The University of Texas at Austin
Center for Identity

TWCF: Trust Weighted Collaborative Filtering based on Quantitative Modeling of Trust

Wenting Song

K. Suzanne Barber

UTCID Report #24-10

December 2024

TWCF: Trust Weighted Collaborative Filtering based on Quantitative Modeling of Trust

Wenting Song
Center For Identity, Electrical and Computer Engineering
The University of Texas at Austin
Austin, Texas, USA
wentingsong@utexas.edu

K. Suzanne Barber
Center For Identity, Electrical and Computer Engineering
The University of Texas at Austin
Austin, Texas, USA
sbarber@identity.utexas.edu

***Abstract*—Trust-based collaborative filtering methods exploit trust information to alleviate cold-start and data sparsity problems and improve recommendation performance. Instead of modeling trust itself, most existing approaches rely on trust connections sourced from online social networking platforms. Typically, users engage with products on rating platforms and interact with friends on social networking platforms, exhibiting distinct behavior patterns. Therefore, simply aggregating information and misinterpreting user preferences may lead to information misuse and recommendation bias. To overcome these limitations, this paper introduces three trust measurement models to capture and quantify trust relationships directly from user-provided ratings. Utilizing the trust-weighted scheme, we propose a hybrid collaborative filtering approach called Trust Weighted Collaborative Filtering (TWCF). Experiments on three real-world datasets show that the trust information extracted from rating data and the trust-weighted scheme can significantly improve the performance of original neighborhood collaborative filtering. TWCF achieves an average improvement of 7.12% in prediction accuracy over the trust-based collaborative filtering baselines. Furthermore, the interpretability and scalability of TWCF provide opportunities for further improvement by incorporating more appropriate trust measurement models.**

***Index Terms*—trust, trust models, social networks, collaborative filtering, recommendation**

I. Introduction

Nowadays, people tend to rely on online social networks and online community applications for daily social interactions. Online community applications have simplified how users participate online and removed the restrictions on when and where users can create and share content. However, with the rapid growth of online content, users are often overwhelmed by the sheer volume of information. Fortunately, trust models are emerging as valuable tools for measuring the trustworthiness of online entities and content. Trust models help filter out low-quality or unreliable sources and provide users with relevant, high-quality information [1], [2]. Trust models add a layer of accountability and quality control, improving the reliability and quality of online systems.

Collaborative filtering (CF) is a well-known recommendation algorithm [3], [4]. It leverages the power of collaboration to predict unknown ratings and make recommendations. Traditional collaborative filtering suffer from two inherent problems: data sparsity and cold start. Trust-based collaborative filtering methods exploit trust information to alleviate these two problems and improve recommendation performance [5], [6]. However, it also brings some problems. One problem is that these methods directly use trust information from online social network platforms for product recommendations. Generally, users interact with products on rating platforms rather than with social friends on social networking platforms. Hence, simply aggregating information and misreading user preferences may lead to information misuse and recommendation bias. Another problem is the modeling and representation of online trust relationships. Existing trust-based collaborative filtering methods represent trust relationships as simple binary values, which cannot reflect the essence of trust relationships. In this paper, we detail our efforts to define, measure, and quantify trust in collaborative filtering. We use numerical values to represent different levels of trust between entities.

In summary, this paper focuses on the quantitative modeling of trust in collaborative filtering. Trust measurement models are proposed to calculate the trust relationship between entities (users/items) based on the ratings provided by users. The calculated trust scores are then used as weights to distinguish the influence of neighbor preferences in the rating prediction process of collaborative filtering. The main contributions of this paper are as follows:

- We introduce three trust measurement models to capture and quantify trust relationships directly from user-provided ratings.
- We propose a hybrid collaborative filtering approach named Trust Weighted Collaborative Filtering (TWCF).
- We conduct empirical experiments on three real-world datasets to compare TWCF with collaborative filtering baselines. Experiments show that the trust information extracted from rating data and the trust-weighted scheme can significantly improve the performance of original neighborhood collaborative filtering.
- Relying solely on user-provided ratings, TWCF achieves an average improvement of 7.12% compared to the trust-based collaborative filtering baselines. Minimal data collection requirements make TWCF adaptable to wider applications across various datasets and scenarios.

The rest of the paper is organized as follows. Section II provides an overview of recent research related to our work, and Section III describes the motivations. Section IV presents the structure of the proposed TWCF approach. The empirical experimental details are explained in Section V. Finally, Section VI discusses the experimental results and findings, and Section VII concludes the paper.

II. Related Work

Online Trust Quantification Modeling. Numerous trust models have been developed and put into practice to depict and quantify trust in online social networks [7]–[9]. In reputation-based trust models, trust is derived from user reputation scores, which are accumulated through past interactions and behaviors [7]. The TrustRank algorithm [10] was adapted from the PageRank algorithm [11] to measure trust. By propagating trust from a small set of trusted nodes, TrustRank can help identify and prioritize trustworthy nodes in online social networks. In addition, subjective logic and fuzzy logic deal with the uncertainty and ambiguity of trust relationships [12], [13]. In subjective logic trust models [12], [14], [15], trust is represented as a belief function that combines evidence from different sources. Probability-based trust models [12], [16], [17] provide a numerical representation of trust, making it easier to compare or aggregate trust values across interactions. However, many existing trust quantification models fail to demonstrate their effectiveness in real-world applications or show how much they can improve the reliability and quality of online systems.

