A multi-criteria trust-enhanced collaborative filtering algorithm for personalized tourism recommendations

Qusai Y. Shambour[1], Mahran M. Al-Zyoud[2], Adeeb M. Alsaaidah[2], Mosleh M. Abualhaj[2], Ahmad A. Abu-Shareha[3]

[1]Department of Software Engineering, Faculty of Information Technology, Al-Ahliyya Amman University, Amman, Jordan
[2]Department of Networks and Cybersecurity, Faculty of Information Technology, Al-Ahliyya Amman University, Amman, Jordan
[3]Department of Data Science and Artificial Intelligence, Faculty of Information Technology, Al-Ahliyya Amman University, Amman, Jordan

Article Info

Article history:

Received Jun 29, 2024
Revised Aug 23, 2024
Accepted Aug 31, 2024

Keywords:

Collaborative filtering
Data sparsity
Multi-criteria
Prediction accuracy
Recommender systems
Tourism
Trust relationships

ABSTRACT

The exponential growth of online information has LED to significant challenges in navigating data overload, particularly in the tourism industry. Travelers are overwhelmed with choices regarding destinations, accommodations, dining, and attractions, making it difficult to select options that best meet their needs. Recommender systems have emerged as a promising solution to this problem, aiding users in decision-making by providing personalized suggestions based on their preferences. Traditional collaborative filtering (CF) methods, however, face limitations, such as data sparsity and reliance on single rating scores, which do not fully capture the complexity of user preferences. This study proposes a hybrid multi-criteria trust-enhanced CF (HMCTeCF) algorithm to improve the accuracy and robustness of tourism recommendations. HMCTeCF improves the quality of recommendations by integrating multi-criteria user preferences with trust relationships among users and between items. Experimental results using real-world datasets, including Restaurants-TripAdvisor and Hotels-TripAdvisor, demonstrate that HMCTeCF outperforms benchmark CF-based recommendation methods. It achieves higher prediction accuracy and coverage rate, effectively addressing the data sparsity problem. This innovative algorithm facilitates a more personalized and enriching travel experience, particularly in scenarios with limited user data. The findings highlight the importance of considering multiple criteria and trust relationships in developing robust recommendation systems for the tourism industry.

Corresponding Author:

Qusai Y. Shambour
Department of Software Engineering, Faculty of Information Technology, Al-Ahliyya Amman University
Amman, Jordan
Email: q.shambour@ammanu.edu.jo

1. INTRODUCTION

Digital technology and the digital age have changed information access across fields with new complexities in both opportunities for consumers of this information as well as challenges. This is particularly true in the tourism industry where travelers are surrounded by a plethora of options regarding destinations, accommodations, dining choices at attractions, or means of transportation. Attempting to navigate this information also requires users to devote a great deal of time and effort to sifting through the options until they find something that meets their unique needs or preferences. These systems can build models of individual user preferences utilizing both, explicit and implicit feedback provided by the users and therefore recommend items or services adapted to such preferences learned from past choices. In the context of tourism, recommender systems offer recommendations customized to each individual based on their preferences for hotels, restaurants, attractions, museums, and other travel-related services [1], [2].

Collaborative filtering (CF) approaches, while widely adopted in recommender systems, struggle to yield satisfying tourism recommendations due to their limitations. Firstly, their use of single-criterion ratings makes it difficult to reflect adequately user preferences. A hotel stay, for instance, is comprised of various aspects such as location, amenities, quality of service, and overall ambiance each with greater or lesser importance to different travelers [3]. Secondly, CF-based recommender systems often struggle with data sparsity, particularly when recommending to users where there is very little or no user-item interactional history available. This sparsity challenge leads to lower accuracy of predictions and therefore, less personalization in the recommendations [4].

To overcome these limitations, a more robust approach that takes into account both the multi-dimensional aspect of individual preferences and complementary sources of information might be needed. Recent research has explored two promising streams: multi-criteria (MC) recommender systems and the integration of social trust networks. MC recommender systems extend beyond single ratings by incorporating user feedback on specific attributes of items. By taking these fine-grained preferences into account, such systems can make recommendations that are closer to the actual user priorities [5]-[7]. Alongside, social trust networks have been explored to address data sparseness. In addition to discovering latent user preferences through social interactions, these networks expose trust relationships between users and items that sunlight potentially influences configurations. However, the explicit trust relationships within these networks can also be sparse, requiring the reflection of implicit connections derived from user-item interactions [8].

Recently, recommender systems have become indispensable in the tourism domain by alleviating information overload and providing personalized recommendations to travelers. Researchers have widely explored various approaches to enhance the performance of these systems, particularly for hotel and restaurant recommendations. In the hotel domain, recommender systems are crucial for travelers who need assistance in picking accommodations that suit them best. Forouzandeh *et al.* [9] proposed a novel approach combining the artificial bee colony algorithm and the fuzzy TOPSIS model, leveraging data from TripAdvisor for tuning hotel recommendations depending on user specifications. Similarly, Cui *et al.* [10] developed a hotel recommendation algorithm that utilizes online reviews and probabilistic linguistic term sets (PLTS). Users often express their sentiments with a great deal of ambiguity and uncertainty in the review text, which makes it difficult for an automatic system to capture all such expressions; therefore, their proposed approach translates user sentiment expressed through natural language into PLTSs that represent more accurate as well as reflective recommendations. Xia *et al.* [11] introduced a deep neural network model that combines pictures, text, and scoring data from user reviews to provide holistic recommendations for hotels. To improve recommendation accuracy, Jose *et al.* [12] proposed a hybrid approach based on dilated multichannel convolutional neural networks (CNN) and bidirectional gated recurrent units (BiGRU) with an attention mechanism that can capture long-term semantic characteristics.

