



HAL
open science

Trustworthy User Recommendation Using Boosted Vector Similarity Measure

Dhanalakshmi Teekaraman, Sendhilkumar Selvaraju, Mahalakshmi
Guruvayur Suryanarayanan

► **To cite this version:**

Dhanalakshmi Teekaraman, Sendhilkumar Selvaraju, Mahalakshmi Guruvayur Suryanarayanan. Trustworthy User Recommendation Using Boosted Vector Similarity Measure. 3rd International Conference on Computational Intelligence in Data Science (ICCIDS), Feb 2020, Chennai, India. pp.43-58, 10.1007/978-3-030-63467-4_4 . hal-03434778

HAL Id: hal-03434778

<https://inria.hal.science/hal-03434778v1>

Submitted on 18 Nov 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

Trustworthy User Recommendation using Boosted Vector Similarity Measure

Dhanalakshmi Teekaraman ¹ [0000-0001-8059-3120], Sendhilkumar S ^{*2} [0000-0001-6006-1866] and Mahalakshmi G.S ³ [0000-0002-6286-2058]

^{1,2} Department of Information Science and Technology, Anna University, India - 600025

³ Department of Computer Science and Engineering, Anna University, India – 600025

*Corresponding Author: ssk_pdy@yahoo.co.in

Abstract. An online social network (OSN) is crowded with people and their huge number of post and hence filtering truthful content and/or filtering truthful content creator is a great challenge. The online recommender system helps to get such information from OSN and suggest the valuable item or user. But in reality people have more belief on recommendation from the people they trust than from untrusted sources. Getting recommendation from the trusted people derived from social network is called Trust-Enhanced Recommender System (TERS). A Trust-Boosted Recommender System (TBRS) is proposed in this paper to address the challenge in identifying trusted users from social network. The proposed recommender system is a fuzzy multi attribute recommender system using boosted vector similarity measure designed to predict trusted users from social networks with reduced error. Performance analysis of the proposed model in terms of accuracy measures such as precision@k and recall@k and error measures, namely, MAE, MSE and RMSE is discussed in this paper. The evaluation shows that the proposed system outperforms other recommender system with minimum MAE and RMSE.

Keywords: Recommender System, Provenance, Trust, Social Network, Fuzzy Classifier

1 Introduction

The Social network is congested by a large number of posts such as blog, reviews, opinions, image, video, etc. Extracting the required information from such a congested network is very difficult and time consuming task. An online recommender system helps to retrieve the desired information from this crowded network. For example, in Amazon's recommender system, item-to-item collaborative filtering approach is used for item recommendation. Similarly, Facebook, LinkedIn and other social networking sites to examine the network of connections between a user and their friends to suggest a new group based on interest. The downside of this online recommender system is that, the recommendations are generated based on anonymous people similar to the target user. This recommendation does not guarantee that the recommendation generated is from trusted people. Therefore, people tend to rely more on trusted person's

recommendation than online recommendation [21]. The recommender system designed for the trust network is called trust-based recommender system.

When the trust model becomes potentially vulnerable then the transparency of the trust rating is lost [13]. The critical analysis of content or web resources makes the trust rating transparent, which is made possible only by provenance. Provenance provides Meta information about the creation and processing of content. Thus, in this model, the trust rating is computed using the provenance data derived from W7 model. Also, the trust ratings derived in the models [1, 2, 8, 9, 10, 14] are single rating or single preference. For example, in five rating scale, the trust value '4' represents high trust while trust value '1' represents very low trust. With a single rating or preference, the multiple aspects of the user or item cannot be expressed which will either directly or indirectly reduce the recommendation quality. Therefore, if the trust rating is derived using multiple criteria or features such as 'Originality of the content', 'Timeliness of the post', and 'Relevancy of the content' as (4, 3, 2), then evidently the quality of recommendation is improved.

The issues discussed above are handled by the proposed recommender system. One is multi-dimensional or multi attributes based trust evaluation than single dimension or a single attribute. If multiple aspects of users are analyzed for trust computation, then the impact of recommendation is stronger and positive. Next the attribute information gain is used as weight component and weighted similarity measure is computed. This multiple dimensions are easily represented using vector and hence vector similarity measures such as Jaccard, Dice and Cosine are used. Then this similarity is boosted by users trust degree or trust level and recommendation is made. The contributions of the proposed recommender system are as follows.

- Modeling the user
- Formation of fuzzy vector space
- Finding preference and recommending top-k users

The structure of the paper is as follows. Section 2 briefs about the related research. The proposed recommender system is elaborately discussed in section 3. Performance evaluation is discussed in section 4. Finally, conclusion and future works are stated in section 5.

2 Related Work

The trust-based recommendation techniques depend on two important components, namely recommendation techniques and representation of trust models.

2.1 Trust-Enhanced Recommendation Techniques

The trust enhanced recommendation algorithms are generally an enhancement of standard recommendation techniques such as Simple mean; Pearson weighted mean, Pearson collaborative filtering. The former method receives recommendation from trusted peers, whereas the latter method received recommendation from normal users.

