarMusic dataset with 43 rows, 139 columns, and 4012 rating values. To build a context-similarity matrix based on users, the contextual attributes are used to split users according to specific contexts when drivers choose to listen to the songs. This leads to the
set of original users being split into 782 different contextual users. Therefore, the context-similarity matrix of users is a square matrix with a structure $782 \times 782$. To build a context-similarity matrix based on items, the contextual attributes are used to split items according to specific contexts when the songs are selected by drivers. This leads to the set of original items being split into 2967 different contextual items. Therefore, the context-similarity matrix based on items is a square matrix with the structure $2967 \times 2967$.

2) EVALUATE BASED ON THE PREDICTED RATING VALUE
Similar to the experimental section on the DePaulMovie dataset, in this evaluation, we continue to calculate the error indicators: RMSE, MSE, MAE for four models: CUBCF, CIBCF, UBCF, IBCF on experimental data divided by k-fold method ($k = 5$). The value of the error indicators of the models is presented in Table 2. These results show that the error indicators of the proposed models (CUBCF and CIBCF) are lower than those of the two traditional CF models (UBCF and IBCF).

3) EVALUATE BASED ON RECOMMENDATION RESULTS
Similar to the experimental section on the DePaulMovie dataset, we continue to calculate indicators for measuring the accuracy of models (CUBCF, CIBCF, UBCF, IBCF) on the dataset divided by the k-fold method ($k = 5$). Figure 6 presents the comparison results of the three above indicators of the models. This comparison results show that the values of the indicators for measuring the accuracy of both models (CUBCF and CIBCF) are higher than those of two traditional CF models (UBCF and IBCF). In particular, CIBCF models have the highest accuracy compared to the remaining three experimental models.

4) COMPARE ACCURACY OF MODELS BASED ON THE PREC/REC RATIO
Similar to the experimental scenario 1, to compare the effectiveness of the two proposed models with two traditional CF models on the InCarMusic dataset, we continue to graph the Prec/Rec ratio to compare each pair of models CUBCF-UBCF, CIBCF-IBCF. Specifically, when the proposed models execute with the number of songs recommended for each user of the models increases gradually from 1 to 10. Figure 7 presents a chart comparing the accuracy of model pairs. This chart shows that CIBCF’s Prec/Rec ratio is always higher than IBCF’s Prec/Rec ratio. For the CUBCF model, the Prec/Rec ratio is equal to the UBCF model’s Prec/Rec ratio when the number of songs recommended for each user of the models increases from 1 to 2. On the contrary, when the experimental models with the number of songs increase from 3 to 10, the Prec/Rec ratio of CUBCF models is always higher than the Prec/Rec of UBCF models.

VII. CONCLUSION
In this article, we propose a method to improve the accuracy of the CF models by integrating contextual information of users or items when developing the recommender models. In these models, the recommendation results are calculated...
based on an integrated similarity matrix. The integrated similarity matrix is built by integrating the similarity matrix based on rating values and context-similarity matrix based on contextual attributes. Our experimental results reveal that the proposed models can offer significant improvements in recommendation tasks based on the two real context datasets. In particular, these models can alleviate the sparse data problem and offer more improvements.

