IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000.

Digital Object Identifier 10.1109/ACCESS.2017.DOI

ACCF: Learning Attentional Conformity for Collaborative Filtering

BIN LIANG[1], CHAOFENG SHA[1], DONG WU[1], BO XU[2], YANGHUA XIAO[1], and WEI WANG[1].
[1]School of Computer Science, Fudan University, Shanghai, China (e-mail: liangbin, cfsha, dongwu14, shawyh, weiwang1@fudan.edu.cn)
[2]School of Computer Science and Technology, Donghua University, Shanghai, China (e-mail: xubo@dhu.edu.cn)

ABSTRACT In recent years, Collaborative Filtering (CF) methods have yielded immense success on recommender systems. They mainly use the similarity between users and items, or the interactions between users and items to predict the unknown ratings. However the social conformity phenomenon received little notice, which means (1) individuals in a social network can have multiple characteristics and hence tend to belong to multiple overlapping groups or communities; (2) when confronted with conformity pressure, people often adjust their responses to conform to others' opinions to obtain social approval and belonging in the community. In this paper, we propose a new collaborative filtering based recommendation framework, called ACCF, which explicitly exploits social conformity of users. We incorporate such social conformity phenomenon into the latent factor model, using a weighted average of community preference profiles as the adjusting factor, and learn the weight of each community's influence through an attention network. Compared with seven state-of-the-art methods on three real-world datasets, our method achieves the best performance.

INDEX TERMS Recommender Systems, Collaborative Filtering, Community Detection, Attention Mechanism, Neural Networks.

I. INTRODUCTION

Collaborative Filtering (CF) methods have yielded immense success on recommender systems. They usually rely on the users' explicit or implicit feedbacks on the items, and predict the unknown ratings of users or items after training on the historical rating behavior data. In recent years, many works have been focus on the use of social relation to enrich the user ratings in recommender systems, such as [8], [9], [12], [23], [25], [33].

However, the exploration of *Social Conformity* phenomenon on users has received relatively less attention. Social conformity phenomenon [6], [31] that occurs when an individual's values, beliefs, behaviors, and attitude are influenced by either one person (minority influence), or by a group of people (majority influence) who establish norms. During conformity one changes the way they behave in response to social pressures. The existence of the social conformity phenomena has been verified on large social networks such as Flickr and Gowalla [31]. In the recommender system, this phenomenon is manifested by the fact that users are more likely to accept recommendations from the same community than all users.

For example in Figure 1, The User belongs to two communities, and different community members like to watch different types of movies. (i.e., Community 1 prefers Romance while Community 2 prefers Thriller). The user tend to watch different types of movies when accompanied by friends from different communities. Similarly, the user can

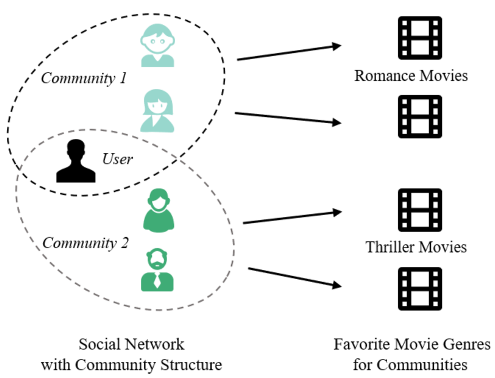

FIGURE 1. An example of a user's social conformity phenomenon. The `User` tend to watch different genres of movies when accompanied by friends from different `Communities`.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2019.2921853, IEEE Access

see his/her friend's rating of a movie on the social network, and may be unconsciously affected when he/she evaluates the movie.

Based on this observation, we propose a novel recommendation method, called ACCF, which incorporates social conformity phenomenon into the collaborative filtering model. At first, we group users in a social network into overlapping communities. Next, we model each user's inherent interest as a latent vector and each community's latent preference vector. As conformity is a group phenomenon, factors such as group size and cohesion help determine the level of social conformity that an individual displays. Thus, to incorporate social conformity phenomenon into the latent factor model, we model each user's feature vector as the inherent user interest adjusted by a weighted average of community preference vectors. The influence or weight of each community on the users' rating behavior might vary with different items. To model this varying influence, we employ an attention network to learn the dynamic weight.

Compared to the previous approach that takes advantage of the community structure of networks [33], the proposed approach is more intuitive, interpretable, and effective. The advantages of our new approach over the existing ones are confirmed by extensive experiments on three real-world datasets. The main contributions of our paper can be summarized as follows:

- Firstly, we exploit the *Social Conformity* phenomenon on users for recommender systems, getting community precise information from social network structure.
- Secondly, we propose a novel recommendation method, called ACCF, which incorporates social conformity phenomenon into the collaborative filtering model with attention network.
- Finally, compared with seven state-of-the-art methods on three real-world datasets, our method achieves the best performance.

The rest of this paper is organized as follows. In Section II, we review the related work. In Section IV, we formalize the problem and describe our proposed system ACCF. In Section V, we describe our experimental settings for evaluation and discuss the results. Finally, we conclude this work in Section VI.

II. RELATED WORK

The Collaborative Filtering which establishes the connection between users and items, is commonly used in recommender systems. Two popular approaches are the neighborhood models (user or item based) [29], and latent matrix factorization such as probabilistic matrix factorization [28], singular value decomposition [13], [14].

The social relationships between users have been explored to improve the performance of the recommender system [8], [25]. The authors of [11] propose Trustwalker based on a random walk model, which combines the collaborative filtering and the trust-based approach for recommendation. The authors of [22] propose to fuse the users' tastes and their trusted friends' favors together. They propose a social trust ensemble to represent the formulation of the social trust restrictions on the recommender systems. In [24], the authors introduce a social regularization term between user preference vectors into the latent matrix factorization model. The authors of [40] develop a circle-based recommendation model through inferring category-specific social trust circles from available rating data and social link data. Through the extensive experiments, the author of [21] studies the advantages of using explicit social information in recommendation models.

