Kybernetika

Khadidja Belmessous; Faouzi Sebbak; M'hamed Mataoui; Walid Cherifi

A new uncertainty-aware similarity for user-based collaborative filtering

Kybernetika, Vol. 60 (2024), No. 4, 446–474

Persistent URL: http://dml.cz/dmlcz/152614

KYBERNETIKA — VOLUME *60* (2024), NUMBER 4, PAGES 446–474

A NEW UNCERTAINTY-AWARE SIMILARITY FOR USER-BASED COLLABORATIVE FILTERING

Khadidja Belmessous, Faouzi Sebbak, M'hamed MAtaoui, and Walid Cherifi

User-based Collaborative Filtering (UBCF) is a common approach in Recommender Systems (RS). Essentially, UBCF predicts unprovided entries for the target user by selecting similar neighbors. The effectiveness of UBCF greatly depends on the selected similarity measure and the subsequent choice of neighbors. This paper presents a new Uncertainty-Aware Similarity measure "UASim" which enhances CF by accurately calculating how similar, dissimilar, and uncertain users' preferences are. Uncertainty is a key factor of "UASim" that is managed in the neighborhood selection step of CF. Extensive experimental evaluation, conducted on Flixter, Movielens-100K, and Movielens-1M datasets, indicates that "UASim" shows better performance compared to many representative predefined similarity measures. The proposed measure demonstrates enhancements across various performance indicators, namely: Mean Absolute Error (MAE), Root Mean Square Error (RMSE), coverage, and the F-score.

Keywords: collaborative filtering, similarity, subjective logic, uncertainty

Classification: 68P10, 68P20, 68T37

1. INTRODUCTION

Recommender Systems (RS) [33] have been developed to provide users with intelligent automated guidance in the era of information overload. RS have been successfully used in various domains such as movies (Netflix), music (Spotify), and social networks (Facebook), as well as in other fields including news [22] and scientific paper research [28]. Collaborative Filtering (CF) is one of the most successful recommendation approaches [23]. It is based on the idea that friends and neighbors frequently influence purchasing decisions.

There are two main classes of CF approaches [1]: model-based approaches and memory-based approaches. Model-based approaches use user-item rating data to generate a model that can provide recommendations. Memory-based approaches, also known as neighborhood-based approaches, work directly on the rating matrix and include user-based and item-based classes. The UBCF

DOI: 10.14736/kyb-2024-4-0446

algorithm is primarily used for preference prediction and item recommendation. UBCF achieves this by mining a specific user's historical behavior data, identifying similar neighbors in the user's set, synthesizing their evaluations of specific items, generating a preference prediction for the items, and finally recommending interesting items to the target user.

Accurately identifying the neighbors is a crucial element of CF [7, 27]. Neighbor selection deals with two important parameters: the size of the neighborhood and the similarity measure choice. Generally, the accuracy of predictions improves as the similarity measure provides better results [34]. Since the introduction of CF, numerous similarity measures have been proposed and investigated [23]. The similarity measures can be grouped into four main categories: predefined similarities, learned similarities, preference-based similarities, and mixed similarities [15]. The first category includes measures like Cosine and Pearson, where similarity is calculated only between users or items that share ratings. The second category permits the calculation of similarity between users or items even without common ratings. The third category focuses on optimizing a user preference assumption as an objective function to calculate similarities, and the fourth category is a hybrid approach, mixing elements from the first three categories.

Predefined similarity measures are of greatest importance in CF due to their simple structure and easy combination [15, 23]. Despite their extensive use and continuous improvements, predefined similarity measures still present multiple issues that require further research and enhancement. The recognized issues include providing high or low similarity scores, ignoring co-rated item proportions, ignoring rating values, and complex formulations [15, 26]. These challenges are predominantly pronounced in the context of data sparsity. The problem of sparsity is fundamental in CF due to the limited number of items rated by users compared to the large number of existing items and users [19, 25]. In the context of sparse datasets, numerous users do not share any items, lowering the effectiveness of predefined similarity measures. Moreover, these measures provide a limited variety of strategies for filtering appropriate neighbors to aggregate their preferences in order to provide predictions [1]. This limitation primarily results from the fact that predefined measures quantify the resemblance between users via a single-valued numerical score [11], which leads to neighbor selection strategies that are mainly based on threshold values or the k-nearest neighbors (k-NN) approaches [37].

Therefore, it's important to develop a straightforward, predefined similarity measure that not only addresses the limitations of existing similarity measures but also permits varied strategies for neighbor selection. In this research, we aim to transform a simple intuition into a three-valued predefined similarity measure for UBCF, designed to effectively handle the challenge of data sparsity. We offer an alternative viewpoint on the similarity of users' preferences built on the intuition that users' preferences simultaneously exhibit characteristics of both similarity and dissimilarity, along with an inherent uncertainty due to insufficient ratings. These characteristics change according to the number of shared ratings. Our Uncertainty-Aware Similarity measure (UASim) offers a more com-

prehensive and precise representation of the relationship between users' preferences by understanding their multifaceted nature. By relying on the flexibility of binomial subjective logic [21], UASim effectively integrates similarity, dissimilarity, and uncertainty into a normalized similarity measure. This, in turn, improves the quality of recommendations, particularly in scenarios where data is sparse. Considering that each neighbor provides a three-valued UASim, selecting the most appropriate neighbors becomes a Multi-Criteria Decision-Making (MCDM) challenge. To address this complexity, we recommend choosing neighborhoods through the Technique for Order Preference based on Similarity to the Ideal Solution (TOPSIS), a widely recognized method for decision support. Our contributions are summarized below:

1. We formally model a simple intuition about the resemblance of users' preferences into an uncertainty-aware similarity measure by using binomial subjective logic. This method produces a measure that offers three separate types of knowledge.
2. We propose a new strategy for neighborhood selection in memory-based collaborative filtering;
3. We experimentally test and compare the performance of the proposed CF approach with CF based on traditional and recent predefined similarities on three well-known benchmark datasets.

This article is organized as follows: In Section 2, we review existing literature mainly related to predefined similarity measures. In section 3 we outline the theoretical background of our proposition. Section 4 delves into the components of our proposition. Section 5 outlines our experimental methodology, presents the results, and offers a comprehensive discussion of our findings. Lastly, Section 6 concludes the paper, summarizing our key conclusions and proposing avenues for future investigation.

2. RELATED WORK

Recommender systems (RS) can be categorized into three main approaches: Content-based filtering (CBF) [6], which provides recommendations based on user profiles that are generally difficult to acquire; Collaborative filtering (CF) [30], which generates recommendations by using the preferences of the most similar users; and hybrid filtering [35], a combination of both CBF and CF. Along with these primary approaches, there are advanced approaches such as the Mobile-Based approach (MBRS) [14], which leverages mobile device capabilities to provide personalized suggestions based on user preferences and contextual data, and the Context-Aware approach (CARS) [32] which specifically enhances recommendations by incorporating real-time situational and environmental information, thus focusing on the relevance and timeliness of the recommendations.

Compared to CBF, CF has made significant progress due to the ease with which real-world information about users' preferences on items may be obtained

[12, 40]. In reality, user preferences are stored as an $M * N$ user-item rating matrix R, where M represents the total number of users and N represents the total number of items. Each row of this matrix represents the rating vector of a particular user u, whereas each column provides ratings received by an item i from all users. Every entry r_{ui} of the user-item matrix represents the rating given by user u to item i. Table 1 shows an example, in which evaluations of five users related to four items are captured. When a user has not given a rating to an item, the missing rating is denoted by the symbol "•". However, due to the enormous number of users who only evaluate a small number of available items in the recommender system, a significant portion of rating inputs are empty. The proportion of empty entries to the overall rating matrix size indicates the level of sparsity of the user-item matrix [40]. Despite the high research volume in the CF domain, sparsity remains a fundamental issue that hinders recommendation performance. Consequently, several solutions have been proposed to reduce the negative effects of sparsity, including improving the similarity computation, which is the heart of CF [5, 19, 36].

	i_1	i_2	i_3	i_4
u_1	4	3	5	4
u_2	5	3	•	•
u_3	4	3	3	4
u_4	2	1	•	•
u_5	4	2	•	•

Tab. 1: An example of a User-Item rating matrix.

Type	**Numerical**			**Structural**
	Rating-based	**Ratio-based**	**Weighted**	
Examples	COS, MSD, PCC	RACF	JMSD, SPCC, TANJ, HSMD, PIP, NHSM, OS, TAN	JACCARD, SMD

Tab. 2: Summary of predefined similarity groups.

Numerical similarity measures depend entirely on shared rating information. Rating-based similarities use ratings or differences in ratings between users to determine similarity. The most popular traditional rating-based similarity measures include Cosine (COS), Pearson Correlation Coefficient (PCC), and Mean Squared Differences (MSD) [40]. Traditional rating-based measures are crucial in CF but have notable weaknesses largely documented in the literature [5, 15, 24]. For instance, COS fails to capture a user's tendency to rate an item highly (4 or 5) or poorly (1 or 2), while PCC assigns low or high similarity scores irrespective of how similarly or differently users rate various items [26], and MSD is unable to consider the proportion of common ratings [26] which sometimes leads to inconsistent similarity values. Ratio-based similarity measures, such as the Ratio-based Similarity for Collaborative Filtering (RACF)

[41], emphasize the proportional relationships in user ratings. These measures focus on comparing the ratios of ratings between two users. Despite their potential to address issues of misleading similarity values [26], ratio-based methods like RACF have not received as much attention as the rating-based measures. A summary of predefined similarity groups is given in Table 2.