The most common trust enhanced recommender strategy is asking the users to explicitly mention the trust statements about other users. For instance the Moleskiing recommender system [3] uses FOAF files that contain trusted information scale ranging from 1 to 9. The Trust model proposed by A. Abdul Rahman and S. Hailes [1] for virtual communities grounded in real-world social trust characteristics, reputation or word-of-mouth. Falcone et.al proposed a fuzzy cognitive map model [8] to derive the trust based on belief value of an agent. This model shows how different component (belief) may change and how their impact can change depending on the specific situation and from the agent personality. The aim of a Golbeck's trust model [9] is, to determine how much one person in the network should trust another person to whom they are not directly connected. This algorithm accurately analyses the opinions of the people in the system. TidalTrust algorithm works based on trust-based weighted mean which uses the trust value of users as a weight for the ratings of other users.

Hang et al. [10] used a graph-based approach to recommend a node in a social network using similarity in trust networks. Massa and Aversani [14] proposed a trust-based recommendation system where it is possible to search for trustable users by exploiting trust propagation over the trust network. Andersen et al. [2] explored an axiomatic approach for trust-based recommendation and propose several recommendation models, some of which are incentive compatible. In MoleTrust method the similarity weight attributed to ratings by user. A trust-filtered collaborative filtering technique is used by O'Donovan and Smith in [4]. Here the trust value is used as a filtering mechanism to choose only, the item raters who are trusted above a certain threshold. An Ensemble trust technique is proposed by victor et.al [17] aims to take into account all possible ways to obtain a positive weight for a rater of an item while favoring trust over similarity.

2.2 Trust Model representation

Trust representations can be classified from three different perspectives, namely (i) Probabilistic vs. gradual trust (ii) Single vs. multi-dimensional trust and (iii) Trust vs. distrust. Probabilistic representations use probabilities to indicate how much trust is placed by a user to another [17] Stronger trust corresponds to a higher probability. Gradual representations [17] use continuous values to represent trust. The values can be any values so they cannot be explained as probabilities. The values directly indicate trust strengths. Here, (u, v, t) denotes that the trust value from u to v is t . Trust is a complex concept with multiple dimensions (i) Multifaceted trust and (ii) Trust evolution. It is an extension of single trust representations of multi-dimensional trust representations [11]. Trust is context dependent. Trusting someone on one topic does not necessarily mean he will be trusted by others. The trust value is represented with $\langle u, v, f, p \rangle$, where u trust v with probability p in the facet f . Also author suggests that trust evolves as humans interact over time T . Josang's subjective logic explores the probabilistic model [12] that considers both trust and distrust simultaneously. A gradual trust model for both trust and distrust can be found in [5, 8, 16]. Guha et.al use a pair (t, d) [18] with trust degree t and distrust degree d and final suggested trust value is obtained by subtracting d from t i.e. $t-d$.

3 Proposed Recommender System

The proposed recommender system is built to recommend the top-k reviewers in a book-based social network. For this, the data about the reviewer and the review is collected from Goodreads, Google Books and Amazon using ad hoc-API and scraping HTML pages. The fields collected from the social network are given in table 1. More than 61,000 reviews and associated reviewer's data available from 2007 to 2015 is collected and details of the dataset collected is given in table 2. Number of reviews and number of reviewers are not same always. But this dataset has single review from each reviewer.

Table 1. Fields collected from the social network

| S.No | Fields | S.No | Fields |
|------|-----------------------|------|-----------------|
| 1 | Review_Text | 9 | Reviewer_ID |
| 2 | Year_Month_of_Joining | 10 | Review_ID |
| 3 | Review_Postdate | 11 | Reviewer_URL |
| 4 | Rate_of_Review | 12 | Review_URL |
| 5 | Number_of_Likes | 13 | Commenter_URL |
| 6 | Month_Last_Active | 14 | Date_of_Comment |
| 7 | Number_of_Comments | 15 | Comments |
| 8 | Number_of_Reply | 16 | Profile_Status |

Table 2. Dataset Details

| Title of the book | No. of Reviews | Overall Rating | No. of years Of review |
|---------------------------------------|----------------|----------------|------------------------|
| Adventures of Huckleberry Finn | 10,153 | 3.78 | 8(2008-2015) |
| Good Night, Mr.Tom | 1,055 | 4.25 | 9(2007-2015) |
| Harry Potter and the Sorcerer's Stone | 49,955 | 4.40 | 9(2007-2015) |

The collected data is preprocessed and from this the trust score of each reviewer is computed using W7 provenance model [22]. Then, using DoT pruned Fuzzy Decision Tree (FDT) classifier [7] the reviewers are classified and fuzzy rules were generated. Finally, fuzzy rules are combined with a target user's request to perform recommendation. The major components of the proposed recommender system are as follows.

1. Provenance Based Trust Assessment
2. Fuzzy Decision Tree Based Classifier
3. Trust-Boosted Recommender System

3.1 Provenance Based Trust Assessment

This section briefs about how trust assessment is made using W7 provenance model. For the experiment the data are collected from Goodreads.com book-based social network. The provenance elements defined in W7 model are a 7-tuple: (WHAT,

WHEN, WHO, HOW, WHY, WHICH, WHERE). The description of provenance elements in the context of trust is shown in table 3.