In [32], the authors exploit local and global social relations in recommendation model. They assume that the user preferences of two socially connected users are correlated locally. Besides, they use the global user reputation score to weight the importance of their ratings. The authors of [19] propose to combine contextual information and social information to improve recommendation performance. They apply matrix factorization method on the sub-matrix which are partitioned based on various context. They also leverage the social information in a social regularization term. Motivated by the heuristic that individuals will affect each other during the process of reviewing, a truster model and a trustee model are proposed in [37] to map users into the same latent feature spaces but with different implications that can explicitly describe the feedback how users affect or follow the opinions of others. They synthesize the two models to one fusing model simultaneously which fits available ratings and trust ties. The authors of [9] extend SVD++ [14] into TrustSVD with the available social trust information. The trust matrix is decomposed into trust-feature matrix and trustee-feature matrix. TrustSVD is extended into Hell-TrustSVD in [30] to incorporate both the extracted implicit social relations and users' ratings.

Some recent work exploit user dependencies or tie strength [7] when designing social recommendation models. In [20], the authors introduce the assumption that users' latent features follow a matrix variate normal distribution. Based on this assumption, they propose to incorporate the learnt positive and negative dependencies between users into a probabilistic relational matrix factorization model. In [36], the authors study the effects of distinguishing strong and weak ties in social recommendation. They use Jaccard similarity between users' neighborhoods to measure the tie strength and incorporate the distinction of strong and weak ties into the Bayesian Personalized Ranking model. In the following work of [35], they propose a probabilistic to learn the personalized preference of strong and weak ties and incorporate these preference into probabilistic matrix factorization for recommendation. The user's social status and homophily are fused into a social matrix factorization-based method proposed [5], where the degree of trust is computed by trust propagation method and the PageRank algorithm. The social information has also been combined with temporal information to improve the recommendation performance, which has been advocated by [3]. The SPMC model proposed

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2019.2921853, IEEE Access

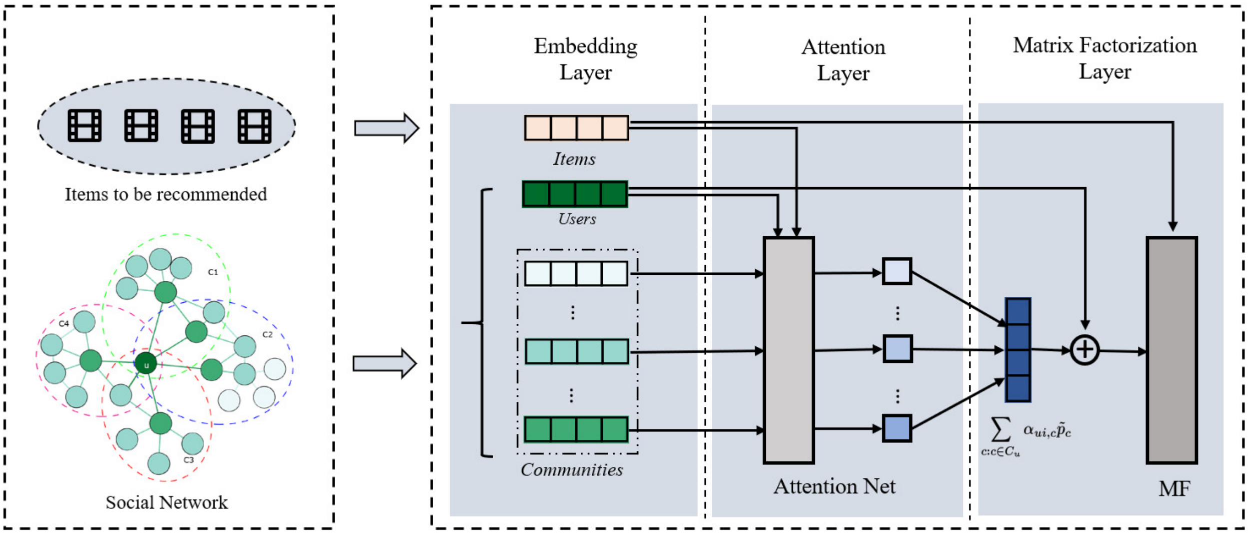

FIGURE 2. The architecture of ACCF. Communities are detected from social network in which each user may belong to several communities. The attention net, whose inputs are embeddings for users/items/communities, is applied to control the weights for each user to the communities of interests. After the attention mechanism, the embedding of each user is justified and extracts both the information of user and corresponding communities. The performance of matrix factorization will be improved with the guidance of communities of interests.

in [3] combines Factorized Personalized Markov Chains [26] and Social Bayesian Personalized Ranking (SBPR) [41] and is effective to deal with the sparsity and cold start problem.

Different from the above work using the micro-structure of users, some social recommendation models explore community structure found in the users' social networks. In [17], they propose two social recommendation models in which a regularization term between overlapping community is incorporated into the matrix factorization model. The framework SoDimRec proposed in [33] which will be detailed in the next section employs social dimensions to capture heterogeneity of social relations and weak dependence. Our method follows this direction, but, we incorporate the communities' preference into the prediction rule rather as a regularization term in the optimization problem. Based on the assumption that the influence strength is dependent on the social roles of users, the authors of [18] propose an incremental clustering algorithm to detect dynamic social roles. Then the social influence network and the user preferences are simultaneously inferred in a matrix factorization framework.

There is a body of recent work for social influence analysis based on conformity. The interplay between influence and conformity is studied in [16], and the conformity in a social group is evaluated by the CASINO algorithm. In [31], the authors define several. major types of conformity in individual, peer, and group levels.

III. PRELIMINARY

In this section, we give a brief descriptions of latent matrix factorization [15] and SoDimRec [33], which are related to our proposed method.

A. LATENT MATRIX FACTORIZATION

Matrix factorization technique is the most pervasive method employed in recommender systems, which maps both users and items to a shared latent factor space of some dimensionality d, such that user-item interactions are modeled as inner products in that space. Accordingly, each user u is associated with a vector $p_u \in R^d$, and each item i is associated with a vector $q_i \in R^d$.