Structural similarity measures account for the proportion of co-rated items. One such example is JACCARD distance (JACC) [23], which only takes the number of co-rated items into consideration and disregards rating values. The Jaccard measure may not have emerged as the most dominant measure, but it has shown its significance as a valuable factor for enhancing various numerical measures [5]. To improve the accuracy of prediction in CF, numerical similarity measures have been weighted using structural measures [38]. In the context of CF recommender systems, JMSD, SPCC, PIP, NHSM, and Bhattacharyya are typical examples of weighted measures [40]. JMSD and SPCC were proposed to address the limitations of MSD and PCC, as mentioned in [15]. The PIP measure [3], addresses the limitations of rating-based similarity measures. It calculates the similarity between users u and v by summing up the PIP values for their shared items. PIP is computed as the product of three components: proximity, which measures the absolute difference in common ratings and agreement; impact, quantifying users' preferences for specific items; and popularity, counting the number of items rated in common. Later came NHSM [26], a similarity measure based on PIP and trying to cover its weaknesses. Then came Bhattacharyya [2, 31] that uses all rating information to improve the reliability of recommendations in sparse datasets. These measures are popular in the CF domain even if they have complex formulas and higher complexity [15]. Research in the field of predefined similarity is continuously evolving, with recent years witnessing the proposal of simple and highly competitive new measures [5, 15]. Notable among these are the OS and TAN similarity measures. The recent proposals have effectively transformed basic intuitions into mathematical formulations, resulting in highly effective measures of similarity, primarily focused on weighted rating-based similarities. However, in the context of ratio-based measures, there is still some work required for improvement.

Our research introduces a new ratio-based similarity measurement in CF, where we emphasize the role of uncertainty in assessing the credibility of similarity [26, 39]. This concept is not entirely new and has been a part of weighted similarity measures in previous studies [23, 42]. However, our contribution lies in the innovative way we handle this information: we quantify the credibility of similarity as a measure of uncertainty. By employing the Technique for Order Preference based on Similarity to the Ideal Solution (TOPSIS) [18], this uncertainty is strategically incorporated into the neighborhood selection process in CF, alongside similarity and dissimilarity. Recently, a notable application of TOPSIS in CARS is demonstrated in [13], where authors integrated TOPSIS fuzzy model with an Artificial Bee Colony (ABC) algorithm to optimize personalized tourism recommendations. This approach utilizes TOPSIS to define ideal and negative ideal solutions, permitting the systematic evaluation and ranking of tourist destinations. Furthermore, in [10], authors investigated the

rank consistency of TOPSIS for mobile-based applications where new data inputs frequently occur. The MBRS developed in this research assists users in selecting sports venues by evaluating multiple criteria, such as location, cost, and facilities, which are typical considerations for such venues. In addition, in one of the most representative work about integrating TOPSIS in memory-based CF [4], authors incorporated the TOPSIS technique, as a new prediction score method, alongside the similarity values of the top-N users to generate the top-N recommended items. This approach contrasts with our proposition that relies on aggregation to predict user ratings, where the emphasis is on devising new Uncertainty-Aware Similarity measure that permits the use TOPSIS in the neighborhood selection step.

3. BACKGROUND

In this section, we will expound on the theoretical underpinnings that will help to understand the formulation and evaluation of our proposition. Specifically, we will draw upon user-based collaborative filtering, subjective logic and the TOPSIS approach.

3.1. User-based collaborative filtering

When it comes to UBCF, the objective is to filter the incoming stream of items based on the ratings given by community members who have previously rated them. If a user finds an item appealing, it will be automatically recommended to other similar users. To achieve this goal, the system must create a matrix that records the similarity scores between users. Hence, the potential rating given to an item by the target user will be determined based on its similarity with its neighboring users. This process involves five primary steps, namely: data preprocessing, similarity computation, neighborhood selection, preference prediction, and recommendation [8]. In this paper, we focus on similarity computation and neighborhood selection. Similarity computation has been detailed in the related work section 2.

3.2. Binomial subjective logic

Subjective Logic (SL) is a type of probabilistic logic that incorporates uncertainty. It is particularly useful for modeling and analyzing uncertain and unreliable source conditions [20]. In SL, arguments are opinions about propositions, and there are three types of opinions:

- Binomial opinions, which are applied to a binomial variable and are represented by a beta probability distribution function;
- Multinomial opinions, which are applied to a state variable with multiple potential values. They are represented by a Dirichlet probability distribution function;

- Hyper opinions, which generalize multinomial opinions and are applied to hyper-domains consisting of composite values represented by a hyper-Dirichlet probability distribution function.

In SL, the notation ω_x^A is used to denote opinions. x represents the target variable or proposition to which the opinion is applied, and A represents the subject agent who holds the opinion. The opinion itself is a composite function that consists of the belief vector $\overrightarrow{b_x}$, the uncertainty mass u_x, and the base rate vector $\overrightarrow{a_x}$. In the case of binary frames, the opinion is binomial. Binomial opinions correspond to a Beta probability distribution function (Beta pdf) [20].

A binary domain consists of only two values, and it can be formally stated as $\mathbb{X} = \{x, \overline{x}\}$. If a binomial random variable $X \in \mathbb{X}$ can be set to $X = x$, then opinions on a binomial variable are referred to as binomial opinions. A specific notation is used for their mathematical representation. Let $\mathbb{X} = \{x, \overline{x}\}$ be a binary domain with a binomial random variable $x \in \mathbb{X}$. A binomial opinion regarding the truth or presence of the value x is the ordered quadruplet $\omega_x = (b_x, d_x, u_x, a_x)$ that satisfies the additivity requirement. The respective parameters are defined as follows:

$$b_x + d_x + u_x = 1. \tag{1}$$

The mass of belief in favor of x being true is denoted as b_x. The mass of disbelief in favor of x being false is represented by d_x. The mass of uncertainty representing the emptiness of evidence is u_x. Finally, the prior probability of x without any evidence is called the base rate and is represented by a_x.

A binomial opinion can be expressed using the Beta probability distribution function (pdf), denoted as Beta (p — α, β), where α and β represent the two evidence parameters. The Beta pdf is given by Equation 2:

$$Beta(p|\alpha, \beta) = \frac{\tau(\alpha + \beta)}{\tau(\alpha) + \tau(\beta)} p^{\alpha-1}(1-p)^{\beta-1}. \tag{2}$$

Here, $0 \leq p \leq 1, \alpha > 0, \beta > 0$. Let r_x denote the number of observations of x, and s_x denote the number of observations of $\overline{x}$. The parameters α and β can be derived from observations (r_x, s_x), the non-informative prior weight W, and the base rate a_x, as shown in Equation 3:

$$\begin{cases} \alpha = r_x + W(a_x) \\ \beta = s_x + W(1 - a_x). \end{cases} \tag{3}$$

W is typically set to 2, which ensures that the prior Beta pdf (i. e., when $r_x = s_x = 0$) with the default base rate $a_x = 0.5$ is a uniform pdf. The parameters of a binomial opinion $\omega_x = (b_x, d_x, u_x, a_x)$ can be mapped to the parameters of a Beta pdf $Beta(p|r_x, s_x, a_x)$ using Equation 4:

$$\begin{cases} b_x = \frac{r_x}{r_x + s_x + W} \\ d_x = \frac{s_x}{r_x + s_x + W} \\ u_x = \frac{W}{r_x + s_x + W}. \end{cases} \tag{4}$$

This method provides a way to quantify a subjective opinion on a binomial variable in terms of a Beta distribution.

3.3. TOPSIS

The Technique for Order Preference based on Similarity to the Ideal Solution (TOPSIS) [18] is a Multi-Criteria Decision-Making (MCDM) method that identifies the best options among a set of alternatives based on the concept of the ideal solution. The core principle of TOPSIS lies in determining the geometric distance of each alternative to an ideal solution that maximizes the positive criteria and minimizes the negative criteria. The TOPSIS method is executed through a sequence of steps, described as follows:

1. Forming the normalized decision matrix: this initial step transforms the available options into normalized matrix with alternatives as rows and criteria as columns, enabling comparable assessment;

2. Weighting the criteria: weights are assigned to each criterion reflecting their relative importance to the decision-making process. These weights are subjective and can be derived from expert opinions or other quantitative methods;

3. Identification of ideal and anti-ideal solutions: the ideal solution is the hypothetical alternative having the best performance on all criteria, while the anti-ideal solution has the worst;

4. Distance calculation: for each alternative, the Euclidean distance to both the ideal and anti-ideal solutions is calculated. This step quantifies how close or far each alternative is from the most and least desirable conditions;

5. Relative closeness to the ideal solution: calculating the relative closeness of each alternative to the ideal solution is done by dividing the distance to the anti-ideal solution by the sum of the distances to both the ideal and anti-ideal solutions;

6. Ranking the alternatives: alternatives are ranked based on their relative closeness to the ideal solution, with the best option being the one closest to the ideal solution and furthest from the anti-ideal solution.

TOPSIS enables rapid identification of the optimal choice, is easy to apply, requires minimal input from decision-makers, and generates easily understandable information. Only the weight values associated with each criterion serve as input parameters [29].

4. PROPOSITION

Before exploring the two key components of our proposition, which are the formulation of UASim and the neighborhood selection strategy, it is crucial to highlight the main contributions of our research.

4.1. Contributions

Our method is unique in incorporating the uncertainty arising from users' shared data into the decision-making process. This approach marks a significant advancement in the field, offering a more realistic way of evaluating user similarities in CF systems. Our proposed measure aims to provide a more comprehensive and realistic assessment of similarity by accounting for:

- The inherent uncertainty in assessing user preference resemblance;
- The simultaneous existence of similarity and dissimilarity in user preferences;
- The complete uncertainty in similarity when no co-rated items exist between users;
- The potential of using similarity, dissimilarity, and uncertainty to improve the prediction of missing ratings through enhanced neighborhood selection.

Our work proposes a theoretical foundation based on subjective logic for a novel similarity measure in user-based collaborative filtering. We opted for user-based CF because of its effectiveness in dynamically capturing user similarities. Subjective logic, as described in [21], is employed for its capability to integrate affirmative, negative, and uncertain judgments within a cohesive framework.

4.2. Formulation of UASim

The proposed Uncertainty-Aware Similarity measure for UBCF is called "UASim". The components of UASim are described in depth below.