REFERENCES

[1] A. K. Dey and G. D. Abowd, “Providing architectural support for building context-aware applications,” Georgia Inst. Technol., vol. 1, pp. 1–34, Nov. 2000.
[2] A. Karatzoglou, X. Amatriain, L. Baltrunas, and N. Oliver, “Multiverse recommendation: N-dimensional tensor factorization for context-aware collaborative filtering,” in Proc. 4th ACM Conf. Recommender Syst. (RecSys), 2010, pp. 79–86.
[3] L. Baltrunas, M. Kaminskas, B. Ludwig, O. Moling, F. Ricci, A. Aydin, K.-H. Lüke, and R. Schwager, “InCarMusic: Context-aware music recommendations in a car,” in E-Commerce and Web Technologies. Berlin, Germany, 2011, pp. 89–100.
[4] C. A. Gomez-Uribe and N. Hunt, “The netflix recommender system: Algorithms, business value, and innovation,” ACM Trans. Manage. Inf. Syst., vol. 5, no. 4, 2015, Art. no. 13.
[5] E. Ashley-Dejo, S. M. Ngwira, and T. Zuva, “A context-aware proactive recommender system for tourists,” in Proc. Int. Conf. Adv. Computing Commun. Eng. (ICACCE), Nov. 2016, pp. 271–275.
[6] F. Zhang, T. Gong, V. E. Lee, G. Zhao, C. Rong, and G. Qu, “Fast algorithms to evaluate collaborative filtering recommender systems,” Knowl.-Based Syst., vol. 96, pp. 96–103, Mar. 2016.
[7] F. Hidouid, B. Frihik, and B. Ouhbi, “Multi-criteria recommender systems based on multi-attribute decision making,” in Proc. Int. Conf. Inf. Integr. Web-Based Appl. Services (IWAS), 2013, pp. 203–226.
[8] G. Adomavicius and A. Tuzhilin, “Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions,” IEEE Trans. Knowl. Data Eng., vol. 17, no. 6, pp. 734–749, Jun. 2005.
[9] G. Adomavicius, B. Mobasher, F. Ricci, and A. Tuzhilin, Context-Aware Recommender Systems. Menlo Park, CA, USA: Association for the Advancement of Artificial Intelligence, 2011, pp. 67–80.
[10] G. Adomavicius, R. Sankaranarayanan, S. Sen, and A. Tuzhilin, “Incorporating contextual information in recommender systems using a multi-dimensional approach,” ACM Trans. Inf. Syst., vol. 23, no. 1, pp. 103–145, Jan. 2005.
[11] G. Linden, B. Smith, and J. York, “Amazon.com recommendations: Item-to-item collaborative filtering,” IEEE Internet Comput., vol. 7, no. 1, pp. 76–80, Jan. 2003.
[12] A. Gunawardana and G. Shani, “A survey of accuracy evaluation metrics of recommendation tasks,” J. Mach. Learn. Res., vol. 10, pp. 2935–2962, Dec. 2009.
[13] M. Hahsler. (2011). Recommenderlab: A Framework for Developing and Testing Recommendation Algorithms. Intelligent Data Analysis Lab at SMU. [Online]. Available: http://lyle.smu.edu/IDA/recommenderlab/
[14] M. Chen, Q. Zhao, P. Du, C. Zaniolo, and M. Gerla, “Demand-driven cache allocation based on context-aware collaborative filtering,” in Proc. 18th ACM Int. Symp. Mobile Ad Hoc Netw. Comput. (MobiHoc), 2018, pp. 302–303, doi:10.1145/3209582.3225198.
[15] R. Cai, C. Zhang, C. Wang, L. Zhang, and W.-Y. Ma, “MusicSense: A multi-agent system for music recommendation based on multi-attribute collaborative filtering,” in Proc. 19th Int. Conf. Multimedia (MULTIMEDIA), 2011, pp. 523–526, Dec. 2011.
[16] R. Burke, The Adaptive Web: Hybrid Web Recommender Systems. P. Brusilovsky, A. Kobsa, and W. Nejdl, Eds. Berlin, Germany: Springer-Verlag, 2007, pp. 377–408.
[17] R. Cai, C. Zhang, C. Wang, L. Zhang, and W.-Y. Ma, “MusicSense: Contextual music recommendation using emotional allocation modeling,” in Proc. 15th Int. Conf. Multimedia (MULTIMEDIA), 2007, pp. 553–556.
[18] R. Ma and X. Zeng, “Collaborative filtering recommendation algorithm based on knowledge graph,” Int. J. Math. Problems Eng., vol. 2018, pp. 1–11, Jul. 2018, doi:10.1155/2018/9617410.
[19] S. M. M. Seyednezhad, K. N. Cozart, J. A. Bowllan, and A. O. Smith, “A review on recommendation systems: Context-aware to social-based,” 2018, arXiv:1811.11866. [Online]. Available: http://arxiv.org/abs/1811.11866
[20] T. T. S. Nguyen, H. Yan Lu, and J. Lu, “Web-page recommendation based on social network analysis and domain knowledge,” IEEE Trans. Knowl. Data Eng., vol. 26, no. 10, pp. 2574–2587, Oct. 2014.
[21] V. S. Dixit and P. Jain, “An improved similarity measure to alleviate sparsity problem in context-aware recommender systems,” in Towards Extensible Adaptible Methods Computing. Singapore: Springer, 2018, pp. 281–295, doi:10.1007/978-981-13-2348-5_21.
[22] W. Woendrll, M. Brocco, and R. Eigner, “Context-aware recommender systems in mobile scenarios,” Int. J. Inf. Technol. Web Eng., vol. 4, no. 1, pp. 67–85, Jan. 2009.
[41] X. Su and T. M. Khoshgoftaar, “A survey of collaborative filtering techniques,” Adv. Artif. Intell., vol. 2009, pp. 1–19, Oct. 2009.