Trust-based Collaborative Filtering. Collaborative filtering is a popular recommendation method that recommends items to users based on the preferences of similar users. Trust-based collaborative filtering leverages trust information to alleviate data sparsity and cold start problems and improve recommendation performance [6], [18]–[20]. Existing trust-based collaborative filtering methods require additional trust-related social information to work, which limits their application in practice [5], [6], [21], [22]. In addition, ITRA [23] performed implicit trust computation based on a pre-assumed trust connection graph. SICF [19] improved collaborative filtering by identifying hidden correlations between users through social influence propagation. Shen et al. [18] utilized reputation measurement models to enhance collaborative filtering, but only considered user-based reputation. AgreeRelTrust [24] relies on user agreement to infer trust, while facing the challenge of computational inefficiency. NSCR [25] integrated user-item interactions and user-user connections to improve recommendation performance but requires complex data collection requirements. Deep learning-based collaborative filtering has achieved excellent performance, but has problems in model interpretability and computational cost [26]–[28]. In this paper, we present our efforts to define, measure, and quantify trust within the context of collaborative filtering. Our approach relies solely on user-provided ratings to calculate the trust relationships between entities (users/items). Furthermore, we incorporate the calculated trust scores as weights in the aggregation and rating prediction process to improve collaborative filtering.

III. Motivations

A. Shortcomings of Similarity-Based Collaborative Filtering

Similarity-based collaborative filtering methods suffer from several shortcomings. First, similarity-based collaborative filtering methods rely on common ratings. However, the user-item rating matrix is usually sparse. When common ratings are limited due to data sparsity, the similarity measures cannot capture the relationships between entities. Second, the "information cocoon" problem. Similarity measure is too uniform and simplistic to be the only criterion for identifying nearest neighbors. Similarity-based collaborative filtering methods tend to emphasize items similar to those the user has previously engaged with, which may lead to limited diversity and fewer unexpected recommendations. Third, the curse of dimensionality phenomenon. In high-dimensional space, the distances between all points become similar, making it challenging to distinguish between "near" and "distant" data points or to identify "nearest neighbors". Even worse, as the number of users and items grows, the computational complexity of similarity calculation rises dramatically.

B. Problems with Existing Trust-based Collaborative Filtering

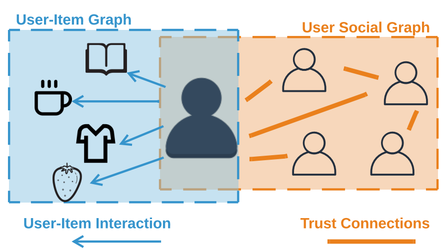

Fig. 1: Users engage with products on rating platforms and interact with friends on social networking platforms, exhibiting distinct behavior patterns [29].

Existing trust-based collaborative filtering methods leverage trust connections from user social networks to improve performance, but this also introduces various problems. The main problem stems from the fact that the user behavior patterns in the social graph and the user-item graph are completely different (Figure 1). For example, suppose a user is connected to another user through a mutual friend on a social network. Despite the connection, these users are unlikely to highly trust or rely on each other's opinions when deciding to purchase products with similar features. Therefore, simply combining information and misinterpreting user preferences may introduce information bias. Nonetheless, most existing "trust-based" collaborative filtering methods directly aggregate these two graphs without further transformation. Another problem is the modeling and representation of trust relationships in collaborative filtering. Existing approaches represent trust relationships as binary values and treat all trust relationships equally. However, the reality is that social networks are full of complexities and varying levels of connection strength. They used the concept of trust but did not model trust itself, nor did they delve into the role that trust modeling could play in collaborative filtering.

C. Calculated Trust Scores as Ranking Criteria

The main idea is to use self-computed trust scores as ranking criteria for collaborative filtering and recommend items based on the target user's trusted neighbors. Utilizing self-computed trust scores as ranking criteria can provide benefits such as improved recommendation accuracy and enhanced resilience against attacks. First, trust scores are calculated directly from user-provided ratings, eliminating bias in the trust information. These calculated trust scores are then used as weights when aggregating the preferences of trusted neighbors, resulting in more personalized and relevant recommendations. Second, trust models enable the system to identify users who may be untrustworthy or engaged in malicious activities, thereby improving its resistance to shilling attacks, where fraudulent or biased ratings are generated to manipulate recommendations.

IV. Trust Weighted Collaborative Filtering

This section introduces the structure and each component of the proposed Trust Weighted Collaborative Filtering (TWCF) approach. As shown in Figure 2, the data input is a raw user-item $m \times n$ rating matrix $\mathbf{R}_{m\times n} = [r_{ui}]$, where m is the number of users and n is the number of items. Therefrom, let I denote the given set of items and U denote the given set of users. The preferences of user u are denoted as an n-dimensional vector $\mathbf{I}_u$, where the k-th entry r_{uk} is the rating of user u on item i_k, or null if user u has not rated the item. Correspondingly, the m-dimensional vector $\mathbf{U}_i$ represents the rating vector of item i. The k-th entry r_{ik} indicates the rating of item i given by user u_k, or null if there is none. TWCF has two main modules as parallel structures: user-based trust-weighted rating prediction and item-based trust-weighted rating prediction [30]. Each module comprises three main stages: trust measures computation, trusted nearest neighbors selection, and hybrid trust-weighted rating prediction. As output, TWCF generates a predicted user-item rating matrix $\hat{\mathbf{R}}_{m\times n}$ for recommendation purposes. In the following subsections, we will introduce these stages sequentially.

A. Trust Entropy Measure

Although users are asked to rate on the same rating scale, each user has a unique interpretation of the rating scale and expresses their preferences in different ways. The rating distribution of users can reflect their rating preferences. For example, one user might assign different ratings with equal probability (uniform distribution), while another user might consistently favor specific ratings more frequently (Gaussian distribution). Moreover, some users are easily satisfied and frequently assign ratings of 4 or 5, while others are more critical, often favoring ratings of 1 or 2. The information conveyed through ratings may vary depending on how users interpret and perceive the rating scale.

Shannon entropy [31] quantifies the uncertainty or unpredictability of a random variable and measures the amount of information contained in its distribution. Based on Shannon entropy, Jensen-Shannon divergence $\mathrm{D}_{\mathrm{js}}(P,Q)$ [32] measures the distance between two probability distributions, $P = \{p_1 \cdots p_p\}$ and $Q = \{q_1 \cdots q_q\}$, in a statistical way,

$$D_{\mathrm{js}}(P,Q) = S(\frac{P+Q}{2}) - \frac{1}{2}S(P) - \frac{1}{2}S(Q) \tag{1}$$

where $S(P) = -\sum_{i=1}^{p} p_i ln p_i$ and $S(Q) = -\sum_{i=1}^{q} q_i ln q_i$. Thence, we use the Jensen-Shannon distance to calculate the difference in rating distribution between two entities (users/items), and $\mathrm{Dist}_{\mathrm{js}} = \sqrt{\mathrm{D}_{\mathrm{js}}(P,Q)}$.