Regarding restaurant recommendations, recommender systems play a crucial role in streamlining the dining experience for users. Asani *et al.* [13] introduced a comment-based context-aware recommendation system that extracted food preferences from user comments and recommended restaurants with high precision in their evaluations. Deep learning techniques have also been employed in restaurant recommendation systems. Saelim and Kijsirikul [14] developed a deep neural network model for restaurant recommendations in Thailand. This model integrates different features such as user preferences, location, and dining history to provide more personalized dining recommendations, thus enhancing the overall user experience. On the other hand, Yang *et al.* [15] developed a personalized restaurant recommendation model using deep learning and big data. Their model utilizes broad datasets to recognize customer behavior and preferences, making it possible to offer personalized recommendations that affect customer decision-making. Perumal *et al.* [16] tackled the cold-start problem in restaurant recommender systems using an ontology-based approach. Ontologies can be used to model the hierarchy of relations (e.g., between different types of cuisines, restaurant properties, and user preferences) enabling the system to make informed recommendations even when limited data is available.

Drawing inspiration from the above-mentioned advancements and the success achieved in the use of hybrid methodologies in recommender systems, this study proposes a hybrid multi-criteria trust-enhanced collaborative filtering (HMCTeCF) algorithm for personalized tourism recommendations. The goal of HMCTeCF is to surpass the limitations of traditional recommender systems by integrating multi-criteria user preferences with trust relationships between users and items. This synthesis is proposed to boost both the accuracy and coverage of recommendations, eventually satisfying travel selections.

The proposed HMCTeCF algorithm builds on previous research in tourism recommender systems by integrating multi-criteria approaches with trust networks. Earlier studies, such as [9], [15], focused on specific techniques such as artificial intelligence or deep learning. In contrast, this study combines multi-criteria preferences and trust relationships to address the limitations of traditional CF. By merging these advanced approaches, the HMCTeCF algorithm aims to provide more accurate and personalized recommendations in tourism. It effectively tackles challenges like data sparsity that have hindered earlier systems.

The proposed HMCTeCF algorithm includes three interrelated modules: a user-based multi-criteria trust-enhanced CF, an item-based multi-criteria trust-enhanced CF, and a hybrid recommendation mechanism. Through direct, propagated, and overall trust scores together with multi-criteria ratings, the HMCTeCF is designed to provide personalized yet diverse recommendations dedicated to cumulative user intelligence. Through extensive experiments on real-world datasets from TripAdvisor, including both restaurant and hotel ratings, we evaluate the effectiveness of the HMCTeCF algorithm for restaurant and hotel rating predictions. Additionally, we test the algorithm's performance under different data sparsity scenarios. Our results demonstrate that the HMCTeCF algorithm consistently outperforms benchmark methods in terms of prediction accuracy and recommendation coverage. Particularly, it achieves greater prediction accuracy, demonstrating a closer alignment between recommendations and actual user preferences. Furthermore, the algorithm shows a notable coverage rate, showcasing its ability to recommend a diverse set of tourism-related facilities.

2. ARCHITECTURAL OVERVIEW OF THE HMCTeCF ALGORITHM

This study introduces an effective HMCTeCF algorithm designed specifically for tourism recommendation systems. The algorithm comprises three primary modules: the user-based multi-criteria trust-enhanced collaborative filtering module, the item-based multi-criteria trust-enhanced collaborative filtering module, and the hybrid recommendation module. These modules work in synergy to provide more accurate and personalized tourism-related recommendations.

2.1. The user-based multi-criteria trust-enhanced CF module

This module harnesses the power of trust relationships among users and multi-criteria ratings to produce user-based trust-enhanced predictions by calculating direct, propagated, and overall trust scores. The process unfolds in several steps:

Step 1: assign trust values for users based on their interactions

Trustworthiness between users is assessed based on their reliability in providing accurate recommendations, which correlates with user similarity [17]. To initiate the prediction process, we first employ the following formula to calculate the predicted ratings for each user pair, as (1).

$$P_{x,i} = \bar{r}_x + (U^y(i) - \bar{r}_y) \tag{1}$$

In this formula, $\bar{r}_x$ and $\bar{r}_y$ denote the average ratings provided by users *x* and *y*, respectively. $U^y(i)$ represents user *y*'s overall utility for item *i*. This utility score is defined as (2).

$$U^y(i) = \sum_{c=1}^{k} w_c^y(i) \times r_c^y(i), where \sum_{c=1}^{k} w_c^y(i) = 1 \tag{2}$$

Within this formula, $w_c^y(i)$ represents the weight that user *y* assigns to criterion *c* for item *i*. This weight reflects the importance user *y* places on criterion *c* when evaluating item *i*. $r_c^y(i)$ represents the specific rating assigned by user *y* for criterion *c* of item *i*.