UASim uses SL to formally define similarity, dissimilarity, and uncertainty between pairs of users' preferences in UBCF. To represent the similarity between pairs of users, UASim employs a binomial opinion $\omega_x = (b_x, d_x, u_x, a_x)$, where b_x represents the degree of similarity (S) between two users' preferences, d_x represents the degree of dissimilarity $(\bar{S})$, u_x represents the uncertainty about the similarity of their preferences (U), and a_x is the base rate.

A general binomial opinion matches the Beta probability distribution function expressed in Equation 2. Equation 4 defines the mapping between the parameters of a binomial opinion $\omega_x = (b_x, d_x, u_x, a_x)$ and the parameters of a Beta pdf $Beta(p|r_x, s_x, a_x)$. To define the binomial opinion parameters that correspond to our similarity measure components $(S, \bar{S}, U)$, we introduce the positive interactions counter r_x and the negative interactions counter s_x, as shown in Equation 5.

$$\begin{cases} r_x = \sum_{i \in I_{uv}} (\frac{r_{\min}}{r_{\max}})_i \\ s_x = \sum_{i \in I_{uv}} (1 - (\frac{r_{\min}}{r_{\max}})_i). \end{cases} \tag{5}$$

Here, $r_{\min}$ and $r_{\max}$ represent the minimum and maximum ratings given to an item i by users u and v, respectively. Specifically, if $r_{ui} < r_{vi}$, then

$r_{\max} = r_{vi}$ and $r_{\min} = r_{ui}$, with $r_{\min} < r_{\max}$ and $r \in \{1, 2, 3, 4, 5\}$. If $r_{ui} = r_{vi}$, then $r_{\min} = r_{\max}$.

Example 1. Suppose $r_{ui} = 3$ and $r_{vi} = 5$. In this case, $r_{\min} = r_{ui} = 3$ and $r_{\max} = r_{vi} = 5$.

Example 2. To demonstrate the calculation of these parameters, we use the example presented in Table 1. To determine the value of the positive interactions counter r_x and the negative interactions counter s_x between $u1$ and $u2$, we use Equation 5. The resulting values of r_x and s_x for each item are shown in Table 3.

	i_1	i_2	i_3	i_4
r_x	0.8	1	•	•
s_x	0.2	0	•	•

Tab. 3: Calculation results of parameters r_x and s_x between $u1$ and $u2$.

For the definition of the binomial opinion through the parameters r_x, s_x, and W, we can map these parameters of a binomial opinion $\omega_x = (b_x, d_x, u_x, a_x)$ to those of a Beta probability density function, denoted by $Beta(p|r_x, s_x, a_x)$, as presented in equation 4. After substituting the parameters r_x, s_x and W in the equation 4, we obtain the following equations:

$$\begin{cases} S = \frac{\sum_{i \in I_{uv}} (\frac{r_{\min}}{r_{\max}})_i}{|I_{uv}|+W} \\ \bar{S} = \frac{|I_{uv}|}{|I_{uv}|+W} - \frac{\sum_{i \in I_{uv}} (\frac{r_{\min}}{r_{\max}})_i}{|I_{uv}|+W} \\ U = \frac{W}{|I_{uv}|+W}. \end{cases} \tag{6}$$

The proposed formalization of UASim transforms an intuitive explanation of the relationship between users' preferences into a consistent measure of similarity that can closely approximate reality.

The application of UASim to the data in Table 1 yields the results presented in Table 4. In accordance with [21], we set the non-informative prior weight (W) to 2 during the calculations. This value guarantees that the Beta probability density function with a default base rate of 0.5 and $r_x = s_x = 0$ corresponds to a uniform pdf.

	u2	u3	u4	u5
u1	**[0.45, 0.05, 0.5]**	**[0.66, 0.07, 0.33]**	**[0.2, 0.3, 0.5]**	**[0.41, 0.09, 0.5]**
u2	-	**[0.45, 0.05, 0.5]**	**[0.18, 0.31, 0.5]**	**[0.36, 0.13, 0.5]**
u3	-	-	**[0.2, 0.3, 0.5]**	**[0.41, 0.09, 0.5]**
u4	-	-	-	**[0.25, 0.25, 0.5]**

Tab. 4: Example outcome of UASim.

UASim is a normalized similarity measure that can be easily combined with other similarity measures. Additionally, in the absence of common items, UASim generates complete uncertainty that describes the resemblance of users' preferences. Moreover, in contrast to the ratio-based method utilized by RACF [41], UASim recognizes the complementary value of both similarity and dissimilarity as sources of knowledge. By considering this tangled viewpoint, UASim is able to identify the most suitable neighbors, whose evaluations can be trusted to improve the accuracy of predictions made through collaborative filtering.

The use of the uncertainty function in computing similarity values primarily serves to increase the precision of similarity measures. Previous research has implicitly examined the credibility of similarity but has not attempted to measure it for any specific purpose. In this work, the representative function of uncertainty in the computation of UASim is based on the non-informative prior weight W and $|I_{uv}|$ the number of items co-rated by two users u and v. The uncertainty value U is determined using the formula in Equation 7.

$$U = \frac{W}{|I_{uv}| + W} = \frac{1}{\frac{|I_{uv}|}{W} + 1}. \tag{7}$$

Since W is a constant value, the variation of U is only dependent on the proportion of common items.

- if $|I_{uv}| = 0$ then $U \to 1$ since $W \neq 0$.
- $\frac{|I_{uv}|}{W} \to \infty$, therefore $U \to 0$. Since W is a fixed value, $|I_{uv}| \to \infty$. If the number of co-rated items increases, U decreases.

By modeling this uncertainty, UASim represents the full range of credibility associated with assessing the similarity of other users. Pairs of users with the same number of common items have the same degree of uncertainty.

4.3. Selection of neighbors

In this section, we propose two methods for selecting the most suitable neighbors based on the three pieces of information. These pieces of information are similarity S, dissimilarity $\bar{S}$, and uncertainty U. The aim is to identify neighbors who can provide the best predicted ratings. The suggested methods are:

1. Basic KNN algorithm: typically, the best KNN are those with a high degree of similarity with the target user. Using our novel measure UASim, we can diversify the neighbor selection process. This includes options such as prioritizing neighbors with high similarity scores, selecting those with minimal uncertainty, choosing neighbors with the least dissimilarity, or even identifying those with significant differences between similarity and dissimilarity $(S - \bar{S})$, as well as those with mean values of similarity and dissimilarity $(S, \bar{S})$;

2. TOPSIS neighborhood selection: since every neighbor provides three pieces of information, utilizing TOPSIS in this context transition of neighbor selection from a traditional approach to a robust MCDM paradigm. This methodology allows for a systematic and objective comparison of potential neighbors based on multiple criteria simultaneously, namely: similarity S, dissimilarity $\bar{S}$, and uncertainty U. This inherently adapts neighborhood selection in CF to the complex and often contradictory nature of user preferences.

Employing TOPSIS, a well-established method recognized for its simplicity and effectiveness in ranking alternatives based on multiple criteria, allows us to fully utilize the potential of UASim by incorporating its three key components: similarity S, dissimilarity $\bar{S}$, and uncertainty U. By applying TOPSIS, we systematically optimize our neighborhood selection process, which in turn enhances the accuracy of our predictions and the quality of our recommendations by selecting neighbors who can provide the most relevant ratings, which are essential for generating reliable predictions and consequently recommendations.

In the TOPSIS neighborhood selection strategy, we construct a normalized decision matrix where each row represents a target user's neighbor and each column corresponds to their normalized values of similarity, dissimilarity, and uncertainty, which are the predefined criteria for ranking neighbors. The weights assigned to these criteria are primarily dependent on the provided data. The resulting process ranks potential neighbors according to their proximity to an ideal solution, defined as an abstract neighbor who exhibits maximal similarity, minimal dissimilarity, and least uncertainty.

In the example of applying TOPSIS for neighborhood selection within a collaborative filtering framework, we consider results of Table 4, user 1 (u1) looking to select the most suitable neighbor from Users u2, u3, u4, and u5, based on their UASim scores. We assign weights of 0.7, 0.2, and 0.1 to S, $\bar{S}$, and U respectively. Applying the weights on normalized scores, the ideal solution is identified as the scenario with the highest weighted similarity, lowest weighted dissimilarity, and lowest weighted uncertainty, whereas the anti-ideal solution is the exact opposite. Calculating the Euclidean distance from each user's score to these ideal and anti-ideal points and then determining the relative closeness to the ideal solution, u3 emerges as the most suitable neighbor due to its high similarity and low dissimilarity and uncertainty, closely aligning with the ideal profile, followed by u2, u5, and u4 in descending order of suitability.

To sum up, UASim is a comprehensive similarity measure that can evaluate similarity, dissimilarity, and uncertainty in relation to users' preferences in UBCF. The uncertainty parameter is critical in UASim, as it indicates the credibility of similarity. Uncertainty is primarily determined by the proportion of co-rated items between each pair of users. UASim offers new alternatives for neighborhood selection.

5. EXPERIMENTAL EVALUATION

This section discusses every aspect of the conducted experiments, including the experimental datasets, evaluation metrics, and results. When evaluating RS, the most commonly used method is offline analysis due to its accessibility [9]. This method is based on common practices in machine learning algorithm evaluation. The data is split into two parts: training and testing, and k-fold cross-validation is used for testing, as shown in Figure 1.