For a given user u, each component of p_u measures the extent of interest the user has on the corresponding factor. Similarly, for a given item i, the components of q_i measure the extent to which the item possesses those factors. The resulting inner product, $p_u^T q_i$, captures the interaction between user u and item i, i.e., the user's overall interest in the item's characteristics. Formally, user u's rating of item i, which is denoted by r_{ui}, is approximated by

$$\hat{r}_{ui} = \mu + b_u + b_i + p_u^T q_i \tag{1}$$

where μ is the average rating, and b_u and b_i represent user and item biases, respectively. To estimate the factor vectors p_u and q_i, the regularized squared error is minimized on the itemset with known ratings:

$$\sum_{(u,i)\in R} (r_{ui} - \hat{r}_{ui})^2 + \lambda(\sum_u b_u^2 + \sum_i b_i^2 + \sum_u ||p_u||^2 + \sum_i ||q_i||^2). \tag{2}$$

Here, R is the set of the (u, i) pairs for which r_{ui} is known in the rating matrix. The parameters can be estimated using the stochastic gradient descent optimization algorithm.

B. SODIMREC

SoDimRec is similar to our method. Specifically, it consists of two steps. In the first step, an overlapping community

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2019.2921853, IEEE Access

detection algorithm is applied to extract C different communities. Then those terms modeling user's and communities' preference vector are added into the standard matrix factorization framework as regularization. Their approach aims to minimize the following objective function:

$$\sum_{(u,i)\in R}(r_{ui}-p_u^T q_i)^2+\lambda_1\sum_u\sum_{c:u\in c}S_{uc}||p_u-\hat{p}_c||^2$$
$$+\lambda_2\sum_u\sum_c\sum_{v\in H_{uc}}S_{uvc}||p_u-p_v||^2$$
$$+\alpha(\sum_c||\hat{p}_c||^2+\sum_u||p_u||^2+\sum_i||q_i||^2). \qquad (3)$$

Here, $\hat{p}_c$ is the social dimension preference vector of community c, and S_{uc} is the association strength of user u with community c; H_{uc} is the user set with weak dependence connection to u in community c, i.e., $H_{uc}=\{v\in c:(u,v)\notin E\}$, and S_{uvc} is the strength of the weak dependence connection between u and v in community c. The strength S_{uc} is explicitly defined as a linear combination of similarity score between the ratings of user u and the average ratings of users in community c, with the "embedding" of the user u into community c; the strength S_{uvc} is defined as the cosine rating similarity between users u and v in terms of community c. As be established in equation (13) of [33], the social dimension preference vector is converged to the smoothed average of preference vector of the users in community c as follows:

$$\hat{p}_c=\frac{\sum_{u\in c}S_{uc}p_u}{\alpha+\sum_{u\in c}S_{uc}}. \qquad (4)$$

The reason why we concern about the setting of SoDimRec is two-folds:

- firstly, the similarity between users computed in terms of users' or communities' rating might not be reliable for those cold-start users.
- secondly, the effect of the communities' preference could be incorporated into prediction explicitly rather as a regularization term in the objective function.

In the next section, we will introduce our method to incorporate the community preference into the prediction rule which does not need to compute the rating similarity between users.

IV. PROPOSED METHOD

In this section, we formalize the problem definition and introduce the proposed method that predicts the rating through exploiting users' conformity with an attention network.

A. PROBLEM FORMULATION

We denote the user set as $U=\{u_1,u_2,\cdots,u_n\}$ and the item set as $I=\{i_1,i_2,\cdots,i_m\}$. In the recommender systems that exploit social conformity, we are given a social network $G=(U,E)$, where E is the set of social relationships between users, in addition to a rating matrix $R=[r_{ui}]_{n\times m}$ where r_{ui} is a rating given by user u on item i.

In this paper, we aim to predict the unknown rating given the rating set R. Matrix factorization technique [15] is the most pervasive method employed in recommender systems, which maps both users and items to a shared latent factor space of some dimensionality d, such that user-item interactions are modeled as inner products in that space. Accordingly, each user u is associated with a vector $p_u\in R^d$, and each item i is associated with a vector $q_i\in R^d$. For a given user u, each component of p_u measures the extent of interest the user has on the corresponding factor. Similarly, for a given item i, the components of q_i measure the extent to which the item possesses those factors. The resulting inner product, $p_u^T q_i$, captures the interaction between user u and item i, i.e., the user's overall interest in the item's characteristics. Note that our approach can be extended to the scenario with implicit feedback by feeding the rating score into a ranking component.

B. FRAMEWORK: ACCF

The proposed method is called **ACCF** (**A**ttentional **C**onformity for **C**ollaborative **F**iltering). At first, we run the overlapping community detection algorithm BigCLAM [38] and CESNA [39] on the social network G, which is based on Nonnegative Matrix Factorization. This algorithm is designed to return C overlapping communities, so that we can obtain the communities each user u belongs $C_u=\{c:u\in c\}$. The experimental results are similar when using the algorithm proposed by [34], which is utilized by SoDimRec in its first step.

Next, for each community $c=1,\ldots,C$, we assume that there is a latent community preference vector $\tilde{p}_c\in R^d$ that represents the overall preference pattern of community c. Since social conformity is a group phenomenon, we could model each u's feature vector as $p_u+\sum_{c:u\in c}\alpha_{uc}\tilde{p}_c$ by considering external group influence and the degree of conformity to different groups. It is essentially the inherent user interest p_u adjusted by a weighted average of community preference vectors, where the weight α_{uc} is used to measure the conformity of user u to community c. However, the weights α_{uc} should not be static with respect to each item. We believe that the influence of each community on him/her might vary when an user is rating different items, especially for those cold-start users who might both rated few items and had few social relationships.