Subsequently, the algorithm employs two complementary metrics: the Euclidean distance method [18] and the Relevant Jaccard method [19], as shown by (3) and (4), respectively. The former measures the distance between predicted and actual ratings, while the latter accounts for the confidence in the similarity assessment by considering the proportion of co-rated items between users.

$$Sim_{x,y}^{Euc} = \frac{1}{1+\sqrt{\sum_{i \in I_{x\cap y}} \left|P_{x,i} - U^x(i)\right|^2}} \tag{3}$$

Where $P_{x,i}$ and $U^x(i)$ represent the predicted rating and total utility of item *i* as perceived by user *x*, respectively. $I_{x\cap y}$ denotes the total number of items that both users *x* and *y* have rated.

$$URJac_{x,y} = \frac{1}{1+\left(\frac{1}{|I_{x\cap y}|}\right)+\left(\frac{|I_x|-|I_{x\cap y}|}{1+|I_x|-|I_{x\cap y}|}\right)+\left(\frac{1}{1+|I_x|-|I_{x\cap y}|}\right)} \tag{4}$$

In this formula,$|I_x||I_x|$ and $|I_y|$ symbolize the total number of items rated by users *x* and *y*, respectively. This formula essentially calculates the proportion of items that users *x* and *y* have rated in common relative to the total number of items they have each rated.

Building upon the concepts introduced and following the above calculations, the direct trust score assigned to the relationship between users *x* and *y* is calculated by combining the similarity metrics as shown in the (5).

$$Trust_{x,y}^{Direct} = Sim_{x,y}^{Euc} \times URJac_{x,y} \quad (5)$$

Step 2: infer trust scores for users with no direct connection

While direct trust scores between users provide a valuable starting point, the resulting trust network often suffers from sparsity. This is due to the common phenomenon in recommender systems where users rate only a limited subset of available items. To address this limitation and fully leverage the potential of trust networks, we draw inspiration from social networks and incorporate trust propagation into our approach. This concept enables trust relationships to be spread through intermediary users, thereby establishing novel indirect connections within the trust network. By doing so, we can form more nuanced trust relationships that spread beyond direct connections between users.

To measure the level of trust propagated between users, we propose an aggregation function that integrates confidence weights. Specifically, this function measures the trust from user *x* to user *z*, mediated by a common neighbor *y*, as (6):

$$Trust_{x,z}^{\text{Prop}} = \frac{\sum_{y \in intermediary(xandz)}(Trust_{x,y}^{Direct} \times URJac_{x,y}) + (Trust_{y,z}^{Direct} \times URJac_{y,z})}{\sum_{y \in \text{intermediary}(xandz)} URJac_{x,y} + URJac_{y,z}} \quad (6)$$

Step 3: determine a user's overall trustworthiness

The overall trust score serves as a crucial factor in improving the recommender's capacity to produce accurate and diverse predicted ratings for unknown items, mainly for users who have limited connections (nearest neighbors). This score aims to enhance the accuracy and coverage of recommendation predictions, and can be determined by combining two factors [20]:

- User rating behavior: this factor measures how much a user's ratings deviate from the average ratings for each item. A user who consistently rates items close to the average (low deviation) is considered more trustworthy than a user who consistently deviates from the average (high deviation).
- User connectivity: this factor reflects how well-connected a user is within the trust network. A user with high connectivity (more connections) is considered more trustworthy than a user with low connectivity.

This makes intuitive sense as users who consistently rate items close to the average and have a wider network of connections are likely to be more reliable sources of recommendations. The overall trust score for a user *x* is measured as (7).

$$UOT_x = exp\left(-\frac{\sum_{i \in I_x}|r_{x,i} - \bar{r}_i|}{|I_x|}\right) \times \sqrt{\frac{|U_x|}{|U|}} \quad (7)$$

where $r_{x,i}r_{x,i}$ exemplifies the average rating given by user *x* across all evaluation criteria for item *i*, $\bar{r}_i$ signifies the mean rating across all users for item *i*, considering all evaluation criteria, and $|U_x|$ is the number of users who are connected to user *x* within the users' trust network.

Step 4: compute the user-based trust-enhanced predicted ratings

The final step in this module comprises generating predicted ratings for an active user *x* on a target item *i* by incorporating the direct, propagated, and overall trust values of his nearest neighbors *NN(x)* into the weighted average [21], as (8).

$$P_{x,i}^{User} = \begin{Bmatrix} \overline{r_x} + \frac{\sum_{y \in NN(x)} Trust_{x,y} \times (r_{y,i} - \overline{r_y})}{\sum_{y \in NN(x)} Trust_{x,y}}; \; if\ Trust_{x,y} \neq 0 \\ \overline{r_x} + \frac{\sum_{y \in NN(x)} UOT_y \times (r_{y,i} - \overline{r_y})}{\sum_{y \in NN(x)} UOT_y}; \; if\ Trust_{x,y} \neq 0 \end{Bmatrix} \quad (8)$$

2.2. The Item-based multi-criteria trust-enhanced CF module

To generate recommendations that consider the trust relationships between items, this module leverages both the items' trust network and each item's overall trust score. This process can be broken down into three main steps:

Step 1: assign trust values for items based on their interactions

Continuing with the rationale established in the previous module, this method utilizes two complementary metrics for trust score calculation between items: the Euclidean distance metric and the Relevant Jaccard metric. These metrics are presented in Equations (9) and (10), respectively.

The Euclidean distance method focuses on the magnitude of the difference between predicted ratings and actual user ratings. In simpler terms, it calculates how close the predicted values are to the real user ratings.

$$Sim_{i,j}^{Euc} = \frac{1}{1+\sqrt{\sum_{x\in U_{i\cap j}} \left|P_{x,i} - U^{x}(i)\right|^{2}}} \tag{8}$$

where $U_{i\cap j}$ signifies the number of users that have provided ratings for both items *i* and *j*.

