Jordanian Journal of Computers and Information Technology (JJCIT), Vol. 11, No. 01, March 2025.

A NOVEL EVIDENTIAL COLLABORATIVE FILTERING FRAMEWORK BASED ON DISCOUNTING CONFLICTING PREFERENCES

Khadidja Belmessous, Faouzi Sebbak and M'hamed Mataoui

(Received: 20-Jun.-2024, Revised: 2-Sep.-2024 and 24-Sep.-2024, Accepted: 27-Sep.-2024)

ABSTRACT

This paper presents a novel framework to enhance Evidential Collaborative Filtering (ECF), a critical Recommender System (RS) designed for sensitive domains like healthcare and target tracking. The focus is on refining how user-rating imperfections are handled, particularly in managing conflicting preferences during neighborhood selection to boost recommendation quality. The newly proposed ECF architecture integrates a two-probabilities-focused approach with an advanced conflict-management technique, employing Deng relative entropy and the Best Worst Method. This allows for assigning more accurate reliability weights to each user, improving preference selection and rating prediction in ECF. Experimental evaluations on Movielens-100K and Flixster datasets show that our framework surpasses baselines in prediction error, precision, recall and F-score.

KEYWORDS

Recommender systems, Collaborative filtering, Dempster-Shafer theory, Conflict, Fusion.

1. INTRODUCTION

Recommender systems (RSs) have been categorized in the literature into three main approaches; namely, content-based filtering (CBF) [1], collaborative filtering (CF) [2], and hybrid filtering [3]. CBF provides recommendations based on user profiles, which are generally difficult to acquire. On the other hand, CF generates recommendations by using the preferences of the most similar users. Hybrid filtering is a combination of both CBF and CF. 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 [4].

Collaborative filtering is a leading approach in RS, based on the idea that our purchase decisions are usually influenced by our similar neighbors. Sparsity is a key challenge in CF, representing the proportion of missing ratings to the overall rating-matrix size. In CF, the subjective nature of user ratings and their intrinsic sparsity not only increase the uncertainty, but also affect the trustworthiness of the recommendation outputs [5]-[6]. Evidential Collaborative Filtering (ECF) is a sub-class of CF that addresses the sparsity issue by handling the inherent uncertainty in RS under the framework of Dempster-Shafer Theory (DST), also called evidence theory [7]-[8]. ECF can be categorized into three main types [9]: ECF using evidential fusion to combine multi-source information, ECF offering soft ratings and ECF providing evidential predictions.

This paper primarily focuses on a specific type of ECF that utilizes soft rating systems. This ECF addresses the limitations of traditional hard-rating scores in capturing user preferences, which can sometimes be an inadequate representation [10]. For instance, consider a user who rates two items, i_1 and i_2, with scores of 3 and 4, respectively. If this user wants to rate a third item, i_3, as better than i_1, but not as good as i_2, standard rating scales might not accurately reflect this nuanced preference. The ECF framework discussed here allows for more flexible user ratings, like a range of {4,5}, to better capture these subtle preferences. Essentially, this branch of ECF is designed to account for the subjective and sometimes imprecise nature of user preferences [11].

Imperfections and conflicts in user preferences negatively affect the trustworthiness and effectiveness of ECF systems [5]. These imperfections can arise due to several factors, including uncertainty, ambiguity and contradictions in user feedback. Existing ECF frameworks rely mainly on the use of

Kh. Belmessous (Corresponding Author), F. Sebbak and M. Mataoui are with Department of Computer Science, Ecole Militaire Polytechnique, Bordj El Bahri, P.O.Box 17, Algiers, Algeria. Emails: kh.belmessous@gmail.com, faouzi.sebbak@emp.mdn.dz and mhamed.mataoui@emp.mdn.dz

Dempster's combination rule (DCR) in combining users' preferences [9]. Nguyen and Huynh explored the fusion of information in RS using DST, as detailed in [12]. They concluded that DCR is ineffective for combining user ratings due to its weakness to handle highly conflicting mass functions [13]. Recently, Belmessous et al. [9] highlighted shortcomings in the existing ECF framework, which often overlooks the importance of managing these conflicting preferences through advanced techniques. This paper addresses this gap by proposing a novel framework that integrates recent advancements in DST to better manage conflict, thereby enhancing the overall performance of ECF systems.

Research on ECF has provided limited solutions for managing conflicting user preferences [9], falling behind in the ongoing advancements within DST research. DST is continually evolving to tackle challenges related to dealing with highly conflicting information [13]. Many studies in ECF overlook the discounting factor, which is key in DST for determining the reliability of user ratings. Nguyen and Huynh [12] have explored the integration of information in RS, highlighting the difficulty in combining mass functions that are highly conflicting. In RS, it's quite common to encounter users giving completely opposite ratings to the same item, which leads to frequent conflicts in mass combination.

Although DST research is continually proposing new solutions for conflict management [14], ECF has not fully capitalized on these advancements. The ongoing challenge in DST of effectively combining highly conflicting evidence remains a significant issue. This gap in ECF [15]-[16], where advanced DST conflict-management proposals are underutilized, is a key focus of this paper. We intend to bridge this gap by incorporating recent solutions for conflict management [17] into ECF systems to enhance their performance.

This paper presents the following key contributions:

- Introduction of a novel framework designed to manage imperfections in users' preferences throughout the decision-making process in ECF;
- Proposition of a new neighbourhood-selection strategy in ECF, utilizing optimal discounting weights;
- Proposition of an efficient method for preference-prediction estimation in ECF.

The remainder of this article is structured as follows: In Section 2, we provide a summary of the theoretical concepts underlying our new approach. We then outline our proposed framework and its main components in Section 3. Section 4 describes our experimental design and presents the obtained results. Section 5 includes a discussion of our findings, strengths and limitations. The article concludes with Section 6, where we summarize our work and suggest areas for future research.

2. BACKGROUND AND RELATED WORK

In this section, we explore foundational theories and contributions important to our study, with a particular focus on the Dempster-Shafer Theory (DST) and conflict management. Initially, we introduce DST, known for its ability to manage uncertainty and make decisions based on evidence. This theory is valuable in decision-making, where the quality of information is crucial. Subsequently, we discuss the metrics used to evaluate conflict within this theoretical framework. Additionally, we explore the conflict-management methodology based on discounting optimal weights and present the related research on ECF offering soft ratings.

2.1 Dempster-Shafer Theory

The Dempster-Shafer theory is a flexible method for modeling uncertainty that does not require assigning a probability to every element in a set. The DST was introduced by Arthur P.Dempster in the context of statistical inference [18], and it was further developed by his student Shafer [19].

DST is founded on a number of concepts, including: the frame of discernment, the mass function also called basic probability assignment (BPA) and Dempster's combination rule. Concerning the frame of discernment Θ, it is a finite set representing the problem domain. All propositions of interest are defined by elements in 2^{Θ}. A BPA is defined as a mapping $m(.) \in [0,1]$ that meets the following properties:

$$m(\emptyset) = 0 \qquad \emptyset: \textit{the empty set}$$

Jordanian Journal of Computers and Information Technology (JJCIT), Vol. 11, No. 01, March 2025.

$$\sum_{H \in 2^{\Theta}} m(H) = 1 \qquad H\text{: a subset of } \Theta.$$

The quantity $m(H)$ can be interpreted as a measure of the belief that is committed exactly to H, given the available evidence. All subsets $H \in 2^{\Theta}$ having a positive mass are considered as focal elements of $m(.)$. Concerning the Dempster's combination rule, it is an operation that permits to combine evidence from multiple independent sources under the same frame of discernment. Let m_1 and m_2 be the two BPAs associated with two independent sources of evidence. H_1 and H_2 are the focal elements of m_1 and m_2, respectively. The resulting mass function m is the combination of m_1 and m_2 and is noted by $m = m_1 \oplus m_2$. The Dempster's combination rule (DCR) is defined by Equations 1 and 2, where m_{DS} is the result of Dempster's combination.

$$m_{DS}(H) = \frac{m_{12}(H)}{1 - m_{12}(\emptyset)} \tag{1}$$

where,

$$m_{12}(H) = \sum_{\substack{H_1, H_2 \in 2^{\Theta} \\ H_1 \cap H_2 = H}} m_1(H_1) m_2(H_2) \tag{2}$$

The body of evidence in DST encompasses all BPAs from independent sources, serving as the aggregate evidence for decision-making. It forms the basis for applying Dempster's combination rule, enabling the synthesis of evidence across the problem domain.

