IJMTARC – VOLUME – VI – ISSUE – 24, OCT-DEC, 2018

ISSN: 2320-1363

SHARED FILTERING-BASED RECOMMENDATION OF ONLINE SOCIAL NETWORK VOTING SYSTEM

KAVURI ANITHA LAKSHMI[1], MR.B.SURESH[2]

[1]PG Scholar, Dept of CSE, Srinivasa Institute of Engineering &Technology, Cheyyeru, Amalapuram-A.P, India

[2]Associate Professor, Dept of CSE, Srinivasa Institute of Engineering & Technology, Amalapuram-A.P, India

Abstract- Social voting is a developing new element in online informal organizations. It postures one of a kind difficulties and open doors for suggestion. In this paper, we build up an arrangement of grid factorization (MF) and closest neighbor (NN)- based recommender frameworks (RSs) that investigate client interpersonal organization and gathering connection data for social voting proposal. Through analyses with genuine social voting follows, we exhibit that interpersonal organization and gathering connection data can fundamentally enhance the exactness of prevalence based voting suggestion, and interpersonal organization data rules aggregate association data in NN-based methodologies. We likewise watch that social and gathering data is substantially more significant to chilly clients than to overwhelming clients. In our tests, basic metapathbased NN models outflank calculation serious MF models in hot-voting proposal, while clients' interests for no hot voting can be better mined by MF models. We additionally propose a half and half RS, stowing diverse single ways to deal with accomplish the best k hit rate.

1. INTRODUCTION

ONLINE interpersonal organizations (OSN, for example, Facebook and Twitter, encourage simple data sharing among companions. A client not exclusively can share her updates, in types of content, picture, and video, with her immediate companions, yet additionally can rapidly scatter those updates to a considerably bigger group of onlookers of roundabout companions, utilizing on the rich availability and worldwide reach of well known OSNs. Numerous OSNs currently offer the social voting capacity, through which a client can impart with companions her insights, e.g., like or aversion, on different subjects, running from client statuses, profile pictures, to recreations played, Manuscript got October 22, 2014; modified August 19, 2016; acknowledged January 24, 2017 items acquired, sites went to, et cetera. Taking like– detest kind of voting's above and beyond, some OSNs, e.g., Sina Weibo [20], engage clients to start their own particular voting efforts, on any theme of their interests, with utilize altered voting choices. The companions of a voting initiator can partake in the crusade or rewet the battle to their companions. Other than fortifying social associations, social voting additionally has numerous potential business esteems. Advertisers can start voting's to promote certain brands. Item directors can start voting's to lead statistical surveying. Web based business proprietors can deliberately dispatch voting's to draw in more online clients. The expanding notoriety of social voting quickly delivers the "data over-burden" issue: a client can be effortlessly overpowered by different voting's that were started, taken part, or withdrew by her immediate and roundabout companions. It is basic and testing to display the "right voting's" to the "right clients" in order to enhance client encounter and amplify client commitment in social voting's. Recommender frameworks (RSs) manage data over-burden by proposing to clients the things that are possibly of their interests. In this paper, we show our ongoing exertion on creating RSs for online social voting, i.e., prescribing fascinating voting efforts to clients. Unique in relation to the customary things for proposal, for example, books and motion pictures, social voting's engendering along social connections. A client will probably be presented to a voting if the voting was instated, taken part, or retreated by her companions.

1) Online social voting has not been highly researched as far as anyone is concerned. We create MF-based and NN-based RS models. We appear through analyses with genuine social voting follows that both interpersonal organization data and gathering alliance data can be mined to fundamentally enhance the precision of fame based voting proposal.

2) Our examinations on NN-based models propose that interpersonal organization data commands bunch connection data. Also, social and gathering

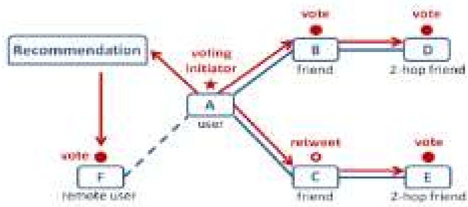
IJMTARC – VOLUME – VI – ISSUE – 24, OCT-DEC, 2018 ISSN: 2320-1363

data is more profitable to frosty clients than to substantial users.

3) We demonstrate that basic metapath-based NN models beat calculation escalated MF models in hot-voting suggestion, while clients' interests for nonhot votings can be better mined by MF models. Whatever remains of this paper is sorted out as takes after. Area II exhibits the related work. We give a brisk outline on the social voting capacity of Sina Weibo and present estimation consequences of our informational index in Section III. In Section IV, we initially build up a multichannel MF demonstrates that at the same time mines client voting, user– client, and client amass data. We at that point propose a few NN models in light of various met ways in the heterogeneous data organize. Exploratory outcomes are displayed in Section V. This paper is finished up in Section VI.