The Relevant Jaccard method, on the other hand, takes into account the confidence level in the similarity assessment between items. It achieves this by considering the proportion of users that have rated both items. A higher proportion of co-rated users suggests a more reliable assessment of item similarity.

$$IRJac_{i,j} = \frac{1}{1+\left(\frac{1}{|U_{i\cap j}|}\right)+\left(\frac{|U_i|-|U_{i\cap j}|}{1+|U_i|-|U_{i\cap j}|}\right)+\left(\frac{1}{1+|U_j|-|U_{i\cap j}|}\right)} \tag{9}$$

Where | U_i | represents the entire amount of users who have rated item *i.*

Following this, the trust score between items *i* and *j* is calculated by combining the similarity metrics as shown in the (11).

$$Trust_{i,j} = Sim_{i,j}^{Euc} \times IRJac_{i,j} \tag{10}$$

Step 2: determine an item's overall trustworthiness

To address the challenge of sparsity, particularly for items with limited ratings, we propose an overall trust score for each item. This score aims to enhance the accuracy and coverage of recommendation predictions. The overall similarity score is calculated by considering two key factors:

– Item rating behavior: this factor measures the dispersion of user ratings for a specific item compared to the average rating across all items. A smaller deviation indicates that the item's ratings tend to be closer to the general user preference.
– Item connectivity: this factor reflects the number of connections an item has within the item-item trust network. A higher number of connections suggests the item is related to a broader range of other items, potentially indicating its overall level of user engagement.

By incorporating these factors, an item with both low deviation and high connectivity receives a higher overall trust score, which can improve the recommendation system's ability to suggest relevant items, especially for those with limited ratings. The overall trust score for an item *i* is measured as (12).

$$IOT_i = exp\left(-\frac{\sum_{x\in U_i}|r_{x,i}-\bar{r}_x|}{|U_i|}\right) \times \sqrt{\frac{|I_i|}{|I|}} \tag{11}$$

In this formula, $\bar{r}_x$ represents the mean rating across all items of user *x*, considering all evaluation criteria, and $|U_i|$ signifies the total number of users who have rated item *i.* Additionally, $|I_i|$ is the number of items connected to item *i* within the items' trust network. Finally, $|I|$ represents the total number of items in the entire dataset.

Step 3: compute the item-based trust-enhanced predicted ratings

This step involves generating predicted ratings for an active user *x* on a target item *i*. This prediction leverages both the direct trust scores and the overall trust scores assigned to a target item *i*'s nearest neighbors *NN(i)* into the weighted average, as (13).

$$P_{x,i}^{Item} = \begin{Bmatrix} \overline{r_i} + \frac{\sum_{j\in NN(i)} Trust_{i,j} \times (r_{x,i}-\overline{r_j})}{\sum_{j\in NN(i)} Trust_{i,j}};\ if\ Trust_{i,j} \neq 0 \\ \overline{r_i} + \frac{\sum_{j\in NN(i)} UOT_j \times (r_{x,i}-\overline{r_j})}{\sum_{j\in NN(i)} UOT_j};\ if\ Trust_{i,j} \neq 0 \end{Bmatrix} \tag{13}$$

2.3. The hybrid recommendation module

Following the success of combining multiple recommendation techniques [22], this module utilizes a dynamic switching approach. This approach selects the most suitable prediction method on the fly, based on its ability to predict ratings for unseen items. The ultimate goal is to improve both the accuracy and coverage of the recommendations. The key factor in this selection process is the ability to predict ratings for items the user hasn't seen before (unseen items). The module employs the root mean square (RMS) metric when both candidate methods can predict unseen items. This metric helps quantify the level of agreement between the two predicted ratings, leading to more robust and reliable recommendations.

$$P_{x,i}^{Final} = \begin{cases} 0; \text{if} P_{x,i}^{User} = 0 \text{and} P_{x,i}^{Item} = 0 \\ P_{x,i}^{User}; \text{if} P_{x,i}^{User} \neq 0 \text{and} P_{x,i}^{Item} = 0 \\ P_{x,i}^{Item}; \text{if} P_{x,i}^{User} = 0 \text{and} P_{x,i}^{Item} \neq 0 \\ \sqrt{\frac{(P_{x,i}^{User})^2 + (P_{x,i}^{Item})^2}{2}}; \text{ if} P_{x,i}^{User} \neq 0 \text{and} P_{x,i}^{Item} \neq 0 \end{cases} \tag{12}$$

3. EXPERIMENTS

3.1. Datasets and evaluation method

In evaluating our proposed algorithm, we utilized two MC rating datasets obtained from the TripAdvisor tourism platform: the restaurants dataset and the Hotels rating dataset [23]. The restaurants MC dataset consists of ratings provided by users for various restaurants. Users rated these restaurants on a scale of 1 to 5 across three criteria: service, food, and value. This dataset contains a total of 14,633 ratings contributed by 1,254 users for 205 restaurants. The Hotels MC dataset, on the other hand, consists of ratings provided by users for numerous hotels. Users were able to rate hotels on a scale of 1 to 5 based on seven criteria: value for money, quality of rooms, location of the hotel, cleanliness of the hotel, quality of check-in, overall quality of services, and quality of business services. This dataset comprises 28,829 ratings provided by 1,039 users for 693 hotels.

To demonstrate the effectiveness of our proposed algorithm, we employ two widely adopted metrics in the field of recommender systems: mean absolute error (MAE) and coverage rate. MAE is a commonly used measure for evaluating prediction accuracy. It quantifies the average absolute deviation between the predicted ratings and the actual ratings of the target items. A lower MAE value indicates a higher accuracy of the recommender system in predicting ratings. Coverage rate, on the other hand, is utilized to assess the diversity of the recommendation results, particularly in situations where data sparsity is a challenge. This metric measures the ability of a recommendation method to generate predictions for a diverse set of items, including those that have not been rated or are newly introduced. It represents the proportion of all available items that can be recommended. A higher coverage rate signifies that the recommendation model is capable of suggesting a wider variety of items within the available data, even including new or unrated items. This enhances the diversity in generating personalized recommendations and mitigates the impact of data sparsity [24].