In the DST framework, decision criteria can include: the maximum of the belief *Bel*(H), which indicates the comprehensive support that evidence lends to a hypothesis *H*; the maximum of the plausibility *Pl(H)*, reflecting the extent to which evidence does not contradict *H*; or the pignistic probability *BetP(H)* [20], which provides a practical way to make decisions under uncertainty by balancing the evidence supporting different hypotheses. The relationship between belief and plausibility is illustrated in Figure 1.

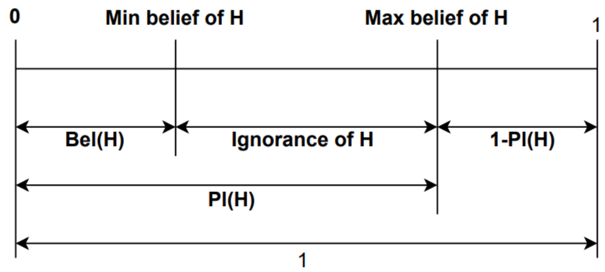

Figure 1. Relationship between belief and plausibility.

Another important tool in the DST framework is the discounting factor proposed by Shafer [19]. The factor α is considered as a discounting rate permitting to control the reliability of the BPA. When α is set to 1, the BPA is deemed fully reliable; conversely, an α value of 0 signifies that the BPA is entirely unreliable. The discounting of the BPA $m(\cdot)$ is defined as follow:

$$\begin{Bmatrix} m'(H) = \alpha . m(H), \qquad \forall H \in 2^{\Theta}, H \neq \Theta \\ m'(\Theta) = (1 - \alpha) + \alpha . m(\Theta) \end{Bmatrix} \tag{3}$$

with $m'(.)$ representing the unreliable source.

2.2 Conflict Metrics in Dempster-Shafer Theory

In scenarios where Dempster's combination rule is applied to fuse evidence from multiple sources, it's possible to reach counter-intuitive conclusions, especially when the evidence conflicts significantly [13], [21]. Consistently, there are novel propositions being introduced for conflict metrics to enhance the accuracy of assessing conflict levels of evidence. Consider two BPAs, m_1 and m_2, defined under the Frame of Discernment (FoD) $H = \{H_1, H_2, \dots, H_i, \dots, H_n\}$. Some representative metrics for evaluating conflict are summarized in Table 1.

In Jousselme et al.'s distance equation, $\widehat{m}_1$ and $\widehat{m}_2$ represent the vector forms of the basic probability assignments m_1 and m_2, respectively and $\overline{D}$ is the Jaccard matrix between all pairwise propositions in m1 and m2. An increased distance in this measure indicates a higher level of conflict among the evidence. Similarly, in Song et al.'s correlation-coefficient equation, $\widehat{m}_1 = m_1.D$ and $\widehat{m}_2 = m_2.D$ are used, where D is the Jaccard matrix applicable to all propositions in m_1 and m_2. In Jiang's correlation-coefficient equation, H_i and H_j serve as focal elements within the power concentration of the frame and their relationship is quantified through a modulus calculation that involves the intersection and union of H_i and H_j.

Table 1. Summary of some representative conflict metrics for DST.

Metric Name	Equation
Jousselme et al.'s evidence distance d [22]	$d(m_1, m_2) = \sqrt{\frac{1}{2}(\widehat{m}_1 - \widehat{m}_2)^T \overline{D}(\widehat{m}_1 - \widehat{m}_2)}$
Song et al.'s correlation coefficient cor [23]	$K_{cor}(m_1, m_2) = 1 - cor(m_1, m_2)$, $cor(m_1, m_2) = \frac{\langle \widehat{m}_1, \widehat{m}_2 \rangle}{\lVert \widehat{m}_1 . \widehat{m}_2 \rVert}$
Jiang's correlation coefficient k_r [24]	$k_r(m_1, m_2) = 1 - \sum_{i=1}^{2^{\lvert n \rvert}} \sum_{j=1}^{2^{\lvert n \rvert}} m_1(H_{i)}\, m_2(H_{j)} \frac{\lvert H_i \cap H_j \rvert}{\lvert H_i \cup H_j \rvert}$
Xiao et al.'s correlation coefficient ECC [25]	$k_{ECC}(m_1, m_2) = 1 - ECC(m_1, m_2)) = 1 - \left[\frac{\langle \widehat{m}_1, \widehat{m}_2 \rangle}{\lVert \widehat{m}_1 . \widehat{m}_2 \rVert}\right]^2$

These methods provide different approaches to understand and quantify the level of agreement or conflict between various BPAs, each with its unique application and implications for decision-making.

2.3 Conflict Management by Considering the Optimal Discounting Weights Using the BWM Method

This sub-section introduces the conflict-management method by considering the optimal discounting weights based on the Best-Worst Method (BWM) [26] to manage evidential conflict in DST. This recent methodology involves selecting the best and worst BPAs to calculate discount weights effectively before the fusion process. The detailed steps of this method are outlined as follows:

1) Evidential distance-matrix establishment: an evidential distance matrix is calculated using Jousselme's distance measure to evaluate the distances between each pair of evidence, helping identify the relative degrees of conflict.
2) Determination of worst and best BPAs:
   - The worst BPA, represents the maximum contribution to overall system conflict.
   - The best BPA is determined based on its relative distance to the worst BPA.
3) Preference calculation for best and worst BPAs: Fei and Deng [27] introduced a new metric called Deng relative entropy to measure the discrepancy between BPAs. Deng relative entropy, as described by formula 4, is specifically designed for mass functions in the context of DST.

   $$r(m_1 \Vert m_2) = \sum_i m_1(L_i) log \frac{m_1(L_i)}{m_2(L_i)} \tag{4}$$

   Deng relative entropy calculates the average logarithmic difference between two BPAs, m1 and m2, thus providing a measure of the informational divergence between them.

   Establishing preference vectors for the best and worst BPAs: by utilizing Deng relative entropy, the preference vector for the best BPA, denoted by m_B, relative to other BPAs is calculated as follows:

   $$M_B = (\sigma(m_B \Vert m_1), \sigma(m_B \Vert m_2), \dots, \sigma(m_B \Vert m_n)) \tag{5}$$

   where $\sigma(m_B \Vert m_j)$ quantifies the relative preference of the best BPA m_B over other BPA j. It is defined such that $\sigma(m_B \Vert m_B) = 1$, indicating the highest self-preference. Similarly, the preference vector for the worst BPA, m_W in relation to other BPAs is given by:

$$M_W = (\sigma(m_1\|m_W), \sigma(m_2\|m_W), \dots, \sigma(m_n\|m_W))^T \quad (6)$$

In this vector, $\sigma(m_{Bi}\|m_W)$ measures the preference of each BPA mi over the worst BPA m_W. This measurement also adheres to the condition $\sigma(m_W\|m_W) = 1$, reflecting maximum self-preference and its role as the most conflict-contributing BPA.

4) Finding the optimal weights for BPAs:

In this phase, optimal weights $(\widetilde{w}_1, \widetilde{w}_2, \dots, \widetilde{w}_n)$ are determined to refine the evidential contributions more effectively. The Consistency Ratio (CR) ζ^* plays an essential role in this process by measuring the consistency of these weights, which is pivotal for evaluating the pairwise comparison's efficacy. Divergences in the expected proportional relationships, such as when $\frac{w_B}{w_j} \neq \sigma\ (m_B\|m_j)$ or $\frac{w_j}{w_W} \neq \sigma\ (m_{Bj}\|m_W)$, necessitate a re-evaluation of CR to ensure the reliability of the weight assignments.

$$\min \max_j \left\{ \left| \frac{w_B}{w_j} - m_{Bj} \right|, \left| \frac{w_j}{w_W} - m_{jW} \right| \right\}$$

$$s.t. \begin{cases} \sum_{\substack{w_j \geq o \\ j=\{1,2,\dots,n\}}}^{\min \zeta} {}_j \quad w_j = 1 \end{cases} \Rightarrow s.t. \begin{cases} \left| \frac{w_B}{w_j} - \sigma(m_B\|m_j \right| \leq \xi \\ \left| \frac{w_j}{w_W} - \sigma(m_j\|m_W \right| \leq \xi \\ \sum_j w_j = 1 \\ w_j \geq 0 \\ j = \{1,2,\dots,n\} \end{cases} \quad (7)$$

Therefore, the optimal weights $(\widetilde{w}_1, \widetilde{w}_2, \dots, \widetilde{w}_n)$ and CR (ζ_*) of BPAs could be calculated. Meanwhile, the CR can be defined as follows:

$$\eta_{CR} = \frac{\zeta_*}{max\{\zeta\}} \quad (8)$$

5) Discounting and fusion: optimal weights obtained from the previous steps are used to discount the BPAs before fusion using the DCR.