II. RELATED WORK

In this project, we characterize clients with under five voting's as cool clients and in excess of ten voting as substantial clients. We characterize voting that draw in no under 1000 clients as hot voting's and under 10 clients as frosty voting. to factorize client thing rating grid and user– client relationship grid together for thing rating forecast. Mama et al. [33] asserted that a client's appraising of a thing is affected by his/her companions. A client's evaluating to a thing comprises of two parts, the client's own particular rating of the thing and the client's companions' evaluations of the thing. The creators at that point proposed to join the two appraisals directly to get a last anticipated rating. Jamal and Ester [31] guaranteed that a client's advantage is affected by his/her companions. Along these lines, a client's idle component is compelled to be like his/her companions' inactive includes during the time spent MF. Yang et al. [30] asserted that a client's advantage is multifaceted and proposed to part the first interpersonal organization into circles. Contrast circles are utilized to anticipate evaluations of things in various classifications. Jiang et al. [3] tended to using data from numerous stages to comprehend client's needs thoroughly. Specifically, they proposed a semi supervised exchange learning strategy in RS to address the issue of cross-stage conduct expectation, which completely misuses the modest number of covered group to connect the data crosswise over various stages. Jiang et al. [39] thought about improving data for precise client thing join forecast by speaking to a social organize as a star-organized crossover chart focused on a social area, which interfaces with other thing spaces to help enhance the forecast precision. In addition, setting mindfulness is likewise a critical measure to encourage suggestion. For illustration, Sun et al. [40] proposed a community oriented now casting model to perform setting mindful proposal in portable computerized associates, which models the convoluted relationship inside relevant signs and amongst setting and plan to address sparsely and heterogeneity of logical signs. examined the substance data on location based informal communities as for purpose of-intrigue properties, client interests, and conclusion signs, which models three kinds of data under a bound together purpose of-intrigue proposal system with the thought of their relationship to registration activities. Interestingly, online social votings are very not quite the same as the customary suggestion things as far as social proliferation. Not the same as the current social-based RSs,



Into NN-based top-*k* RSs. Trust-CF calculates the predicted rating for a candidate item as the weighted average of all observed ratings in the traditional CF neighborhood and social neighborhood. Trust-CF does not work with binary data set, as the weighted average of all observed items is 1. Yang *et al.* [14] proposed Trust-user latent feature space-based collaborative filtering approach (Trust-CF-ULF) to incorporate social network information into top-*k* RSs. *Trust-CF-ULF approach* is the combination of CF-ULF and social network based approach. Using metapath-based approaches, we consider a wider set of neighborhoods than [14], which can be treated as a special case of our hybrid NN approaches. Social voting as a new social network application has not been studied much in the existing literature. Compared with traditional items for recommendation, the uniqueness of online social voting lays in its social propagation along social links. Also, the purpose of initializing a voting is to engage people to express their opinions. Thus, the topics covered in online social votings are generally more engaging than other applications in

ISSN: 2320-1363

OSNs. Section III presents some interesting statistics of our online social voting data trace.

III. SOCIAL VOTING

Weibo (the Chinese word for "microblog") is a cross breed of Twitter and Facebook-like social application propelled by the Sina company, China's greatest Web gateway, in August 2009. Starting at 2013, it had amassed in excess of 600 million enlisted clients and more than 120 million day by day dynamic clients 2016 [21]. Clients on Weibo take after each other. A client can compose posts (tweets) and offer them with his supporters. Clients can likewise join diverse intrigue bunches in view of their geographic/statistic highlights and interests of points. Voting [22] is an inserted highlight of Sina Weibo. More than 92 million clients have taken part in different votes on Weibo as of January 2013. There are in excess of 2.2 million continuous voting's accessible on Sina Weibo every day. As appeared after a client started or took an interest in a voting, all his/her supporters can see the voting; client can likewise pick just rewet a voting to his devotees without cooperation. The other route is through Weibo voting suggestion list, which comprises of prevalent voting and customized suggestion. We have no data about Weibo's voting suggestion calculations.

TABLE I

GENERAL STATISTICS OF WEIBO DATA SET

Users	1,011,389	Social Relations	83,636,677
Votings	185,387	User-Groups	5,643,534
Groups	299,077	User with Votings	525,589
User-Votings	3,908,024	User with Groups	723,913

A. Estimation Study

We got client voting logs straightforwardly from the specialized group of Sina Weibo.2 The informational index covers votings from November 2010 to January 2012. The informational index has definite data about votings every client took an interest in, voting substance, and the end time of each voting. We just know client voting interest, not client voting comes about, i.e., we do not know which voting choice a client picked. The informational index moreover contains social associations amongst clients and gatherings a client joined. The informational index just contains bidirectional social connections, i.e., A takes after B and B takes after A.