1) User-based Trust Computation:

The Trust Entropy measure is calculated between two users u and v. The preference of user u is denoted as I_u, which is the set of items that user u has rated. Similarly, the preference of user v is denoted as I_v. The entropy value of user u is calculated as $S(U) = -\sum_{r\in R} p_r ln p_r$, where $p_r = \frac{N_r}{N_R}$, p_r is the probability of the occurrence of each rating r, N_r is the number of rating r in I_u, and N_R is the total number of ratings in I_u. For the rating scale from one to five in step one, $R = \{1,2,3,4,5\}$. The Jensen-Shannon distance ($\mathrm{Dist}_{\mathrm{js}} \in [0,1]$) between I_u and I_v is calculated as,

$$\begin{aligned} \mathrm{D}_{\mathrm{js}}(I_u,I_v) = S(\frac{U+V}{2}) - \frac{1}{2}S(U) - \frac{1}{2}S(V) \\ \mathrm{Dist}_{\mathrm{js}}(I_u,I_v) = \sqrt{D_{\mathrm{js}}(I_u,I_v)} \end{aligned} \tag{2}$$

The Trust Entropy measure between user u and user v is given as, $\mathrm{Trust}_{\mathrm{js}}(u,v) = 1 - \mathrm{Dist}_{\mathrm{js}}(I_u,I_v)$, and $\mathrm{Trust}_{\mathrm{js}}(u,v) \in [0,1]$.

2) Item-based Trust Computation:

Likewise, we take two columns (U_i and U_j) to calculate the Trust Entropy scores between item pairs i and j. The Jensen-Shannon distance between the rating distributions of item pairs U_i and U_j is calculated as,

$$\begin{aligned} D_{\mathrm{js}}(U_i,U_j) = S(\frac{I+J}{2}) - \frac{1}{2}S(I) - \frac{1}{2}S(J) \\ \mathrm{Dist}_{\mathrm{js}}(U_i,U_j) = \sqrt{D_{\mathrm{js}}(U_i,U_j)} \end{aligned} \tag{3}$$

The Trust Entropy measure between item i and item j is calculated as, $\mathrm{Trust}_{\mathrm{js}}(i,j) = 1 - \mathrm{Dist}_{\mathrm{js}}(U_i,U_j)$, and always lies in the range $[0,1]$. Thereafter, the calculated Trust Entropy scores are used as the criteria for selecting the most trusted nearest neighbors and as weights in the hybrid trust-weighted rating prediction process.

B. Trust Anomaly Measure

Iterative filtering algorithms [33] have been widely used in trust and reputation systems [18], [34]–[36]. The idea behind this algorithm is to adjust the aggregation weight of a data point based on how much it deviates from other data points. Applying this to our scenario, ratings closer to the consensus estimate will be given higher weights, while outliers will be given lower weights and reduced impact. Inspired by the

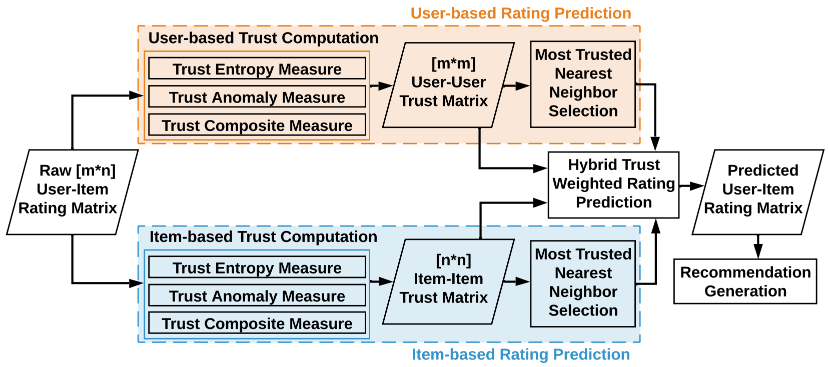

Fig. 2: The structure of the proposed Trust Weighted Collaborative Filtering (TWCF).

iterative filtering algorithms [33], we assign trust scores to entities (users/items) based on how close their ratings are to public ratings.

Anomaly detection is an important technique for detecting abnormal behaviors in online social networks [37], [38]. Abnormal patterns of rating behavior may indicate that the ratings are unreliable, potentially biased, and lack trustworthiness. Since anomalies tend to deviate from typical data patterns, we evaluate how far an entity's rating behavior deviates from the typical rating behavior pattern. This distance can be used as a trust measure to evaluate the trustworthiness of entities. The Trust Anomaly measure is calculated based on an anomaly detection algorithm named Isolation Forest (IF) [39], [40]. We also include a closeness coefficient α that measures the closeness between entities by counting the number of co-rated items.

1) User-based Trust Computation:
Through Min-Max normalization, the IF anomaly scores are scaled to the range $[0, 1]$. The Trust Anomaly measure between user u and user v is calculated as,

$$\alpha_{uv} = \frac{|I_u \cap I_v|}{\max\limits_{a \in U}(|I_u \cap I_a|)}, \quad \text{Trust}_{\text{IF}}(u, v) = \alpha_{uv} \times \text{Anomaly}(v) \tag{4}$$

where α_{uv} is the closeness coefficient, $|I_u \cap I_v|$ is the number of items that users u and v have commonly rated, $\max\limits_{a \in U}(|I_u \cap I_a|)$ is the maximum number of commonly rated items between user u and his potential neighbors in the user set U. $\text{Trust}_{\text{IF}}(u, v)$ is normalized to the range $[0, 1]$ for ease of comparison.