To learn this flexible influence we resort to neural attention mechanism [1] which has been applied extensively in recommendation models such as AFM [4] and NAIS [10]. The overall architecture of ACCF is illustrated in Figure 2. We use the attention layer to learn the influence of each community $c\in C_u$ on user u with respect to item i. Here we employ a fully connected neural network to learn the score $e_{ui,c}$ which could be interpreted as the influence of community c on user u when interacting with item i. Then the scores are normalized through softmax operation. Formally,

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2019.2921853, IEEE Access

Method	Epinions	Ciao	Flixster
PMF	$\lambda = 0.1, \gamma = 0.01$	$\lambda = 0.1, \gamma = 0.01$	$\lambda = 0.1, \gamma = 0.01$
SoRec	$\lambda = 0.001$ $\lambda_c = 1, \gamma = 0.001$	$\lambda = 0.001$ $\lambda_c = 0.01, \gamma = 0.05$	$\lambda = 0.001$ $\lambda_c = 0.001, \gamma = 0.003$
SoReg	$\lambda = 0.0001$ $\beta = 0.1 , \gamma = 0.001$	$\lambda = 0.001$ $\beta = 0.1 , \gamma = 0.0001$	$\lambda = 0.00005$ $\beta = 1 , \gamma = 0.001$
TrustMF	$\lambda = 0.001$ $\lambda_t = 1 , \gamma = 0.05$	$\lambda = 0.001$ $\lambda_t = 1 , \gamma = 0.1$	$\lambda = 0.01$ $\lambda_t = 0.1 , \gamma = 0.003$
TrustSVD	$\lambda = 0.6$ $\lambda_t = 0.5 , \gamma = 0.001$	$\lambda = 0.5$ $\lambda_t = 1 , \gamma = 0.001$	$\lambda = 0.01$ $\lambda_t = 1 , \gamma = 0.001$
MFC$_b$	$\lambda = 0.1 , \lambda_1 = 1$ $\gamma = 0.00001 , C = 300$	$\lambda = 0.1 , \lambda_1 = 1$ $\gamma = 0.001 , C = 200$	$\lambda = 0.1 , \lambda_1 = 1$ $\gamma = 0.00001 , C = 500$
SoDimRec	$\lambda_1 = 5 , \lambda_2 = 10$ $\lambda = 1 , C = 500$	$\lambda_1 = 5 , \lambda_2 = 10$ $\lambda = 1 , C = 200$	$\lambda_1 = 5 , \lambda_2 = 10$ $\gamma = 1 , C = 500$
ACCF	$\lambda = 0.2 , C = 300$ $\gamma = 0.02$	$\lambda = 0.2, C = 200$ $\gamma = 0.02$	$\lambda = 0.2, C = 500$ $\gamma = 0.02$

TABLE 1. Parameter settings of respective methods for three datasets.

the attention coefficients $\alpha_{ui,c}$ are computed as follows:

$$e_{ui,c} = v^T \sigma(W_u p_u + W_c \tilde{p}_c + W_i q_i + b) \tag{5}$$

$$\alpha_{ui,c} = \frac{\exp(e_{ui,c})}{\sum_{c':u \in c'} \exp(e_{ui,c'})} \tag{6}$$

where $W_u, W_c, W_i \in R^{t \times k}$, $b \in R^t$ and $v \in R^t$ are model parameters, $\sigma(x) = ReLU(x) = \max(0, x)$. Here the rectifier is chosen as the activation function thanks to good performance empirically.

After learning the normalized conformity attention weight, we arrive at the following prediction rule to predict the rating of user u to item i:

$$\hat{r}_{ui} = \mu + b_u + b_i + q_i^T (p_u + \sum_{c:c \in C_u} \alpha_{ui,c} \tilde{p}_c). \tag{7}$$

where μ is the average rating, and b_u and b_i represent user and item biases, respectively. This form is inspired by SVD++ proposed by [14], except that we take the attentive influence by social conformity into account, instead of by the set of rated items.

Though both SoDimRec and ACCF make use of overlapping community structure in a network, the key difference is that we explicitly incorporate the community preference into the prediction rule (7) rather than as a regularization term into the objective function ([33]). We believe our approach is more intuitive and interpretable. Compared to TrustSVD [9] where each user has two latent feature vectors p_u and c_u, and the influence of each neighbor $v \in T(u)$ is uniform $(p_u + |T(u)|^{-\frac{1}{2}} \sum_{v \in T(u)} c_v)$, our ACCF learns the influence of each community through an attention network. The flexibility of our design is demonstrated in the experimental evaluation.

C. OBJECTIVE FUNCTION AND PARAMETER ESTIMATION

The model parameters

$$\Theta = \{\mu, b_u, b_i, p_u, \tilde{p}_c, q_i, W_u, W_c, W_i, b\}$$

are estimated by minimizing the following regularized objective function:

$$\begin{aligned} \mathcal{L} = & \sum_{(u,i) \in R} (r_{ui} - \hat{r}_{ui})^2 + \lambda (\sum_u b_u^2 + \sum_i b_i^2 \\ & + \sum_u ||p_u||^2 + \sum_{c \in C} ||\tilde{p}_c||^2 + \sum_i ||q_i||^2) \end{aligned} \tag{8}$$

where $\hat{r}_{ui}$ is defined in equation (7), and λ is a parameter to control model complexity and to avoid overfitting.

We use stochastic gradient descent to estimate the parameters Θ with deep learning platform Theano [2]. The parameters of attention network are initialized with standard normal distribution $N(0, 0.0001)$.

V. EXPERIMENT

In this section, we evaluate the effectiveness of the proposed ACCF framework using three real-world datasets. The experiments are set up as the following.

A. EXPERIMENTAL SETTING

1) Experimental Datasets

Three publicly available datasets are used in our experiments: Epinions[1], Ciao[2], and Flixster[3]. All datasets contain both ratings and social links. Although the social links are directed in Epinions and Ciao, we treat all these links as undirected when performing overlapping community detection. The statistics of the datasets are shown in Table 2.

2) Cross-validation

We use 5-fold cross-validation for training and testing. Each dataset is randomly divided into five folds. In each round, four folds are used as training set and the remaining fold is used as testing set. The average of the five rounds is reported as the final result.