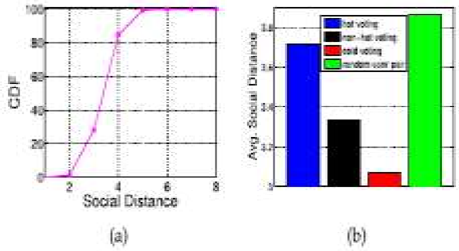


Fig. 3. (a) Distribution of random user pair's social distance. (b) Average social distance between user pairs in hot, nonhot, and cold votings versus random user pairs.

Social separation among the sets between these sources and every single other client. The normal social separation is 3.86 bounces. At that point, we compute the social separations for clients taking an interest in a similar voting. We arrange the votings as indicated by the number of members and separated them into three sorts: hot votings, with no under 1000 clients; no-hit votings with fewer than 1000 clients; and cool votings with fewer than 10 clients. We haphazardly select 100 hot votings, no hot votings, and cool votings independently. The normal social separations between clients in various sorts of votings are appeared in Fig. 3(b), and the normal separations are 3.71, 3.34, and 3.07 for hot, no hot, what's more, chilly votings, separately. Of course, clients took part in a similar voting are socially nearer than haphazardly chosen clients. Better known social votings proliferate further in the fundamental informal community, and their members can be more distant far from each other than less prominent votings. In our Weibo informational collection, for a voting took an interest by a client, there is a 33.9% shot that no less than one of his 1-bounce followee has taken an interest in the voting, and 80.6% possibility that at any rate one of his followees inside two bounces has taken an interest in the voting, and the number for followees inside 3-bounce is 96.4%. For correlation, the relating numbers for normal thing appropriation in the Opinions' informational collection [25] are: 28.7%, 61.9%, also, 76.9%, separately. Pinions is a purchaser sentiment site where clients survey different things, for example, autos, films, books, programming, et cetera, and allocate evaluations to the things. Clients moreover dole out put stock in values (i.e., an estimation of 1) to different clients whose audits as well as evaluations they discover significant. It is additionally fascinating to ponder the relationship between's social votings and social gatherings. We see from the Weibo information that among clients taking an interest in a same voting, 10.40% of the

ISSN: 2320-1363

client sets joined no less than one regular gathering, while just 0.92% of arbitrarily chose client sets joined no less than one normal gathering. It demonstrates that clients in a same gathering share comparable voting interests. In this manner, we will likewise contemplate how much gathering data can enhance social voting proposal in this paper.

IV. SOCIAL VOTING RECOMMENDATION

We consider top-k voting proposal in OSNs. For every client, the RS needs to suggest a modest number, say k, of votings from every single accessible voting. We present execution measurements for top-k proposal in Section IV-A. MF techniques were observed to be extremely proficient when all is said in done best k proposal [1], [2]. Besides, interpersonal organization data can be abused to enhance the exactness of best k proposal [4], [6]. Therefore, we begin with MF approaches utilizing both interpersonal organization data and gather association data. In Section IV-B, we propose a multichannel MF demonstrate, which factorizes client voting between actions, user–user interactions, and user-group interactions simultaneously, gearing to optimize top-*k* hit rate. Other than MF approaches, we also consider NN approaches in Section IV-C. We first construct neighborhoods by traversing different types of metapaths in the Weibo heterogeneous information network. We then explore user neighborhoods in the latent feature space derived from MF models.

A. Performance Metrics

Recall or top-*k* hit rate is widely used in evaluating RSs. To compute the top-*k* hit rate, we rank the items $i \in I$ according to their predicted rating $\hat{R}_{u,i}$ for each user $u \in U$. An item is defined as *relevant* to a user in the test set if she/he finds it appealing or interesting, e.g., the rating value is above a certain threshold. In our experiments with Weibo data, the real rating values are binary *(*0, 1*)*, and we consider 1

As relevant. The top-*k* hit rate or recall of user *u* is defined as the fraction of relevant items in the test set that appear in the top-*k* of the ranking list, denoted by *N(k, u)*, from among all relevant items, *N(u)*. Similar to [12], the recall over all users is computed as follows:

$$\text{recall} = \frac{\sum_u N(k,u)}{\sum_u N(u)}. \qquad (1)$$

Note that a higher top-*k* hit rate or recall is better. We us recall as the evaluation metric in our experiments.