This conflict-management methodology [17] ensures the reliability of the optimal weights by employing a consistency ratio for reference comparisons, guaranteeing that each piece of evidence contributes appropriately to the final fused result.

2.4 Related Work on Evidential Collaborative Filtering Offering Soft Ratings

The pioneering effort in the area of ECF that introduces soft preferences in RS was initiated by Wickramarathne et al. [28]. This approach leverages the DST to effectively handle uncertainties in user preferences for a CF system. Emphasizing prediction accuracy, this evidential RS design accepts higher computational demands. Its sophisticated nature makes it suitable for critical and advanced applications, including those in medical and healthcare services and security-threat evaluations. Subsequently, Nguyen expanded on this foundation by developing an evidential RS that incorporates soft ratings, drawing inspiration from Wickramarathne et al. [28], and tackling the issue of data sparsity by leveraging community context under the DST framework [11]. Nguyen and Huynh further enhanced this system by integrating the reliability of predicted ratings, acknowledging their inherent imprecision compared to real ratings, to refine the recommendations [29]. Later, Nguyen aimed at reducing computational load by proposing an optimization that prioritizes the combination of focal elements with the top two probabilities within their sets [10].

Furthermore, Nguyen et al. extended the application of their ECF to incorporate social-media platforms [30]. In this context, user ratings and community preferences gathered from social networks are represented as mass functions. These are then combined according to Dempster's rule of combination. Moreover, Nguyen and Huynh introduced an innovative approach for combining evidence in their system as described in [29] through [31]. Their technique focuses on discarding focal elements with negligible probabilities, considered as noise in the fusion of information, thereby enhancing the efficiency of computations without sacrificing data integrity. In addition, their research in [12] delves into optimizing evidence combination for DST-based RS. This study establishes the essential parameters for crafting a combination operator that aligns with the requisites of DST-based RS. Within this framework, Nguyen and Huynh unveiled new strategies for executing mixed combinations, showcasing their commitment to refining the DST-based RS.

Nguyen's 2017 study [15] introduced an innovative BPA combination approach named the "Two-probabilities focused-combination method". This method permits to combine belief masses with significant conflicts and offers the advantage of decreased computational time. Although the proposed method is not stable due to the fact that it is non-associative, indicating that the order of inputs can influence the results, the sequence in which inputs are combined has an impact on the outcome. Further, Nguyen and Huynh tackled the challenges of data sparsity and the cold-start problem in [32] through an ECF that incorporates soft ratings alongside community preferences. They also proposed a novel approach for assessing user-user similarity, prioritizing provided over predicted ratings, within a similar system [33]. Dong et al. followed up with a different strategy in [34], introducing the modified rigid coarsening method based on hierarchical decomposition to simplify the frame of discernment in the combination process. Lastly, Bahri et al. presented ECFAR in [16], a rule-based CF system that leverages the DST, marking another contribution to the field.

3. METHODOLOGY

This study presents a novel framework, Conflict-Aware Evidential Collaborative Filtering (CA-ECF), which integrates an advanced conflict-management methodology from recent research [17] into a classical ECF framework [15]. This methodology, aimed to managing evidential conflict within DST, optimizes weights for BPAs using the BWM, as elaborated in sub-section 2.3.

To ensure clarity and consistency of mathematical notations throughout our proposition, we have defined all the used variables and notations in Table 2.

Table 2. Notations' table.

Symbol	Description
R_{MN}	Rating matrix with M and N representing the total number of users and items, respectively. Here, M corresponds to the set of users $U = \{U_1, U_2, \dots, U_M\}$ and N corresponds to the set of items $I = \{I_1, I_2, \dots, I_N\}$.
$\hat{R}_{MN}$	Dense User-Item rating matrix.
Θ	Set of preference levels, denoted by $\Theta = \{\theta_1, \theta_2, \dots, \theta_L\}$, where L is the number of the available preferences.
$r_{i,k}$	Rating of user U_i on item I_k.
C	Set of concepts within the contextual data, denoted by $C = \{C_1, C_2, \dots, C_P\}$, where P is the total number of concepts. Each concept C_p, with $1 \le p \le P$, can consist of at most Q_p groups, indicating that $C_p = \{G_{p,1}, G_{p,2}, \dots, G_{p,Q_p}\}$.
$g_p(U_i)$	Groups within concept C_p that user U_i is interested in.
$g_q(I_k)$	Groups within concept C_q that item I_k is associated with.
$G_{p,q}$	The intersection of user and item interest groups associated with concept C_p.
$m_{i,k}$	BPA corresponding to a rating $r_{i,k}$.
$\uplus$	Two-probabilities focused combination.
$d(U_i, U_j)$	Jousselme's distance measure between users U_i and U_j.
$s(U_i, U_j)$	Similarity score between users U_i and U_j.
m_B, m_W	Best and worst BPAs.
$\sigma(m_1 \Vert m_2)$	Deng relative entropy measuring the conflict between two BPAs m_1 and m_2.
M_B, M_W	Vectors representing the preference of the best BPA and worst BPA over other BPAs using the Deng relative entropy.
N_{U_i}	Set of k closest neighbors for user U_i.
$knn_{i,k}$	Set of neighbors of user U_i that have rated the target item I_k.
ξ	Optimization variable used to minimize the optimal weights.
w_i	Weight assigned to the i^{th} BPA, used in the discounting and fusion processes.
$\hat{r}_{ik}$	Predicted rating.

The proposed CA-ECF framework represents an advanced version of the classical ECF, with its main characteristics detailed in [10][15][28]. CA-ECF innovates by using conflict in user preferences to identify the most similar neighbors. It then adjusts their influence in making predictions based on their optimal weights. The architecture of the proposed framework is depicted in Figure 2.

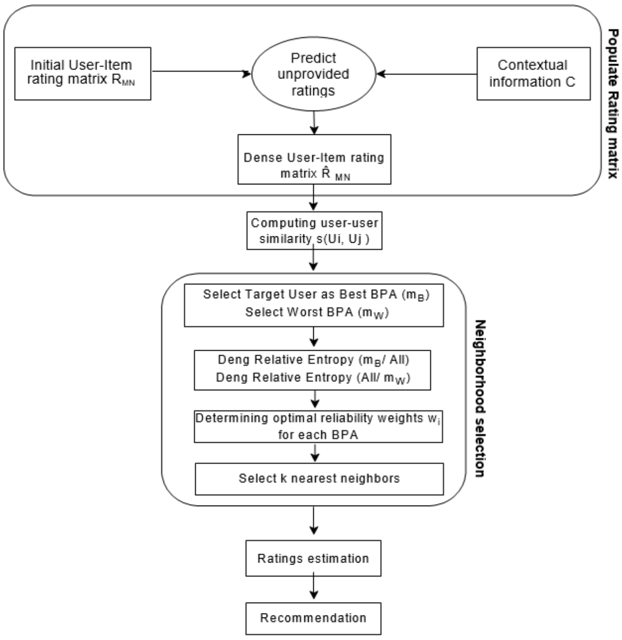

Figure 2. The proposed CA-ECF framework.

The CA-ECF framework, similar to classical ECF, follows five distinct steps, as illustrated in Figure 2. Initially, the unrated entries within the rating matrix R_{MN} are calculated using contextual information C in order to construct a dense rating matrix $\hat{R}_{MN}$. Subsequently, user-user similarities $s(U_i, U_j)$ are calculated using both provided and predicted ratings in $\hat{R}_{MN}$. For each active user U_i, a neighborhood set $knn_{i,k}$ is selected and the user's rating for each item is estimated based on the combined ratings from these selected neighbors. Following this, the estimated ratings for all unrated items are systematically ranked and the most appropriate items are chosen for recommendations to the active user.

In classical ECF, neighborhood sets $knn_{i,k}$ for each unrated item I_k are determined based on similarity scores $s_{i,j}$, which must meet or exceed a specific threshold. This traditional approach, however, does not provide a mechanism to assess the reliability of the selected neighbors. In contrast, the CA-ECF framework selects neighborhood sets based on their corresponding optimal discounting weights. These weights are then utilized to discount the BPAs during the prediction step, thereby refining the accuracy of the recommendations.