[1]http://www.trustlet.org/epinions.html
[2]http://www.jiliang.xyz/trust.html
[3]http://www.cs.ubc.ca/ jamalim/datasets/

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2019.2921853, IEEE Access

Feature	Epinions	Ciao	Flixster
#Users	40,163	7,375	47,109
#Items	139,738	105,114	27,457
#Ratings	664,824	282,650	1,675,545
Density$_1$	**0.012%**	**0.037%**	**0.130%**
#Trusts	442,979	111,781	839,277
Density$_2$	**0.028%**	**0.206%**	**0.038%**

TABLE 2. Statistics of the three datasets. `#Ratings` given by `#Users` to `#Items`, `Density`$_1$ means the rating matrices on the datasets; `#Trusts` is social relationships between users, `Density`$_2$ means the social relation matrices on the datasets.

3) Evaluation Metrics

Two well-known metrics, Mean Absolute Error (MAE) and Root Mean Square Error (RMSE), are used to evaluate prediction performance. They are defined as

$$MAE = \frac{\sum_{u,i} |\hat{r}_{ui} - r_{ui}|}{N_{test}}$$

and

$$RMSE = \sqrt{\frac{\sum_{u,i} (\hat{r}_{ui} - r_{ui})^2}{N_{test}}}$$

where N_{test} is the number of test ratings, r_{ui} denotes the observed rating in the testing data, and $\hat{r}_{ui}$ is the predicted rating. As pointed by [14] and [33], *small improvement in RMSE or MAE terms can have a significant impact on the quality of the top-few recommendation.*

B. COMPETING METHODS

To evaluate the performance of ACCF, we compare our methods with the following seven methods:

- PMF: This method performs matrix factorization on the user-item rating matrix. It only utilizes rating information [28].
- SoRec: This method co-factorizes the user-term rating matrix and user-user social relation matrix under matrix factorization framework [23].
- SoReg: This method is also based on matrix factorization, and social regularization is defined to capture strong dependency connections [24].
- TrustMF: This method factorizes rating matrix and truster network simultaneously [37].
- TrustSVD: This method also factorizes rating matrix and truster network simultaneously based on the SVD++ framework [9].
- MFC$_b$: This method models overlapping community information as regularization terms into matrix factorization framework [17].
- SoDimRec: This method models social networks from heterogeneity of social relations and weak dependency connections based on social dimension, and is under matrix factorization framework [33].

We use the implementations of SoRec, SoReg, TrustMF, and TrustSVD provided by LibRec[4]. The codes of the SoDimRec and MFC$_b$ are not publicly available, so we use our implementations.

[4]http://www.librec.net

1) Parameter Settings

We choose the optimal parameter settings according to either previous work or experiments. Specifically, We set the dimension of latent factors d to 5, 10, and 20. The other settings are given in Table 1. When training our attention network, we initialize the learning rate $\gamma = 0.02$ and decrease to $\gamma = 0.001$ after 20 iterations. For all the results in Table.V-B and Table.V-B1, we use BigCLAM as our community detection algorithm. All the weights and biases in attention networks are initialized randomly from Guassian distribution with $\mu = 0$ and $\sigma = 0.1$.

C. PERFORMANCE

As in [37] and [9], both All and Cold Start views are created for testing. All ratings are used as the test set in the All view, while only the users who rate less than five ratings are used as the test set in the Cold Start view. The performance measures are shown in Table 3 and Table 4, respectively.

Table 3 (in the testing view of All) shows that ACCF outperforms all the competing methods across almost all the settings except one case (TrustSVD@Flixster with $d = 5$@MAE). The results clearly demonstrate the benefit of introducing social group conformity into the prediction model.

In the testing view of Cold Start, as shown in Table 4, ACCF achieves the best accuracy for all settings. Like our models, both MFC$_b$ and SocDimRec leverage the community structure when building the recommendation models. However, both MFC$_b$ and SocDimRec use the community's preference vectors in the regularization terms, that limits their prediction capacity. Through explicitly exploiting the community's preference vectors in the prediction rule, the improved performance of ACCF demonstrates the advantage of our designs. In particular, the out-performance of ACCF demonstrates the advantage of introducing the attention network.

In general, the performance of ACCF is the best and more robust with different latent dimensions.

1) Performance with Different Community Detectors

As shown in Table 5, in both full view and cold start situations, the performance of `ACCF` with `BigClam` is slightly better than the one with `CESNA` in a range of $[0, 0.006)$, except for three experiments settings with a slight decrease of 0.001. It indicates that different community detection algorithms has no prominent impact on the performance of `ACCF`, and the proposed model has shown its robustness on the the way for generating latent communities. In this paper we choose `BigClam` as our algorithm of community detection.

2) Performance with Latent Factors

Figure 3 shows the performance of `ACCF` with metric `RMSE` on Epinions dataset, with d varying in the range of $[5, 30]$.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2019.2921853, IEEE Access

Dataset	d	Metrics	PMF	SoRec	SoReg	TrustMF	TrustSVD	MFC$_b$	SoDimRec	ACCF
Epinions	5	RMSE	1.627	1.131	1.373	1.081	1.054	1.283	1.299	**1.040**
		MAE	1.184	0.927	1.071	0.826	0.809	0.997	1.016	**0.797**
	10	RMSE	1.587	1.143	1.176	1.127	1.053	1.205	1.154	**1.041**
		MAE	1.179	0.912	0.918	0.835	0.807	0.953	0.913	**0.794**
	20	RMSE	1.481	1.222	1.127	1.304	1.112	1.163	1.012	**1.039**
		MAE	1.122	0.917	0.841	0.880	0.810	0.896	0.846	**0.792**
Ciao	5	RMSE	1.883	0.993	1.482	0.981	0.956	1.231	1.247	**0.948**
		MAE	1.394	0.768	1.139	0.753	0.723	0.962	0.973	**0.713**
	10	RMSE	1.750	1.010	1.190	0.997	0.957	1.071	1.095	**0.948**
		MAE	1.324	0.768	0.911	0.753	0.723	0.839	0.856	**0.713**
	20	RMSE	1.532	1.077	1.095	1.107	0.962	1.013	1.030	**0.949**
		MAE	1.181	0.778	0.836	0.763	0.732	0.768	0.781	**0.711**
Flixster	5	RMSE	0.971	0.899	0.965	0.984	0.893	0.955	0.983	**0.880**
		MAE	0.705	0.676	0.725	0.775	**0.655**	0.737	0.753	0.661
	10	RMSE	0.950	0.908	0.929	1.025	0.908	0.942	0.956	**0.877**
		MAE	0.702	0.683	0.694	0.812	0.666	0.720	0.731	**0.661**
	20	RMSE	0.937	0.907	1.033	1.186	0.938	0.914	0.943	**0.876**
		MAE	0.697	0.683	0.756	0.910	0.688	0.697	0.719	**0.660**