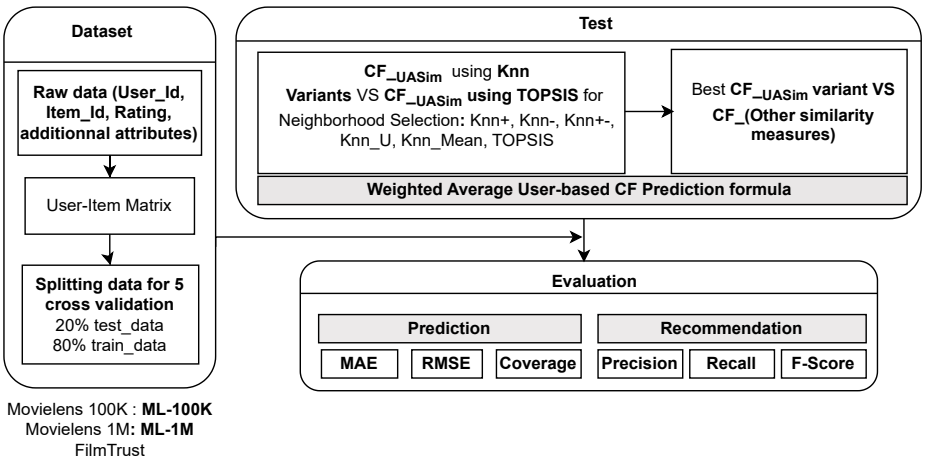

Fig. 1: Overall methodology for the experimental evaluation.

Our research undertakes a rigorous experimental evaluation of memory-based UBCF using UASim. Our tests were carefully organized into two main stages, which can be seen in Figure 1. The purpose of these stages is to compare different neighborhood selection strategies. In the initial phase, we explored multiple neighborhood selection strategies for UBCF using UASim. This stage is crucial for determining the most effective neighborhood selection approach. Advancing to the second phase, the best UBCF using UASim version was benchmarked against UBCF using representative traditional and recent similarity measures [5, 11, 26, 31, 39]. In this research, we used the weighted average prediction formula for UBCF. It is defined as folows:

$$p_{ui} = \overline{r_u} + \frac{\sum_{v \in N_u^i} \mathrm{Sim}_{uv} * (r_{vi} - \overline{r_v})}{\sum_{v \in N_u^i} |\mathrm{Sim}_{uv}|}. \tag{8}$$

In Eq.8, Sim_{uv} is the similarity between users u and v. N_u^i refers to the group of the k-nearest neighbors to the target user u, specifically those who have rated item i. Here, v is one of the users in this nearest neighbors set. Additionally, $\overline{r_u}$ and $\overline{r_v}$ represent the average ratings given by user u and user v, respectively.

In our experimental results, the nomenclature for each UBCF system is designated by the similarity measure it employs. For example, a UBCF utilizing cosine similarity is referred to as 'COS'. This naming convention facilitates a clear and systematic discussion of the various tested methods.

5.1. Datasets

For our experiments, we utilized three benchmark datasets: FilmTrust [16], MovieLens 100K (ML-100K), and MovieLens 1 Million (ML-1M) [17]. Each user in these datasets has rated at least 20 movies. These datasets are popular in the CF domain and are frequently used by academics and developers for their research and applications. The main characteristics of these datasets are summarized in Table 5.

To demonstrate the effectiveness of UASim, each dataset was split into two parts: 20% of all ratings were designated as testing ratings, while the remaining 80% were designated as training ratings. We determined the accuracy of our predictions by performing 5-fold cross-validation, in which we randomly selected 5 different training and test sets.

Dataset	**Ratings**	**Users**	**Items**	**Sparsity**
FilmTrust	35,497	1,508	2,071	98.86%
ML-100K	100,000	943	1,682	93.7%
ML-1M	1,000,209	6,040	3,900	95.8%

Tab. 5: Summary of datasets used in experiments.

5.2. Evaluation metrics

We employ several widely-used evaluation metrics in the research and practice of the RS domain, namely the Mean Absolute Error (MAE), the Root Mean Squared Error (RMSE), F1-score, and coverage. A lower MAE and RMSE value, as well as higher values for F1-score and coverage, are indicative of better algorithmic performance.

The MAE and RMSE metrics are typically used to measure the similarity between the predicted ratings p_{ui} and the actual ratings r_{ui}. They are expressed as follows:

$$\begin{aligned} MAE &= \frac{\sum_{(ui)\in T} |r_{\mathrm{ui}} - p_{\mathrm{ui}}|}{|T|} \\ RMSE &= \sqrt{\frac{\sum_{(ui)\in T} \left(r_{\mathrm{ui}} - p_{\mathrm{ui}}\right)^2}{|T|}}. \end{aligned} \tag{9}$$

Here, T represents the total number of test users involved in the prediction step, r_{ui} is the actual rating provided by user u_i, and p_{ui} is the predicted rating for the same user u_i.

F1-score is a comprehensive metric that combines both Precision and Recall, where Precision and Recall are defined by Equation (10):

$$\text{Precision} = \frac{|L \cap L_{rec}|}{|L_{rec}|}, \quad \text{Recall} = \frac{|L \cap L_{rec}|}{|L|}. \tag{10}$$

Precision represents the proportion of items in the recommendation list L_{rec} that were correctly predicted as liked by the users, compared to the total number of items in L_{rec}. Conversely, Recall measures the proportion of items that were correctly predicted as liked by the users in L_{rec} relative to the total number of liked items in L [39]. According to the approach described in [5, 27], the ratings for items recommended by the system fall within the upper half of the rating scale. For instance, on a scale from 1 to 5, the ratings for recommended items would range from 3 to 5, encompassing the median to the highest possible rating.

The F1-score, which combines precision and recall, is calculated as follows:

$$F1\text{-}score = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}. \tag{11}$$

Coverage, another important metric, measures the proportion of items for which a RS is capable of providing recommendations. It is calculated as the proportion of the number of predicted ratings $|I_{p_u}|$ relative to the total number of ratings to be predicted present in the test set T:

$$\text{Coverage} = \frac{\sum_{u=1}^{m} |I_{p_u}|}{|T|}. \tag{12}$$

5.3. Methods of comparison

We conducted a comparative analysis using UBCF based on both traditional and recent representative predefined similarity measures. Table 6 presents these methods, detailing their formulations and including references for the recent measures.

In this context, I is the set of items rated by both users u and v. The sets of items that user u and user v have rated are called I_u and I_v, respectively. The ratings user u gives to an item i are shown as r_{ui}, and the ratings user v gives to the same item are shown as r_{vi}. The symbols $\bar{r}_u$ and $\bar{r}_v$ represent the average ratings given by user u and user v, respectively. Additionally, N refers to the set of items rated by only one of the two users, and r_m stands for the median of all ratings.

Category	Method	Formula
Traditional	**COS**	$\mathrm{COS}(u,v) = \frac{\sum_{i \in I} r_{ui} r_{vi}}{\sqrt{\sum_{i \in I_u} r_{ui}^2} \sqrt{\sum_{i \in I_v} r_{vi}^2}}$
	MSD	$\mathrm{MSD}(u,v) = 1 - \frac{\sum_{i \in I} (r_{ui} - r_{vi})^2}{\lvert I \rvert}$
	JMSD	$\mathrm{JMSD}(u,v) = \mathrm{JACC}(u,v) \cdot \mathrm{MSD}(u,v)$ where $\mathrm{JACC}(u,v) = \frac{\lvert I_u \cap I_v \rvert}{\lvert I_u \cup I_v \rvert}$
	SPCC	$\mathrm{SPCC}(u,v) = \mathrm{PCC}(u,v) \cdot \frac{1}{1 + \exp(-\frac{\lvert I \rvert}{2})}$ where $\mathrm{PCC}(u,v) = \frac{\sum_{i \in I} (r_{ui} - \bar{r}_u)(r_{vi} - \bar{r}_v)}{\sqrt{\sum_{i \in I} (r_{ui} - \bar{r}_u)^2} \sqrt{\sum_{i \in I} (r_{vi} - \bar{r}_v)^2}}$
Recent	**RACF** [41]	$\mathrm{Sim}(u,v) = \frac{\sum_{i \in I} \frac{\min(r_{ui}, r_{vi})}{\max(r_{ui}, r_{vi})}}{\lvert I \rvert}$
	OS [15]	$\mathrm{sim}_{uv}^{OS} = \mathrm{sim}_{uv}^{PNCR} \cdot \mathrm{sim}_{uv}^{ADF}$ where $\mathrm{sim}_{uv}^{PNCR} = \exp\left(-\frac{N - \lvert I_u \cap I_v \rvert}{N}\right)$ and $\mathrm{sim}_{uv}^{ADF} = \frac{\sum_{i \in I} \exp\left(-\frac{\lvert r_{ui} - r_{vi} \rvert}{\max\{r_{ui}, r_{vi}\}}\right)}{\lvert I_u \cap I_v \rvert}$
	TAN [5]	$\mathrm{TAN}(u,v) = \mathrm{TA}(u,v)$ $u \cdot v = \sum_{i \in I_1 \cap I_2} (r_{ui} - r_m)(r_{vi} - r_m), where$ $\lvert u \rvert = \sqrt{\sum_{i \in I_1 \cap I_2} (r_{ui} - r_m)^2}, and \quad \lvert v \rvert = \sqrt{\sum_{iin I_1 \cap I_2} (r_{vi} - r_m)^2}$

Tab. 6: Summary of methods of comparison for the experimental evaluation.

5.4. Results and discussion

The experimental tests were structured into two main phases. Initially, we focused on identifying best neighborhood selection strategy utilizing UASim within UBCF framework. This was followed by a comparative analysis of UASim against both traditional and recent similarity measures within the framework of UBCF. We experimented with different neighborhood sizes for prediction, specifically k = [40, 50, 60, 70, 80, 90, 100, 120, 140, 160, 180, 200].

5.4.1. Neighborhood selection strategies

To assess the efficacy of different neighborhood selection strategies within the UBCF framework employing UASim, an ablation study was conducted on the TOPSIS variants. This study was crucial for identifying the most effective variant, which was then included in the comprehensive comparison of selection strategies.

Extensive experiments were conducted using five-fold cross-validation across different neighborhood sizes on three datasets. This approach helped determine the best weights for the criteria employed in our TOPSIS variants. The average performance across different neighborhood sizes is summarized in Table 7. In the TOPSIS variant, weights were set as follows: for FilmTrust, weights were set at [0.3, 0.4, 0.3] for similarity, dissimilarity, and uncertainty respectively; for ML-100K and ML-1M, the weights were [0.7, 0.2, 0.1]. For the TOPSIS_Sim_Dissim variant, which considers only similarity and dissimilarity, weights were [0.7, 0.3]

for FilmTrust and [0.8, 0.2] for both ML-100K and ML-1M. Similarly, in the TOPSIS_Sim_Uncertainty variant, where only similarity and uncertainty are considered, the same weighting scheme was applied.