In the following sub-sections, we will explore each step of the CA-ECF recommendation process in detail.

3.1 Constructing Dense User-item Rating Matrix

In the classical ECF architecture, each user evaluation is represented as a BPA (m) that spans the evaluation space Θ, enabling it to capture a wide range of user preferences, for instance: uncertain and ambiguous data. The first step of the CA-ECF framework is to predict all the unrated entries $r_{i,k}$ of the user-item matrix using contextual data C in order to mitigate the sparsity issue of CF. Contextual data consists of a set of concepts $C = \{C_1, C_2, \dots, C_P\}$, where each concept p encompasses a set of groups G_p. Both CA-ECF and ECF consider that users who share an interest in a particular group will also have similar choices with respect to that group. The group preference is defined as follows:

- First, consider a concept Cp. For each group $G_{p,q}$ that intersects with $G_p(U_i)$, which is the users' group of interest and $g_q(I_k)$, the items' group of interest, it is assumed that the group's overall preference for item I_k within $G_{p,q}$ reflects the specific group preference of user U_i for item I_k within the same group. Therefore, the concept preference of user U_i for item I_k related to concept C_p is the result of the combination of all the group preferences, represented as two-probabilities focused-mass functions.

- Second, the overall context preferences are computed as the combination of all concept preferences for the target item, represented as two-probabilities focused-mass functions.
- Finally, the unrated entry $r_{i,k}$ is replaced by the context preference of user U_i for item I_k. If the context information does not allow for making conclusions on the concept preference, then the unrated entry is determined by aggregating the ratings from users who have rated item I_k.

At this stage, all the user-item matrix entries $\hat{R}_{MN}$ are given (provided and predicted) and they all will be used in the subsequent steps.

3.2 Computing User-User Similarity

In contrast to the classical ECF systems, in the CA-ECF we propose to evaluate the similarity between users using Jousselme's distance [22], a decision that directly supports the used conflict-management approach [17].

$$d(U_i, U_j) = \sqrt{\frac{1}{2}(\hat{m}_i - \hat{m}_j)^T \overline{D}(\hat{m}_i - \hat{m}_j)} \tag{9}$$

Since ratings have two sources (provided and predicted), we discount the predicted ratings [29].

$$s(U_i, U_j) = \sum_{k=1} \mu(x_{i,k}, x_{j,k}) * d(U_i, U_j) \tag{10}$$

where $\mu(x_{i,k}, x_{j,k})$ is calculated as follows:

$$\mu(x_{i,k}, x_{j,k}) = 1 - w_1(x_{i,k} + x_{j,k}) - w_2 x_{i,k}, x_{j,k}$$

where w_1 and w_2 are the reliability coefficients [29].

User-user similarities are stored as a matrix. The lower the value of $s(U_i, U_j)$ the more similar user U_i is to user U_j.

3.3 Neighborhood Selection

Consider a target user-item pair, (U_i, I_k). We select a set of the k closest neighbors for U_i, denoted by N_{Ui}, by following four steps, as outlined below:

1) Define best and worst BPA: in order to define those two BPAs, we first define the set of neighbors that have rated the target item I_k, following the equation below:

$$knn_{i,k} = \{U_j \in \mathrm{U} \mid I_k \in R(U_j)\} \tag{11}$$

Then, we set best BPA as the target user $m_B = mU_i$. Additionally, the worst BPA can be determined using the best BPA. The exact definition is provided as follows:

$$\boldsymbol{m}_W = \max_i s(\boldsymbol{m}_B, \boldsymbol{m}_i) \tag{12}$$

2) Compute Deng's relative entropy (best/ others) and (others/ worst): Deng relative entropy is given by the following equation:

$$\sigma = (m_1 \| m_2) = \sum_i m_1(L_i) log \frac{m_1(L_i)}{m_2(L_i)} \tag{13}$$

At this stage, in order to compute the reliability factors, two vectors need to be calculated using the $knn_{i,k}$ set of users.

$$M_B = (\sigma(m_B \| m_1), \sigma(m_B \| m_2), ..., \sigma(m_B \| m_n)) \tag{14}$$

which describes the preference of the best BPA m_B over the other BPAs and

$$MW = (\sigma\ (m1 \| mW\ )\ ,\ \sigma\ (m2 \| mW\ )\ ,\ .\ .\ .\ ,\ \sigma\ (mn \| mW\ ))T \tag{15}$$

which describes the preference of BPAs m_i over the worst BPA.

Determining optimal-reliability factors: this step involves determining the optimal weights for BPAs to improve the process of discounting evidence. The consistency ratio (Equation 8) is crucial in this step, as it assesses the consistency of these weights. This step follows a constrained-optimization approach, as formulated in Equation 7, to establish the weights accurately, relying on BWM. By solving an optimization problem that includes non-linear constraints, the optimal weights for the evidence are obtained.

Select k-nearest neighbors: the selection of neighborhoods is based on reliability factors. We order all members within $knn_{i,k}$ in descending order according to their reliability factors, denoted by w_i. Then, the top K members from this ordered list are chosen to form the neighborhood set N_{Ui}.

3.4 Ratings' Estimation

Rating estimation for each unrated item I_k by an active user U_i is computed using the ratings from the user's neighborhood. Ratings are first adjusted by their respective discounting weights according to Equation 3. Then, the two-probabilities-focused method is used to fuse the evidence to obtain the final fusion result. The steps for preference aggregation are outlined in Algorithm .

Algorithm 1. Preference aggregation for rating estimation in CA-ECF.

```
procedure EstimateRating(U_i, I_k, Neighborhoods N_Ui)
    Initialize r̃_i,k ← 0                         ▷ Initialize the estimated rating for item I_k
    for each neighbor U_j ∈ N_Ui do
        r_j,k ← rating of U_j on I_k
        w_i,j ← discounting weight
        r̃_j,k ← discounted BPA according to Equation 3
        r̃_i,k ← r̃_i,k ⊎ r̃_j,k                  ▷ Fusion of discounted BPAs
    end for
    output the estimated rating r̃_i,k
end procedure
```

This algorithm synthesizes the weighted contributions of a user's neighbors to predict unrated items. By applying discounting weights, which are optimized during the neighborhood-selection phase, the reliability of each contribution is assessed, ensuring that the final-rating estimation for $\hat{r}_{i,k}$ is not only a reflection of collective-neighborhood opinion, but also of its credibility and relevance to user U_i's preferences.

3.5 Recommendation

Notably, ECF systems can produce both hard (rating as singleton) and soft (rating as sub-sets) recommendations. For a hard recommendation, the pignistic-probability method is employed to select the item with the highest likelihood as the preferred choice. Conversely, for a soft recommendation, the system adopts a maximum-belief strategy with an overlapping interval approach (maxBL) [15], [35]. This method selects an item based on its belief being greater than the plausibility of any alternative, ensuring that a decision can still be made when a direct class label is absent by favoring a composite class label that combines the most believable item with those of higher plausibility.

4. EXPERIMENTS AND RESULTS

Our experiments were performed on Movielens-100K [36], and Flixster [11] datasets. The MovieLens-100K dataset consists of 943 users who have provided 100,000 ratings for 1,682 movies. The ratings are given on a five-point scale, represented as $\Theta = \{1,2,3,4,5\}$. Each user in this dataset has rated at least 20 movies. On the other hand, Flixster dataset includes 535,013 ratings from 3,827 users for 1,210 movies. The rating scale in this dataset is composed of ten possible scores, denoted as $\Theta = \{0.5,1.0,1.5,2.0,2.5,3.0,3.5,4.0,4.5,5.0\}$. Each user has provided at least 15 ratings.

Moreover, in Movielens-100K, the information used to categorize users is the genre, which has 19 values.

$C_1 = \{G_{1,1}, G_{1,2}, \ldots, G_{1,19}\}$ = {Unknown, Adventure, Action, Animation, Children's, Comedy, Drama, Documentary, Crime, Musical, Film-Noir, Fantasy, Horror, Western, Sci-Fi, Romance, Thriller, War, Mystery}.