2) Item-based Trust Computation:
Similarly, the Trust Anomaly measure between item i and item j is given by,

$$\alpha_{ij} = \frac{|U_i \cap U_j|}{\max\limits_{b \in U}(|U_i \cap U_b|)}, \quad \text{Trust}_{\text{IF}}(i, j) = \alpha_{ij} \times \text{Anomaly}(j) \tag{5}$$

where α_{ij} is the closeness coefficient and $\text{Anomaly}(j) \in [0, 1]$. The calculated Trust Anomaly scores are normalized to the range $[0, 1]$ and used as the criteria for selecting the most trusted nearest neighbors and as weights in the hybrid trust-weighted rating prediction process.

C. Trust Composite Measure

Trust is a complex concept with multiple dimensions. Trust evaluation scores can vary significantly across different evaluation dimensions because each dimension emphasizes a specific aspect of trust. With this in mind, we propose a composite measure that combines multiple trust measures to evaluate trust from various dimensions. Multi-criteria decision analysis (MCDA) or multi-criteria decision-making (MCDM) methods have been extensively developed to tackle decision-making challenges by evaluating multiple factors and resolving conflicting criteria to determine appropriate solutions [41], [42]. Inspired by this, we utilize MARCOS [41], [42] to combine the similarity measure, the Trust Entropy measure, and the Trust Anomaly measure to derive a multidimensional composite trust measure.

D. Most Trusted Nearest Neighbor Selection

As discussed in Section III-A, traditional collaborative filtering methods face limitations when relying solely on similarity measures as the criteria for determining nearest neighbors. In TWCF, we adopt the "most trusted nearest neighbor" selection to replace the original "nearest neighbor" selection. Instead of relying solely on similarity, we introduce three distinct trust measurement models for neighbor selection and ranking. The term "nearest neighbor" now encompasses a variety of interpretations and selection criteria. Furthermore, using self-computed trust scores as selection criteria, the "most trusted nearest neighbors" are formed by the most trusted neighbors. It helps generate more reliable and higher-quality recommendations, thus improving the performance of collaborative filtering.

E. Hybrid Trust Weighted Rating Prediction

Leveraging the idea of the Pearson correlation coefficient [43], the average rating of each user is subtracted from their rating for normalization. It helps correct for situations where some users may tend to consistently give higher or lower ratings than other users. With the normalization, we can focus on analyzing patterns in user rating behavior rather than absolute ratings. The calculated trust scores serve as weights, and a weighted average of the ratings from the most trusted neighbors is calculated to predict the rating of the target item [44].

1) User-based Rating Prediction:
The predicted rating of user u for item i is calculated as,

$$\tilde{r_{ui}} = r_{ui} - \mu_u, \hat{r_{ui}} = \mu_u + \frac{\sum_{v \in N_u(i)} \text{Trust}(u,v) \cdot \tilde{r_{vi}}}{\sum_{v \in N_u} |\text{Trust}(u,v)|} \quad (6)$$

where Trust(u, v) is the calculated trust score between user u and user v. $N_u(i)$ is the set of users who have rated item i and are included in the "most trusted nearest neighbors" of the target user u.

2) Item-based Rating Prediction:
Similarly, the raw ratings are normalized and the trust-weighted average of the ratings from the "most trusted nearest neighbors" of item i is calculated as,

$$\tilde{r_{ui}} = r_{ui} - \mu_i, \hat{r_{ui}} = \mu_i + \frac{\sum_{j \in N_i(u)} \text{Trust}(i,j) \cdot \tilde{r_{uj}}}{\sum_{j \in N_i} |\text{Trust}(i,j)|} \quad (7)$$

where Trust(i, j) is the calculated trust score between item i and item j, $N_i(u)$ is the set of items that have been rated by user u and also included in the "most trusted nearest neighbors" of the target item i.

V. Empirical Experiments

We conduct offline experiments on three real-world datasets that provide user ratings and trust information: FilmTrust [45], Ciao [46], and Epinions [46]. While the proposed TWCF approach only uses user-provided ratings, the trust-based collaborative filtering baselines require additional trust information to function. Therefore, considering the robustness and consistency of the comparison, we conduct experiments on datasets containing user-provided ratings and trust information. The statistics of the datasets are shown in Table. I.

Datasets	FilmTrust [45]	Ciao [46]	Epinions [46]
#Users	1,508	7,375	49,290
#Items	2,071	99,746	139,738
#Ratings	35,497	278,483	664,824
Scale / Step	[0.5, 4.0] / 0.5	[1, 5] / 0.5	[1, 5] / 1
Density	1.1366%	0.0379%	0.0096%
# Trust Connections	1,853	111,781	487,183

TABLE I: Dataset Statistics.

A. Baselines

First, we compare the proposed TWCF approach with the original k-Nearest Neighbors (kNN) methods.

- original user-based *kNN* with pearson correlation
- original item-based *kNN* with pearson correlation

The comparison is conducted to evaluate the impact of the proposed trust measurement models and the trust-weighted scheme on traditional collaborative filtering. More detailed experimental comparison results are given in Section VI-A.

Next, we compare the proposed TWCF approach with six widely recognized collaborative filtering methods. Many state-of-the-art collaborative filtering methods focus on generating top-k recommendations, which is different from our task of rating prediction [5], [28], [47]. Therefore, these methods are not included as baselines in our experiments. Besides, even a slight improvement in rating prediction accuracy can significantly improve the performance of top-k recommendations [48]. The six baselines are divided into two categories: (1) baselines that rely only on user-provided ratings; (2) baselines that rely on additional trust information to function.

Three baselines that rely only on user-provided ratings:

- PMF [49], utilizes Gaussian distribution to model the latent factors of users and items. It predicts ratings by learning low-dimensional user feature matrix and item feature matrix.
- SVD++ [48], extends the singular value decomposition method by considering implicit ratings.
- EE [50], a collaborative filtering method based on Euclidean embedding.

