A HYBRID RECOMMENDER SYSTEM COMBINING COLLABORATIVE FILTERING WITH UTILITY MINING

Mohammed Ali Fouad*  
Information Systems Department  
Luxor University, Faculty of Computer and Information Sciences  
Luxor, Egypt  
mfouad@fci.svu.edu.eg

Wedad Hussein  
Information Systems Department  
Ain Shams University, Faculty of Computer and Information Sciences  
Cairo, Egypt  
Wedad.Hussein@cis.asu.edu.eg

Sherine Rady  
Information Systems Department  
Ain Shams University, Faculty of Computer and Information Sciences  
Cairo, Egypt  
srady@cis.asu.edu.eg

Philip S. Yu  
Department of Computer Science  
University of Illinois at Chicago  
Chicago, USA  
psyu@uic.edu

Tarek F. Gharib  
Information Systems Department  
Ain Shams University, Faculty of Computer and Information Sciences  
Cairo, Egypt  
tfgharib@cis.asu.edu.eg

Received 2022-06-16; Revised 2022-08-02; Accepted 2022-08-09

Abstract: Based on a variety of information sources, recommender systems can identify specific items for various user interests. Techniques for recommender systems are classified into two types: personalized and non-personalized. Personalized algorithms are based on individual user preferences or collaborative filtering data; as the system learns more about the user, the recommendations will become more satisfying. They do, however, suffer from data sparsity and cold start issues. On the other hand, non-personalized algorithms make recommendations based on the importance of the items in the database; they are very useful when the system has no information about a specific user. Their accuracy, however, is limited by the issue of personalization. In most cases, one of the recommendation categories can be used to make recommendations. Yet, it is a challenge to evaluate the importance of items to the user while simultaneously using personalized and non-personalized preferences functions and ranking a set of candidate items based on these functions. This paper addresses this issue and improves recommendation quality by introducing a new hybrid recommendation technique. The proposed hybrid recommendation technique combines the importance of items to the user obtained by the utility mining method with the similarity weights of items produced by the collaborative filtering technique to make the recommendation process more reasonable and accurate. This technique can provide appropriate recommendations whether or not users have previous purchasing histories. Finally, experimental results show that the proposed hybrid recommendation technique outperforms both implemented collaborative filtering and utility-based recommendation techniques.

Keywords: Recommender systems, Collaborative Filtering, Utility Mining.

* Corresponding author: Mohammed Ali Fouad  
Information Systems Department, Luxor University, Faculty of Computer and Information Sciences, Luxor, Egypt  
E-mail address: mfouad@fci.svu.edu.eg
1. Introduction

The collaborative filtering (CF) technique is a key component of most recommendation systems, in which data from the user's preferences is joined with that of other users to predict what additional items the user may want [1]. This technique is based on observing other users' preferences that are similar to the target user's historical preferences or finding items that are considered similar to the items the user liked previously [2]. CF-based systems suggest items to the user depending on the ratings of those other users. Firstly, users rate the items then the system compares their ratings to those of other users to recommend items based on similar preferences [3]. The core data used to build the CF-based system is users' past ratings. These ratings are used to locate other people with similar interests [4]. The CF technique is classified into two categories. First, the User-Based CF method quantifies the similarities between the users [4]. It first identifies other users who share the target user's characteristics and then makes a recommendation to the target user based on the items liked and rated by other users. Second, the Item-Based CF method quantifies the similarities between the target item and other items rather than users [4]. It looks at the user-item matrix to find items rated by users and then uses this specific item to determine how similar they are to the target item.

Although the CF technique has a good application as a standard recommendation technology, some issues still need to be addressed. This technique is unstable when there is not enough information about the user or the item to recommend, as it suffers from data sparsity and cold start issues [3]. Moreover, most CF-based algorithms ignore the sequential characteristics of historical transactions and are unable to display users' short-term preferences [2]. These issues eventually lead to a reduction in the quality or effectiveness of the outcomes.