Table 3. Description of Provenance Elements

| Provenance Element | Description |
|--------------------|---|
| P _{WHAT} | Describes the review content that is relevant to the topic |
| P _{WHEN} | Represents the effective time spent by the reviewer. |
| P _{WHERE} | Refers to the location (IP_Address, Domain_name) from where review is posted. |
| P _{WHO} | Refers to the reviewer who is an author or creator of the review. |
| P _{HOW} | Describes how review content is deviated from the rating given by the reviewer. |
| P _{WHICH} | Refers to the application or device used to post the review. |
| P _{WHY} | Describes the intention behind the post of review content |

Since, the data for P_{WHERE} and P_{WHICH} is not provided by the domain, these two elements cannot be modeled. Therefore, the core provenance elements taken for trust quantification are P_{WHAT}, P_{WHEN}, P_{WHO}, P_{HOW} and P_{WHY}. Trust assessment algorithm quantifies these five provenance elements. This trust value is then given to the learning model to classify the users with various levels.

3.2 Fuzzy Decision Tree Based Classification

The learning model takes the quantified provenance value obtained using W7 model as a trust input. This is fuzzified using Triangular Membership Function (TMF) and rule base is constructed using Mamdani's 'If... Then' interpretation. Fuzzy Decision Tree (FDT) [19] takes the rule base and generates decision trees using a fuzzy ID3 [6] algorithm. To construct FDT, two criteria need to be evaluated, one is splitting criterion and the other is stopping criterion. The former one helps to choose the root node and child nodes. The latter one controls the growth of the tree.

In FDT, provenance element having highest information gain is assigned as the root node and leaf node denotes trust decision. Each distinct path from root to a leaf produces distinct rule. Each generated rule is assigned Degrees of Truth (DoT) [15] to state that how much truth value it holds. If DoT=1, then the rule is absolutely true and if DoT=0 then the rule is absolutely false. Sample fuzzy rules are shown in figure 1. Here, reviewers are classified into 5 different trust levels as VHGT (Very High Trust), HT (High Trust), MT (Moderate Trust), LT (Low Trust) and VLWT (Very Low Trust). The abbreviation for the linguistic terms present in figure 1 is as follows. MSM (Moderately Same), HD (Highly Deviated), HSM (Highly Same), MD (Moderately Deviated), MITM (Moderately Ineffective Time Spent), HITM (Highly Ineffective Time Spent), HUTR (Highly Untruthful), HR (Highly Relevance), NTR (Neutral Truthful), MIR (Moderately Irrelevance) and MDSML (Moderately Dissimilar).

| |
|--|
| Rule 1: If P_{HOW} is MSM and P_{WHEN} is MITM and P_{WHAT} is HR Then Trust is HT (DTH=1) |
| Rule 2: If P_{HOW} is HD and P_{WHY} is HUTR Then Trust is LT (DTH=0.93) |
| Rule 3: If P_{HOW} is HSM and P_{WHAT} is HR Then Trust is VHGT (DTH=1) |
| Rule 4: If P_{HOW} is MD and P_{WHY} is NTR and P_{WHEN} is MITM Then Trust is MT (DTH=0.98) |
| Rule 5: If P_{HOW} is MD and P_{WHY} is NTR and P_{WHEN} is HITM and P_{WHAT} is MIR and P_{WHO} is MDSML Then Trust is VLWT (DTH=1) |

Fig. 1. Fuzzy rules

Table 4. Rules Generated

| β | Total rules generated | No. of rules | | | |
|---------|-----------------------|--------------|-----|-----|-----|
| | | 2 | 3 | 4 | 5 |
| 0.9 | 721 | 16 | 161 | 343 | 201 |
| 1 | 477 | 8 | 61 | 207 | 201 |

In order to get better accuracy with minimum number of rules, the stopping criterion (β) is used. The value of β chosen is 0.9 and 1 and lengths of rules ranges from 2 to 5. The table 4 shows the number of rules generated. For example, rule #2 and #3 in figure 1 has length 2 and rule #5 has length 5. These fuzzy rules are taken as input to build a trust-boosted recommender system.

3.3 Trust-Boosted Recommender System

The proposed trust-boosted recommender system recommends the trustworthy users to the target user U_T is shown in figure 2. The major components of this recommender system model are:

- User Profile Learning
- Recommendation

The target user or requester (U_T) sends a query as a request (R_q) asking for recommendations from the trust network. This query is sent to the trusted network and it checks whether the U_T is new user or not. If U_T is the existing user, then recommend the highly trusted users. Otherwise the request is sent to the profile learner where profile data (P_{data}) are updated based on the query and existing profile information. Then this updated P_{data} is sent to the trust network. In the trust network, each reviewer is grouped based on trust levels $\langle \text{VHGT}, \text{HT}, \text{MT}, \text{LT}, \text{VLWT} \rangle$. From the set of fuzzy rules, extract the conditional attribute and the decision attributes. For each conditional attribute, generate fuzzy vector space (FVSP). The FVSP consists of a tuple $\langle \text{Attribute}, \text{Preference based Fuzzy Number} \rangle$. The vector similarity measures such as Jaccard, Dice and Cosine is carried out to find how much target user is similar to the others in trust network.