Another three baselines that rely on additional trust information to function:

- RSTE [21], combines the preferences of target users and trusted neighbors for rating prediction.
- SocialMF [22], incorporates trust information and trust propagation into the matrix factorization model to achieve higher performance.
- SREE [51], incorporates additional social relations into Euclidean embedding-based collaborative filtering to improve recommendation accuracy.

Experimental results and discussions are in Section VI-B.

B. Evaluation Metrics

Our experiments focus on recommendation methods that predict explicit user ratings. To evaluate the accuracy of predicted ratings versus actual ratings, we employ two standard metrics, mean absolute error (MAE) and root mean squared error (RMSE). Top-k recommendation evaluation metrics such as recall, precision, and coverage are not included in our experiments.

1) Mean Absolute Error (MAE):

$$\text{MAE} = \frac{1}{|\Gamma|} \sum_{(u,i) \in \Gamma} |\hat{r_{ui}} - r_{ui}| \quad (8)$$

where Γ is the user-item set, $\hat{r_{ui}}$ is the predicted ratings, and r_{ui} is the actual ratings. The lower the MAE value, the more accurate the prediction.

2) Root Mean Squared Error (RMSE):

$$\text{RMSE} = \sqrt{\frac{1}{|\Gamma|} \sum_{(u,i) \in \Gamma} (\hat{r_{ui}} - r_{ui})^2} \quad (9)$$

Likewise, lower RMSE values indicate more accurate predictions.

VI. Experimental Results and Discussions

In this section, we present the experimental results and discuss our observations and findings from the experiments.

A. Improvements to Traditional Collaborative Filtering

Methods	FilmTrust		Ciao		Epinions	
	MAE	RMSE	MAE	RMSE	MAE	RMSE
original user*kNN*	0.6326	0.8232	0.8222	1.1238	1.0111	1.3664
trust entropy user*kNN*	0.6075	0.7946	0.8109	1.1026	**1.0089**	**1.3628**
trust anomaly user*kNN*	0.6025	0.7967	0.8111	1.1047	1.0098	1.3645
trust composite user*kNN*	**0.6040**	**0.7946**	**0.8109**	**1.1026**	1.0090	1.3644
original item*kNN*	0.7245	0.9305	0.7882	1.0753	0.9108	1.1285
trust entropyy item*kNN*	0.6354	0.8696	**0.7670**	**1.0458**	0.9081	1.1272
trust anomaly item*kNN*	0.6751	0.8922	0.7709	1.0577	**0.9079**	**1.1270**
trust composite item*kNN*	**0.6214**	**0.8433**	0.7771	1.0634	0.9083	1.1273

(a) Experimental results of MAE and RMSE metrics.

Methods	FilmTrust		Ciao		Epinions	
	MAE	RMSE	MAE	RMSE	MAE	RMSE
trust entropy user*kNN*	3.97%	3.48%	1.38%	1.89%	**0.23%**	**0.27%**
trust anomaly user*kNN*	4.76%	3.22%	1.36%	1.70%	0.13%	0.14%
trust composite user*kNN*	**4.52%**	**3.48%**	**1.38%**	**1.89%**	0.21%	0.15%
trust entropy item*kNN*	12.30%	6.55%	**2.69%**	**2.75%**	0.31%	0.13%
trust anomaly item*kNN*	6.82%	4.12%	2.20%	1.64%	**0.32%**	**0.14%**
trust composite item*kNN*	**14.23%**	**9.38%**	1.42%	1.12%	0.29%	0.11%

(b) Improvement ratio of the TWCF approach over the original *kNN* methods.

TABLE II: Comparison of the original *kNN* methods with the proposed TWCF approach.

First, we evaluate the performance of the methods in the TWCF approach (Trust Entropy *kNN*, Trust Anomaly *kNN* and Trust Composite *kNN*) compared to the original k-nearest neighbors method. We analyze the impact of the proposed trust measurement models and the trust-weighted scheme on the original k-nearest neighbors method. Table. II shows the experimental comparison results between the original *kNN* methods and the proposed TWCF approach.

Overall, TWCF outperforms traditional collaborative filtering methods in recommendation performance. The prediction accuracy of both user-based and item-based *kNN* methods is significantly improved. On the FilmTrust dataset, Trust Composite item*kNN* shows the most significant improvement, with a 14.23% increase in MAE metric and a 9.38% increase in RMSE metric compared to the original item*kNN*. Meanwhile, on the Ciao dataset, Trust Entropy item*kNN* outperforms the original item*kNN* by 2.69% and 2.75% in MAE and RMSE metrics, respectively. However, the improvement on the Epinions dataset is not as significant as on the other two datasets. The difference in improvement rate is attributed to the different degrees of data sparsity in the original dataset. The data sparsity of the FilmTrust dataset is 1.1366%, so the TWCF method achieves the most significant improvement on it. On the other hand, Epinions has a data sparsity of 0.0096%, showing a relatively small improvement. Hence, our observation is that the improvement potential of TWCF is affected by the data sparsity of the rating dataset. This is in line with common sense since the user-item rating matrix is the only source of information for the trust measurement model calculation and trust-weighted rating prediction. In conclusion, the comparison between the TWCF approach and the original *kNN* method highlights the benefits of incorporating trust information derived from rating data into original collaborative filtering. The trust measures computation and the trust-weighted scheme improve the selection of nearest neighbors and the aggregation of ratings, leading to more accurate predictions and enhanced recommendation performance. However, the potential for improvement may be affected by the sparsity of rating data.

B. Comparison with Collaborative Filtering Baselines

Furthermore, we compare the experimental performance of the proposed TWCF approach with six collaborative filtering baselines, where PMF [49], SVD++ [48], and EE [50] rely only on user-provided ratings, whereas RSTE [21], SocialMF [22], and SREE [51] rely on additional trust information to function.

1) Baselines that Rely Only on User-provided Ratings:

To begin with, we compare the experimental results of the TWCF approach with baselines PMF [49], SVD++ [48] and EE [50]. They all only utilize the user-item rating matrix to generate recommendations. Table. III shows the experimental comparison results.