Utility mining is a novel data mining approach that investigates the relative importance of items and has emerged as one of the most thriving research topics [5]. The utility indicator (i.e., importance, satisfaction, or interest) of each item can be predefined in the utility mining techniques based on a user's preferences or background knowledge. The utility is a quantitative representation of user preference [6]. The utility-based system uses a function that maps how satisfied a user is. It defines the users' satisfaction and computes the utility of the candidate items to the users [7]. The significance of utility-based recommender systems is that they can contribute non-product attributes to the utility computation, such as provider reliability and stock availability, making it possible, for example, to trade off price versus delivery schedule for a user with an immediate need [8]. Utility-based methods can address the shortcomings of new user issues, data sparsity, and new items that collaborative filtering cannot address [7].

As recommender systems grew, a pressing need for more qualified and relevant recommendations appeared, resulting in the emergence of new hybrid models. Hybrid systems combine recommendation techniques that inherit the benefits of different techniques while eliminating their drawbacks [9]. Many techniques combine CF with other techniques [4]. The majority of current hybrid based recommendation
approaches are a combination of CF and content-based recommendation methods to overcome the disadvantages of each by leveraging their respective strengths [10]. Despite the advantages of hybrid methods over single approach applications, these methods have some drawbacks. Using a hybrid approach in real-time systems, such as session-based e-commerce websites, can be challenging because two or more models must be trained, and the recommendation of hybrid model interrogation is time-consuming [8]. As a result, returning a recommendation list may take too long [11].

Hybrid models can help to increase the quality of recommendations and provide various types of interest. In order to achieve good results, recommender systems have recently employed more evaluation indicators to meet the individualized needs of users [12]. This field of academic research commonly makes recommendations based on information about the user or items, such as the user's preferences or the properties of the items [10]. However, e-commerce recommendations are more likely to require the user to have additional domain knowledge, such as risk, profit, or cost. As a result, more qualified and accurate recommendations are required to consider domain knowledge with the collaborative filtering systems.

This paper investigates a recommender system based on a hybrid algorithm of utility mining and item-based collaborative filtering to address the issues mentioned above and better serve users. Its core algorithm addresses the problem of data sparsity and ensures recommendation quality. The proposed solution assesses the importance of items to the user using three interestingness criteria: profit, stability, and frequency. The rest of this paper is structured as follows. Section 2 discusses the major published works related to hybrid recommendation techniques. The proposed hybrid recommendation technique is introduced in Section 3. Section 4 shows experiments to validate the effectiveness of the proposed technique. The conclusions and future work are outlined in Section 5.

2. Related Work

It became more difficult as the amount of information increased for users to choose the items or services they wanted to buy or subscribe to. This is where recommender systems can help [4]. Many researchers proposed various techniques and methods and focused on the issue of developing a high-quality recommender system. Collaborative filtering [13], Session-based [14], Markov-chain [15], Content-based [4], Utility-based [16], and Sequential-based [17] recommendation techniques are frequently used to recommend items for users to use/purchase based on their preferences, ratings, or purchasing history.

Moreover, researchers have recently used a variety of implicit measurements to determine the importance of items to users, such as frequency [18], novelty [19], utility [20], popularity [13], occupancy [21], reliability [5], trust [22], stability [23], diversity [12], consistency [24], convergence [12], and so on. These algorithms can present a variety of recommender systems; however, the issue of personalization limits their accuracy. As these algorithms can be classified as non-personalized recommender algorithms because they do not rely on the user to provide recommendations.
This section will go into the hybrid methods for implementing recommender systems as well as the methodologies that these methods employ. Since many factors may impact a user’s decision at the same time, the hybrid method is thought to improve existing recommendation algorithms. Hybrid baseline algorithms, such as SFUI-UF [25], GP-TopFreq [18], SPP-Growth[23], and HUOPM [21], use more than one objective to produce more accurate purchasing patterns. For example, the skyline query was used by SFUI-UF to find all non-dominated patterns based on frequency and utility. Nonetheless, the HUOPM algorithm takes into account items’ relative importance in terms of frequency, occupancy, and utility. These methods use a combination of the techniques to increase the accuracy and decrease the limitations of standard techniques. Moreover, we can combine two or more recommendation techniques to construct a more accurate algorithm or system for generating a recommendation. For example, deep learning models have been combined with collaborative filtering as well as content-based information to improve the accuracy of recommendations [11].