$m_{i,k} = \{\alpha_{i,k}(1 - \alpha_{i,k}) \quad for\ A = \theta_l$

$$m_{i,k} = \begin{cases} \alpha_{i,k}(1-\alpha_{i,k}), & for\ A = \theta_l \\ \alpha_{i,k}\sigma_{i,k}, & for\ A = B; \\ 1-\alpha_{i,k} & forA = \Theta; \\ 0, & otherwise \end{cases} \text{ with } B = \begin{cases} (\theta_1, \theta_2), & if\ l = 1; \\ (\theta_{L-1}, \theta_L), & if\ l = L \\ (\theta_{l-1}, \theta_l, \theta_{l+1}), otherwise \end{cases} \quad (16)$$

Movielens-100K was transformed into an evidential dataset using Equation 16 as proposed in [15], where $\alpha_{i,k} \in [0, 1]$ and $\sigma_{i,k}$ are trust actor and dispersion factor, respectively. Also, given the absence of specific information regarding the genres that a user prefers, it is presumed that a user's interest spans all genres associated with any item having been rated.

In the Flixster dataset, every hard rating $r_{i,k}$ was converted into a soft rating $m_{i,k}$ using the Dempster-Shafer modeling function [11], as explained below:

$$m_{i,k}(A) = \begin{cases} \alpha_{i,k}(1-\sigma_{i,k}), & for\ A = \{\theta_l\}; \\ \frac{3}{5}\alpha_{i,k}\sigma_{i,k}, & for\ A = B; \\ \frac{2}{5}\alpha_{i,k}\sigma_{i,k}, & for\ A = C; \\ 1-\alpha_{i,k}, & for\ A = \Theta; \\ 0, & otherwise. \end{cases} \quad (17)$$

where $B = \begin{cases} (\theta_1, \theta_2), & if\ l = 1; \\ (\theta_{L-1}, \theta_L), & if\ l = L; \\ (\theta_{l-1}, \theta_l, \theta_{l+1}), & otherwise; \end{cases}$ and $C = \begin{cases} \{\theta_1, \theta_2, \theta_3\}, & if\ l = 1; \\ \{\theta_1, \theta_2, \theta_3, \theta_4\}, & if\ l = 2; \\ \{\theta_{L-3}, \theta_{L-2}, \theta_{L-1}, \theta_L\}, & if\ l = L-1; \\ \{\theta_{L-2}, \theta_{L-1}, \theta_L\}, & if\ l = L; \\ \{\theta_{l-2}, \theta_{l-1}, \theta_l, \theta_{l+1}, \theta_{l+2}\}, & otherwise; \end{cases}$

The available genres in Flixster dataset are as follows:

Genre = {Drama, Comedy, Action & Adventure, Television, Mystery & Suspense, Horror, Science Fiction & Fantasy, Kids & Family, Art House & International, Romance, Classics, Musical & Performing Arts, Anime & Manga, Animation, Western, Documentary, Special Interest, Sports & Fitness, Cult Movies}.

It's important to highlight that the selection of parameters within these systems is primarily influenced by the outcomes analyzed and reported in the published literature [10], [15].

In our study, the choice of baseline for comparison is carefully considered within the context of ECF offering soft ratings, where diversity in baseline methods is limited. Specifically, we have selected the two- probability-focused ECF [15] as our baseline. This ECF variant has not only performed well in prior studies, but also exceeds the performance of earlier baselines, making it a pertinent choice for comparative analysis. The two-probability-focused ECF represents a more advanced iteration, reflecting both the evolution that addresses conflicting preferences in ECF [9] and the state-of-the-art in ECF research.

Additionally, a 10-fold cross-validation approach was adopted for the experiments. Initially, the ratings within the dataset were divided into 10 distinct folds, with each fold comprising a random selection of 10% of each user's ratings. The experimental process was repeated ten times; during each iteration, one fold was designated as the test dataset, while the other ratings were utilized for training purposes. The mean outcomes from these ten iterations are detailed in the subsequent part of this section.

In the field of ECF offering soft ratings, researchers have developed new evaluation methods capable of assessing their performance. These include DS-MAE, DS-Precision, DS-Recall and DS-Fscore [9], [15], [28]. Let $\hat{r}_{i,k}$be the final estimated rating for user U_i and item I_k and $\widehat{Bp}_{i,k}$ represent the pignistic-probability distribution of the mass function $\hat{r}_{i,k}$. The selected evaluation metrics are defined as follows:

$$DS - MAE(\theta_j) = \frac{1}{|D_j|} \sum_{(i,k)\in D_j, \theta_l \in \Theta} \widehat{Bp_{\iota,k}}(\theta_l)|\theta_j - \theta_i|$$

$$DS - Precision\ (\theta_j) = \frac{TP(\theta_j)}{TP(\theta_j) + FP(\theta_j)}$$

$$DS - Recall\ (\theta_j) = \frac{TP(\theta_j)}{TP(\theta_j) + FN(\theta_j)}$$

$$DS - F_i(\theta_j) = \frac{(i^2+1)(DS - Precision\ (\theta_j)\left(DS - Recall\ (\theta_j)\right)}{i^2(DS - Precision\ (\theta_j) + \left(DS - Recall\ (\theta_j)\right)}$$

where D_j is the test set identifying user-item pairs whose true evaluation is $\theta_j \in \Theta$ and:

$$TP(\theta_j) = \sum_{(i,k)\in D_j} \widehat{Bp_{\iota,k}}\ (\theta_j)$$

$$FP(\theta_j) = \sum_{(i,k)\in D_j, j\neq 1} \widehat{Bp_{\iota,k}}\ (\theta_l)$$

$$FN(\theta_j) = \sum_{(i,k)\in D_j} \widehat{Bp_{\iota,k}}\ (\theta_l)$$

4.1 Results for Movielens-100K Dataset

Tables 3 and 4 provide a comprehensive comparison of the CA-ECF and baseline method across various rating values, assessing their performance through hard metrics, such as MAE, precision, recall and F-score and soft metrics, such as DS-MAE, DS-precision, DS-recall and DS-F-score. The CA-ECF method demonstrates superior precision and recall in both soft and hard recommendations, particularly notable in the middle rating values, where it significantly outperforms the baseline. This trend is consistent across the precision and F-score metrics as well, with CA-ECF showing enhanced accuracy.

Table 3. Comparison in hard decisions for CA-ECF and baseline on Movielens-100K dataset.

Metrics	Rating value					Global
	1	2	3	4	5	
			CA-ECF			
MAE	**2.4011**	**1.5072**	**0.7286**	**0.3495**	**1.0153**	**0.8326**
Precision	**0.182**	0.2208	**0.3221**	**0.3935**	**0.4648**	**0.3752**
Recall	**0.0177**	0.0912	**0.3167**	**0.6657**	**0.1863**	**0.3828**
F-score	**0.0322**	0.129	**0.3193**	**0.4946**	**0.2659**	**0.3789**
			Baseline			
MAE	2.4075	1.5087	0.7382	0.369	1.0157	0.8343
Precision	0.177	**0.2242**	0.3206	0.3919	0.4484	0.3641
Recall	0.0152	**0.0924**	0.3158	0.6642	0.1851	0.3718
F-score	0.0649	**0.1434**	0.3175	0.4923	0.2592	0.3468

Table 4. Comparison in soft decisions for CA-ECF and baseline on Movielens-100K dataset.

DS-Metrics	Rating value					Global
	1	2	3	4	5	
			CA-ECF			
DS-MAE	**2.4057**	**1.4897**	**0.7337**	**0.3702**	**1.0159**	**0.8243**
DS-Precision	**0.1756**	**0.2380**	**0.3191**	**0.3916**	0.4452	**0.36015**
DS-Recall	**0.0159**	**0.0962**	**0.3177**	**0.6612**	0.1787	**0.3722**
DS-F-score	**0.0292**	**0.1370**	**0.3184**	**0.4919**	0.2550	**0.3315**
			Baseline			
DS-MAE	2.4066	1.4918	0.7344	0.3713	1.0175	0.8327
DS-Precision	0.1749	0.2300	0.3175	0.3908	**0.4462**	0.3609
DS-Recall	0.0156	0.0949	0.3164	0.6605	**0.1815**	0.3702
DS-F-score	0.0267	0.1329	0.3161	0.4903	**0.2560**	0.3315

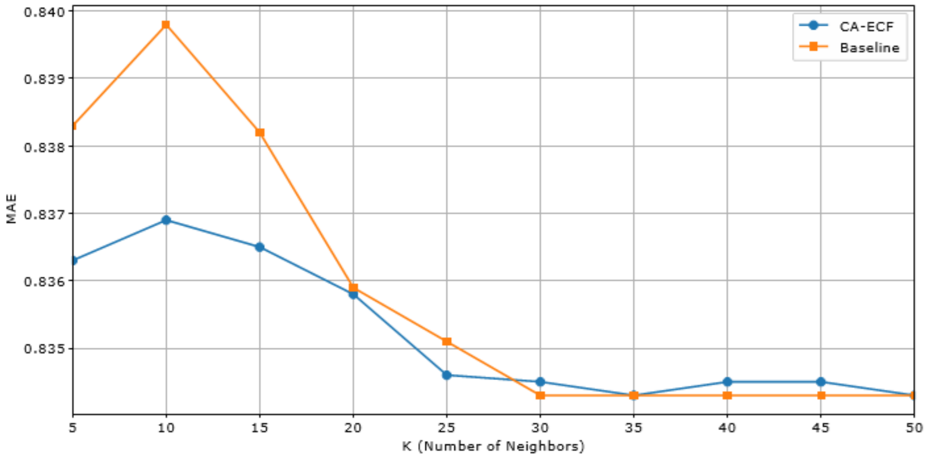

Figure 3. Overall MAE for CA-ECF *versus* baseline on Movielens-100K dataset.