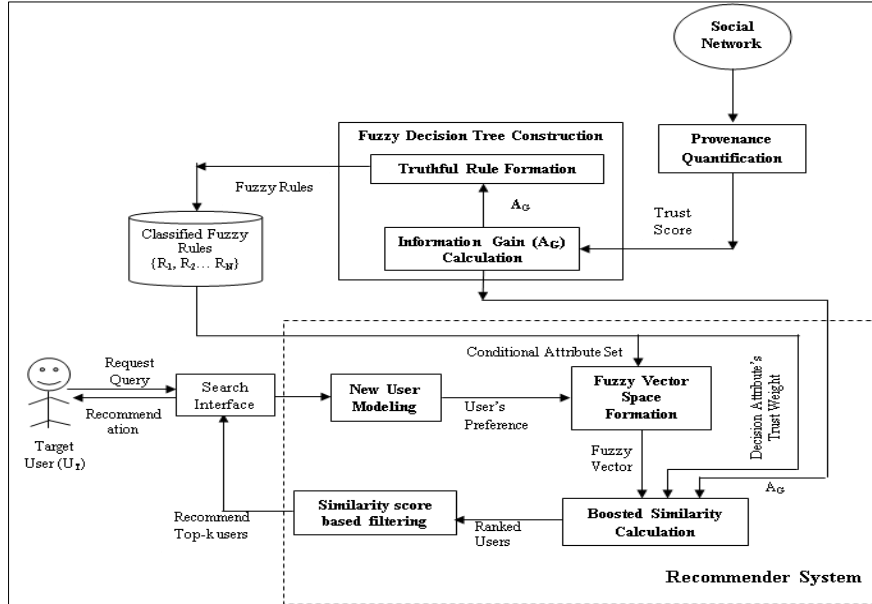


Fig. 2. Proposed Trust Boosted Recommender System

The gain value of each attribute (A_G) is assigned as a weight component and it is applied to the above mentioned three similarity measures and weighted similarity value is calculated. Then this similarity is boosted by the corresponding decision attribute's trust degree (trust level). Then, based on the boosted similarity value, the trusted users are ranked from highest to lowest. Finally, top-k users are recommended to the target user (U_T). After a recommendation, the target user's feedback is collected and profile learner will update the P_{data} accordingly. To collect the feedback, set of feedback query (FD_{qry}) is formulated based on five attributes P_{WHAT} , P_{WHEN} , P_{WHO} , P_{HOW} and P_{WHY} . For each FD_{qry} , users are asked to provide a quantitative value in the scale 0 to 1. This recommendation process is repeated for each user request with the updated profile.

User Profile Learning Phase.

In user modeling phase, the fuzzy rules are extracted from the rule database derived using fuzzy decision tree (discussed in section 3.2). Each user may have one or more than one rule as shown in figure 1. Using rule matching algorithm, each users is assigned matched rule(s). Then the users are grouped based on above mentioned five different trust levels. For example, if the user U_{109} ; U_{169} is classified as a Low Trust (LT) then under LT these users are grouped. For the rest of the users similar procedure is carried out and users are grouped accordingly. If U_T is an existing user then the details (profile) of the user are known already and can directly access the trust

network. If U_T is new user then profile of the user needs to be learned prior to access the trust network. The profile learning is depicted in figure 3.

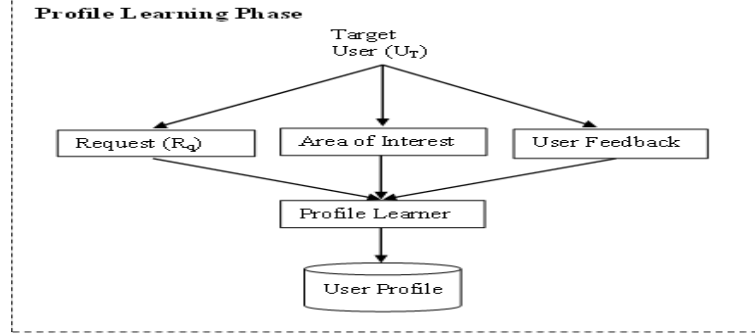


Fig. 3. Profile Learning Model

Initially U_T 's field of interest and training examples or already labeled items are collected and this forms a basic profile (P_b) of the user. The set of feedback (F_b) provided by the U_T for the items is also collected. Finally the (P_b, F_b, R_q) are combined and sent to the profile learner. Then the learned user profile is given as input to the recommendation phase.

Recommendation Phase.

The recommendation is carried out in two steps. One is creation of FVSP and the other is recommendation of the user.

Formation of FVSP.

The rules extracted from the trust network are partitioned into conditional attributes set (CAS) and decision attributes set (DAS). The CAS consists of all the trust attributes $\langle P_{WHAT}, P_{WHEN}, P_{WHO}, P_{HOW}, P_{WHY} \rangle$. The DAS consists of trust decision $\langle VLWT, LT, MT, HT, VHGT \rangle$.