In our evaluation, we assessed the performance of proposed HMCTeCF algorithm by comparing it with four benchmark CF-based approaches: i) the multi-criteria user-based CF approach (MC-UCF) [18], which employs the similarity-based approach to incorporate and leverage multi-criteria rating between users to improve recommendation accuracy; ii) the multi-criteria item-based CF algorithm (MC-ICF) [18], which similarly uses the similarity-based approach but focuses on incorporating and leveraging multi-criteria ratings between items to improve recommendation accuracy; iii) the multi-criteria user-item based CF approach (MC-UICF), a hybrid approach that combines the principles of MC-UCF and MC-ICF to leverage user and item multi-criteria ratings for enhanced recommendations; and iv) the MC trust-enhanced CF (MC-TeCF) approach [25], which seeks to improve predictive accuracy and address issues of data sparsity and cold-start users by utilizing multi-criteria ratings in conjunction with inferred trust relationships among users.

3.2. Comparison of recommendation approaches

This section provides a detailed comparative analysis of the proposed algorithm's performance, comparing it to benchmark approaches across two key metrics: prediction accuracy and coverage rate. The analysis is based on a series of experiments designed to thoroughly evaluate the algorithm's effectiveness.

3.2.1. Evaluation of prediction accuracy using tourism-related datasets

The evaluation of the proposed HMCTeCF algorithm's prediction accuracy on two real-world datasets from the TripAdvisor tourism platform, the Restaurants dataset and the Hotels dataset showcases its effectiveness and superiority when compared to benchmark approaches including MC-UCF, MC-ICF, MC-UICF, and MC-TeCF. This evaluation covers various sizes of nearest neighbors.

The results on the restaurants dataset, as shown in Figure 1, demonstrate the superiority of the proposed HMCTeCF algorithm in terms of prediction accuracy, as measured by the MAE metric. Across different sizes of nearest neighbors, from 5 to 70, the HMCTeCF consistently outperforms the benchmark approaches, including MC-UCF, MC-ICF, MC-UICF, and MC-TeCF. The trend of superior performance continues across all neighbor sizes, with HMCTeCF achieving the lowest MAE of 0.975 at 25 neighbors. Overall, the average MAE of HMCTeCF is 0.982, representing a 16% improvement over MC-UCF, a 10% improvement over MC-ICF, a 6% improvement over MC-UICF, and a 4% improvement over MC-TeCF. These improvements highlight the algorithm's efficiency in the restaurant recommendation domain.

The results on the Hotels dataset, as shown in Figure 2, further reinforce the superior performance of the proposed HMCTeCF algorithm. Across all neighborhood sizes, HMCTeCF exhibits the lowest MAE values, outperforming the benchmark approaches by a substantial margin. This highlights the effectiveness of the HMCTeCF algorithm in providing more accurate recommendations. The HMCTeCF algorithm consistently outperformed the benchmarks across various sizes of nearest neighbors, with the most notable gains at smaller neighbor sizes, crucial for addressing data sparsity, where traditional CF-based approaches often struggle. As the number of neighbors increases, the MAE of HMCTeCF continues to remain the lowest, with values such as 0.722 for 10 neighbors and 0.710 for 20 neighbors. The closest competitor, MC-TeCF, has corresponding MAEs of 0.843 and 0.827, respectively. On average, the HMCTeCF shows the highest percentage improvement in MAE compared to MC-UCF (40%), followed by MC-ICF (29%), MC-UICF (18%), and MC-TeCF (14%). These improvements highlight the effectiveness and accuracy of HMCTeCF in predicting user preferences for hotel recommendations.

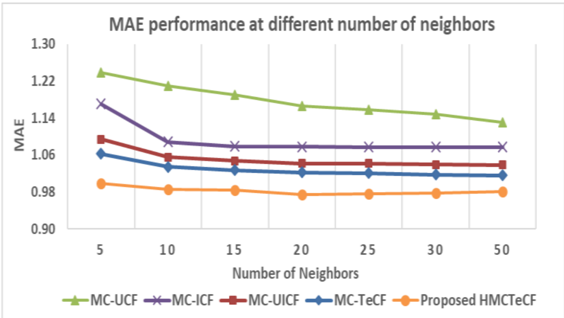

Figure 1. The MAE performance on the restaurant's dataset with varying nearest neighbor sizes

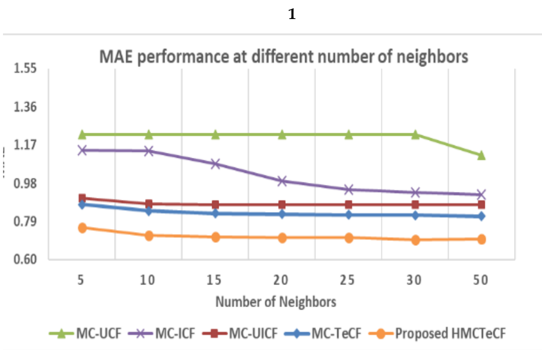

Figure 2. The MAE performance on the Hotels dataset with varying nearest neighbor sizes

3.2.2. Evaluation of prediction accuracy and coverage rate across varying levels of sparsity

To assess the robustness of the proposed HMCTeCF algorithm in handling data sparsity, a common challenge in recommender systems, we evaluated its performance on six datasets with varying levels of sparsity, ranging from 99.8% to 98.0%. The main goal was to assess the capacity of the compared approaches to handle varying levels of data sparsity and to evaluate their respective performance in that context.