TABLE 3. Performance comparison in All view. All the results of the proposed model are smaller than other models, which means the effect is better, except for `TrustSVD`@Flixster with $d = 5$@MAE.

Dataset	d	Metrics	PMF	SoRec	SoReg	TrustMF	TrustSVD	MFC$_b$	SoDimRec	ACCF
Epinions	5	RMSE	2.213	1.163	2.149	1.136	1.116	1.672	1.885	**1.110**
		MAE	1.776	0.937	1.893	0.902	0.878	1.365	1.618	**0.850**
	10	RMSE	2.234	1.188	1.536	1.184	1.119	1.378	1.492	**1.108**
		MAE	1.825	0.872	1.312	0.856	0.880	1.065	1.272	**0.851**
	20	RMSE	2.158	1.310	1.233	1.359	1.120	1.273	1.217	**1.105**
		MAE	1.789	0.856	0.862	0.864	0.886	0.840	0.863	**0.834**
Ciao	5	RMSE	2.09	1.030	2.796	1.042	0.979	2.116	1.994	**0.973**
		MAE	1.795	0.842	2.603	0.856	0.724	1.931	1.794	**0.700**
	10	RMSE	2.197	0.991	2.340	1.045	0.986	1.493	1.683	**0.977**
		MAE	1.946	0.715	2.100	0.742	0.726	1.311	1.488	**0.705**
	20	RMSE	2.091	1.220	1.173	1.227	**0.980**	1.156	1.135	**0.980**
		MAE	1.874	0.781	0.954	0.724	0.726	0.844	0.738	**0.675**
Flixster	5	RMSE	1.97	1.152	1.505	1.157	1.169	1.419	1.649	**1.054**
		MAE	1.475	0.913	1.258	0.948	0.949	1.117	1.412	**0.842**
	10	RMSE	1.862	1.167	1.287	1.165	1.147	1.323	1.366	**1.055**
		MAE	1.4	0.910	1.059	0.928	0.897	1.020	1.338	**0.835**
	20	RMSE	1.693	1.201	1.277	1.278	1.199	1.247	1.241	**1.055**
		MAE	1.291	0.932	1.010	0.996	0.938	0.957	0.973	**0.833**

TABLE 4. Performance Comparison in Cold-start View. `ACCF` gets the best performance at all datasets with all metrics, except for being equal with `TrustSVD`@Ciao with $d = 20$@RMSE.

When the dimension of latent factors d grows, the amount of parameters becomes larger, but the performance of `ACCF` is not significantly better, but randomly varies in the range of $[1.037, 1.044]$. However, in many neural networks, larger amount of parameters usually means better performance. In this way, the performance of `ACCF` does not significantly depends on the amount of parameters on neural networks, which indicates that our model has converged in our problem settings.

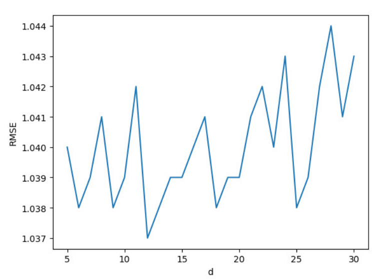

FIGURE 3. Performance for ACCF on Epinions@RMSE with d varying

D. IMPACT OF THE SOCIAL CONFORMITY

In this subsection, we study the effect of a user's social conformity to his communities. When performing community detection, we find many outlier users who do not belong to any discovered communities. We compare the prediction error on those outlier users with the prediction error on all users as well as on those users who belong to some community. Here we only report the results of ACCF on Ciao dataset (Figure 4), since the results both on Epinions and Flixster are similar.

As Figure 4 shows, both MAE and RMSE computed on

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2019.2921853, IEEE Access

Datasets	d	Metrics	Full View		Cold Start	
			ACCF@BigCLAM	ACCF@CESNA	ACCF@BigCLAM	ACCF@CESNA
Epinions	5	RMSE	**1.040**	1.044	**1.110**	1.200
		MAE	**0.797**	0.803	**0.850**	0.855
	10	RMSE	**1.041**	0.042	**1.108**	1.114
		MAE	**0.794**	0.798	**0.851**	0.856
	20	RMSE	**1.039**	1.043	**1.105**	1.107
		MAE	**0.792**	0.795	0.834	**0.833**
Ciao	5	RMSE	**0.948**	0.951	**0.973**	0.975
		MAE	**0.713**	0.716	**0.700**	0.703
	10	RMSE	**0.948**	0.949	**0.977**	0.979
		MAE	**0.713**	0.717	**0.705**	0.710
	20	RMSE	**0.949**	0.952	**0.980**	0.984
		MAE	**0.711**	0.711	**0.675**	0.679
Flixster	5	RMSE	0.880	**0.879**	**1.054**	1.060
		MAE	**0.661**	0.665	**0.842**	0.845
	10	RMSE	**0.877**	0.883	**1.055**	1.057
		MAE	**0.661**	0.664	**0.835**	0.839
	20	RMSE	**0.876**	0.879	**1.055**	1.056
		MAE	**0.660**	0.660	0.833	**0.832**

TABLE 5. Performance comparison in `All View` and `Cold Start` with two different algorithms of community detection. All the results are smaller, which means the effect is better. Performance of `ACCF` with `BigClam` is better than the one with `CESNA` in a range of $[0, 0.006)$.