Step 1: For each attribute in the conditional attribute set, assign attribute grade. This is based on the position of the triangular fuzzy function and is given in table 5. The linguistics space of each attribute is given below.

$$Linguistic\ Space = \left(\begin{array}{l} P_{WHAT} = [HIR, MIR, NR, MR, HR] \\ P_{HOW} = [HSM, MSM, NSM, MD, HD] \\ P_{WHEN} = [HITM, MITM, NETM, METM, HETM] \\ P_{WHY} = [HTR, MTR, NTR, MUTR, HUTR] \\ P_{WHO} = [HDSML, MDSML, NDSML, MSML, HSML] \end{array} \right)$$

Table 5. Fuzzy Number for the attribute P_{WHAT}

| Grade | Fuzzy Number |
|-------|--------------------|
| 1 | (0.0, 0.0, 0.25) |
| 2 | (0.0, 0.25, 0.50) |
| 3 | (0.25, 0.50, 0.75) |
| 4 | (0.50, 0.75, 1.0) |
| 5 | (0.75, 1.0, 1.0) |

For example, in P_{WHAT} attribute the position of 'HIR' has low grade, i.e. 1 and 'HR' has high grades, i.e. 5. Similarly, in P_{WHY} attribute the position of 'MTR' (Moderately Truthful") has medium grade, i.e. 3.

Step 2: Assign the fuzzy number for each linguistic term based on the grade. Since it follows the triangular fuzzy logic, the fuzzy number assigned for each grade is shown below. For example, the fuzzy number for the linguistic term for the attribute P_{WHAT} is shown in table 6. For other attributes, fuzzy number is same as that of shown in table 5.

Table 6. Fuzzy Number for the attribute P_{WHAT}

| P_{WHAT} Linguistic Term | Fuzzy Number |
|-----------------------------------|--------------------|
| HIR (Highly Irrelevant) | (0.0, 0.0, 0.25) |
| MIR (Moderately Irrelevant) | (0.0, 0.25, 0.50) |
| NR (Neutrally Relevant) | (0.25, 0.50, 0.75) |
| MR (Moderately Relevant) | (0.50, 0.75, 1.0) |
| HR (Highly Relevant) | (0.75, 1.0, 1.0) |

Step 3: The fuzzy number for each attribute is now represented as a vector in FVSP. The FVSP for each rule is represented as a pair $\{<A_K, FN_{AK}>\}$. where,

- K refers to a number of attributes, here $K=5$.
- A_K represents the current attribute and
- FN_{AK} refers to the fuzzy number for the specified attribute A_K . That is $FVSP = \{<A_1, (a_{11}, a_{12}, a_{13})>, <A_2, (a_{21}, a_{22}, a_{23})> \dots <A_5, (a_{51}, a_{52}, a_{53})>\}$. Here (a_{11}, a_{12}, a_{13}) is a triplet used in TMF to define the fuzzy number where $0 \leq a_{11} \leq a_{12} \leq a_{13} \leq 1$.

For example FVSP for the rule1 shown in figure 1 is given below. $\{<P_{\text{HOW}}, (0.25, 0.50, 0.75)>, <P_{\text{WHY}}, (0.50, 0.75, 1.0)>, <P_{\text{WHEN}}, (0.50, 0.75, 1.0)>, <P_{\text{WHAT}}, (0.50, 0.75, 1.0)>, <P_{\text{WHO}}, (0.75, 1.0, 1.0)>\}$. This FVSP is taken as input to calculate the vector similarity and to suggest the top-k trustworthy users.

Recommendation of the top-k users.

In the vector space there are some similarity measures between two vectors which have been successfully applied in fields such as pattern recognition, classification of

complex objects and other decision making problems. The vector similarity measures chosen in the proposed recommendation system are the Cosine similarity. Using this FVSP, above vector based similarity measure is carried out to find how much U_T is similar to the other users in trust network. The gain value of each attribute (A_G) is taken as a weight component and it is applied to the above mentioned measures and similarity value is calculated.

Let $X = U_T = (a_1, a_2, a_3)$ and $Y = U_N = (b_1, b_2, b_3)$ is the fuzzy number of the target user (U_T) and the other user (U_N) from the trust network respectively, then the cosine similarity measure is given in equation (1) is as follows.

$$S = \text{Cosine}(U_T, U_N) = \sum_{k=1}^5 A_{Gk} \frac{\sum_{f=1}^3 (FN_{ATkf} \cdot FN_{ANKf})}{\sqrt{\sum_{f=1}^3 (FN_{ATkf}^2)} \cdot \sqrt{\sum_{f=1}^3 (FN_{ANKf}^2)}} \quad (1)$$

where,

- A_G represents the attribute gain,
- f represents the fuzzy number of values in each fuzzy number,
- a_1, a_3, b_1 and b_3 are the endpoints of fuzzy numbers,
- a_2 and b_2 are the peak point of fuzzy numbers

After finding the similarity (S), boost this value by a corresponding trust score of the user U_N given in equation (2). Boosting is linear, since it is done with associated trust level. Using this boosted similarity (S_b), prediction of the target user's trust score is carried out. The prediction formula is given in equation (3).