Dataset	TOPSIS		TOPSIS_Sim_Dissim		TOPSIS_Sim_Uncertainty	
	MAE	RMSE	MAE	RMSE	MAE	RMSE
FilmTrust	**0.6055**	**0.7951**	0.6077	0.7991	0.6088	0.8008
ML-100K	**0.7447**	**0.9484**	0.7501	0.9573	0.7491	0.9561
ML-1M	**0.7320**	**0.9372**	0.7539	0.9609	0.7576	0.9654

Tab. 7: Average MAE and RMSE performance of TOPSIS variants across FilmTrust, ML-100K and ML-1M datasets.

The experiments demonstrated that the model using all three criteria (similarity, dissimilarity, and uncertainty) achieves the best performance, as indicated by the values in Table 7, which are **bolded and underlined**. This model not only aligns with theoretical expectations but also empirically outperforms models with two criteria.

The average MAE and RMSE values for different neighborhood sizes, as mentionned above, reported in Table 8 reflect the performance of various neighborhood selection strategies implemented within UBCF framework using UASim. The strategies include: Knn+, which selects neighbors based on the highest similarity values; Knn-, which opts for neighbors with the lowest dissimilarity values; Knn+-, focusing on minimizing the difference between similarity and dissimilarity; Knn_U, which prioritizes minimal uncertainty in neighbor selection; Knn_Mean, that considers an average of similarity and dissimilarity for neighbor selection; and TOPSIS, applying the TOPSIS method that integrates simultanuously similarity, dissimilarity, and uncertainty.

Dataset	Knn+		Knn-		Knn+-		Knn_U		Knn_Mean		TOPSIS	
	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE
FilmTrust	0.6083	0.8000	0.7053	0.9175	**0.6067**	**0.7966**	0.6932	0.8895	0.7811	1.1029	**0.6055**	**0.7951**
ML-100K	**0.7497**	**0.9569**	0.8738	1.1128	0.7610	0.9708	0.7528	0.9632	0.9823	1.2471	**0.7447**	**0.9484**
ML-1M	0.7559	0.9633	0.9426	1.1959	0.7900	1.0060	**0.7343**	**0.9395**	1.0072	1.2805	**0.7320**	**0.9372**

Tab. 8: Knn Variants VS TOPSIS for neighborhood selection in UBCF using UASim.

Performance rankings in the table are indicated by the formatting: **bold and underlined** for the best values, **bold** for the second-best, and underlined for the third-best. The TOPSIS neighborhood selection strategy consistently showed the best performance across all three datasets. The optimal weight for each experiment was determined empirically based on the best performance outcomes observed; they are detailed in Table 9. The top three performing variants for each dataset are depicted in Figure 2 for FilmTrust, Figure 3 for ML-100K, and Figure 4 for ML-1M.

Dataset	Similarity weight	Dissimilarity weight	Uncertainty weight
FilmTrust	0.3	0.4	0.3
ML-100K	0.7	0.2	0.1
ML-1M	0.7	0.2	0.1

Tab. 9: TOPSIS weights for experimental evaluation.

Figure 2 presents a comparison of neighborhood selection strategies using MAE and RMSE metrics across various neighborhood sizes for the FilmTrust dataset. Three strategies are compared: Knn+, Knn+-, and TOPSIS. In terms of MAE, all strategies exhibit slight increase as neighborhood size expands, suggesting a decrease in prediction accuracy with larger neighborhoods. Notably, TOPSIS outperforms the other methods for neighborhood sizes up to 130, after which its performance converges with that of Knn+-. For RMSE, a similar trend is observed, with a gradual increase in error rates as the neighborhood size increases. TOPSIS demonstrates superior performance across all neighborhood sizes, particularly evident at k=130 where it closely aligns with Knn+-.

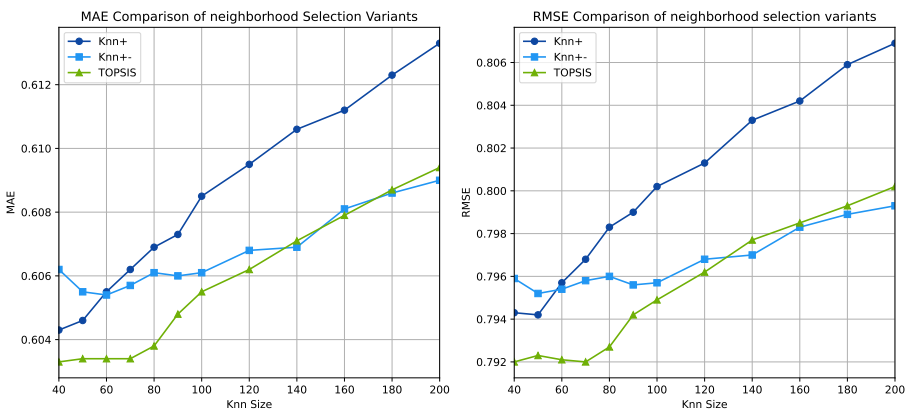

Fig. 2: Performance comparison of neighborhood selection variants for UBCF based on UASim for FilmTrust dataset.

Figure 3 presents the MAE and RMSE variation for three neighborhood selection strategies, namely: Knn+, Knn_U, and TOPSIS across various neighborhood sizes for the ML-100K dataset. The figure demonstrates a decline in both MAE and RMSE with increasing neighborhood sizes, suggesting enhanced prediction accuracy as more neighbors are included. Notably, TOPSIS consistently achieves the lowest error rates, underscoring its effectiveness in balancing similarity, dissimilarity, and uncertainty.

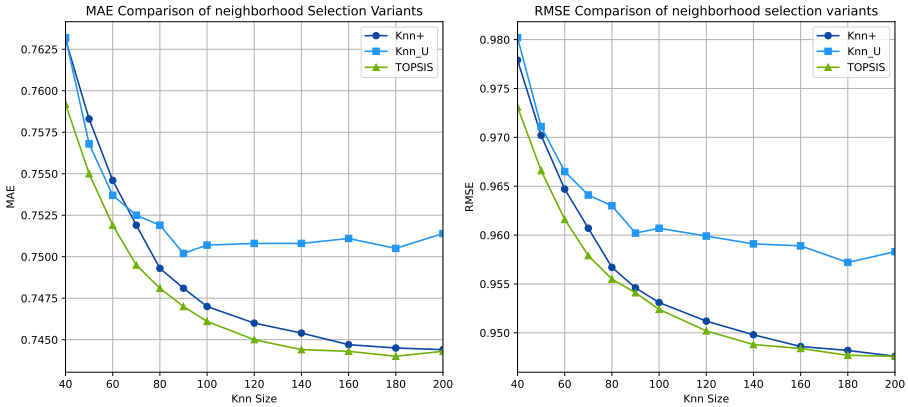

Fig. 3: Performance comparison of neighborhood selection variants for UBCF based on UASim for ML-100K dataset.

Figure 4 showcases the variations in MAE and RMSE across different neighborhood sizes for the ML-1M dataset, comparing three neighborhood selection strategies: Knn+, Knn_U, and TOPSIS. The graph reveals a general trend of decreasing MAE and RMSE as the neighborhood size increases, suggesting improved prediction accuracy with a larger pool of neighbors. Notably, TOPSIS exhibits consistently lower error rates.

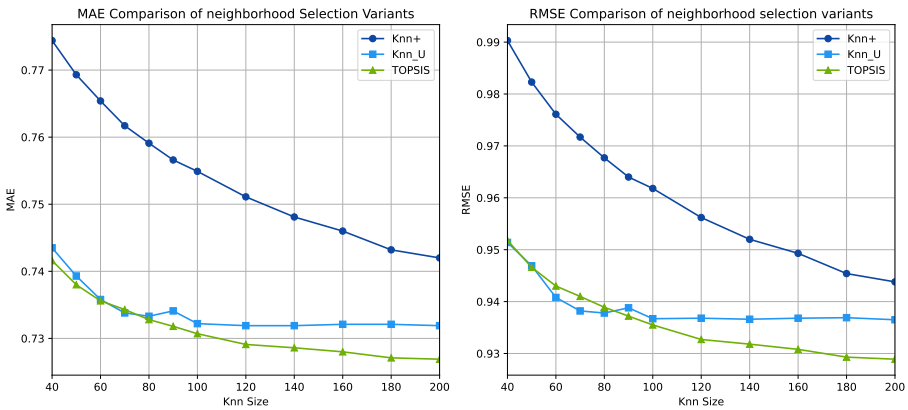

Fig. 4: Performance comparison of neighborhood selection variants for UBCF based on UASim for ML-1M dataset.

Based on the results obtained for the different neighborhood selection strategies across the three datasets, the TOPSIS strategy effectively employs the dis-

criminative power of the UASim measure, which distinctly separates similarity from dissimilarity while incorporating uncertainty. Consequently, it yields the best results in predicting user ratings. In the second part of the experimental evaluation, we will compare the UBCF based on UASim and using TOPSIS for neighborhood selection, designed as "UASim_TOPSIS", with various other UBCF that utilize representative traditional and recent predefined similarity measures.

5.4.2. UASim VS other similarity measures

In terms of MAE and RMSE metrics, in both Table 10 and Table 11, the UBCF using UASim_TOPSIS consistently demonstrates superior performance when compared against UBCF using both traditional and recent similarity measures. UASim_TOPSIS achieves the lowest values in both MAE and RMSE metrics across all considered datasets, namely: FilmTrust, ML-100K, and ML-1M. This trend is consistent across the entire spectrum of neighborhood sizes, ranging from k=40 to k=180. Furthermore, Table 11 quantitatively underscores the improvement percentage of UASim_TOPSIS in relation to other methods [7]. These figures further confirm the position of UASim_TOPSIS as a highly effective approach, yielding significant enhancements in recommendation accuracy, as evidenced by its consistent outperformance across varying neighborhood sizes and datasets.