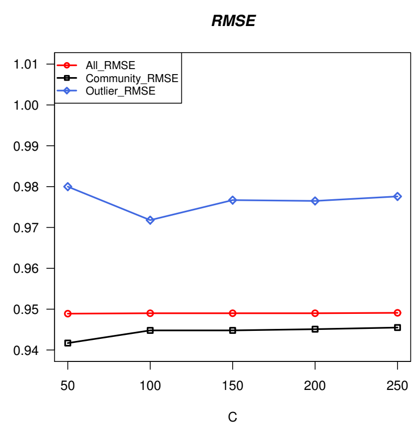

FIGURE 4. Impact of Social Conformity (Ciao dataset). `All` measures the overall prediction error on all users; `Community` measures the prediction error on those users who belong to some community; `Outlier` measures the prediction error on those users who do not belong to any community.

the users within some community (Community-MAE and Community-RMSE) are lower than the ones computed on outlier users (Outlier-MAE and Outlier-RMSE). This suggests we can improve the prediction accuracy by incorporating the users' social conformity to communities. The same effect can also be observed on the Epinions and Flixster datasets.

VI. CONCLUSION

In this paper, we propose a recommendation framework ACCF which exploits social conformity of users to their communities. We first group users into overlapping communities, and then incorporate the community preference vectors into the latent factorization model, and learn the weight of each community's influence through an attention network. When compared with seven state-of-the-art baseline methods on three real-world datasets, ACCF achieves the best performance. In the future, we will explore the problem of item recommendation with no explicit ratings. A different direction is to exploit the method proposed in [27] which infers the strengths of social ties based on community structure into the social recommendation.

REFERENCES

[1] Dzmitry Bahdanau, Kyunghyun Cho, and Yoshua Bengio. Neural machine translation by jointly learning to align and translate. CoRR, abs/1409.0473, 2014.

[2] James Bergstra, Olivier Breuleux, Frédéric Bastien, Pascal Lamblin, Razvan Pascanu, Guillaume Desjardins, Joseph Turian, David Warde-Farley, and Yoshua Bengio. Theano: a cpu and gpu math expression compiler. In Proceedings of the Python for scientific computing conference (SciPy), volume 4. Austin, TX, 2010.

[3] Chenwei Cai, Ruining He, and Julian McAuley. SPMC: socially-aware personalized markov chains for sparse sequential recommendation. In Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence, IJCAI 2017, Melbourne, Australia, August 19-25, 2017, pages 1476–1482, 2017.

[4] Jingyuan Chen, Hanwang Zhang, Xiangnan He, Liqiang Nie, Wei Liu, and Tat-Seng Chua. Attentive collaborative filtering: Multimedia recommendation with item- and component-level attention. In Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval, Shinjuku, Tokyo, Japan, August 7-11, 2017, pages 335–344, 2017.

[5] Rui Chen, Qingyi Hua, Bo Wang, Min Zheng, Weili Guan, Xiang Ji, Quanli Gao, and Xiangjie Kong. A novel social recommendation method fusing user's social status and homophily based on matrix factorization techniques. IEEE Access, 7:18783–18798, 2019.

[6] R. B. Cialdini and N. J. Goldstein. Social influence: Compliance and conformity. Annual Review of Psychology, 55:591–621, 2004.

[7] David A. Easley and Jon M. Kleinberg. Networks, Crowds, and Markets - Reasoning About a Highly Connected World. Cambridge University Press, 2010.

[8] Jennifer Golbeck. Generating predictive movie recommendations from trust in social networks. In Trust Management, 4th International Conference, iTrust 2006, Pisa, Italy, May 16-19, 2006, Proceedings, pages 93–104, 2006.

[9] Guibing Guo, Jie Zhang, and Neil Yorke-Smith. Trustsvd: Collaborative filtering with both the explicit and implicit influence of user trust and of item ratings. In Proceedings of the Twenty-Ninth AAAI Conference on

Artificial Intelligence, January 25-30, 2015, Austin, Texas, USA., pages 123–129, 2015.

[10] Xiangnan He, Zhankui He, Jingkuan Song, Zhenguang Liu, Yu-Gang Jiang, and Tat-Seng Chua. NAIS: neural attentive item similarity model for recommendation. IEEE Trans. Knowl. Data Eng., 30(12):2354–2366, 2018.

[11] Mohsen Jamali and Martin Ester. TrustWalker: a random walk model for combining trust-based and item-based recommendation. In Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Paris, France, June 28 - July 1, 2009, pages 397–406, 2009.

[12] Mohsen Jamali and Martin Ester. A matrix factorization technique with trust propagation for recommendation in social networks. In Proceedings of the 2010 ACM Conference on Recommender Systems, RecSys 2010, Barcelona, Spain, September 26-30, 2010, pages 135–142, 2010.

[13] Y. Koren and R. Bell. Advances in collaborative filtering. Recommender Systems Handbook, pages 145–186, 2011.

[14] Yehuda 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, Las Vegas, Nevada, USA, August 24-27, 2008, pages 426–434, 2008.

[15] Yehuda Koren, Robert Bell, and Chris Volinsky. Matrix factorization techniques for recommendation systems. IEEE Computers, 2009.

[16] Hui Li, Sourav S. Bhowmick, and Aixin Sun. CASINO: towards conformity-aware social influence analysis in online social networks. In Proceedings of the 20th ACM Conference on Information and Knowledge Management, CIKM 2011, Glasgow, United Kingdom, October 24-28, 2011, pages 1007–1012, 2011.

[17] Hui Li, Dingming Wu, Wenbin Tang, and Nikos Mamoulis. Overlapping community regularization for rating prediction in social recommender systems. In Proceedings of the 9th ACM Conference on Recommender Systems, RecSys 2015, Vienna, Austria, September 16-20, 2015, pages 27–34, 2015.

[18] Dugang Liu, Jie Huang, and Chen Lin. Recommendation with social roles. IEEE Access, 6:36420–36427, 2018.