$$S_b = S * S^{Twt} \quad (2)$$

$$\text{Pred}(U_T, I_j) = \begin{cases} tr_{U_T}, & \text{if } S_b = 0 \text{ or if } tr_{U_N, I_j} = \overline{tr_{U_N}} \\ tr_{U_T} + \frac{\sum_{U_N \in NB} S_b(U_N, U_T) \times (tr_{U_N, I_j} - \overline{tr_{U_N}})}{\sum_{U_N \in NB} |S_b(U_N, U_T)|}, & \text{else} \end{cases} \quad (3)$$

where,

- T_{wt} refers to trust weight assigned based on the trust level of user U_N . (For e.g., VHGT has T_{wt} of 1 and MT has T_{wt} of 0.6 and VLWT has T_{wt} of 0.2)
- tr_{U_T} represents the trust value of the Target User U_T presented in fuzzy number format as shown in Table 6.
- I_j represents items (books) which are not given any review
- NB represents the number of neighbors chosen

Consider the randomly chosen reviewer say reviewer 72 (R_{72}) requesting for the recommendation of k users (Let $k=10$). The similarity (S) between the requester and the rest of the users is calculated. Then it is boosted using equation (2). The table 7 shows the similarity and boosted similarity (S_b) score of the top- k reviewer. The reviewers are sorted based on similarity from highest to lowest. Though both similarities show the highest score for the top reviewers, the trust level differs. The trust level of highly matched reviewer with R_{72} is 'HT'. Therefore, the top-10 reviewers are

expected to have the trust level of ‘HT’. But, in case of without boosting, top 4th, 6th and 10th reviewers have other trust level (‘MT’) instead of ‘HT’. Similarly, in case of boosting, the top 10th reviewer has different trust level.

Table 7. Similarity and Boosted similarity score of top-k reviewer.

| Without Boost | | | With Boost | | | Top-k reviewer |
|-----------------|-------|-------------|-----------------|----------------|-------------|----------------|
| Reviewer Number | S | Trust Level | Reviewer Number | S _b | Trust Level | |
| 722 | 1 | HT | 28 | 1 | HT | 1 |
| 600 | 1 | HT | 29 | 1 | HT | 2 |
| 335 | 1 | HT | 722 | 1 | HT | 3 |
| 251 | 1 | MT | 600 | 1 | HT | 4 |
| 216 | 1 | HT | 335 | 1 | HT | 5 |
| 212 | 1 | MT | 165 | 1 | HT | 6 |
| 194 | 1 | HT | 177 | 1 | HT | 7 |
| 177 | 1 | HT | 194 | 1 | HT | 8 |
| 165 | 1 | HT | 216 | 1 | HT | 9 |
| 633 | 0.998 | MT | 158 | 0.998 | MT | 10 |

Therefore, the prediction error is more in without boosting and lesser in boosted method. Boosting the similarity appropriately ranks the top-k reviewers. This way the proposed system gets a reduced MAE and RMSE.

4 Experiments and Result Analysis

To evaluate the performance of gain weighted trust boosted recommender system, experiments are conducted on the popular book based social network Goodreads data set. The aim of these experiments is to present a comparative study of proposed recommendation strategy in fuzzy trust concept. Also proposed trust boosted model is evaluated against other weight strategies. The performance of the proposed recommendation strategy is measured with respect to quality of predictions and quality of recommendations. The quality of prediction is done by measuring Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE). Similarly the quality of recommendation is done by measuring precision@k and recall@k and Average Precision (AP). Leave-one-out method is used to evaluate recommendation systems [14]. This technique involves withholding one rating and trying to predict it with remaining ratings. Then the predicted rating can be compared with the actual rating and the difference will be considered as the prediction error.

4.1 Evaluation of different weight strategies

The different weight strategies considered for evaluation are expected weight method; preference based method [20] and proposed trust boosted method. The MAE, RMSE and AP measures are evaluated for the above mentioned weight strategies. The figure 4 shows the MAE value obtained for the cosine similarity method. From this

figure, it is observed that the proposed trust boosted method shows the less prediction error (MAE) than the other two methods. Similarly, figure 5 shows the RMSE value obtained for the cosine similarity measure. From the figure it is observed that the proposed trust boosted method shows the less prediction error (RMSE) when compared with the expected weight method. The preference based method shows more error rate than the other two methods.