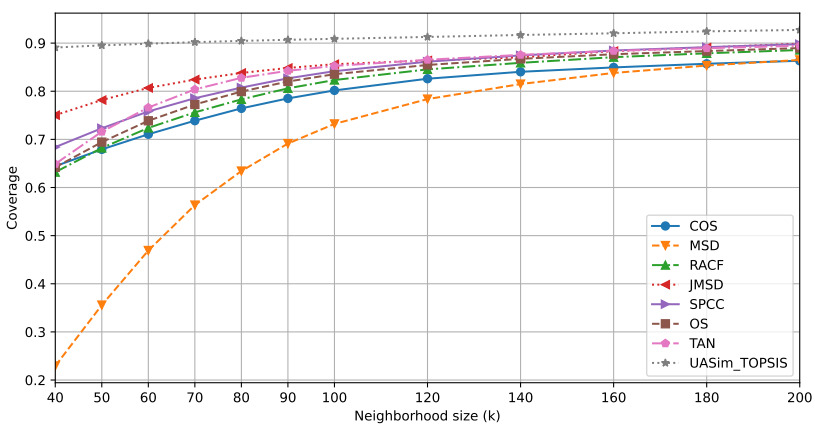

Fig. 5: Coverage results for the FimTrust dataset.

Method	Dataset	k=40		k=50		k=60		k=70		k=80		k=90	
		MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE
COS	FilmTrust	0.7187	0.9212	0.7098	0.911	0.702	0.9025	0.694	0.8926	0.686	0.8823	0.6778	0.8729
	ML-100K	0.9389	1.2037	0.9198	1.1773	0.9034	1.1569	0.887	1.1351	0.8725	1.1156	0.8554	1.0937
	ML-1M	0.9768	1.2453	0.9687	1.2349	0.9608	1.2257	0.9529	1.2166	0.9443	1.2053	0.9369	1.1959
MSD	FilmTrust	0.7928	1.005	0.7665	0.9768	0.7302	0.9305	0.7062	0.9059	0.6874	0.8839	0.6726	0.8671
	ML-100K	0.8975	1.1485	0.8805	1.1271	0.8651	1.106	0.8546	1.0923	0.8417	1.0739	0.8303	1.0587
	ML-1M	0.9647	1.2132	0.9559	1.2024	0.9441	1.1892	0.9376	1.1825	0.9291	1.1713	0.9242	1.167
JMSD	FilmTrust	0.7067	0.9187	0.7182	0.9271	0.7099	0.9168	0.6993	0.9029	0.6853	0.8849	0.6744	0.8749
	ML-100K	0.9415	1.2041	0.9212	1.1777	0.9092	1.1613	0.9004	1.1495	0.8932	1.1409	0.8837	1.1277
	ML-1M	0.8251	1.0508	0.8138	1.0365	0.8046	1.0244	0.7971	1.0147	0.7923	1.0084	0.7886	1.004
SPCC	FilmTrust	0.6869	0.8939	0.6791	0.8849	0.672	0.8773	0.6629	0.8667	0.6575	0.8593	0.652	0.8524
	ML-100K	0.933	1.194	0.9139	1.1701	0.8943	1.1439	0.8732	1.1183	0.8554	1.0949	0.842	1.0782
	ML-1M	0.9647	1.2322	0.9544	1.2183	0.9487	1.21	0.9418	1.1998	0.9342	1.1897	0.9272	1.1814
RACF	FilmTrust	0.6921	0.8964	0.6823	0.8834	0.6731	0.8718	0.6639	0.8631	0.6555	0.852	0.6484	0.8427
	ML-100K	0.8972	1.1433	0.8764	1.116	0.86	1.0944	0.8455	1.0757	0.8325	1.059	0.823	1.0464
	ML-1M	0.9609	1.2202	0.9527	1.2093	0.9441	1.1985	0.9368	1.1903	0.9304	1.1809	0.9236	1.1711
OS	FilmTrust	0.6866	0.8921	0.6756	0.877	0.6649	0.8631	0.6568	0.8552	0.6462	0.8406	0.6398	0.8324
	ML-100K	0.8365	1.0732	0.8198	1.0503	0.8058	1.0309	0.795	1.0163	0.7871	1.0058	0.78	0.9962
	ML-1M	0.8835	1.1326	0.8793	1.126	0.8724	1.1174	0.8678	1.1107	0.8638	1.1052	0.8578	1.0971
TAN	FilmTrust	0.6907	0.8936	0.6704	0.8681	0.6564	0.8517	0.6443	0.8379	0.6351	0.8262	0.6284	0.8199
	ML-100K	0.8642	1.1026	0.8418	1.0731	0.8245	1.0504	0.811	1.0331	0.8023	1.0226	0.7943	1.0119
	ML-1M	0.9507	1.2087	0.9401	1.1936	0.9322	1.1823	0.9235	1.1719	0.9155	1.162	0.9076	1.1516
UASim _TOPSIS	FilmTrust	**0.6033**	**0.792**	**0.6034**	**0.7923**	**0.6034**	**0.7921**	**0.6034**	**0.792**	**0.6038**	**0.7927**	**0.6048**	**0.7942**
	ML-100K	**0.7532**	**0.9659**	**0.7503**	**0.9612**	**0.7482**	**0.958**	**0.7465**	**0.9555**	**0.7456**	**0.954**	**0.7449**	**0.9523**
	ML-1M	**0.7416**	**0.9518**	**0.738**	**0.9466**	**0.7356**	**0.943**	**0.7343**	**0.941**	**0.7328**	**0.9389**	**0.7318**	**0.9372**

MAE and RMSE of different similarity measures for different neighborhood sizes.

Method	Dataset	k=100		k=120		k=140		k=160		k=180		Mean	
		MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE
COS	FilmTrust	0.6711	0.8652	0.6575	0.8493	0.6456	0.8352	0.636	0.825	0.6292	0.8177	0.6752	0,8704
	ML-100K	0.8423	1.0774	0.8176	1.0453	0.7981	1.0199	0.7831	1.0007	0.7719	0.9866	0.8536	1.0920
	ML-1M	0.9309	1.1872	0.9228	1.176	0.9132	1.164	0.9041	1.152	0.8952	1.1397	0.9369	1.1947
MSD	FilmTrust	0.6613	0.8554	0.6468	0.8384	0.638	0.8275	0.6296	0.8176	0.6242	0.8118	0.6868	0.8836
	ML-100K	0.8208	1.0461	0.8028	1.0229	0.7888	1.0051	0.7789	0.9919	0.7729	0.9841	0.8303	1.0596
	ML-1M	0.9176	1.1596	0.9078	1.1453	0.8965	1.1321	0.8889	1.1236	0.8818	1.1151	0.9225	1.1637
JMSD	FilmTrust	0.6722	0.8711	0.6619	0.8584	0.6536	0.85	0.654	0.8508	0.6481	0.8452	0.6803	0.8818
	ML-100K	0.8757	1.117	0.8632	1.1008	0.8534	1.0881	0.843	1.0738	0.8341	1.0625	0.8835	1.1275
	ML-1M	0.785	0.9999	0.7787	0.9916	0.775	0.9867	0.7705	0.9801	0.7674	0.9758	0.79073	1.0066
SPCC	FilmTrust	0.6455	0.8449	0.6342	0.83	0.6256	0.8198	0.619	0.8129	0.6153	0.809	0.65	0.8500
	ML-100K	0.8279	1.0599	0.8063	1.0317	0.7905	1.0106	0.7768	0.9921	0.7677	0.9809	0.8437	1.0795
	ML-1M	0.9224	1.1751	0.9112	1.1602	0.9008	1.1477	0.8893	1.1326	0.8799	1.1207	0.9249	1.1788
RACF	FilmTrust	0.644	0.8365	0.6333	0.823	0.6269	0.8154	0.6231	0.8119	0.621	0.8096	0.6512	0.8459
	ML-100K	0.8136	1.0346	0.798	1.0138	0.7857	0.9982	0.7773	0.9873	0.7712	0.9802	0.8254	1.0499
	ML-1M	0.918	1.1643	0.909	1.1507	0.8995	1.1388	0.8913	1.1283	0.8823	1.1163	0,9226	1,1698
OS	FilmTrust	0.6347	0.8261	0.6263	0.8157	0.6221	0.8106	0.6184	0.8074	0.6163	0.8051	0.6443	0.8386
	ML-100K	0.7748	0.9889	0.7666	0.9774	0.7605	0.9688	0.7573	0.9649	0.7543	0.9606	0,7852	1,0030
	ML-1M	0.8533	1.0905	0.846	1.0797	0.8385	1.0694	0.8313	1.0591	0.8235	1.0484	0.8561	1.0941
TAN	FilmTrust	0.6234	0.8137	0.6184	0.8081	0.6169	0.8064	0.6154	0.8043	0.6145	0.8028	0.6376	0.8302
	ML-100K	0.7886	1.0046	0.7799	0.992	0.7736	0.9844	0.7678	0.9763	0.7643	0.9718	0.8011	1.020
	ML-1M	0.8988	1.1398	0.8868	1.1239	0.8747	1.1093	0.8623	1.0936	0.8529	1.0817	0.9041	1.1471
UASim _TOPSIS	FilmTrust	**0.6055**	**0.7949**	**0.6062**	**0.7962**	**0.6071**	**0.7977**	**0.6079**	**0.7985**	**0.6087**	**0.7993**	**0.6052**	**0.7947**
	ML-100K	**0.7445**	**0.9512**	**0.7437**	**0.9492**	**0.7438**	**0.9487**	**0.7437**	**0.948**	**0.744**	**0.948**	**0.7462**	**0.9538**
	ML-1M	**0.7307**	**0.9355**	**0.7291**	**0.9327**	**0.7286**	**0.9318**	**0.728**	**0.9308**	**0.7271**	**0.9293**	**0.7325**	**0.9380**

Tab. 10: MAE and RMSE of different similarity measures for different neighborhood sizes.