B. Multichannel Matrix Factorization

The social network information is represented by a matrix $S \in R^{u0\times u0}$, where *u*0 is the number of users. The directed and weighted social relationship of user *u* with user *v* (e.g., user *u* trusts/knows/follows *v*) is represented by a positive value $S_{u,v} \in (0, 1]$. An absent or unobserved social relationship is reflected by $S_{u,v} = s_m$, where typically $s_m = 0$. The user-group affiliation information is represented by matrix $G \in R^{u0\times n0}$, where $G_{u,n}$ is binary and takes value 1 if user *u* joins group *n*, and 0 otherwise.

1) Weibo-MF Model: The graphic model of Weibo-MF is shown in Fig. 4. The user-voting interaction $R_{u,i}$ is determined by user latent feature Q_u and voting latent feature P_i, usergroup interaction $G_{u,n}$ is determined by user latent feature Q_u and group latent feature Y_n, and user–user interaction $S^*_{u,v}$ is determined by user latent feature Q_u and factor feature Z_v .

Similar to [2], we normalize the social network matrix *S* to incorporate local authority and local hub values

$$S^*_{u,v} = S_{u,v}\sqrt{\frac{d_v^-}{d_u^+ + d_v^-}}$$

Where d^+_u is the out-degree of user *u* in the social network (i.e., the number of users whom *u* follows/trusts), and d^-_v is the in-degree of user *v* in the network (i.e., the number of users who follow/trust user *v*). The predicted rating of user *u* for item *i* is a function of *u*'s latent feature Q_u and *i* 's latent feature P_i

$$\hat{R}_{u,i} = r_m + Q_u P_i^T \qquad (2)$$

with matrices $P \in \mathbb{R}^{i_0\times j_0}$ and $Q \in \mathbb{R}^{u_0\times j_0}$, where $j_0 \ll i_0$, u_0 is the rank; and $r_m \in \mathbb{R}$ is a (global) offset. Besides the rating data, the social network information is also used in model training. The social relationships between users are predicted as follows:

$$\hat{S}^*_{u,v} = s_m + Q_u Z_v^\top \qquad (3)$$

where $Z \in R^{u0\times j0}$ is a third matrix in this model, besides *P* and *Q*. The row vector Z_v denotes factor specific latent feature vector of user *v*. Ma *et al.* [32] provide more detailed description of matrix *Z*. Note that the matrix *Z* is not needed predicting rating values, and, hence, may be discarded after the matrices *P* and *Q* have been learned.Another matrix *G* is used for factorization. $G_{u,n}$ is the affinity of user *u* to group *n*. Typically, the affinity

IJMTARC – VOLUME – VI – ISSUE – 24, OCT-DEC, 2018

ISSN: 2320-1363

value is binary, i.e., user u belongs to a group n or not. Group affinity values are predicted as

C. Nearest-Neighbor Methods

Other than MF approaches, NN-based recommendations have also been studied. NN methods are widely used in RSs [3], [4], [6]. Thus, it is very intriguing to study the performance of NN models on social voting recommendation problem. In NN-based approaches, the neighborhood of a user can be calculated using collaborative filtering, or it can be a set of directly or indirectly connected friends in a social Other than MF approaches, NN-based recommendations have also been studied. NN methods are widely used in RSs [4], [4], [6]. Thus, it is very intriguing to study the performance of NN models on social voting recommendation problem. In NN-based approaches, the neighborhood of a user can be calculated using collaborative filtering, or it can be a set of directly or indirectly connected friends in a social.

Algorithm 1 Algorithm of Weibo-MF Model

```
Data: Sina Weibo voting dataset
Result: Top-k Hit Rate
// Training part
Load sina weibo voting training data;
Initialize latent feature matrices Q and P;
// Update latent features by ALS
while Not Converge & Iteration Number is less than Iter_Num do
    Update Q by fixing P and minimizing Eq. (5);
    Update P by fixing Q and minimizing Eq. (5);
end
// Testing part
for each user u in Sina Weibo voting dataset for testing do
    for each voting i in test dataset for user u do
        Calculate the predicted rating of user u on voting i
        as R̂_{u,i} = r_m + Q_u P_i^T ;
        Put R̂_{u,i} into the queue recomm_pool;
    end
    Sort recomm_pool in an decreasing order according
    to the value of R̂_{u,i};
    Select foremost K votings with largest R̂_{u,i} from
    recomm_pool as the items for recommendation;
    Calculate top-k hit rate for user u;
end
Return average top-k hit rate for entire system;
```

Network or just a set of users with similar interests in a same group. This makes it convenient to incorporate social trust and user-group interaction into NN-based top-k recommendation. In this section, we try different approaches to construct nearest neighborhood for a target user.