To deliver more accurate recommendations, different hybrid recommendation techniques have recently been integrated. The authors in [4] have attempted to combine CF with content-based techniques in order to improve recommendation quality. Tsikrika et al. [26] developed a hybrid-based recommendation approach that works by merging different recommendation techniques. For example, a hybrid approach is employed while selecting the most similar neighbors by combining content-based and ratings-based evidence. However, to improve both diversity and accuracy, the authors in [27] suggest combining a neighborhood-based method with a session-based recommendation method. Recommender systems can address data sparsity and cold start issues using a hybrid collaborative filtering recommendation algorithm that combines the KNN model with the XGBoost model and employs the scores predicted by the model-based personalized recommendation algorithm as features [28]. While incorporating interestingness measurements into the similarity matrix is a good method, the relevance of the items helps to reinforce the accuracy of CF-based algorithms [13]. The UP-CF method, for example, merges user-based CF and user-wise popularity approaches. Furthermore, the IP-CF method integrates item-based CF and user-wise popularity techniques [13]. Because these methods have been shown to be effective, it may be preferable to incorporate as many interestingness criteria as possible into the similarity matrix simultaneously. Finally, the goal of the proposed hybrid recommendation technique is to realize that the utility-based recommendation method is used to optimize the results of collaborative filtering recommendations. Still, this area has yet to be extensively researched.

3. The Proposed Hybrid Recommendation Technique

Traditional CF recommendation algorithms generate recommendations based on the correlation or similarity of items or users in order to improve recommendation accuracy. Many academics, however, believe that the CF technique is ineffective due to the low quality of recommendations [29], [30]. As a result, this section proposes a new extension to the traditional CF recommendation technique. This technique combines the importance of items to the user obtained by the utility-based recommendation technique with the similarity score of items produced by the item-based CF algorithms to make the recommendation process more efficient and accurate.
The proposed technique not only takes advantage of the benefits of utility mining or the utility-based recommendation approach, but it can also perform similarity matching filtering on all items to identify the items that can be filtered out and recommended to users. In a recommender system, when the number of users, items, and evaluation levels is large, the user-item data matrix becomes rather dense, which might cause a sparsity problem and complicate collaborative filtering. In the case that items are not evaluated by many users, we employ domain knowledge about the items, such as profit. This technique may be used to address the shortcomings of data sparsity, new users, and new items that the collaborative filtering method cannot handle. System performance will be considerably improved as a result of integrating the similarity matrix with the utility of the items to the users.

A hybrid algorithm based on utility mining and item-based CF is presented based on the enhanced collaborative filtering method. A user-item matrix is built first. Then it reinforces or supports using the items’ utility by combining the user rating of the item if found with utility features or simply uses the item’s utility in the event of new items or new users. Thus, the proposed technique can address the standard collaborative filtering algorithm's data sparsity issue. Simultaneously, for new items, it can predict users who may be interested in new items based on the score of the temporal utility of the item at this time, effectively solving the problem of cold start. It is expected that the proposed hybrid recommendation technique plays an important role in tackling the speed bottleneck problems of data sparsity and cold start and ensuring superior recommendation quality.

Figure 1 depicts the schema of the proposed hybrid recommendation technique, which is based on utility mining and collaborative filtering recommendation methodologies. The entire recommendation process is divided into two modules: utility based recommendation and collaborative filtering recommendations, both of which are beneficial to users and lead to higher accuracy.
The recommendation technique is prepared in the following manner: first, the user's interest is extracted from the user's history data. Second, the data processing is established based on the utility recommendation module using domain knowledge. After that, based on the user's purchasing and browsing data, a collaborative filtering-based recommendation module is built by examining the similarity between all pairs of items. The weight of the items is then calculated as the total similarity with the target user's current access sequence of items. Then, it combines the resulting weight with the item's actual utility in the domain to provide the user with the highest interesting items.