In Figure 3, the data comparing CA-ECF and baseline across varying values of K reveal that both recommendation frameworks show an improvement in DS-MAE as the number of neighbors increases up to K = 20, beyond which the improvement in error rates stabilizes. CA-ECF consistently performs better than baseline at lower values of K, indicating its superior efficiency in scenarios with fewer neighbors. Both methods reach their optimal performance around K = 20. This indicates that increasing K beyond 20 offers no significant benefit, possibly leading to over-specialization and unnecessary computational overhead.

In Figure 4, both CA-ECF and baseline show a trend where the DS-MAE values generally decrease as K increases from 5 to 20. Around K = 20, both CA-ECF and baseline achieve their minimum DS-MAE values, indicating an optimal point for the Movielens dataset. Post this point, both frameworks stabilize, with slight fluctuations in DS-MAE values, suggesting that increasing K beyond this point does not significantly enhance the accuracy. CA-ECF appears to be more robust at lower neighborhood sizes, which could be advantageous in scenarios where the data is sparse or when it is computationally preferable to consider fewer neighbors.

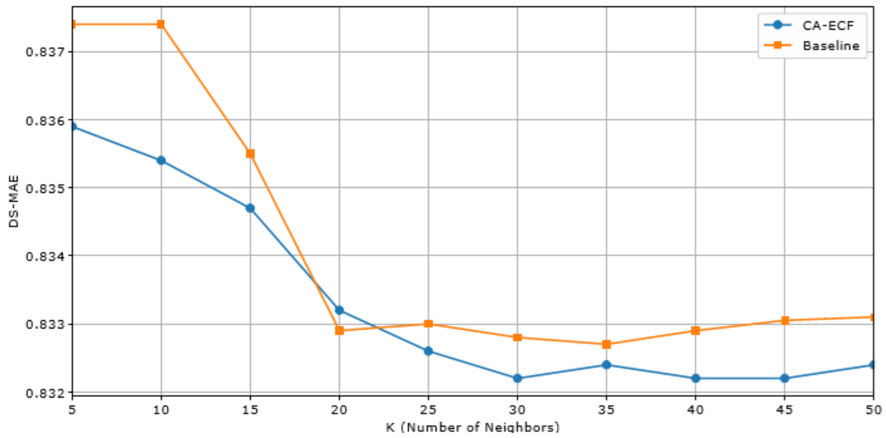

Figure 4. Overall DS-MAE for CA-ECF *versus* baseline on Movielens-100K dataset.

4.2 Results for Flixster Dataset

Based on the presented data from the hard and soft decision comparisons between CA-ECF and the baseline on the Flixster dataset, several insights emerge. As shown in Table 5, for hard decisions, CA-ECF exhibits consistently lower MAE across all rating values compared to the baseline, showcasing its superior accuracy in prediction. Notably, the global MAE for CA-ECF stands at 0.8281, which is lower than the baseline's 0.8503, underscoring the enhanced precision of CA-ECF in handling diverse rating scales from 0.5 to 5.0.

Jordanian Journal of Computers and Information Technology (JJCIT), Vol. 11, No. 01, March 2025.

Table 5. Comparison in hard decisions for CA-ECF and baseline on Flixster dataset.

Metrics	**Rating value**										**Global**
	0.5	**1.0**	**1.5**	**2.0**	**2.5**	**3.0**	**3.5**	**4.0**	**4.5**	**5.0**	
CA-ECF											
MAE	**3.2204**	**2.6843**	**2.1821**	**1.7530**	**1.2887**	**0.7529**	**0.4068**	**0.1341**	**0.5169**	**0.9056**	**0.8281**
Precision	**0.8790**	0	0	**0.1857**	**0.2727**	**0.2283**	**0.1976**	0.2091	**0.1769**	0.3856	**0.2860**
Recall	**0.0252**	0	0	**0.0017**	**0.0036**	**0.0703**	**0.1489**	0.7767	**0.0287**	0.0852	**0.3160**
F-score	**0.0489**	0	0	**0.0033**	**0.0071**	**0.1074**	**0.1698**	0.3294	**0.0493**	0.1395	**0.3002**
Baseline											
MAE	3.2708	2.7865	2.3006	1.7741	1.3163	0.7806	0.4204	0.1360	0.5264	0.9081	0.8503
Precision	0.8521	0	0	0.1697	0.2435	0.1975	0.1886	**0.2150**	0.1747	**0.3921**	0.2404
Recall	0.0242	0	0	0.0015	0.0031	0.0652	0.1478	**0.7812**	0.0280	**0.0867**	0.3114
F-score	0.0470	0	0	0.0029	0.0061	0.0980	0.1657	**0.3371**	0.0482	**0.1420**	0.2713

Similarly, in the soft-decision results of Table 6, CA-ECF maintains its edge over the baseline, with a global DS-MAE of 0.8190 against the baseline's 0.8381. This precision is further reflected in the metrics of DS-Precision, DS-Recall and DS-F-score, where CA-ECF consistently outperforms the baseline across most rating values, particularly in the mid to high range. These metrics confirm the robustness of CA-ECF in synthesizing evidential data to produce reliable and nuanced recommendations, highlighting its applicability in systems where user preferences are particularly conflicting. In the Flixster dataset, ratings of 1.0 and 1.5 are significantly less frequent compared to higher ratings. Consequently, the columns for ratings 1.0 and 1.5 in the comparison tables sometimes show values as zero, indicating sparse data in these categories.

Table 6. Comparison in soft decisions for CA-ECF and baseline on Flixster dataset.

DS-Metrics	**Rating value**										**Global**
	0.5	**1.0**	**1.5**	**2.0**	**2.5**	**3.0**	**3.5**	**4.0**	**4.5**	**5.0**	
CA-ECF											
DS-MAE	**3.2137**	**2.6702**	**2.1787**	**1.7467**	**1.2861**	**0.7491**	**0.4002**	**0.1291**	**0.5126**	**0.9014**	**0.8190**
DS-Precision	**0.8702**	**0**	**0**	**0.1767**	**0.2543**	**0.2056**	**0.1998**	0.2165	**0.2036**	0.3891	**0.2561**
DS-Recall	**0.0282**	0	**0**	**0.0018**	**0.0045**	**0.0710**	**0.1502**	0.7689	**0.0312**	0.0821	**0.3189**
DS-F-score	**0.0546**	**0**	**0**	**0.0035**	**0.0088**	**0.1055**	**0.1714**	0.3378	**0.0541**	0.1355	**0.2840**
Baseline											
DS-MAE	3.2360	2.7653	2.2781	1.7482	1.2909	0.7665	0.4187	0.1322	0.5202	0.9061	0.8381
DS-Precision	0.8562	0	0	0.1710	0.2435	0.1998	0.1906	**0.2172**	0.1872	**0.3987**	0.2468
DS-Recall	0.0253	0	0	0.0016	0.0037	0.0667	0.1491	**0.7851**	0.0290	**0.0892**	0.3114
DS-F-score	0.0491	0	0	0.0031	0.0072	0.1000	0.1673	**0.3402**	0.0502	**0.1457**	0.2753

Figure 5 depicts the MAE performance of CA-ECF and the baseline across varying neighborhood sizes (*K*) on the Flixster dataset. For CA-ECF, there is a consistent enhancement in performance across all K values, showcasing its robustness in managing different neighborhood sizes effectively. In contrast, the baseline exhibits a reduction in MAE as the number of neighbors increases, reaching a plateau at $K = 35$. Beyond this point, no significant gains are observed, indicating that larger neighborhoods do not further contribute to accuracy improvements. This data highlights CA-ECF's superior efficiency, particularly notable at smaller neighborhood sizes.