1) Metapath Neighborhoods:

In heterogeneous information networks, objects are of multiple types and are linked via different types of relations or sequences of relations, forming a set of metapaths [15]. Metapath is a path that connects objects of different types via a sequence of relations. Different metapaths have different semantics. Sun *et al.* [16] employ metapaths for clustering task in heterogeneous information networks. In this paper, we use metapaths for recommendation task. In this paper, we leverage the idea of metapath [15] to construct nearest neighborhoods for target users. Different from [15], the starting object type in a metapath is user, and the ending object type is voting. Fig. 5(a) shows the schema of Weibo heterogeneous information network. It contains three types of objects, namely, user (U), voting (V), and group (G). Links exist between a user and a voting by the relation of “vote” and “voted by,” between a user and a group by “join” and “joined by,” between a user and another user by “follow” and “followed by.” We consider a set of different metapaths for the purpose of NN voting recommendation. Fig. 5(b)–(d) shows different metapaths. The solid lines between users social connections; the dashed lines between users and groups are user-group interactions, i.e., a user joins a group; the dashed lines between users and votings are user-voting activities, i.e., a user participates in a voting. In Fig. 5(b)–(d), the red highlighted lines compose the metapaths, and the starting object of metapaths is $U1$.

a) UGUV metapath: As shown in Fig. 5(b), the semantic of using $U - G - U - V$ metapath for recommendation is finding users that in a same group with the target user, then recommending their votings to the target user. More specifically, *UGUV* works as follows.

1) For a target user u, UGUV searches for all the groups that u has joined. Denote the set of groups as Gu.

2) For each joined group $g \in Gu$, search for all the users that belong to group g.

3) Users in group g report their relevant votings.

4) Combine the reports of all groups. The score for a candidate voting i to the target user u is computed as

$$\mathrm{Score}_{u,i} = \sum_{g\in G^u}\sum_{v\in g}\sum_{i} w(g)\delta_{i\in I_v} \tag{9}$$

where δ is the Kronecker delta, Iv denotes the set of user v’s relevant votings, and $w(g)$ is the weight of users in group g. In our later experiments, we try $w(g)$ as a function of group size. We found that the best function of $w(g)$ is to simply set $w(g) = 1$.

5) Rank recommended votings according to their scores, and return the top-*k* votings.

b) UUV(m-hop) metapath: As shown in Fig. 5(c), the semantic of *U −U −V (m −hop)* metapath-based recommendation is to recommend a target user the relevant votings of his followees within m-hops. *UUV* approach employs the BFS in social network to find users similar to the target user *u*. The scoring scheme is similar to the scheme employed in *UGUV*

$$\text{Score}_{u,i} = \sum_{v\in N_u^{(s)}} \sum_{i} w_s(u,v)\delta_{i\in I_v} \qquad (10)$$

Where *N(s) u* is the set of neighbors of *u* in social networks and *ws(u, v)* is the weight of user *v*. We set *ws (u, v)* = *ws (dv)*, where *dv* is the depth of user *v* in the BFS tree rooted at user *u*.

By fixing 1-hop followees' weight at *ws (*1*)* = 1, we tune the weight of 2-hop users. In our later experiments, we found the best value is *ws (2)* = 0.1. Votings are ranked according to their scores to form the recommendation list.

Algorithm 2 Algorithm of UGUV Metapath

Data: Sina Weibo voting dataset
Result: Top-k votings for recommendation
Initialization;
for *each target user* u **do**
  Find all groups g's that user u has joined and put them in a set G^u;
  for *each joined group* $g \in G^u$ **do**
    Find all user v's in group g;
    for *each user* v *in group* g **do**
      User v reports its relevant votings and put them in a set I_v;
      for *each candidate voting* $i \in I_v$ **do**
        $Score_{u,i} += w(g)$;
      end
    end
  end
  Sort $\{Score_{u,i}\}$ in a decreasing order;
  Return and recommend top k votings with highest scores to user u;
end

c) UVUV metapath: As shown in Fig. 5(d), the semantic of *U − V − U − V* metapath-based recommendation is to find users that share votings with the target user, and then recommend their relevant votings to the target user. For a target user *u*, *UVUV* works as follows.

1) Find all votings that *u* has participated in, and denote this voting set as *Iu*.

2) For each of the voting *j* ∈ *Iu*, find the set of users who have participated in *j* . Denote the set of users as *Nj* .

3) Each user *v* ∈ *Nj* reports all the votings that he has participated in.