Traditional CF recommendation algorithms consider \( B_x \) a collection of transactions in which item \( x \) appears in the database. The similarity between two items \( x \) and \( j \) is described by the asymmetric cosine similarity [13], \( \text{sim}(x, j) \), and calculated using Eq. (1).

\[
\text{sim}(x, j) = \frac{|B_x \cap B_j|}{|B_x||B_j|} \quad (1)
\]

The similarity between item \( x \) and user \( u \) is calculated using Eq. (2) as \( s(u, x) \), where \( I^u \) is a collection of items bought or used by user \( u \) and \( |I^u| \) is the number of these items.

\[
s(u, x) = \frac{\sum_{j \in I^u} \text{sim}(x, j)}{|I^u|} \quad (2)
\]

Utility mining is based on the assumption that users will act independently in a way that minimizes costs while maximizing benefits [31]. This paper investigates the profit, stability, and frequency metrics in order to determine the utility or importance of items to users. The utility function will extract all items having a high utility value from a transaction database. Given a transactional database \( D \), which contains a set of transactions, \( D = \{T_1, T_2, \ldots, T_n\} \). The utility of the item is the total of its internal and external utility. The utility of an item \( i \) in a transaction database \( D \) is termed as \( u(i, D) \), \( u(i, D) = p(i) + q(i, D) + s(i, D) \), where \( p(i) \) is the predefined external utility or profit unit of item \( i \). The internal utility include the frequency or quantity of item \( i \) in \( D \) is indicated as \( q(i, D) \) and the stability of item \( i \) in \( D \) is measured using the maximum periodicity function [32] and indicated as \( s(i, D) \). As a result, the item's importance is calculated by dividing the sum of its importance or utility according to these criteria by the number of criteria. The importance or utility of item \( x \) at time \( t \) is calculated using Eq. (3).

\[
w_t(x) = \frac{\sum_{c=1}^{n} f_t^c(x)}{n} \quad (3)
\]

where \( n \) is the number of criteria, \( f_t^c(x) \) reflects the rank of item \( x \) based on criterion \( c \) at time \( t \), and \( c=\{\text{profit, stability, frequency}\} \). The proposed technique defines the importance of item \( x \) to user \( u \) as \( r_x^u \), which is estimated using the formula in Eq. (4).

\[
r_x^u = s(u, x) + w_t(x) \quad (4)
\]
Algorithm 1 shows the pseudo-code for the solution's processes that receive a set of transactions. Based on the criteria (profit, stability, frequency), algorithm 1 extracts a list of weights for each item. Consequently, the weights of items created by the utility algorithm are added to the similarity weights. The last step of the process is the items with the highest weights for similarity and utility were chosen as the likely next basket. To conclude, this algorithm employs two techniques for recommending the most important items to a user. By evaluating the relevance of items to users and exploiting the similarity value of items, this technique can provide a logical interpretation of the findings. As a result, more qualified and relevant recommendations can be made.

Algorithm 1. Recommender Algorithm

**input:** $B$: a set of transactions

$u$: desirable user

$pid$: desirable period

**output:** list of recommended items

1: \textbf{foreach} item $x \in I^{pid}$

$I^{pid}$: a set of items used in the period $pid$.

2: \textbf{foreach} criterion $c \in C$

3: $w_{pid}(x) = c + f_{pid}^c(x)$

$f_{pid}^c(x)$: the rank of $x$ up to $c$ in the period $pid$.

4: $w_{pid}(x)$: the total utility of $x$ in the period $pid$.

5: end for

6: $w_{pid}(x) = w_{pid}(x) / n$

$n$: the number of criteria

7: \textbf{foreach} item $j \in I^u$

$I^u$: a set of items used by the user $u$

8: \textbf{foreach} item $x \in I^{pid}$

9: $s(u, x) = \text{sim}(x, j)$

10: \textbf{end for}

11: $r_x^u = w_{pid}(x) + s(u, x)$

12: $I^* \leftarrow$ the list of items $I$ arrange by $r_x^u$, $x \in I^{pid}$

13: \textbf{return} Top-N ($I^*$).
4. Experiments and Results

The proposed technique is evaluated in this experimental study. Four real-world datasets were used to test various recommendation algorithms, namely Foodmart, E-commerce, TaFeng, and Dunnhumby. These datasets contain transactions that show which items were purchased by which user or customer and when. Table 1 displays the datasets’ statistics including names, the number of transactions, products, and customers or users, the mean of items in a transaction, the mean of transactions made by a customer, and the rate of the density of the dataset.