Figures 3 and 4 illustrate the superior performance of the HMCTeCF algorithm in terms of MAE and Coverage rate compared to the benchmark approaches. As shown by Figure 3, as the level of sparsity decreases (from 99.8% to 98.0%), HMCTeCF maintains the lowest MAE values compared to other approaches. At 99.8% sparsity, HMCTeCF achieves an MAE of 1.302, significantly better than the next best (MC-TeCF) with 2.731. This trend continues across all sparsity levels, with HMCTeCF achieving the lowest MAE of 0.668 at 98.0% sparsity. The proposed algorithm shows substantial improvements, especially in high-sparsity scenarios. For example, at a sparsity level of 99.8%, the proposed algorithm shows an improvement of around 67% over the MC-UCF, MC-ICF, and MC-UICF approaches, and approximately 52% over the MC-TeCF approach. The large percentage improvements at the 99.8% sparsity level advise the proposed approach is particularly well-suited for applications with very limited data availability.

In terms of coverage rate, as depicted in Figure 4, the HMCTeCF excels in terms of coverage across all sparsity levels. At 99.8% sparsity, it achieves a coverage of 89%, far exceeding the next best (MC-TeCF) with 41%. This high coverage rate is maintained even at lower sparsity levels, indicating the algorithm's ability to provide relevant recommendations despite data limitations. The coverage rate improvement continues at lower sparsity levels, with HMCTeCF achieving 99.570 coverage rate at 98.0% sparsity, compared to MC-TeCF's 98.460, showing a consistent advantage in recommendation comprehensiveness.

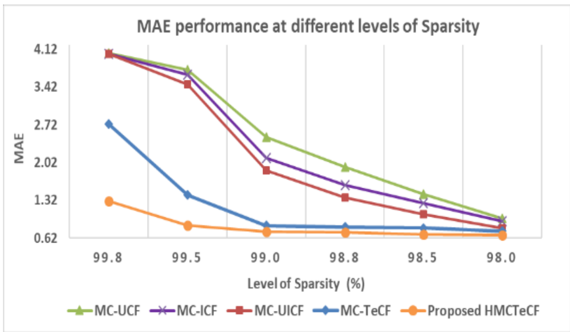

Figure 3. Analyzing MAE performance across various sparsity levels

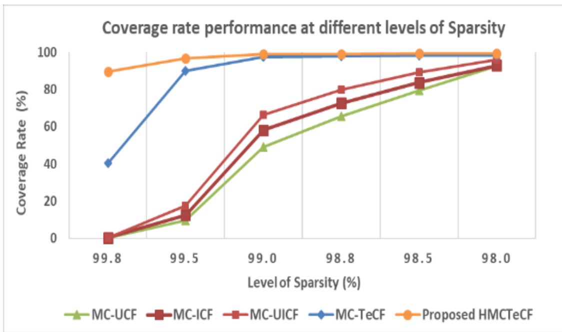

Figure 4. Analyzing coverage rate performance across various sparsity levels

Overall, the HMCTeCF algorithm demonstrated superior performance in prediction accuracy across multiple datasets and evaluation scenarios. When tested on the Restaurants dataset, HMCTeCF achieved an average MAE of 0.982, showing improvements of 4-16% over benchmark approaches. The algorithm's effectiveness was even more evident in the Hotels dataset, with MAE improvements ranging from 14-40% compared to other benchmarks. These results were consistent across various sizes of nearest neighbors, indicating the algorithm's robust performance in different recommendation contexts. The HMCTeCF excelled in addressing data sparsity, a common issue in recommender systems. Across sparsity levels from 99.8% to 98.0%, the algorithm maintained the lowest MAE values compared to benchmarks. At the highest sparsity level of 99.8%, HMCTeCF achieved the lowest MAE of 1.302, significantly outperforming other benchmarks. In terms of coverage rate, HMCTeCF consistently provided more comprehensive recommendations, achieving 89% coverage at 99.8% sparsity, far surpassing other benchmark approaches. This consistently superior performance in both prediction accuracy and coverage rate, particularly in high-sparsity scenarios, underscores HMCTeCF's potential for real-world applications where data availability is limited.

To conclude, the HMCTeCF algorithm represents a significant advancement in tourism recommender systems, effectively addressing the challenges of information overload, data sparsity, and diverse user preferences. By integrating trust relationships and multi-criteria ratings, HMCTeCF delivers more accurate, personalized, and reliable recommendations, paving the way for enhanced decision-making and more satisfying travel experiences. As the tourism industry evolves within the digital landscape, HMCTeCF stands out as a promising solution, offering significant benefits for both travelers and businesses. It empowers travelers by delivering tailored recommendations that meet their unique needs, thereby improving satisfaction and engagement. For businesses, the algorithm provides a tool to better understand and cater to customer preferences, potentially increasing loyalty and revenue.

4. CONCLUSION

In the era of information abundance, where online resources related to travel and tourism are proliferating at an unprecedented rate, advanced tools are imperative to assist users in navigating this vast landscape and making well-informed decisions. This study introduces the HMCTeCF algorithm, a novel approach designed to enhance the effectiveness of recommender systems in the tourism domain. By seamlessly integrating multi-criteria ratings and trust-based mechanisms, HMCTeCF offers a comprehensive solution that addresses the intricate nature of travel preferences and the complexities posed by the overwhelming volume of available information, ultimately leading to more satisfying and personalized tourism experiences for both tourists and industry stakeholders.