4) Aggregate the reports of all users to assign scores to votings as follows:

$$\text{Score}_{u,i} = \sum_{j\in I_u}\sum_{v\in N_j}\sum_{i} w(v)\delta_{i\in I_v}. \qquad (11)$$

In our later experiments, we set *w(v)* = 1 for all users. Finally, we summarize the algorithm details of UGUV, UUV (m-hop), and UVUV metapath approaches in Algorithms 2–4, respectively.

2) Neighborhoods in Latent Feature Space: Other than neighborhoods visited through metapaths, we also explore

Algorithm 3 Algorithm of UUV(m-Hop) Metapath

Data: Sina Weibo voting dataset
Result: Top-k votings for recommendation
Initialization;
for *each target user* u **do**
  Find all followees v's within m-hops by BFS;
  Put all those v's in a set $N_u^{(s)}$;
  for *each user* $v \in N_u^{(s)}$ **do**
    User v reports its relevant votings and put them in a set I_v;
    Set weight parameter $w_s(u,v)$ according to the depth of user v in the BFS tree rooted at user u;
    for *each voting* $i \in I_v$ **do**
      $Score_{u,i} += w_s(u,v)$;
    end
  end
  Sort $\{Score_{u,i}\}$ in a decreasing order;
  Return and recommend top k votings with highest scores to user u;
end

Algorithm 4 Algorithm of UVUV Metapath

Data: Sina Weibo voting dataset
Result: Top-k votings for recommendation
Initialization;
for *each target user* u **do**
  Find all votings j's that user u has participated;
  Put all those voting j's into a set I_u;
  for *each voting* $j \in I_u$ **do**
    Find all users v's who ever participated in voting j and put them in a set N_j;
    for *each user* $v \in N_j$ **do**
      Find all votings i's that user v has participated and put them in a set I_v;
      for *each voting* $i \in I_v$ **do**
        $Score_{u,i} += w(v)$;
      end
    end
  end
  Sort $\{Score_{u,i}\}$ in a decreasing order;
  Return and recommend top k votings with highest scores to user u;
end

IJMTARC – VOLUME – VI – ISSUE – 24, OCT-DEC, 2018 ISSN: 2320-1363

Neighborhoods in the user latent feature space derived from MF models. Note that, previous works show that *PureSVD* [10] and *AllRank* [12] perform better than neighborhood-based approaches in user-item space directly when used in top-k recommendation. Yang *et al.* [14] shows that neighborhood in latent feature space approach is comparable with *AllRank*; therefore, we study neighborhood in latent feature space in this section.

a) UNN: *UNN* uses MF (i.e., *AllRank* [12]) to obtain the user latent features. Users are then clustered in the user latent feature space using the Pearson correlation coefficient. Users nearest to the source user u are identified and denoted as Nu . The relevant votings of these nearest users are scored and ranked to form the top-k recommendation list. The score of a candidate voting i is calculated as follows:

$$\text{Score}_{u,i} = \sum_{v \in N_u} \sum_{i} \text{sim}(u, v)\delta_{i \in I_v} \tag{12}$$

Where Nu is the set of NNs of user u in the user latent feature space, and the NNs of user u are weighted according to their similarity *sim(u, v)* with user u, measured in terms of the Pearson correlation between user u and v.

b) VNN: This approach works similarly as UNN, except we cluster votings in the voting latent feature space

$$\text{Score}_{u,i} = \sum_{x \in I_u} \sum_{i} sim(x, i)\delta_{i \in N_x} \tag{13}$$

where Iu is the set of votings participated by user u and Nx is the set of NNs of voting x in the voting latent feature space.

3) Combined Neighborhoods: Hybrid Approach is the combination of UGUV, UUV(m-hop), UVUV, and UNN approaches. We integrate the four recommenders by combining their voting results. Basically, for a target user u, we consider a set of neighboring users that either share the same group with u, or have short social distances to u, or share similar tastes in votings. The score of a potential relevant voting i for user u am calculated as

$$\begin{aligned}\text{Score}_{u,i} = {} & \rho_1 \times \sum_{g \in G^u} \sum_{v \in g} \sum_{i} w(g)\delta_{i \in I_v} \\ & + \rho_2 \times \sum_{v \in N_u^{(s)}} \sum_{i} w_s(u, v)\delta_{i \in I_v} \\ & + \rho_3 \times \sum_{j \in I_u} \sum_{v \in N_j} \sum_{i} w(v)\delta_{i \in I_v} \\ & + \rho_4 \times \sum_{v \in N_u} \sum_{i} sim(u, v)\delta_{i \in I_v}\end{aligned} \tag{14}$$

Where $\rho 1$, $\rho 2$, $\rho 3$, and $\rho 4$ are the weights of UGUV, UUV (m-hop), UVUV, and UNN approaches, respectively.