<table>
<thead>
<tr>
<th>Dataset</th>
<th>#Transactions</th>
<th>#Elements (I)</th>
<th>#Users</th>
<th>#avg. Len(A)</th>
<th>#avg. Transactions/user</th>
<th>#Density (A/I)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Foodmart¹</td>
<td>19050</td>
<td>1559</td>
<td>4110</td>
<td>4.25</td>
<td>4.64</td>
<td>0.27%</td>
</tr>
<tr>
<td>E-commerce²</td>
<td>20567</td>
<td>3938</td>
<td>2845</td>
<td>24.27</td>
<td>7.23</td>
<td>0.62%</td>
</tr>
<tr>
<td>TaFeng³</td>
<td>107700</td>
<td>23383</td>
<td>20388</td>
<td>6.69</td>
<td>5.28</td>
<td>0.03%</td>
</tr>
<tr>
<td>Dunnhumby⁴</td>
<td>276481</td>
<td>92339</td>
<td>2497</td>
<td>9.39</td>
<td>110.73</td>
<td>0.01%</td>
</tr>
</tbody>
</table>

The proposed recommendation technique is compared to item-based CF and utility-based recommendation algorithms in this experimental study. Users who have only one transaction are removed from the data in the preprocessing task. Each user’s transactions were organized by date and time. The dataset was then split into two sections: training and testing. The latest purchased transaction of each user is included in the testing set. The training set includes the remaining user transactions. This strategy is referred to as the leave-one-out strategy [33]. Recall is the most widely used metric for assessing the quality of recommendation systems, and we use it to assess the performance of each algorithm [18].

Table 2 compares the proposed hybrid technique to an item-based CF algorithm called IP-CF [13] and utility mining algorithms called SFUI-U U [25] and HUOPM [21], as well as demonstrates how the proposed hybrid recommendation technique outperforms state-of-the-art methods. Table 2 highlights the best and second-best results in each dataset in a bold and underlined font, respectively. Moreover, the improvement rate is calculated by the relative delta method as computing the percentage change using the equation 100* [(our score) - (best state-of-the-art score)] / (best state-of-the-art method score). Note, Top@10 means the number of items recommended to the user equals ten.

The experiments in this section provide the following observations. First, as in E-commerce, the proposed hybrid technique gains competitive performance precisely with a high-density rate. Second, the proposed hybrid technique’s performance is proportional to the average number of items in a dataset transaction. Because narrowing the user’s search space can help identify similarity patterns between items and recommend them to users. Third, utility mining-based algorithms are usually worse than collaborative filtering-based algorithms. It implies that personalized recommendation approaches, such

---

A Hybrid Recommender System Combining Collaborative Filtering with Utility Mining

as collaborative filtering algorithms, produce more satisfactory results that are closely related to the needs of the user. Fourth and most importantly, in terms of recall across all datasets, the proposed hybrid recommendation technology based on utility mining and collaborative filtering outperforms the state-of-the-art technologies. It implies that combining the importance of items obtained by the utility mining method with the user-item similarity matrix generated by the collaborative filtering method is critical in providing practical recommendations. Finally, the results across experimental datasets show that, the proposed hybrid recommendation technique is at least 15.18%, 137.73%, 10.36%, and 55.40% more accurate than other algorithms in datasets Foodmart, E-commerce, TaFeng, and Dunnhumby, respectively.