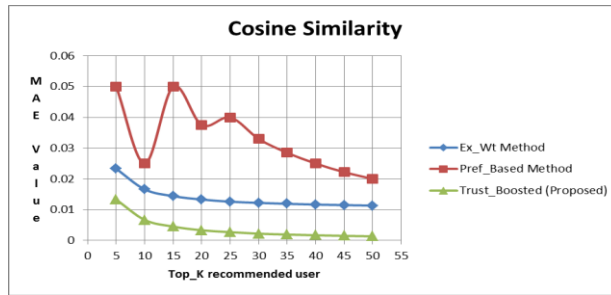


Fig. 4. MAE for Cosine Similarity Measure

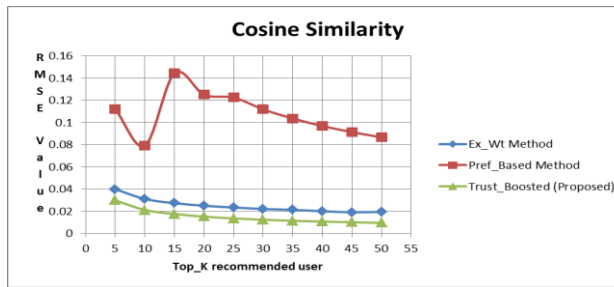


Fig. 5. RMSE for Cosine Similarity Measure

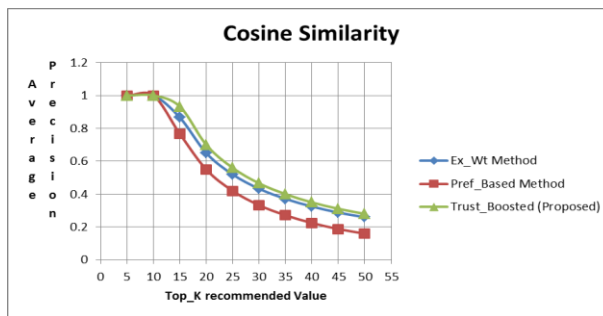


Fig. 6. Average Precision for Cosine Similarity Measure

The AP value is shown in figure 6 for the above similarity measure. The precision value for the proposed method is higher than the other two methods. The AP is almost

same for top-5 and top-10 users. Up to top 20 users precision value is greater than or equal to 0.90. After that the precision value is start decreasing and for top-50 user, the precision value is very less in preference based method.

4.2 Comparison With Other Trust-Based Recommender System

The proposed recommender system is compared with other trust-based recommender system. The evaluation is done on MAE and RMSE. First, the proposed method (boosted similarity) is compared against without boosting the similarity. The MAE of this is shown in figure 7. The experiment is carried out with Jaccard, Dice and Cosine similarity measure. All these three measures show the lesser prediction error while boosting the similarity than without boosting the similarity. In Jaccard repetition of a word does not reduce the similarity but Cosine measure reduce the similarity.

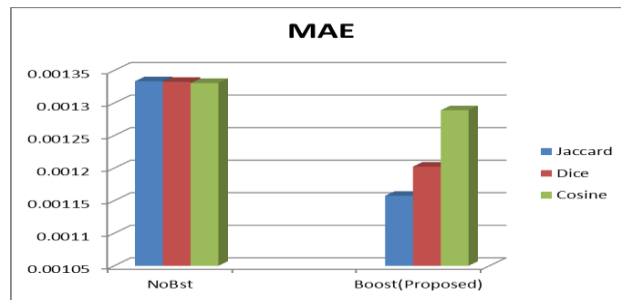


Fig. 7. MAE for with and without boost

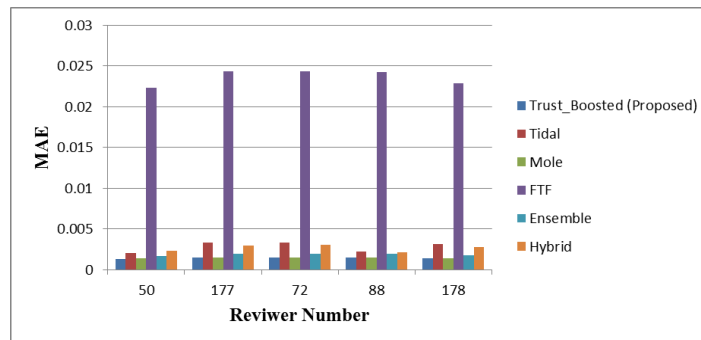


Fig. 8. MAE compared with existing trust-based recommender system

The figure 8 and 9 shows the MAE and RMSE values of the proposed approach with existing trust-based recommender system respectively. The reviewers are chosen through random sampling. The existing approaches considered are Tidal trust, Mole trust, Fuzzy Trust Filtering, Ensemble and Hybrid. The MAE value is checked for few randomly selected reviewers. The graph shows the reduced prediction error in the proposed method.

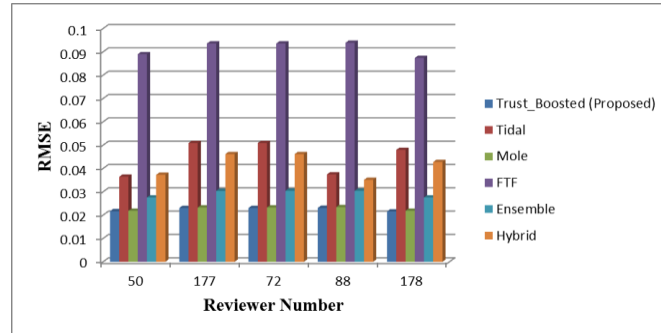


Fig. 9. RMSE compared with existing trust-based recommender system

Similar to MAE, the RMSE value is checked with few randomly selected reviewers. The graph shows the reduced prediction error in the proposed method.