V. EXPERIMENTS

In this section, we evaluate the proposed MF models and NN models using Sina Weibo voting data set.

A. Methodology

We evaluate the performance of a set of voting RSs using the same trace. We use a simple popularity-based RS as the baseline model.

• *MostPop:* This RS recommends the most popular items to users, i.e., the votings that have been voted by the most numbers of users. For the Weibo-MF model proposed in (5), we evaluate several variants by setting different weights for social and group information.

1) *Voting-MF:* By setting $\gamma s = 0$ and $\gamma g = 0$ in (5), we only consider user-voting matrix and ignore social and group information. Note that Voting-MF is essentially the same as All Rank model, which is proposed in [12]. All Rank was found to be the best model of optimizing top-k hit ratio on various data sets according to [10] and [12].

2) *Voting + Social-MF*: By setting $\gamma s > 0$ and $\gamma g = 0$, we additionally consider social network information on top of Voting-MF.

3) *Voting + Group-MF:* By setting $\gamma s = 0$ and $\gamma g > 0$, we additionally consider user-group matrix information on top of Voting-MF.

4) *Weibo-MF:* By setting $\gamma s > 0$ and $\gamma g > 0$, we add both social and group information to Voting-MF. For NN-based RSs, we evaluate UGUV metapath and UUV(mhop) metapath (with $m = 1$, 2) described in Section IVC1; UNN, VNN described in Section IV-C2; and the hybrid approach described in Section IV-C3 by setting different weights in (14). We randomly choose 80% of the data set as training set and the remaining 20% as test set. The random selection

ISSN: 2320-1363

was carried out five times independently, and we report the average statistics. We conducted our experiments on a Linux server with four E5640 Intel Xeon CPUs. Each CPU has four cores with 2.67 GHz, and each core has 12.3-MB cache. The shared memory size is 36 GB.

B. MF-Based Approaches

We tune the regularization constant λ and the optimal value is 0.5. For the dimensionality, we choose $j0$ = 10. We tune the remaining parameters to optimize top-20 hit rate. The performance of MF-based RSs is compared in Table II. In Voting-MF model, the parameters that lead to the best top-20 hit rate are: wm = 0.01 and rm = 0. As expected, ***Voting-MF significantly outperforms the naive popularity based*** *RS*. Since user-voting data are binary, impute the Missing value of user-voting as $rm < 1$, leading to the same result as $rm = 0$. In Voting + Group-MF, the optimal parameters are γg = 0.1,$w(G)\ m$ = 0.001, and gm = 0. In Voting + Social- MF, the optimal parameters are γs = 0.1,$w(S)\ m$ = 0.00005, and sm = 0. Due to the computation constraints, we only present the results of $j0$ = 10 for all different MF models here. It is evident that Weibo-MF outperforms all other MF-based approaches, since more information used in the Model leads to more prediction power. Regarding the results between Voting-MF and Voting + Social-MF, it is noticed that Voting-MF model is good to represent and mine the data with 40.6%–60.6% relative improvement over MostPop. Adding social information to Voting-MF leads to additional ten plus percent relative gain, which validates that explicitly reinforcing the social influence in MF model can further improve the performance at certain level. Another interesting observation is that Voting + Group-MF and Weibo-MF almost cannot or can Only bring limited improvement over Voting + Social-MF approach.

C. Hot-Voting-Only Recommendation

As mentioned in Section III, it is very intriguing to study hot-voting recommendation as it propagates through both social networks and global channels, such as headline news. In this section, we focus on recommending hot votings only. To study hot-voting recommendation, we filter out a hotvoting-data set that only contains hot votings. We choose votings with no less than 1000 participants as hot votings. In the training set, we pick out all the hot votings and only keep hot-voting related tuples. In the testing set, we only keep hot-votings related tuples for testing. We further get rid of users in the testing set who do not appear in the training set. In the resulting hot-voting data set, there are 290 184 users and 329 votings, 700 628 user-voting tuples in the training set, and 138 682 user-voting tuples in the testing set. In hot-voting experiments, we report top-5 to top-50 results. We tune all parameters to optimize top-10 hit rate. The optimal regularization constant λ is 0.5. We try different value of dimensionality and the best value is $j0$ = 20. As we can see that, due to much less number of votings, the optimal $j0$ is much smaller than in the whole data set. In Voting-MF model, the optimal parameters are: wm = 0.06 and rm = 0. The optimal parameters of UUV(2-hop) are $ws(2)$ = 0.1, the same as the whole data set. The optimal weights of UUV(2-hop) + UNN are $\rho 2$ = 1 and $\rho 4$ = 1. The optimal weights of UUV(2-hop) + UVUV are $\rho 2$ = 1 and $\rho 3$ = 0.1. In Table VI, Voting-MF is better than UNN, and UVUV is better than Voting-MF. From Section V-D, we know that UVUV favors hot-voting recommendation. It might just because UVUV tends to recommend more hot votings than other methods. From these hot-voting-only experiments, we can see that *UVUV can indeed recommend hot votings more accurately than other methods, even if all the methods are only focused on hot votings.* One explanation is that UVUV approach's neighborhood size is very large. Through a hot voting, a user is connected to more than 1000 other