Table 2 A Comparison of the Performance of State-of-the-art Algorithms

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Algorithm</th>
<th>Top@10</th>
<th>Top@20</th>
<th>Top@30</th>
<th>Top@40</th>
<th>Top@50</th>
<th>Top@60</th>
<th>Top@70</th>
<th>Top@80</th>
<th>Top@90</th>
<th>Top@100</th>
</tr>
</thead>
<tbody>
<tr>
<td>Foodmart</td>
<td>Utility mining (SFUI-UF)</td>
<td>0.0092</td>
<td>0.0185</td>
<td>0.0254</td>
<td>0.0332</td>
<td>0.0398</td>
<td>0.0487</td>
<td>0.0565</td>
<td>0.0640</td>
<td>0.0702</td>
<td>0.0775</td>
</tr>
<tr>
<td></td>
<td>Utility mining (HUOPM)</td>
<td>0.0078</td>
<td>0.0156</td>
<td>0.0217</td>
<td>0.0285</td>
<td>0.0352</td>
<td>0.0414</td>
<td>0.0481</td>
<td>0.0558</td>
<td>0.0630</td>
<td>0.0702</td>
</tr>
<tr>
<td></td>
<td>Item-based CF</td>
<td>0.0329</td>
<td>0.0576</td>
<td>0.0797</td>
<td>0.0972</td>
<td>0.1148</td>
<td>0.1304</td>
<td>0.1458</td>
<td>0.1592</td>
<td>0.1735</td>
<td>0.1860</td>
</tr>
<tr>
<td></td>
<td>Hybrid CF + Utility mining</td>
<td>0.0379</td>
<td>0.0674</td>
<td>0.0948</td>
<td>0.1195</td>
<td>0.1439</td>
<td>0.1649</td>
<td>0.1855</td>
<td>0.2050</td>
<td>0.2251</td>
<td>0.2445</td>
</tr>
<tr>
<td></td>
<td>improvement %</td>
<td>15.18%</td>
<td>16.96%</td>
<td>18.97%</td>
<td>23.00%</td>
<td>25.36%</td>
<td>26.43%</td>
<td>27.19%</td>
<td>28.75%</td>
<td>29.69%</td>
<td>31.47%</td>
</tr>
<tr>
<td>E-commerce</td>
<td>Utility mining (SFUI-UF)</td>
<td>0.0280</td>
<td>0.0541</td>
<td>0.0712</td>
<td>0.0878</td>
<td>0.1042</td>
<td>0.1242</td>
<td>0.1401</td>
<td>0.1556</td>
<td>0.1661</td>
<td>0.1789</td>
</tr>
<tr>
<td></td>
<td>Utility mining (HUOPM)</td>
<td>0.0081</td>
<td>0.0164</td>
<td>0.0345</td>
<td>0.0417</td>
<td>0.0536</td>
<td>0.0654</td>
<td>0.0781</td>
<td>0.0955</td>
<td>0.1045</td>
<td>0.1131</td>
</tr>
<tr>
<td></td>
<td>Item-based CF</td>
<td>0.0336</td>
<td>0.0595</td>
<td>0.0805</td>
<td>0.0986</td>
<td>0.1146</td>
<td>0.1312</td>
<td>0.1492</td>
<td>0.1655</td>
<td>0.1830</td>
<td>0.2003</td>
</tr>
<tr>
<td></td>
<td>Hybrid CF + Utility mining</td>
<td>0.0940</td>
<td>0.1576</td>
<td>0.2107</td>
<td>0.2564</td>
<td>0.2984</td>
<td>0.3373</td>
<td>0.3742</td>
<td>0.4095</td>
<td>0.4430</td>
<td>0.4763</td>
</tr>
<tr>
<td></td>
<td>improvement %</td>
<td>179.70%</td>
<td>161.12%</td>
<td>161.76%</td>
<td>160.11%</td>
<td>160.51%</td>
<td>157.15%</td>
<td>150.74%</td>
<td>147.46%</td>
<td>142.05%</td>
<td>137.73%</td>
</tr>
<tr>
<td>TaFeng</td>
<td>Utility mining (SFUI-UF)</td>
<td>0.0229</td>
<td>0.0577</td>
<td>0.0755</td>
<td>0.0879</td>
<td>0.0972</td>
<td>0.1087</td>
<td>0.1186</td>
<td>0.1284</td>
<td>0.1367</td>
<td>0.1444</td>
</tr>
<tr>
<td></td>
<td>Utility mining (HUOPM)</td>
<td>0.0006</td>
<td>0.0009</td>
<td>0.0066</td>
<td>0.0070</td>
<td>0.0073</td>
<td>0.0249</td>
<td>0.0262</td>
<td>0.0293</td>
<td>0.0309</td>
<td>0.0335</td>
</tr>
<tr>
<td></td>
<td>Item-based CF</td>
<td>0.0528</td>
<td>0.0777</td>
<td>0.1017</td>
<td>0.1206</td>
<td>0.1362</td>
<td>0.1498</td>
<td>0.1629</td>
<td>0.1745</td>
<td>0.1853</td>
<td>0.1956</td>
</tr>
<tr>
<td></td>
<td>Hybrid CF + Utility mining</td>
<td>0.0583</td>
<td>0.0953</td>
<td>0.1261</td>
<td>0.1538</td>
<td>0.1796</td>
<td>0.2042</td>
<td>0.2274</td>
<td>0.2495</td>
<td>0.2707</td>
<td>0.2909</td>
</tr>
<tr>
<td></td>
<td>improvement %</td>
<td>10.36%</td>
<td>22.54%</td>
<td>23.93%</td>
<td>27.53%</td>
<td>31.93%</td>
<td>36.28%</td>
<td>39.60%</td>
<td>43.02%</td>
<td>46.13%</td>
<td>48.67%</td>
</tr>
<tr>
<td>Dunnhumby</td>
<td>Utility mining (SFUI-UF)</td>
<td>0.0098</td>
<td>0.0295</td>
<td>0.0476</td>
<td>0.0670</td>
<td>0.0849</td>
<td>0.1012</td>
<td>0.1077</td>
<td>0.1146</td>
<td>0.1222</td>
<td>0.1270</td>
</tr>
<tr>
<td></td>
<td>Utility mining (HUOPM)</td>
<td>0.0052</td>
<td>0.0054</td>
<td>0.0054</td>
<td>0.0056</td>
<td>0.0056</td>
<td>0.0060</td>
<td>0.0060</td>
<td>0.0061</td>
<td>0.0062</td>
<td>0.0063</td>
</tr>
<tr>
<td></td>
<td>Item-based CF</td>
<td>0.0415</td>
<td>0.0608</td>
<td>0.0761</td>
<td>0.0879</td>
<td>0.0986</td>
<td>0.1101</td>
<td>0.1214</td>
<td>0.1313</td>
<td>0.1391</td>
<td>0.1461</td>
</tr>
<tr>
<td></td>
<td>Hybrid CF + Utility mining</td>
<td>0.0644</td>
<td>0.0973</td>
<td>0.1228</td>
<td>0.1450</td>
<td>0.1651</td>
<td>0.1824</td>
<td>0.1976</td>
<td>0.2125</td>
<td>0.2248</td>
<td>0.2375</td>
</tr>
<tr>
<td></td>
<td>improvement %</td>
<td>55.40%</td>
<td>59.89%</td>
<td>61.46%</td>
<td>64.97%</td>
<td>67.50%</td>
<td>65.61%</td>
<td>62.77%</td>
<td>61.88%</td>
<td>61.58%</td>
<td>62.57%</td>
</tr>
</tbody>
</table>
5. Conclusions and Future Work