Metric	Dataset	COS	MSD	JMSD	SPCC	RACF	OS	TAN
MAE	FilmTrust	10.37%	11.89%	11.04%	6.89%	7.06%	6.07%	5.08%
	ML-100K	12.58%	10.13%	15.54%	11.56%	9.60%	4.97%	6.85%
	ML-1M	21.82%	20.60%	7.36%	20.81%	20.60%	14.44%	18.98%
RMSE	FilmTrust	8.70%	10.06%	9.88%	6.51%	6.06%	5.24%	4.28%
	ML-100K	12.66%	9.99%	15.41%	11.64%	9.15%	4.91%	6.51%
	ML-1M	21.49%	19.39%	6.81%	20.40%	19.82%	14.27%	18.23%

Tab. 11: Percentage of improvement in MAE and RMSE for UASim_TOPSIS compared to different methods.

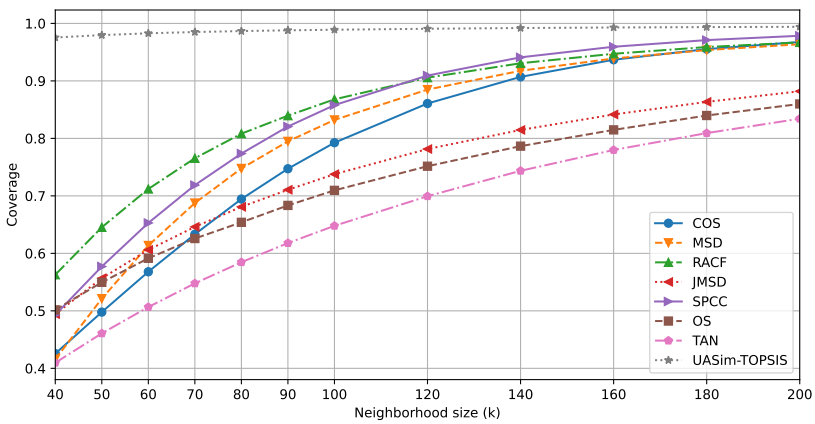

Fig. 6: Coverage results for the ML-100K dataset.

Figures 5, 6, and 7 depict the coverage analysis across the three datasets: FilmTrust, ML-100K, and ML-1M. Notably, UASim_TOPSIS stands out in all datasets, exhibiting the highest coverage and demonstrating its comprehensive ability to offer accurate recommendations across a wide array of items. In the FilmTrust dataset, JMSD and SPCC also show strong coverage, indicating their suitability for this particular dataset. However, MSD's coverage remains notably lower, suggesting a more selective recommendation scope. In the ML-100K dataset (Figure 6), while UASim_TOPSIS maintains its leading position, the SPCC and OS methods show significant improvement. Conversely, TAN and COS, despite their moderate performance, still contribute valuable insights. The ML-1M dataset (Figure 7), being the largest, underscores the dominance of UASim_TOPSIS with a very hight coverage, closely followed by JMSD. This dataset also exposes the most pronounced disparities in performance, with MSD trailing significantly. Overall, UASim_TOPSIS consistently offers broad coverage in the three experimental datasets.

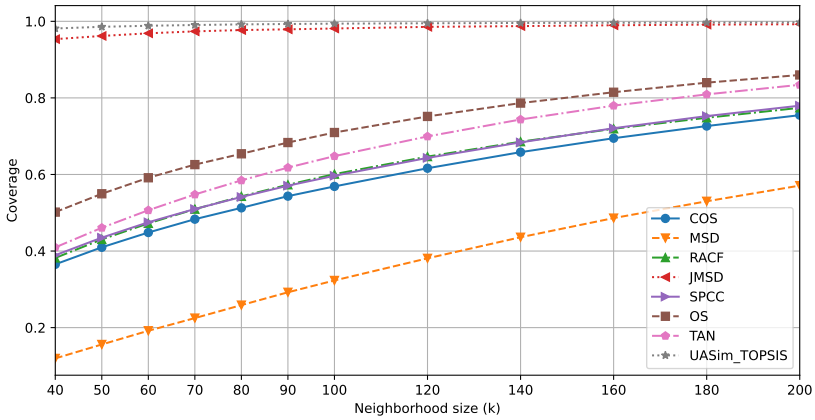

Fig. 7: Coverage results for the ML-1M dataset.

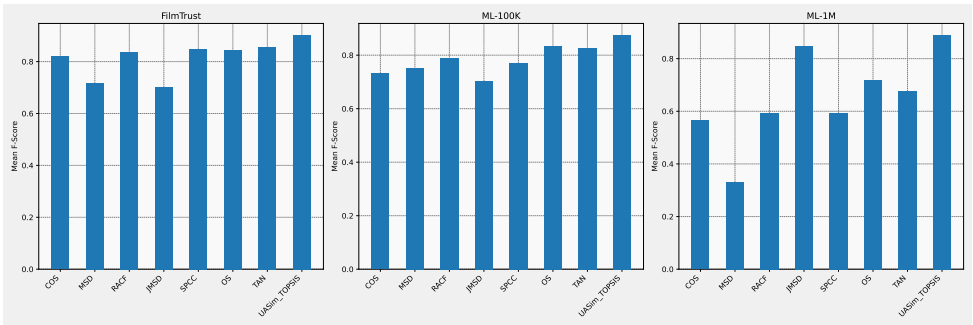

Fig. 8: Average F-score results for the FimTrust, ML-100K and ML-1M datasets.

Analyzing Figure 8 that shows the F1-scores across three datasets, our proposed method UASim_TOPSIS consistently emerges as the top-performing method, achieving the highest scores in the FilmTrust, ML-100K, and ML-1M datasets. This underlines its robustness and effectiveness in balancing precision and recall. Other methods like TAN, SPCC, and OS also demonstrate good performance. Notably, JMSD shows a significant improvement in the ML-1M dataset compared to others. In contrast, methods like COS and MSD exhibit varying degrees of effectiveness, with particularly lower scores in the ML-1M dataset.

5.4.3. Discussion

Our UASim measure employs subjective logic to effectively model user interactions within CF systems, capturing not only similarity and dissimilarity but also the inherent uncertainty in user preferences. Unlike traditional similarity measures, UASim excels in differentiating between various aspects of user

preferences, which significantly enhances the adaptability and accuracy of neighborhood selection strategies. This capability is particularly showcased in our implementation of the TOPSIS-based strategy for neighborhood selection, which systematically integrates similarity, dissimilarity, and uncertainty to select the most suitable neighbors.

UASim stands out as a ratio-based method that takes into account both the proportion of common ratings and individual user preferences. This is evident from our experimental results, where UASim surpasses the performance of established ratio-based similarity measures like RACF. This confirms the efficacy of UASim in selecting more appropriate neighborhoods by simultaneously considering the similarity, dissimilarity, and uncertainty aspects of user preferences.

However, the dependency of our approach on precise parameter tuning within the TOPSIS framework poses a limitation, as the performance significantly depends on the specific weights assigned to similarity, dissimilarity, and uncertainty, which are currently determined empirically based on the best performance outcomes observed. To address this, future research should explore adaptive methods for weight determination, aiming to enhance the robustness and applicability of UASim in diverse CF scenarios.

Regarding the computational efficiency of our proposed method, it is governed by two primary components: the similarity measure computation and the TOPSIS-based neighborhood selection process. Our similarity measure involves identifying co-rated items and calculating their similarity and dissimilarity, which leads to a computational complexity of $O(n)$ per user pair, where n is the number of items co-rated by both users. This complexity is comparable to traditional methods such as COS and PCC, which also operate with $O(n)$ complexity. The TOPSIS method employed for neighborhood selection introduces further steps such as: weight application, ideal solution determinations, and distance calculations, resulting in a complexity of $O(m \times c)$, where m is the number of users and c is the number of criteria. Although this is higher than basic KNN algorithm, the additional computational overhead is justified by the enhanced decision-making capability provided by considering multiple criteria, which is especially valuable in handling the complex dynamics of user preferences in recommender systems.

In summary, UASim provides a nuanced and sophisticated similarity measure for UBCF, marking a significant advancement over ratio-based similarity measures. The success of UASim in our experiments highlights its potential to refine recommendation systems further, offering a promising direction for subsequent research in this field.

6. CONCLUSION

In this study, we introduced a new Uncertainty-Aware Similarity (UASim) measure, which effectively integrates similarity, dissimilarity, and uncertainty under the framework of subjective logic.

UASim provides various strategies for neighborhood selection within the UBCF framework. Among these methods, the TOPSIS strategy proved to

be the most effective, consistently surpassing other approaches across multiple datasets, including FilmTrust, ML-100K, and ML-1M. The superiority of TOPSIS was clear for both MAE and RMSE metrics, underscoring its ability to leverage the comprehensive capabilities of UASim to effectively select the most appropriate neighbors. Our comparative analysis further highlighted the robustness of UASim_TOPSIS against both traditional and recent predefined similarity measures, demonstrating its significant potential to enhance prediction accuracy and recommendation quality.

Looking forward, the reliance of UASim_TOPSIS on precise parameter tuning within the TOPSIS method presents an essential area for future research. Developing adaptive methods for weight determination could substantially improve the robustness and applicability of our approach across various collaborative filtering scenarios, offering promising directions for enhancing the performance of recommendation systems.

SOURCE CODE AVAILABILITY

The codes used in this study are available upon request from the corresponding author.

ACKNOWLEDGEMENTS

We express our gratitude to Editor in Chief, Prof. Sergej Čelikovský, and the anonymous reviewers for their dedication and insightful feedback on our manuscript.

(Received February 4, 2024)

REFERENCES

[1] Ch. C. Aggarwal et al.: Recommender systems, volume 1. 2016.