[42] Y. Zheng and A. A. Jose, “Context-aware recommendations via sequential predictions,” in Proc. 34th ACM/SIGAPP Symp. Appl. Comput. (SAC), Limassol, Cyprus, Apr. 2019, pp. 2525–2528, doi: 10.1145/3297280.3297639.

[43] Y. Shi, M. Larson, and A. Hanjalic, “Mining contextual movie similarity with matrix factorization for context-aware recommendation,” ACM Trans. Intell. Syst. Technol., vol. 14, no. 1, 2013, Art. no. 16.

[44] Z. D. Champiri, S. R. Shahamiri, and S. S. B. Salim, “A systematic review of scholar context-aware recommender systems,” Expert Syst. Appl., vol. 42, no. 3, pp. 1743–1758, Feb. 2015.

[45] Y. Zheng, B. Mobasher, and R. Burke, “CARSKit: A java-based context-aware recommendation engine,” in Proc. IEEE Int. Conf. Data Mining Workshop (ICDMW), Atlantic City, NJ, USA, Nov. 2015, pp. 1668–1671.

[46] M. Ali, L. H. Son, N. D. Thanh, and N. V. Minh, “A neutrosophic recommender system for medical diagnosis based on algebraic neutrosophic measures,” Appl. Soft Comput., vol. 71, pp. 1054–1071, Oct. 2018.

[47] N. D. Thanh, L. H. Son, and M. Ali, “Neutrosophic recommender system for medical diagnosis based on algebraic similarity measure and clustering,” in Proc. IEEE Int. Conf. Fuzzy Syst. (FUZZ-IEEE), Jul. 2017, pp. 1–6.

[48] L. H. Son, “Dealing with the new user cold-start problem in recommender systems: A comparative review,” Inf. Syst., vol. 58, pp. 87–104, Jun. 2016.

[49] L. H. Son, “HU-FCP++: A novel hybrid method for the new user cold-start problem in recommender systems,” Eng. Appl. Artif. Intell., vol. 41, pp. 207–222, May 2015.

[50] L. H. Son and N. T. Thong, “Intuitionistic fuzzy recommender systems: An effective tool for medical diagnosis,” Knowl.-Based Syst., vol. 74, pp. 133–150, Jan. 2015.

HIEP XUAN HUYNH received the B.E. degree from Can Tho University (CTU), the master’s degree from the Institut de la Francophonie pour l’Informatique (IFI), the Ph.D. degree from the Politechnics School of Nantes University (Polytech’Nantes), and the Habilitation à Diriger des Recherches (HDR) degree from the Université de Bretagne Occidentale (UBO), all in computer science. He is currently an Associate Professor of computer science with the College of Information and Communication Technology, Can Tho University, Vietnam. His research interests include modeling of decisions with interestingness measure and fuzzy integral, recommender system with statistical implicational analysis, a cyber-physical system with cellular automata and wireless sensor networks, metrics for computer vision and population dynamics, agriculture, aquaculture, and environment issues.