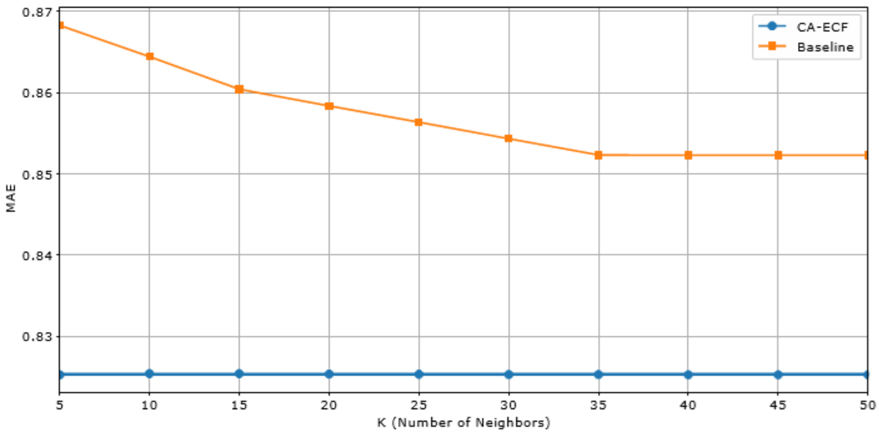

Figure 5. Overall MAE for CA-ECF *versus* baseline on Flixster dataset.

Figure 6 demonstrates the performance trend of CA-ECF and the baseline as the number of neighbors (K) increases within the Flixster dataset. Both frameworks exhibit distinct performance trends. CA-ECF demonstrates stable performance with consistently low DS-MAE values across all K values. In contrast, the baseline framework shows a decrease in DS-MAE from $K = 5$ to $K = 15$, suggesting that accuracy improves with a larger neighborhood up to this point. At $K = 15$, the baseline reaches its lowest DS- MAE, indicating an optimal balance between neighborhood size and predictive accuracy. Beyond $K = 15$, the baseline exhibits negligible improvements and slight fluctuations in DS-MAE, signaling that further increases in K do not yield substantial benefits and may lead to diminishing returns.

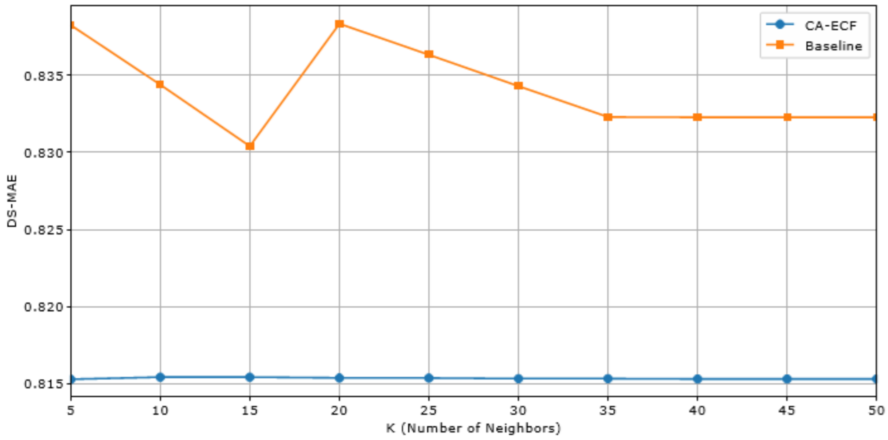

Figure 6. Overall MAE for CA-ECF *versus* baseline on Flixster dataset.

5. Discussion

In the evaluation of the CA-ECF framework, our experiments reveal that the framework demonstrates notably improved performance for rating values that exhibit higher density within the dataset. This enhanced performance can be attributed to the framework's use of reliability factors that judiciously discount ratings. By adjusting the influence of these ratings in the evidential fusion step, the framework not only refines the prediction accuracy, but also effectively manages the inherent uncertainty associated with sparse data. Such a mechanism ensures that the contributions of the neighbors are weighted, which is particularly crucial in sparse datasets where every rating can significantly influence the outcome. This approach underscores the ability of CA-ECF to deliver more reliable and precise recommendations by effectively capturing and utilizing the underlying patterns in user-item interactions.

However, the CA-ECF framework introduces additional computational complexities, primarily from the calculation of Deng's relative entropy and the optimization of reliability factors. The computation of Deng's relative entropy within the neighborhood set $knn_{i,k}$ presents a quadratic complexity, $O(k^2)$, where k represents the number of neighbors who have rated the target item. Further complexity arises during the optimization step to determine optimal reliability factors, potentially extending to $O(k^3)$ depending on the algorithm used. In contrast, classical ECF methodologies typically involve linear operations based on similarity scores, resulting in a lower overall time complexity of $O(n)$. Thus, while the CA-ECF framework incurs a higher computational cost, it leverages this complexity to enhance the accuracy and reliability of recommendations, which is particularly advantageous in applications where the quality of recommendations is critical.

6. Conclusions

This research introduces a novel Conflict-Aware Evidential Collaborative Filtering framework that significantly advances the management of conflicts in user ratings. By integrating a two-probabilities-focused approach with the advanced conflict-resolution technique based on the Best Worst Method, the framework refines the weighting of user preferences. This precision in handling ratings leads to discernibly improved performance across key metrics, including DS-MAE, DS-precision, DS-recall and DS-F-score, outperforming existing methodologies. While our framework enhances recommendation accuracy and reliability, especially in handling uncertain, imprecise or incomplete user preferences, it introduces complexities related to its computational demands. The detailed calculations required by

Dempster-Shafer theory, along with those needed to optimize reliability factors, can lead to increased computational time, potentially limiting its immediate practicality in real-world scenarios. Looking forward, we plan to explore the potential of distributed computing and Monte Carlo approximations to manage the computational overhead effectively. These techniques aim to reduce the computational intensity while maintaining accuracy, offering scalable solutions for large datasets. We are also keen on investigating alternative fusion rules that can further enhance the framework's ability to handle conflicts. These steps are aimed at extending the scalability and practicality of our framework.

DECLARATION OF COMPETING INTERESTS

The authors declare that they don't have any competing financial interests or personal connections that could have influenced the work reported in this paper.

ACKNOWLEDGEMENTS

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

REFERENCES

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

[2] H. Papadakis, A. Papagrigoriou, C. Panagiotakis, E. Kosmas and P. Fragopoulou, "Collaborative Filtering Recommender Systems Taxonomy," Knowledge and Information Systems, vol. 64, no. 1, pp. 35–74, 2022.

[3] R. Seth and A. Sharaff, "A Comparative Overview of Hybrid Recommender Systems: Review, Challenges and Prospects," Chapter 3 in Book: Data Mining and Machine Learning Applications, pp. 57–98, DOI: 10.1002/9781119792529.ch3, 2022.

[4] 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 Systems with Applications, vol. 166, p. 114074, 2021.

[5] R. Abdelkhalek, I. Boukhris and Z. Elouedi, "Towards More Trustworthy Predictions: A Hybrid Evidential Movie Recommender System," JUCS: Journal of Universal Computer Science, vol. 285, no. 10, pp. 1003–1029, 2022.

[6] K. Belmessous, F. Sebbak, M. Mataoui and W. Cherifi, "A New Uncertainty-aware Similarity for User-based Collaborative Filtering," Kybernetika, vol. 60, no. 4, pp. 446–474 , 2024.

[7] N. Idrissi and A. Zellou, "A Systematic Literature Review of Sparsity Issues in Recommender Systems," Social Network Analysis and Mining, vol. 10, no. 1, pp. 1–23, 2020.

[8] K. Belmessous, F. Sebbak, M. Mataoui and A. Batouche, "Co-rating Aware Evidential User-based Collaborative Filtering Recommender System," Proc. of the Int. Conf. on Computing Systems and Applications, Advances in Computing Systems and Applications (CSA 2022), vol. 513, pp. 51–60, Springer, 2022,.

[9] K. Belmessous, F. Sebbak, M. Mataoui, M. R. Senouci and W. Cherifi, "Dempster-Shafer Theory in Recommender Systems: A Survey," Int. Journal of Uncertainty, Fuzziness and Knowledge-based Systems, vol. 32, no. 5, pp. 747–780, 2024.