The goal of recommender systems is to increase the efficiency of e-commerce systems by making it easier for customers to find relevant products based on their purchasing and browsing history. However, there is a need to produce recommendations on whether or not users have purchasing histories. This paper aims to improve the quality of the recommendation process by assessing the importance of items to the user through the use of personalized and non-personalized preferences functions. We proposed a hybrid recommendation technique based on combining the advantages of utility mining and collaborative filtering, and we thoroughly investigated the hybrid recommendation technique's workflow and data processing algorithm. The hybrid recommendation technique's performance is validated. In that order, we designed experiments based on real-world datasets to evaluate the system performance of hybrid recommendation technology, item-based collaborative filtering recommendation technology, and utility-based recommendation technology. The experimental results show that the proposed hybrid recommendation technology based on utility mining and collaborative filtering outperforms the two technologies. The proposed technique is at least 10.36% more accurate than several state-of-the-art recommendation models. In the future, we plan to use query intent understanding to enhance the quality of recommendations. The submitted technique only considers the users' purchasing data for generating recommendations. Understanding the intent behind a user's query, on the other hand, can help search engines obtain particularly relevant recommendations, resulting in significantly higher user satisfaction. Therefore, combining the query intent understanding method's item weights with the weights of items obtained by the proposed recommendation technique can result in more satisfactory results. Therefore, we intend to consider existing work in the field of query intent understanding to improve the evaluation process of the importance of items to the user.

References