5 Conclusion and Future Work

In this work, trust-boosted recommender system is designed to recommend top-k reviewers of the book based social network. The use of provenance based trust computation from multiple aspects has improved the recommendation quality. Also performance of proposed trust boosted (the gain as weight) measure is compared with other weights such as expected value, and preference based method. The analysis shows that the precision@k is increased 10.166% when compared to the expected weight method and 2.186% when compared with preference based weight method. Also proposed approach is compared with other trust based methods and the results shows that the prediction is achieved with minimum MAE and RMSE. The future work is to recommend the top-k reviewers to a group of users. That is to develop a group recommender system

Acknowledgements

This Publication is an outcome of the R&D work undertaken in the project under the Visvesvaraya PhD Scheme (Unique Awardee Number: VISPHD-MEITY-2959) of Ministry of Electronics & Information Technology, Government of India, being implemented by Digital India Corporation (formerly Media Lab Asia).

References

1. Abdul-Rahman, A., Hailes, S.: Supporting trust in virtual communities. In: Proceedings of the 33rd Hawaii International Conference on System Sciences, pp. 1769-1777, IEEE, Maui, HI, USA (2000).
2. Andersen, R., Borgs, C., Chayes, J., Feige, U., Flaxman, A., Kalai, A., Mirrokni, V., Tenenholz, M.: Trust-based recommendation systems: An axiomatic approach. In: Proceed-

- ing of the 17th International Conference on World Wide Web, pp.199-208, ACM, New York (2008).
3. Avesani, P., Massa, P., Tiella, R.: Moleskiing.it a trust-aware recommender system for ski mountaineering. *International Journal for Infonomics*, 1-10, (2005).
 4. B.Smyth and J.O'Donovan : Trust in recommender systems. In: 10th international conference on intelligent user interfaces, pp.167-174, ACM, San Diego, California, USA (2005).
 5. C.N. Ziegler and G. Lausen: Propagation models for trust and distrust in social networks. *Information Systems Frontiers*, 7 (4-5), 337-358, (2005).
 6. Cristina Olaru, Louis Wehenkel: A Complete Fuzzy Decision Tree Technique. *Journal of Fuzzy set and Systems*, 138, 221-254, (2003).
 7. Dhanalakshmi Teekaraman, Sendhilkumar S, Mahalakshmi G.S.: Semantic Provenance Based Trustworthy User Classification on Book-Based Social Network using Fuzzy Decision Tree. *International Journal of Uncertainty, Fuzziness and Knowledge- Based Systems*, 28 (1), 47-77, (2020).
 8. Falcone, R., Pezzulo, G., Castelfranchi, C: A fuzzy approach to a belief-based trust computation. In: Eder, J., Haav, H-M., Kalja, A., Penjam, J. (eds.) *Lecture Notes in Artificial Intelligence*, vol.2631, pp.73-86, 2003.
 9. Golbeck, J: Computing and applying trust in web-based social networks, PhD Dissertation, University of Maryland at College Park College Park, MD, USA, 2005.
 10. Hang, W. C , Singh, M. P.: Trust based recommendation based on graph similarities, 2010, <http://www.csc.ncsu.edu/faculty/mpsingh/papers/mas/aamas-trust-10-graph.pdf>.
 11. 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, pp.93-102, ACM, Seattle, Washington, USA (2012)
 12. Josang, A.: A logic for uncertain probabilities. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 9, pp.279-311 (2001).
 13. Krzysztof Janowicz : Trust and Provenance You Can't Have One without the Other. Technical Report, Institute for Geoinformatics, University of Muenster, 2009.
 14. Massa, P., Avesani, P.: Trust-aware recommender systems. In: proceedings of the first ACM Conference on Recommender Systems, pp.17-24, ACM, Minneapolis, USA (2007).
 15. Nicholas, J.J.Smith: Vagueness and degrees of truth. Oxford scholarship, 2009.
 16. P. Victor, C. Cornelis, M. De Cock, and P. Pinheiro da Silva: Gradual trust and distrust in recommender systems. *Fuzzy Sets and Systems*, 160(10), 1367-1382 (2009).
 17. P.Victor, C.Cornellis, M.D.Cock and A.M.Teredesai: Trust and distrust-based recommendations for controversial reviews. *IEEE intelligent systems*, 26, 48-55 (2011).
 18. R. Guha, R. Kumar, P. Raghavan, and A. Tomkins: Propagation of trust and distrust. In proceedings of the 13th international conference on World Wide Web, pp.403-412, ACM, New York, USA (2004).
 19. S. Mitra, K.M. Konwar, S.K. Pal: Fuzzy decision tree, linguistic rules and fuzzy knowledge-based network generation and evaluation. *IEEE Transaction on System, Man, and Cybernetics, Part C (Applications and Reviews)*, 32(4), 328-339 (2002).
 20. Samia Boulkrinat, Allel Hadjali, Aicha Mokhtari: Towards recommender systems based on a fuzzy preference aggregation. In: 8th conference of the European Society for Fuzzy Logic and Technology, Atlantis Press, University of Milano-Bicocca, Milan, Italy (2013).
 21. Sinha, R.R., Swearingen, K.: Comparing recommendations made by online systems and friends. In: DELOS Workshop, Personalization and Recommender Systems in Digital Libraries, Dublin, Ireland (2001).
 22. Sudha Ram, Jun Liu: Book on Active conceptual modeling of learning. pp.17-29, Springer-Verlag Berlin, Heidelberg (2007).