[2] F. J. Aherne, N. A. Thacker, and P. I. Rockett: The bhattacharyya metric as an absolute similarity measure for frequency coded data. Kybernetika *34* (1998), 4, 363–368. DOI:10.1016/S0932-4739(98)80004-7

[3] H. J. Ahn: A new similarity measure for collaborative filtering to alleviate the new user cold-starting problem. Inform. Sci. *178* (2008), 1, 37–51. DOI:10.1016/j.ins.2007.07.024

[4] H. Al-Bashiri, M. A. Abdulgabber, A. Romli, and H. Kahtan: An improved memory-based collaborative filtering method based on the topsis technique. PloS one *13* (2018), 10. e0204434. DOI:10.1371/journal.pone.0204434

[5] A. A Amer, H. I. Abdalla, and L. Nguyen: Enhancing recommendation systems performance using highly-effective similarity measures. Knowledge-Based Systems *217* (2021), 106842. DOI:10.1016/j.knosys.2021.106842

[6] P. B. Anand and R. Nath: Content-based recommender systems. In: Recommender System with Machine Learning and Artificial Intelligence: Practical Tools and Applications in Medical, Agricultural and Other Industries. 2020, pp. 165–195.

[7] Y. Ar, Ş. E. Amrahov, N. A. Gasilov, and S. Y.-Sert: A new curve fitting based rating prediction algorithm for recommender systems. Kybernetika *58* (2022), 3, 440–455. DOI:10.14736/kyb-2022-3-0440

[8] K. Belmessous, F. Sebbak, A. Batouche, et al.: Co-rating aware evidential user-based collaborative filtering recommender system. In: International Conference on Computing Systems and Applications, Springer 2022, 51–60. DOI:10.1007/978-3-031-12097-8_5

[9] M. Chen and P. Liu: Performance evaluation of recommender systems. Int. J. Performability Engrg. *13* (2017), 8, 1246. DOI:10.1055/s-0036-1591686

[10] R. K. Dewi, A. W. Widodo, Y. A. Sari, and N. I. M. Aziz: Rank consistency of topsis in mobile based recommendation system. In: Proc. 5th International Conference on Sustainable Information Engineering and Technology, ACM Digital Library 2020, pp. 107–112.

[11] J. Feng, X. Fengs, N. Zhang, and J. Peng: An improved collaborative filtering method based on similarity. PLoS One *13* (2018), 9, e0204003. DOI:10.1371/journal.pone.0204003

[12] Fethi Fkih: Similarity measures for collaborative filtering-based recommender systems: Review and experimental comparison. Journal of King Saud University-Computer and Information Sciences, 2021.

[13] S. Forouzandeh, M. Rostami, and K. Berahmand: A hybrid method for recommendation systems based on tourism with an evolutionary algorithm and topsis model. Fuzzy Inform. Engrg. *14* (2022), 1, 26–50. DOI:10.1080/16168658.2021.2019430

[14] D. Gavalas, Ch. Konstantopoulos, K. Mastakas, and G. Pantziou: Mobile recommender systems in tourism. J. Network Computer Appl. *39* (2014), 319–333. DOI:10.1016/j.jnca.2013.04.006

[15] A. Gazdar and L. Hidri: A new similarity measure for collaborative filtering based recommender systems. Knowledge-Based Syst. *188* (2020), 105058. DOI:10.1016/j.knosys.2019.105058

[16] G. Guo, J. Zhang, and N. Yorke-Smith: A novel bayesian similarity measure for recommender systems. In ACM, editor, *Proceedings of the 23rd International Joint Conference on Artificial Intelligence (IJCAI)*, pages 2619–2625, 2013.

[17] F. M. Harper and J. A. Konstan: The movielens datasets: History and context. ACM Trans. Int. Intell. Systems (TIIS) *5* (2015), 4, 1–19.

[18] Ch. L. Hwang and K. Yoon: Multiple Attribute Decision Making: Methods and Applications A State-Of-The-Art Survey, volume 186. Springer Science Business Media, 2012.

[19] N. Idrissi and A. Zellou: A systematic literature review of sparsity issues in recommender systems. Social Network Anal. Mining *10* (2020), 1, 1–23.

[20] A. Jøsang: A logic for uncertain probabilities. Int. J. Uncertainty, Fuzziness Knowledge-Based Syst. *9* (2001), 3, 279–311. DOI:10.1016/S0218-4885(01)00083-1

[21] A. Jøsang: Subjective Logic, volume 4. 2016.

[22] M. Karimi, D. Jannach, and M. Jugovac: News recommender systems–survey and roads ahead. Inform. Process. Management *54* (2018), 6, 1203–1227. DOI:10.1016/j.ipm.2018.04.008

[23] H. Khojamli and J. Razmara: Survey of similarity functions on neighborhood-based collaborative filtering. Expert Syst. Appl. *185* (2021), 115482. DOI:10.1016/j.eswa.2021.115482

[24] K. Kim: A new similarity measure to increase coverage of rating predictions for collaborative filtering. Appl. Intell. *53* (2023), 23, 28804–28818. DOI:10.1007/s10489-023-05041-1

[25] V. L. Chetana and H. Seetha: Handling massive sparse data in recommendation systems. J. Inform. Knowledge Management (2024), 2450021. DOI:10.1142/s0219649224500217

[26] H. Liu, Z. Hu, A. Mian, H. Tian, and X. Zhu: A new user similarity model to improve the accuracy of collaborative filtering. Knowledge-Based Syst. *56* (2014), 156–166. DOI:10.1016/j.knosys.2013.11.006

[27] S. Manochandar and M. Punniyamoorthy: A new user similarity measure in a new prediction model for collaborative filtering. Appl. Intell. *51* (2021), 1, 586–615. DOI:10.1007/s10489-020-01811-3

[28] M. Mataoui, F. Sebbak, A. H. Sidhoum, T. E. Harbi, M. R. Senouci, and K. Belmessous: A hybrid recommendation system for researchgate academic social network. Social Network Anal. Mining *13* (2023), 1. 53. DOI:10.1007/s13278-023-01056-1

[29] D. L. Olson: Comparison of weights in topsis models. Math. Comput. Modell. *40* (2004), 7–8, 721–727. DOI:10.1016/j.mcm.2004.10.003

[30] H. Papadakis, An. Papagrigoriou, C. Panagiotakis, E. Kosmas, and P. Fragopoulou: Collaborative filtering recommender systems taxonomy. Knowledge Inform. Syst. *64* (2022), 1, 35–74. DOI:10.1007/s10115-021-01628-7

[31] B. K. Patra, R. Launonen, V. Ollikainen, and S. Nandi: A new similarity measure using bhattacharyya coefficient for collaborative filtering in sparse data. Knowledge-Based Syst. *82* (2015), 163–177. DOI:10.1016/j.knosys.2015.03.001

[32] F. Ricci, L. Rokach, and B. Shapira: Context-aware recommender systems: recommender systems handbook. In: Recommender Systems Handbook, Springer, 2011, pp. 217–253.

[33] D. Roy and M. Dutta: A systematic review and research perspective on recommender systems. J. Big Data *9* (2022), 1, 59. DOI:10.1186/s40537-022-00592-5

[34] P. Sánchez and A. Bellogín: Building user profiles based on sequences for content and collaborative filtering. Inform. Process. Management *56* (2019), 1, 192–211. DOI:10.1016/j.ipm.2018.10.003

[35] R. Seth and A. Sharaff: A comparative overview of hybrid recommender systems: Review, challenges, and prospects. Data Mining Machine Learning Appl. (2022), 57–98. DOI:10.1002/9781119792529.ch3

[36] M. Shojaei and H. Saneifar: MFSR: A novel multi-level fuzzy similarity measure for recommender systems. Expert Systems Appl. *177* (2021), 114969. DOI:10.1016/j.eswa.2021.114969

[37] D. Valcarce, J. Parapar, and Á. Barreiro: Finding and analysing good neighbourhoods to improve collaborative filtering. Knowledge-Based Syst. *159* (2018), 193–202. DOI:10.1016/j.knosys.2018.06.030

[38] D. Wang, Y. Yih, and M. Ventresca: Improving neighbor-based collaborative filtering by using a hybrid similarity measurement. Expert Syst. Appl. *160* (2020), 113651. DOI:10.1016/j.eswa.2020.113651

[39] Y. Wang, J. Deng, J. Gao, and P. Zhang: A hybrid user similarity model for collaborative filtering. Inform. Sci. *418* (2017), 102–118. DOI:10.1016/j.ins.2017.08.008

[40] Y. Wang, P. Wang, Z. Liu, and L. Y. Zhang: A new item similarity based on α-divergence for collaborative filtering in sparse data. Expert Syst. Appl. *166* (2021), 114074. DOI:10.1016/j.eswa.2020.114074

[41] X. Wu, B. Cheng, and J. Chen: Collaborative filtering service recommendation based on a novel similarity computation method. IEEE Trans. Services Comput. *10* (2015), 3, 352–365. DOI:10.1109/tsc.2015.2479228

[42] X. Wu, Y. Huang, and S. Wang: A new similarity computation method in collaborative filtering-based recommendation system. In: 2017 IEEE 86th Vehicular Technology Conference (VTC-Fall), IEEE, 2017, pp. 1–5. DOI:10.1109/vtcfall.2017.8288359

Khadidja Belmessous, Ecole Militaire Polytechnique, P. O. Box 17, Bordj El Bahri, 16111 Algiers. Algeria
e-mail: kh.belmessous@gmail.com

Faouzi Sebbak, Ecole Militaire Polytechnique, P. O. Box 17, Bordj El Bahri, 16111, Algiers. Algeria.
e-mail: Faouzi.SEBBAK@emp.mdn.dz

M'hamed Mataoui, Ecole Militaire Polytechnique, P. O. Box 17, Bordj El Bahri, 16111, Algiers. Algeria.
e-mail: Mhamed.MATAOUI@emp.mdn.dz

Walid Cherifi, Ecole Militaire Polytechnique, P. O. Box 17, Bordj El Bahri, 16111, Algiers. Algeria.
e-mail: wa.cherifi@gmail.com