Methods	FilmTrust		Ciao		Epinions	
	MAE	RMSE	MAE	RMSE	MAE	RMSE
PMF [49]	0.6713	0.8984	0.8603	1.1668	0.9430	1.2377
SVD++ [48]	0.6369	0.8531	0.7781	1.0524	0.9265	1.2356
EE [50]	0.6687	0.8943	0.8066	1.0610	0.9211	1.1661
TWCF	**0.6040**	**0.7946**	**0.7670**	**1.0458**	**0.9079**	**1.1270**

(a) Experimental results of MAE and RMSE metrics.

Methods	FilmTrust		Ciao		Epinions	
	MAE	RMSE	MAE	RMSE	MAE	RMSE
PMF [49]	10.03%	11.56%	10.85%	10.38%	3.73%	8.95%
SVD++ [48]	5.17%	6.86%	1.43%	0.64%	2.01%	8.79%
EE [50]	9.68%	11.15%	4.91%	1.44%	1.44%	3.36%

(b) Improvement ratio of the TWCF approach over baselines.

TABLE III: Comparison with baselines that rely only on user-provided ratings.

TWCF achieves the best performance among all baselines in terms of recommendation accuracy metrics MAE and RMSE. On average, across three real-world datasets, it outperforms PMF by 6.26%, SVD++ by 4.16%, and EE by 5.34%. This can be interpreted as a significant improvement for recommendation systems based on rating prediction [52]. Note that even a slight improvement in rating prediction accuracy can significantly improve the performance of top-k recommendations [48]. Furthermore, since the TWCF methods are modified from the traditional *kNN* methods, they retain excellent model interpretability compared to the matrix factorization-based collaborative filtering methods PMF and SVD++. The excellent transparency and interpretability of TWCF provide opportunities for further improvement by incorporating more appropriate trust measurement models.

2) Baselines that Rely on Trust Information to Function: Next, we compare the experimental results of the TWCF approach with baselines RSTE [21], SocialMF [22] and SREE [51]. Most existing trust-based recommendation methods rely on user-provided ratings sourced from user-item interaction and trust information sourced from user social connections to function. These methods, including RSTE [21], SocialMF [22], and SREE [51], have specific requirements for datasets and impose limitations on data collection and practical implementation.

Methods	FilmTrust		Ciao		Epinions	
	MAE	RMSE	MAE	RMSE	MAE	RMSE
RSTE [21]	0.6589	0.8859	0.8417	1.1166	0.9652	1.2984
SocialMF [22]	0.6197	0.8302	0.8633	1.1559	1.0086	1.3507
SREE [51]	0.6396	0.8388	0.7913	1.0871	0.9145	1.1567
TWCF	**0.6040**	**0.7946**	**0.7670**	**1.0458**	**0.9079**	**1.1270**

(a) Experimental results of MAE and RMSE metrics.

Methods	FilmTrust		Ciao		Epinions	
	MAE	RMSE	MAE	RMSE	MAE	RMSE
RSTE [21]	8.34%	10.31%	8.88%	6.35%	5.94%	13.21%
SocialMF [22]	2.54%	4.30%	11.16%	9.53%	9.99%	16.57%
SREE [51]	5.57%	5.27%	3.08%	3.81%	0.73%	2.58%

(b) Improvement ratio of the TWCF approach over baselines.

TABLE IV: Comparison with baselines that rely on additional trust information to function.

According to the experimental results in Table. IV, TWCF outperforms all the trust-based collaborative filtering baselines on the three real-world datasets. Averaging over three real-world datasets, TWCF achieves 8.84% higher prediction accuracy than RSTE, 9.02% higher than SocialMF, and 3.51% higher than SREE. Across three real-world datasets, TWCF achieves an average improvement of 7.12% in rating prediction accuracy over the trust-based collaborative filtering baselines. This can be considered a significant improvement in prediction performance [48], [52]. Furthermore, TWCF relies entirely on user-item ratings to compute trust relationships and make recommendation predictions. Despite the limited data requirements, it still demonstrates commendable recommendation performance. Furthermore, the minimal data collection requirements make TWCF applicable to a wider range of datasets and application contexts.

VII. Conclusions and Future Work

Online social networks enable users to create and share information and ideas, but they also generate a large amount of content of varying quality. Trust models are emerging as valuable tools for measuring the trustworthiness of online entities and content. Meanwhile, trust-based collaborative filtering leverages trust information to improve recommendation performance. However, most existing approaches rely on trust connection information from social networking platforms rather than modeling trust itself. In this paper, we present our efforts to define, measure, and quantify trust in the context of collaborative filtering. More specifically, we introduce three trust measurement models: the Trust Entropy measure, the Trust Anomaly measure, and the Trust Composite measure to capture trust relationships directly from user-provided ratings. Utilizing the trust-weighted scheme, we propose a hybrid collaborative filtering approach named Trust Weighted Collaborative Filtering (TWCF). TWCF comprises four stages: trust measures computation, trusted nearest neighbors selection, hybrid trust-weighted rating prediction, and recommendation generation. Next, we conduct empirical experiments on three real-world datasets to compare TWCF with collaborative filtering baselines. Experimental results show that extracting trust information from rating data and incorporating trust-weighted schemes can substantially improve the performance of original neighborhood collaborative filtering. Furthermore, relying solely on user-provided ratings, TWCF achieves an average improvement of 7.12% compared to the trust-based collaborative filtering baselines. The minimal data collection prerequisites make it flexible to be applied to a wide range of datasets and application scenarios. Moreover, TWCF's excellent transparency and interpretability open opportunities for further improvement by incorporating more appropriate trust measures.

In future work, we plan to explore more practical models for quantitatively modeling and measuring online trust. By leveraging and optimizing available information, we can enhance collaborative filtering while minimizing the need for extensive user data collection.

References

[1] T.-C. Huang, R. N. Zaeem, and K. S. Barber, "It is an equal failing to trust everybody and to trust nobody: Stock price prediction using trust filters and enhanced user sentiment on twitter," *ACM Transactions on Internet Technology (TOIT)*, vol. 19, no. 4, pp. 1–20, 2019.