[10] V.-D. Nguyen and V.-N. Huynh, "Evidence Combination Focusing on Significant Focal Elements for Recommender Systems," Proc. of the Int. Symposium on Integrated Uncertainty in Knowledge Modeling and Decision Making, vol. 9376, pp. 290–302, Springer, 2015.

[11] V.-D. Nguyen and V.-N. Huynh, "A Community-based Collaborative Filtering System Dealing with Sparsity Problem and Data Imperfections," Proc. of Pacific Rim Int. Conf. on Artificial Intelligence, Trends in Artificial Intelligence (PRICAI 2014), vol. 8862, pp. 884–890, Springer, 2014.

[12] V.-D. Nguyen and V.-N. Huynh, "On Information Fusion in Recommender Systems Based on Dempster-Shafer Theory," Proc. of the 2016 IEEE 28th Int. Conf. on Tools with Artificial Intelligence (ICTAI), pp. 78–85, San Jose, CA, USA, 2016.

[13] F. Sebbak, M. R. Senouci, F. Benhammadi et al., "Towards Cardinality-aware Evidential Combination Rules in Dempster–Shafer Theory," KI-Künstliche Intelligenz, pp. 1–16, DOI: 10.1007/s13218-024-00859-4, 2024.

[14] K. Zhao, L. Li, Z. Chen, R. Sun, G. Yuan and J. Li, "A Survey: Optimization and Applications of Evidence Fusion Algorithm Based on Dempster–Shafer Theory," Applied Soft Computing, vol. 124, p. 109075, 2022.
[15] V.-D. Nguyen and V.-N. Huynh, "Two-probabilities Focused Combination in Recommender Systems," Int. Journal of Approximate Reasoning, vol. 80, pp. 225–238, 2017.
[16] N. Bahri, M. A. Bach Tobji and B. Ben Yaghlane, "ECFAR: A Rule-based Collaborative Filtering System Dealing with Evidential Data," Proc. of the Int. Conf. on Intelligent Systems Design and Applications, Intelligent Systems Design and Applications (ISDA 2021), vol. 418, pp. 944–955, Springer, 2022.
[17] L. Zhou, H. Cui, X. Mi, J. Zhang and B. Kang, "A Novel Conflict Management Considering the Optimal Discounting Weights Using the BWM Method in Dempster-Shafer Evidence Theory," Information Sciences, vol. 612, pp. 536–552, 2022.
[18] A. P. Dempster, "Upper and Lower Probabilities Induced by a Multi-valued Mapping," Chapter in Book: Classic Works of the Dempster-Shafer Theory of Belief Functions, pp. 57–72, Springer, 2008.
[19] G. Shafer, A Mathematical Theory of Evidence, vol. 42, DOI: 10.2307/j.ctv10vm1qb, Princeton University Press, 1976.
[20] P. Smets, "The Combination of Evidence in the Transferable Belief Model," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 12, no. 5, pp. 447–458, 1990.
[21] L. A. Zadeh, "A Simple View of the Dempster-Shafer Theory of Evidence and Its Implication for the Rule of Combination," AI Magazine, vol. 7, no. 2, pp. 85–85, 1986.
[22] A.-L. Jousselme, D. Grenier and É. Bossé, "A New Distance between Two Bodies of Evidence," Information Fusion, vol. 2, no. 2, pp. 91–101, 2001.
[23] Y. Song, X. Wang, J. Zhu and L. Lei, "Sensor Dynamic Reliability Evaluation Based on Evidence Theory and Intuitionistic Fuzzy Sets," Applied Intelligence, vol. 48, no. 11, pp. 3950–3962, 2018.
[24] W. Jiang, "A Correlation Coefficient for Belief Functions," Int. Journal of Approximate Reasoning, vol. 103, pp. 94–106, 2018.
[25] F. Xiao, Z. Cao and A. Jolfaei, "A Novel Conflict Measurement in Decision-making and Its Application in Fault Diagnosis," IEEE Transactions on Fuzzy Systems, vol. 29, no. 1, pp. 186–197, 2020.
[26] J. Rezaei, "Best-worst Multi-criteria Decision-making Method," Omega, vol. 53, pp. 49–57, 2015.
[27] L. Fei and Y. Deng, "A New Divergence Measure for Basic Probability Assignment and Its Applications in Extremely Uncertain Environments," Int. Journal of Intelligent Systems, vol. 34, no. 4, pp. 584–600, 2019.
[28] T. L. Wickramarathne, K. Premaratne, M. Kubat and D. Jayaweera, "CoFIDS: A Belief-theoretic Approach for Automated Collaborative Filtering," IEEE Transactions on Knowledge and Data Engineering, vol. 23, no. 2, pp. 175–189, 2010.
[29] V.-D. Nguyen and V.-N. Huynh, "A Reliably Weighted Collaborative Filtering System," Proc. of the European Conf. on Symbolic and Quantitative Approaches to Reasoning and Uncertainty (ECSQARU 2015), vol. 9161, pp. 429–439, Springer, 2015.
[30] V.-D. Nguyen and V.-N. Huynh, "Integrating with Social Network to Enhance Recommender System Based on Dempster-Shafer Theory," Proc. of the Int. Conf. on Computational Social Networks (CSoNet 2016), pp. 170– 181, Springer, 2016.
[31] V.-D. Nguyen and V.-N. Huynh, "Noise-Averse Combination Method," Proc. of the 2016 IEEE 28th Int. Conf. on Tools with Artificial Intelligence (ICTAI), pp. 86–90, San Jose, CA, USA, 2016.
[32] V.-D. Nguyen, S. Sriboonchitta and V.-N. Huynh, "Using Community Preference for Overcoming Sparsity and Cold-start Problems in Collaborative Filtering System Offering Soft Ratings," Electronic Commerce Research and Applications, vol. 26, pp. 101–108, 2017.
[33] V.-D. Nguyen, V.-N. Huynh and S. Sriboonchitta, "Integrating Community Context Information into a Reliably Weighted Collaborative Filtering System Using Soft Ratings," IEEE Transactions on Systems, Man and Cybernetics: Systems, vol. 50, no. 4, pp. 1318–1330, 2017.
[34] Y. Dong, X. Li and Y. Liu, "A Fast Combination Method in DSMT and Its Application to Recommender System," PloS One, vol. 13, no. 1, p. e0189703, 2018.
[35] I. Bloch, "Some Aspects of Dempster-Shafer Evidence Theory for Classification of Multi-modality Medical Images Taking Partial Volume Effect into Account," Pattern Recognition Letters, vol. 17, no. 8, pp. 905–919, 1996.
[36] F. M. Harper and J. A. Konstan, "The Movielens Datasets: History and Context," ACM Transactions on Interactive Intelligent Systems (TIIS), vol. 5, no. 4, pp. 1–19, 2015.

ملخص البحث:

تقدّم هذه الورقة إطارَ عملٍ مُبتكراً لتحسين التّصفية التّعاونية المستندة إلى الأدلّة، يمثّل نظام توصيةٍ حاسماً مصمَّماً خِصّيصاً للمجالات الحسّاسة، مثل الرّعاية الصّحية وتتبُّع الأهداف وغيرهما.

يركّز إطار العمل المقترح على تصفيةِ الكيفية الّتي يتمُّ بها التّعامل مع التّفْضيلات المتناقضة للمستخدمين الّتي تشتمل على مواضِع خللٍ، وبخاصّةٍ من حيث إدارة التّفْضيلات المتناقضة في أثناء انتقاء تفْضيلات الجِوار، وذلك بُغية تحسين جودة التّوصْية.

يجمع النّظام المقترح بين نهجٍ قائمٍ على احتمالين وتقنية متقدّمةٍ لإدارة التّناقض، الأمر الّذي من شأنه تخصيص وزْنٍ موثوقيةٍ أكثر دقّةً لكلّ مُسْتخدم، وذلك عبر استخدام طريقة (Deng) النّسبية وطريقة "الأفضل-الأسوأ". والجديرُ بالذّكر أنّ ذلك يحسِّن من اختيار التّفْضيلات في نظام التّصْفية التّعاونية المستندِ إلى الأدلة.

تمّ تجريب إطار العمل المقترح على عددٍ من مجموعات البيانات، وبيّنت النّتائج أنّه يتفوّقُ على مجموعةٍ من أُطُر العمل المشابهة الواردة في أدبيات الموضوع، وذلك عند مقارنته معها من حيث خطأ التّوقُّع والدّقة والضّبط وغير ذلك من مقاييس الأداء.