Through the integration of trust relationships among users and items, HMCTeCF leverages the collective wisdom of the user community, identifying reliable neighbors and improving the reliability of recommendations. Concurrently, the incorporation of multi-criteria ratings provides deep insights into users' diverse preferences, enabling HMCTeCF to offer precise and personalized recommendations that align with their unique needs and priorities. Extensive experimental evaluations, conducted on real-world multi-criteria rating datasets from the well-known TripAdvisor tourism platform, validate the efficacy of HMCTeCF in comparison to benchmark recommendation approaches. Across multiple performance metrics, including prediction accuracy and coverage rate, HMCTeCF consistently outperforms these approaches, demonstrating superior accuracy and coverage of recommended items. Moreover, HMCTeCF effectively addresses the challenge of data sparsity. By leveraging trust relationships and multi-criteria ratings, the algorithm can generate meaningful recommendations even in scenarios with limited user interaction data, ensuring a robust and reliable performance in real-world tourism applications. Future research on the HMCTeCF algorithm could develop hybrid models that merge multi-criteria collaborative filtering with content-based filtering or deep learning methods to improve recommendation accuracy. Incorporating sentiment analysis from user reviews could facilitate ongoing refinement and personalization. Integrating data from multiple platforms might lead to more complete user profiles, further enhancing recommendation relevance. These improvements aim to build a more advanced system that delivers highly relevant recommendations across different domains, ultimately enhancing users' decision-making processes and experiences.

REFERENCES

[1] F. Ricci, "Recommender systems in tourism," in *Handbook of e-Tourism*, Cham: Springer International Publishing, 2022, pp. 457–474.

[2] J. L. Sarkar, A. Majumder, C. R. Panigrahi, S. Roy, and B. Pati, "Tourism recommendation system: a survey and future research directions," *Multimedia Tools and Applications*, vol. 82, no. 6, pp. 8983–9027, Mar. 2023, doi: 10.1007/s11042-022-12167-w.

[3] Q. Y. Shambour, A. A. Abu-Shareha, and M. M. Abualhaj, "A hotel recommender system based on multi-criteria collaborative filtering," *Information Technology and Control*, vol. 51, no. 2, pp. 390–402, Jun. 2022, doi: 10.5755/j01.itc.51.2.30701.

[4] H. Ko, S. Lee, Y. Park, and A. Choi, "A survey of recommendation systems: recommendation models, techniques, and application fields," *Electronics (Switzerland)*, vol. 11, no. 1, p. 141, Jan. 2022, doi: 10.3390/electronics11010141.

[5] Q. Y. Shambour, M. M. Al-Zyoud, A. H. Hussein, and Q. M. Kharma, "A doctor recommender system based on collaborative and content filtering," *International Journal of Electrical and Computer Engineering*, vol. 13, no. 1, pp. 884–893, Feb. 2023, doi: 10.11591/ijece.v13i1.pp884-893.

[6] Q. Y. Shambour, N. M. Turab, and O. Y. Adwan, "An effective e-commerce recommender system based on trust and semantic information," *Cybernetics and Information Technologies*, vol. 21, no. 1, pp. 103–118, Mar. 2021, doi: 10.2478/cait-2021-0008.

[7] Q. Y. Shambour, A. H. Hussein, Q. M. Kharma, and M. M. Abualhaj, "Effective hybrid content-based collaborative filtering approach for requirements engineering," *Computer Systems Science and Engineering*, vol. 40, no. 1, pp. 113–125, 2022, doi: 10.32604/CSSE.2022.017221.

[8] F. Wang, H. Zhu, G. Srivastava, S. Li, M. R. Khosravi, and L. Qi, "Robust collaborative filtering recommendation with user-item-trust records," *IEEE Transactions on Computational Social Systems*, vol. 9, no. 4, pp. 986–996, Aug. 2022, doi: 10.1109/TCSS.2021.3064213.

[9] S. Forouzandeh, K. Berahmand, E. Nasiri, and M. Rostami, "A hotel recommender system for tourists using the artificial bee colony algorithm and fuzzy TOPSIS model: a case study of TripAdvisor," *International Journal of Information Technology and Decision Making*, vol. 20, no. 1, pp. 399–429, Jan. 2021, doi: 10.1142/S0219622020500522.

[10] C. Cui, M. Wei, L. Che, S. Wu, and E. Wang, "Hotel recommendation algorithms based on online reviews and probabilistic linguistic term sets," *Expert Systems with Applications*, vol. 210, p. 118503, Dec. 2022, doi: 10.1016/j.eswa.2022.118503.

[11] H. Xia, W. An, G. Liu, R. Hu, J. Z. Zhang, and Y. Wang, "Smart recommendation for tourist hotels based on multidimensional information: a deep neural network model," *Enterprise Information Systems*, vol. 17, no. 4, Apr. 2023, doi: 10.1080/17517575.2021.1959651.

[12] G. J. A. Jose, M. Mastan, and L. A. H. Al-Nuaimy, "Aspect based hotel recommendation system using dilated multichannel CNN and BiGRU with hyperbolic linear unit," *International Journal of Machine Learning and Cybernetics*, May 2024, doi: 10.1007/s13042-024-02184-6.

[13] E. Asani, H. Vahdat-Nejad, and J. Sadri, "Restaurant recommender system based on sentiment analysis," *Machine Learning with Applications*, vol. 6, p. 100114, 2021, doi: 10.1016/j.mlwa.2021.100114.

[14] A. Saelim and B. Kijsirikul, "A deep neural networks model for restaurant recommendation systems in Thailand," in *ACM International Conference Proceeding Series*, Feb. 2022, pp. 103–109, doi: 10.1145/3529836.3529925.