[2] ——, "Identifying real-world credible experts in the financial domain," *Digital Threats: Research and Practice*, vol. 2, no. 2, pp. 1–14, 2021.

[3] B. Sarwar, G. Karypis, J. Konstan, and J. Riedl, "Item-based collaborative filtering recommendation algorithms," in *Proceedings of the 10th international conference on World Wide Web*, 2001, pp. 285–295.

[4] G. Linden, B. Smith, and J. York, "Amazon. com recommendations: Item-to-item collaborative filtering," *IEEE Internet computing*, vol. 7, no. 1, pp. 76–80, 2003.

[5] F. Wang, H. Zhu, G. Srivastava, S. Li, M. R. Khosravi, and L. Qi, "Robust collaborative filtering recommendation with user-item-trust records," *IEEE Transactions on Computational Social Systems*, vol. 9, no. 4, pp. 986–996, 2021.

[6] B. Yang, Y. Lei, J. Liu, and W. Li, "Social collaborative filtering by trust," *IEEE transactions on pattern analysis and machine intelligence*, vol. 39, no. 8, pp. 1633–1647, 2016.

[7] D. D. S. Braga, M. Niemann, B. Hellingrath, and F. B. D. L. Neto, "Survey on computational trust and reputation models," *ACM Computing Surveys (CSUR)*, vol. 51, no. 5, pp. 1–40, 2018.

[8] S. M. Ghafari, A. Beheshti, A. Joshi, C. Paris, A. Mahmood, S. Yakhchi, and M. A. Orgun, "A survey on trust prediction in online social networks," *IEEE Access*, vol. 8, pp. 144 292–144 309, 2020.

[9] J. Wang, X. Jing, Z. Yan, Y. Fu, W. Pedrycz, and L. T. Yang, "A survey on trust evaluation based on machine learning," *ACM Computing Surveys (CSUR)*, vol. 53, no. 5, pp. 1–36, 2020.

[10] Z. Gyongyi, H. Garcia-Molina, and J. Pedersen, "Combating web spam with trustrank," in *Proceedings of the 30th international conference on very large data bases (VLDB)*, 2004.

[11] L. Page, S. Brin, R. Motwani, and T. Winograd, "The pagerank citation ranking: Bringing order to the web." Stanford InfoLab, Tech. Rep., 1999.

[12] G. Liu, Q. Yang, H. Wang, and A. X. Liu, "Trust assessment in online social networks," *IEEE Transactions on Dependable and Secure Computing*, vol. 18, no. 2, pp. 994–1007, 2019.

[13] Z. Lv and H. Song, "Trust mechanism of feedback trust weight in multimedia network," *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)*, vol. 17, no. 4, pp. 1–26, 2021.

[14] J.-H. Cho, "Dynamics of uncertain and conflicting opinions in social networks," *IEEE transactions on computational social systems*, vol. 5, no. 2, pp. 518–531, 2018.

[15] A. Jøsang, *Subjective logic*. Springer, 2016, vol. 3.

[16] G. Lingam, R. R. Rout, D. V. Somayajulu, and S. K. Das, "Social botnet community detection: a novel approach based on behavioral similarity in twitter network using deep learning," in *Proceedings of the 15th ACM Asia Conference on Computer and Communications Security*, 2020, pp. 708–718.

[17] C. Naumzik and S. Feuerriegel, "Detecting false rumors from retweet dynamics on social media," in *Proceedings of the ACM web conference 2022*, 2022, pp. 2798–2809.

[18] L. Shen, S. Huang, and X. Mao, "A reputation-enhanced model for trust-based collaborative filtering recommender system," in *2020 International Joint Conference on Neural Networks (IJCNN)*. IEEE, 2020, pp. 1–7.

[19] Y. Zhou, L. Liu, K. Lee, B. Palanisamy, and Q. Zhang, "Improving collaborative filtering with social influence over heterogeneous information networks," *ACM Transactions on Internet Technology (TOIT)*, vol. 20, no. 4, pp. 1–29, 2020.

[20] Z. Hu, G. Xu, X. Zheng, J. Liu, Z. Li, Q. Z. Sheng, W. Lian, and H. Xian, "Ssl-svd: semi-supervised learning–based sparse trust recommendation," *ACM Transactions on Internet Technology (TOIT)*, vol. 20, no. 1, pp. 1–20, 2020.

[21] H. Ma, I. King, and M. R. Lyu, "Learning to recommend with social trust ensemble," in *Proceedings of the 32nd international ACM SIGIR conference on Research and development in information retrieval*, 2009, pp. 203–210.

[22] M. Jamali and M. Ester, "A matrix factorization technique with trust propagation for recommendation in social networks," in *Proceedings of the fourth ACM conference on Recommender systems*, 2010, pp. 135–142.

[23] Y. Li, J. Liu, J. Ren, and Y. Chang, "A novel implicit trust recommendation approach for rating prediction," *IEEE Access*, vol. 8, pp. 98 305–98 315, 2020.

[24] A. Zahir, Y. Yuan, and K. Moniz, "Agreereltrust—a simple implicit trust inference model for memory-based collaborative filtering recommendation systems," *Electronics*, vol. 8, no. 4, p. 427, 2019.

[25] X. Wang, X. He, L. Nie, and T.-S. Chua, "Item silk road: Recommending items from information domains to social users," in *Proceedings of the 40th International ACM SIGIR conference on Research and Development in Information Retrieval*, 2017, pp. 185–194.

[26] X. He, L. Liao, H. Zhang, L. Nie, X. Hu, and T.-S. Chua, "Neural collaborative filtering," in *Proceedings of the 26th international conference on world wide web*, 2017, pp. 173–182.

[27] S. Rendle, W. Krichene, L. Zhang, and J. Anderson, "Neural collaborative filtering vs. matrix factorization revisited," in *Proceedings of the 14th ACM Conference on Recommender Systems*, 2020, pp. 240–248.