VI. RELATED OUTPUTS

Admin login page:

Admin home page:

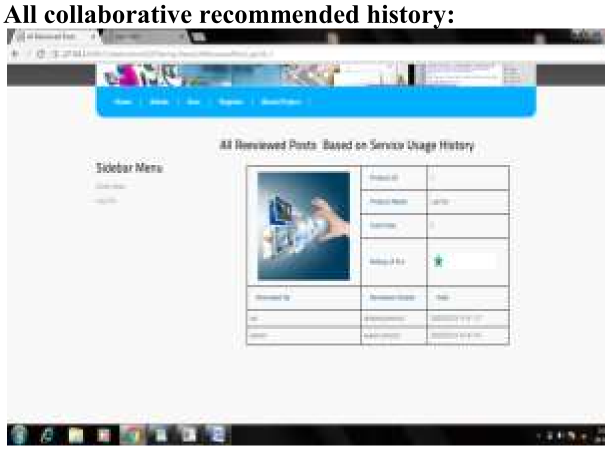
IJMTARC – VOLUME – VI – ISSUE – 24, OCT-DEC, 2018

ISSN: 2320-1363

Add product:

All post with Rank:

All collaborative recommended history:



6. CONCLUSION

In this project, we present a set of MF-based and NN-based RSs for online social voting. Through experiments with real data, we found that both social network information and group affiliation information can significantly improve the accuracy of popularity-based voting recommendation, especially for cold users, and social network information dominates group affiliation information in NN-based approaches. This paper demonstrated that social and group information is much more valuable to improve recommendation accuracy for cold users than for heavy users. This is due to the fact that cold users tend to participate in popular votings. In our experiments, simple metapath-based NN models outperform computation intensive MF models in hot-voting recommendation, while users' interests for nonhot votings can be better mined by MF models. This paper is only our first step toward thorough study of social voting recommendation. As an immediate future work item, we would like to study how voting content information can be mined for recommendation, especially for cold votings. We are also interested in developing voting RSs customized for individual users, given the availability of multichannel information about their social neighborhoods and activities

7. REFERENCES

[1] R. M. Bond *et al.*, "A 61-million-person experiment in social influence and political mobilization," *Nature*, vol. 489, pp. 295–298, Sep. 2012.

[2] G. Adomavicius and A. Tuzhilin, "Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions," *IEEE Trans. Knowl. Data Eng.*, vol. 17, no. 6, pp. 734–749, Jun. 2005.

[3] X. Su and T. M. Khoshgoftaar, "A survey of collaborative filtering techniques," *Adv. Artif. Intell,* vol. 2009, Aug. 2009, Art. No. 421425, doi: 10.1155/2009/421425.

[4] Y. Koren, "Factorization meets the neighborhood: A multifaceted collaborative Filtering model," in *Proc. ACM KDD*, 2008, pp. 426–434.

[5] Y. Koren, "Collaborative filters with temporal dynamics," in *Proc. KDD*, Paris, France, 2009, pp. 447–456.

[6] A. Paterek, "Improving regularized singular value decomposition for collaborative filtering," in *Proc. KDDCup*, 2007, pp. 39–42.

[7] R. Salakhutdinov and A. Mnih, "Probabilistic matrix factorization," in *Proc. NIPS*, vol. 20. 2008, pp. 1257–1264.

About authors:

KAVURI ANITHA LAKSHMI:- B.Tech in C.S.E from affiliated to J.N.T.U. Kakinada. She is pursuing M.Tech in the stream of C.S.E in, Srinivasa Institute of Engineering & Technology an affiliated to J.N.T University Kakinada, A.P, Cheyyeru (v), Amalapuram-53322.

IJMTARC – VOLUME – VI – ISSUE – 24, OCT-DEC, 2018

ISSN: 2320-1363

Guide:

MR.B.SURESH:- Working as Associate professor, Srinivasa Institute of Engineering & Technology an affiliated to J.N.T University Kakinada, A.P, Cheyyeru (v), Amalapuram-53322. His Qualification is M.tech in C.S.E he has 2 years experience teaching in CSE he has published 1 papers.