[19] Xin Liu and Karl Aberer. Soco: a social network aided context-aware recommender system. In 22nd International World Wide Web Conference, WWW '13, Rio de Janeiro, Brazil, May 13-17, 2013, pages 781–802, 2013.

[20] Yong Liu, Peilin Zhao, Xin Liu, Min Wu, Lixin Duan, and Xiao-Li Li. Learning user dependencies for recommendation. In Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence (IJCAI-17), pages 2379–2385, 2017.

[21] Hao Ma. An experimental study on implicit social recommendation. In The 36th International ACM SIGIR conference on research and development in Information Retrieval, SIGIR '13, Dublin, Ireland - July 28 - August 01, 2013, pages 73–82, 2013.

[22] Hao Ma, Irwin King, and Michael R. Lyu. Learning to recommend with social trust ensemble. In Proceedings of the 32nd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR 2009, Boston, MA, USA, July 19-23, 2009, pages 203–210, 2009.

[23] Hao Ma, Haixuan Yang, Michael R. Lyu, and Irwin King. Sorec: social recommendation using probabilistic matrix factorization. In Proceedings of the 17th ACM Conference on Information and Knowledge Management, CIKM 2008, Napa Valley, California, USA, October 26-30, 2008, pages 931–940, 2008.

[24] Hao Ma, Dengyong Zhou, Chao Liu, Michael R. Lyu, and Irwin King. Recommender systems with social regularization. In Proceedings of the Forth International Conference on Web Search and Web Data Mining, WSDM 2011, Hong Kong, China, February 9-12, 2011, pages 287–296, 2011.

[25] Paolo Massa and Paolo Avesani. Trust-aware recommender systems. In Proceedings of the 2007 ACM Conference on Recommender Systems, RecSys 2007, Minneapolis, MN, USA, October 19-20, 2007, pages 17–24, 2007.

[26] Steffen Rendle, Christoph Freudenthaler, and Lars Schmidt-Thieme. Factorizing personalized markov chains for next-basket recommendation. In Proceedings of the 19th International Conference on World Wide Web, WWW 2010, Raleigh, North Carolina, USA, April 26-30, 2010, pages 811–820, 2010.

[27] Polina Rozenshtein, Nikolaj Tatti, and Aristides Gionis. Inferring the strength of social ties: A community-driven approach. In Proceedings of the 23rd ACM SIGKDD international conference on Knowledge discovery and data mining, pages 1017–1025, 2017.

[28] R. Salakhutdinov and A. Mnih. Probabilistic matrix factorization. In Advances in neural information processing systems, pages 1257–1264, 2008.

[29] X. Su and T.M. Khoshgoftaar. A survey of collaborative filtering techniques. Advances in Artificial Intelligence, 2009(12), 2009.

[30] Seyed Mohammad Taheri, Hamidreza Mahyar, Mohammad Firouzi, Elahe Ghalebi K., Radu Grosu, and Ali Movaghar. Extracting implicit social relation for social recommendation techniques in user rating prediction. In Proceedings of the 26th International Conference on World Wide Web Companion, Perth, Australia, April 3-7, 2017, pages 1343–1351, 2017.

[31] Jie Tang, Sen Wu, and Jimeng Sun. Confluence: conformity influence in large social networks. In The 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD 2013, Chicago, IL, USA, August 11-14, 2013, pages 347–355, 2013.

[32] Jiliang Tang, Xia Hu, Huiji Gao, and Huan Liu. Exploiting local and global social context for recommendation. In IJCAI 2013, Proceedings of the 23rd International Joint Conference on Artificial Intelligence, Beijing, China, August 3-9, 2013, pages 2712–2718, 2013.

[33] Jiliang Tang, Suhang Wang, Xia Hu, Dawei Yin, Yingzhou Bi, Yi Chang, and Huan Liu. Recommendation with social dimensions. In AAAI, pages 251–257, 2016.

[34] Fei Wang, Tao Li, Xin Wang, Shenghuo Zhu, and Chris H. Q. Ding. Community discovery using nonnegative matrix factorization. Data Min. Knowl. Discov., 22(3):493–521, 2011.

[35] Xin Wang, Steven C. H. Hoi, Martin Ester, Jiajun Bu, and Chun Chen. Learning personalized preference of strong and weak ties for social recommendation. In Proceedings of the 26th International Conference on World Wide Web, WWW 2017, Perth, Australia, April 3-7, 2017, pages 1601–1610, 2017.

[36] Xin Wang, Wei Lu, Martin Ester, Can Wang, and Chun Chen. Social recommendation with strong and weak ties. In Proceedings of the 25th ACM International Conference on Information and Knowledge Management, CIKM 2016, Indianapolis, IN, USA, October 24-28, 2016, pages 5–14, 2016.

[37] Bo Yang, Yu Lei, Dayou Liu, and Jiming Liu. Social collaborative filtering by trust. In IJCAI 2013, Proceedings of the 23rd International Joint Conference on Artificial Intelligence, Beijing, China, August 3-9, 2013, pages 2747–2753, 2013.

[38] Jaewon Yang and Jure Leskovec. Overlapping community detection at scale: a nonnegative matrix factorization approach. In Sixth ACM International Conference on Web Search and Data Mining, WSDM 2013, Rome, Italy, February 4-8, 2013, pages 587–596, 2013.

[39] Jaewon Yang, Julian McAuley, and Jure Leskovec. Community detection in networks with node attributes. In 2013 IEEE 13th International Conference on Data Mining, pages 1151–1156. IEEE, 2013.

[40] Xiwang Yang, Harald Steck, and Yong Liu. Circle-based recommendation in online social networks. In The 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '12, Beijing, China, August 12-16, 2012, pages 1267–1275, 2012.

[41] Tong Zhao, Julian J. McAuley, and Irwin King. Leveraging social connections to improve personalized ranking for collaborative filtering. In Proceedings of the 23rd ACM International Conference on Conference on Information and Knowledge Management, CIKM 2014, Shanghai, China, November 3-7, 2014, pages 261–270, 2014.