[28] R. Yu, D. Ye, Z. Wang, B. Zhang, A. M. Oguti, J. Li, B. Jin, and F. Kurdahi, "Cffnn: Cross feature fusion neural network for collaborative filtering," *IEEE Transactions on Knowledge and Data Engineering*, vol. 34, no. 10, pp. 4650–4662, 2021.

[29] J. Wu, W. Fan, J. Chen, S. Liu, Q. Li, and K. Tang, "Disentangled contrastive learning for social recommendation," in *Proceedings of the 31st ACM International Conference on Information & Knowledge Management*, 2022, pp. 4570–4574.

[30] Q. Shambour and J. Lu, "An effective recommender system by unifying user and item trust information for b2b applications," *Journal of Computer and System Sciences*, vol. 81, no. 7, pp. 1110–1126, 2015.

[31] C. E. Shannon, "A mathematical theory of communication," *The Bell system technical journal*, vol. 27, no. 3, pp. 379–423, 1948.

[32] J. Lin, "Divergence measures based on the shannon entropy," *IEEE Transactions on Information theory*, vol. 37, no. 1, pp. 145–151, 1991.

[33] C. De Kerchove and P. Van Dooren, "Iterative filtering in reputation systems," *SIAM Journal on Matrix Analysis and Applications*, vol. 31, no. 4, pp. 1812–1834, 2010.

[34] M. Rezvani, A. Ignjatovic, E. Bertino, and S. Jha, "A collaborative reputation system based on credibility propagation in wsns," in *2015 IEEE 21st International Conference on Parallel and Distributed Systems (ICPADS)*. IEEE, 2015, pp. 1–8.

[35] M. Rezvani and M. Rezvani, "A randomized reputation system in the presence of unfair ratings," *ACM Transactions on Management Information Systems (TMIS)*, vol. 11, no. 1, pp. 1–16, 2020.

[36] J. Saúde, G. Ramos, L. Boratto, and C. Caleiro, "A robust reputation-based group ranking system and its resistance to bribery," *ACM Transactions on Knowledge Discovery from Data (TKDD)*, vol. 16, no. 2, pp. 1–35, 2021.

[37] J. Wen, N. Jiang, L. Li, J. Zhou, Y. Li, H. Zhan, G. Kou, W. Gu, and J. Zhao, "Ta-detector: A gnn-based anomaly detector via trust relationship," *ACM Transactions on Multimedia Computing, Communications and Applications*, 2024.

[38] Y. Li, J. Zhu, C. Zhang, Y. Yang, J. Zhang, Y. Qiao, and H. Wang, "Thgnn: An embedding-based model for anomaly detection in dynamic heterogeneous social networks," in *Proceedings of the 32nd ACM International Conference on Information and Knowledge Management*, 2023, pp. 1368–1378.

[39] F. T. Liu, K. M. Ting, and Z.-H. Zhou, "Isolation forest," in *2008 eighth ieee international conference on data mining*. IEEE, 2008, pp. 413–422.

[40] ——, "Isolation-based anomaly detection," *ACM Transactions on Knowledge Discovery from Data (TKDD)*, vol. 6, no. 1, pp. 1–39, 2012.

[41] Ž. Stević, D. Pamučar, A. Puška, and P. Chatterjee, "Sustainable supplier selection in healthcare industries using a new mcdm method: Measurement of alternatives and ranking according to compromise solution (marcos)," *Computers & Industrial Engineering*, vol. 140, p. 106231, 2020.

[42] A. Ulutaş, D. Karabasevic, G. Popovic, D. Stanujkic, P. T. Nguyen, and Ç. Karaköy, "Development of a novel integrated ccsd-itara-marcos decision-making approach for stackers selection in a logistics system," *Mathematics*, vol. 8, no. 10, p. 1672, 2020.

[43] I. Cohen, Y. Huang, J. Chen, J. Benesty, J. Benesty, J. Chen, Y. Huang, and I. Cohen, "Pearson correlation coefficient," *Noise reduction in speech processing*, pp. 1–4, 2009.

[44] P. Resnick, N. Iacovou, M. Suchak, P. Bergstrom, and J. Riedl, "Grouplens: An open architecture for collaborative filtering of netnews," in *Proceedings of the 1994 ACM conference on Computer supported cooperative work*, 1994, pp. 175–186.

[45] G. Guo, J. Zhang, and N. Yorke-Smith, "A novel bayesian similarity measure for recommender systems." in *IJCAI*, vol. 13, 2013, pp. 2619–2625.

[46] J. Tang, H. Gao, and H. Liu, "mtrust: Discerning multi-faceted trust in a connected world," in *Proceedings of the fifth ACM international conference on Web search and data mining*, 2012, pp. 93–102.

[47] L. Xia, C. Huang, Y. Xu, J. Zhao, D. Yin, and J. Huang, "Hypergraph contrastive collaborative filtering," in *Proceedings of the 45th International ACM SIGIR conference on research and development in information retrieval*, 2022, pp. 70–79.

[48] Y. Koren, "Factorization meets the neighborhood: a multifaceted collaborative filtering model," in *Proceedings of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining*, 2008, pp. 426–434.

[49] A. Mnih and R. R. Salakhutdinov, "Probabilistic matrix factorization," *Advances in neural information processing systems*, vol. 20, 2007.

[50] M. Khoshneshin and W. N. Street, "Collaborative filtering via euclidean embedding," in *Proceedings of the fourth ACM conference on Recommender systems*, 2010, pp. 87–94.

[51] W. Li, M. Gao, W. Rong, J. Wen, Q. Xiong, R. Jia, and T. Dou, "Social recommendation using euclidean embedding," in *2017 International Joint Conference on Neural Networks (IJCNN)*. IEEE, 2017, pp. 589–595.

[52] W. Fan, Y. Ma, Q. Li, Y. He, E. Zhao, J. Tang, and D. Yin, "Graph neural networks for social recommendation," in *The world wide web conference*, 2019, pp. 417–426.

WWW.IDENTITY.UTEXAS.EDU