[15] S. Yang, Q. Li, D. Jang, and J. Kim, "Deep learning mechanism and big data in hospitality and tourism: developing personalized restaurant recommendation model to customer decision-making," *International Journal of Hospitality Management*, vol. 121, p. 103803, Aug. 2024, doi: 10.1016/j.ijhm.2024.103803.

[16] S. Perumal, S. Rawal, and Richa, "Handling cold-start problem in restaurant recommender system using ontology," in *Proceedings of Emerging Trends and Technologies on Intelligent Systems*, 2023, pp. 319–329, doi: 10.1007/978-981-19-4182-5_25.

[17] J. Lu, Q. Zhang, and G. Zhang, "Recommender systems," in *World Scientific*, 2020.

[18] G. Adomavicius and Y. O. Kwon, "New recommendation techniques for multicriteria rating systems," *IEEE Intelligent Systems*, vol. 22, no. 3, pp. 48–55, May 2007, doi: 10.1109/MIS.2007.58.

[19] S. Bag, S. K. Kumar, and M. K. Tiwari, "An efficient recommendation generation using relevant Jaccard similarity," *Information Sciences*, vol. 483, pp. 53–64, May 2019, doi: 10.1016/j.ins.2019.01.023.

[20] H. Song, Q. Pei, Y. Xiao, Z. Li, and Y. Wang, "A novel recommendation model based on trust relations and item ratings in social networks," *Proceedings - 2017 International Conference on Networking and Network Applications, NaNA 2017*, vol. 2018-January, pp. 17–23, 2017, doi: 10.1109/NaNA.2017.17.

[21] P. Resnick, N. Iacovou, M. Suchak, P. Bergstrom, and J. Riedl, "GroupLens: an open architecture for collaborative filtering of netnews," *Proceedings of the 1994 ACM Conference on Computer Supported Cooperative Work, CSCW 1994*, pp. 175–186, 1994, doi: 10.1145/192844.192905.

[22] R. Burke, "Hybrid web recommender systems," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 4321 LNCS, Berlin, Heidelberg: Springer Berlin Heidelberg, 2007, pp. 377–408.

[23] D. Jannach, M. Zanker, and M. Fuchs, "Leveraging multi-criteria customer feedback for satisfaction analysis and improved recommendations," *Information Technology and Tourism*, vol. 14, no. 2, pp. 119–149, 2014, doi: 10.1007/s40558-014-0010-z.

[24] E. Zangerle and C. Bauer, "Evaluating recommender systems: survey and framework," *ACM Computing Surveys*, vol. 55, no. 8, pp. 1–38, Aug. 2022, doi: 10.1145/3556536.

[25] Q. Shambour, "A user-based multi-criteria recommendation approach for personalized recommendations," *International Journal of Computer Science and Information Security (IJCSIS)*, vol. 14, no. 12, pp. 657–663, 2016.

BIOGRAPHIES OF AUTHORS

Qusai Y. Shambour holds a B.Sc. degree in Computer Science from Yarmouk University, Jordan, in 2001, an M.Sc. degree in computer networks from the University of Western Sydney, Australia, in 2003, and a Ph.D. degree in software engineering from the University of Technology Sydney, Australia, in 2012. Currently, he is a Professor at the Department of Software Engineering, Al-Ahliyya Amman University, Jordan. His research interests include information filtering, recommender systems, machine learning, and data science. He can be contacted at email: q.shambour@ammanu.edu.jo.

Mahran M. Al-Zyoud is an assistant professor of Networks and Cybersecurity at Al-Ahliyya Amman University. He received a B.Sc. in Computer Science and a M.Sc. in Computer Information Systems from The University of Jordan in 2004 and 2012. He received a Ph.D. in Computer Science from The University of Alabama, USA, in 2019. His research interests include data privacy and IoT security. He can be reached by email at: m.zyoud@ammanu.edu.jo.

Adeeb M. Alsaaidah received the bachelor's degree in Computer Engineering from the Faculty of Engineering, AL-Balqa Applied University, the master's degree in Networking and Computer Security from NYIT University, and the Ph.D. degree in Computer Network from USIM, Malaysia. He is currently an Assistant Professor in Network and Cybersecurity department at Al-Ahliyya Amman University (AAU). His research interests include network performance, multimedia networks, network quality of service (QoS), the IoT, network modeling and simulation, network security, and cloud security. He can be contacted at email: a.alsaaidah@ammanu.edu.jo.

Mosleh M. Abualhaj is a senior lecturer at Al-Ahliyya Amman University. He received his first degree in Computer Science from Philadelphia University, Jordan, in 2004, his master degree in Computer Information System from the Arab Academy for Banking and Financial Sciences, Jordan, in 2007, and his Ph.D. in Multimedia Networks Protocols from Universiti Sains Malaysia in 2011. His research area of interest includes VoIP, multimedia networking, and congestion control. He can be contacted at email: m.abualhaj@ammanu.edu.jo.

Ahmad Adel Abu-Shareha received his first degree in Computer Science from Al Al-Bayt University, Jordan, in 2004, Master degree from Universiti Sains Malaysia (USM), Malaysia, in 2006, and his Ph.D. degree from USM, Malaysia, in 2012. His research focuses on data mining, artificial intelligent, and multimedia security. He investigated many machine learning algorithms and employed artificial intelligence in a variety of fields, such as network, medical information processing, knowledge construction, and extraction. He can be contacted at email: a.abushareha@ammanu.edu.jo.