NGHIA QUOC PHAN received the master’s degree in information system from Can Tho University (CTU), Vietnam, in 2010, and the Ph.D. degree in computer science from the University of Da Nang (UDN), Vietnam, in 2019. He is currently a Lecturer and a Researcher with the School of Engineering and Technology, Tra Vinh University, Tra Vinh, Vietnam. His current research interests include artificial intelligence, recommender system with statistical implicational analysis, and similarity measures.

NGHI MONG PHAM received the B.E. degree in information technology from Mekong University (MKU), Vietnam, in 2010, and the master’s degree in information technology from Tra Vinh University (TVU), Vietnam, in 2019. She is currently a Teacher with the Thampuoi Vocational School, Vietnam. Her current research interests include context-aware recommender systems and objective interestingness measures.

VAN-HUY PHAM received the M.Sc. degree in computer science from the Ho Chi Minh City University of Science, Ho Chi Minh City, Vietnam, in 2007, and the Ph.D. degree in computer science from University of Ulsan, South Korea, in 2015. Since 2015, he has been a Lecturer and a Researcher with the Faculty of Information Technology, Ton Duc Thang University, Ho Chi Minh City. His current research interests include artificial intelligence, image processing, and computer vision.

LE HOANG SON received the Ph.D. degree in mathematics and informatics from the VNU University of Science, Vietnam National University (VNU) in conjunction with the Politecnico di Milano, Italy, in 2013. He has been promoted to an Associate Professor of information technology in Vietnam, since 2017. He worked as a Senior Researcher and the Vice Director at the Center for High Performance Computing, VNU University of Science, Vietnam National University, from 2007 to 2018. Since August 2018, he has been a Senior Researcher with the Department of Multimedia and Virtual Reality, VNU Information Technology Institute. His major fields include artificial intelligence, data mining, soft computing, fuzzy computing, fuzzy recommender systems, and geographic information systems. He is a member of the Vietnam Journalists Association, the International Association of Computer Science and Information Technology, the Vietnam Society for Applications of Mathematics, and the Key Laboratory of Geotechnical Engineering and Artificial Intelligence, University of Transport Technology. He is also an Associate Editor of the Journal of Intelligent & Fuzzy Systems (SCIE), IEEE Access (SCIE), Data Technologies and Applications (SCIE), the International Journal of Data Warehousing and Mining (SCIE), Neutrosophic Sets and Systems (ESCI), the Research and Development on Information and Communication Technology, Vietnam, the VNU Journal of Science: Computer Science and Communication Engineering, and Frontiers in Artificial Intelligence. He serves on the Editorial Board of Applied Soft Computing (SCIE), PLOS One (SCIE), the International Journal of Web and Grid Services (SCIE), the International Journal of Ambient Computing and Intelligence (ESCI), and the Journal of Computer Science and Cybernetics, Vietnam. He is also the Guest Editor of several Special Issues at the International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems (SCIE), and the Journal of Ambient Intelligence and Humanized Computing.
MOHAMED ABDEL-BASSET received the B.Sc. and M.Sc. degrees from the Faculty of Computers and Informatics, Zagazig University, Egypt, and the Ph.D. degree from the Faculty of Computers and Informatics, Menoufia University, Egypt. He is currently an Associate Professor with the Faculty of Computers and Informatics, Zagazig University. His current research interests are optimization, operations research, data mining, computational intelligence, applied statistics, decision support systems, robust optimization, engineering optimization, multiobjective optimization, swarm intelligence, evolutionary algorithms, and artificial neural networks. He is also working on the application of multiobjective and robust meta-heuristic optimization techniques. He is also an Editor/Reviewer in different international journals and conferences. He holds the Program Chair in many conferences in the fields of decision making analysis, big data, optimization, complexity, and the Internet of Things, and editorial collaboration in some journals of high impact.

MAHMOUD ISMAIL received the B.Sc., M.Sc., and Ph.D. degrees from the Faculty of Computers and Informatics, Zagazig University, Egypt. He is currently working on the application of metaheuristic optimization techniques. His current research interests are optimization, operations research, computational intelligence, decision support, swarm intelligence, and evolutionary algorithms. He is also a reviewer in different international journals and conferences. * * *