# Exploring Users' Internal Influence from Reviews for Social Recommendation

Guoshuai Zhao, Xiaojiang Lei, Xueming Qian, *Member, IEEE*, and Tao Mei, *Senior Member, IEEE*

***Abstract*—In recent years, we have witnessed a flourish of social review websites. Internet users can easily share their experiences on some products and services with their friends. Therefore, measuring interpersonal influence becomes a popular method for recommender systems. However, traditional works are all based on external tangible activities, such as following, retweeting, mentioning, etc. In this paper, we explore user internal factors to measure his/her influence on a specific domain, namely, the social network on local businesses. The proposed user internal factors include user sentimental deviations and the review's reliability. The internal factors are not from explicit behavior but could help us to understand users. In addition, we utilize an attention mechanism that could auto-learn the weights of different factors. Through a case study on the Yelp dataset, we found that the proposed user internal factors on influence, that is, the proposed user sentimental deviations and the review's reliability, are effective in improving the accuracy of rating predictions.**

***Index Terms*—Data mining, interpersonal influence, recommender system, review sentiment, social network.**

## I. Introduction

SOCIAL networking sites have become an influential platform for people to share their experiences, reviews, ratings, photos, videos, check-ins, and moods. Such information brings new opportunities for recommender systems and provides us a broad spectrum for mining users' preferences. The first generation of recommender systems with traditional collaborative filtering algorithms [10], [12], [27] mostly focus on personalized recommendations by predicting user preferences and ratings.

Manuscript received December 23, 2016; revised July 10, 2017, April 13, 2018, and June 8, 2018; accepted July 19, 2018. Date of publication August 6, 2018; date of current version February 21, 2019. This work was supported in part by the NSFC under Grants 61732008, 61772407, 61332018, and u1531141, and in part by the National Key R&D Program of China under Grant 2017YFF0107700. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Benoit Huet. *(Corresponding author: Xueming Qian.)*

G. Zhao is with the Key Laboratory for Intelligent Networks and Network Security, Ministry of Education, Xi'an Jiaotong University, Xi'an 710049, China, also with the SMILES Laboratory, Xi'an Jiaotong University, Xi'an 710049, China (e-mail: zgs2012@stu.xjtu.edu.cn).

X. Lei is with SMILES Laboratory, Xi'an Jiaotong University, Xi'an 710049, China (e-mail: xiaolei3439@stu.xjtu.edu.cn).

X. Qian is with the Key Laboratory for Intelligent Networks and Network Security, Ministry of Education, Xi'an Jiaotong University, Xi'an 710049, China, also with the SMILES Laboratory, Xi'an Jiaotong University, Xi'an 710049, China, and also with Zhibian Technology Co. Ltd., Taizhou 317000, China (e-mail: qianxm@mail.xjtu.edu.cn).

T. Mei is with JD AI Research, Beijing 100105, China (e-mail: tmei@live.com).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TMM.2018.2863598


Fig. 1. An example of user reviews on the Yelp website. The reviews contain three pieces of information: User ratings and reviews, overall product rating, and user's social information.

Recently, more researchers have begun to pay close attention to social information [34], [35], [46], [47], [63], [65], [66] and the interaction between users [23], [27], [30], [68]. Many approaches around interpersonal influence in social networks have proven their good performance in social-based recommendations. These methods can solve the "cold start" problems effectively. However, the existing approaches [2]– [5], [25], [26], [32], and [39] mainly take advantage of items' category information or tag information to study interpersonal trust. However, category and tag information are not always available on websites.

At the same time, there is much personal information in online textual reviews, which plays a very important role in decision processes. In Fig. 1, we show an example of social user reviews on the Yelp website. From Fig. 1, we can obtain three kinds of information. First, for a specific product, we can know the overall ratings by all customers. Second, for a specific user, we can obtain the profile and the rating information, including the reviews. Finally, we have the social information for all users. The key information from the review website can provide us with more opportunities in mining the interaction between users. For example, a customer will decide what to buy if he/she sees valuable reviews posted by others, especially by his/her trusted friends. Hence, how to mine review information and the reviewer's influence in social networks on local businesses has become an important issue in recommender systems.

On review websites, both positive reviews and negative reviews are valuable as references. In Fig. 1, there are two representative reviews. There are some positive words in the 5-star review, such as "*great*" and "*friendly*". However, in the 1-star review, we find negative words, such as "*noisy*" and "*rudeness*". Users may find polarized comments more convenient because they are quicker to read and understand, while users are not willing to spend lot of time to see very detailed information. In other words, polarized comments are more accessible, so they may have more influence on the users who are quickly browsing the reviews. Thus, in terms of accessibility, polarized comments may have more influence on users. If a reviewer always has the same opinion on different items, we think he/she does not show any biases and he/she may have a low influence on other users. Therefore, if a user provides polarized comments and variant sentiment on different items, he/she may have more influence on others. This is an internal factor for user influence.

Additionally, users have different ways to write reviews. Someone likes to write a long review, while others may like to write a short review. Then, we have some questions: does a long review really reflect the quality of the product/service? Are we willing to believe a long review more than a short one? In our opinion, different writing patterns could reveal different reliabilities of users. This is another internal factor for user influence.

In this paper, we propose a Review-based Recommendation Model (RRM). Our target is to improve the accuracy of rating predictions. The proposed method explores the impact of several factors on user influence, particularly, user internal factors on influence, including user sentimental deviations and review reliability. Additionally, we utilize an attention mechanism that could auto-learn the weights of different factors. That is, we explore combining some factors of user influence to enhance the accuracy of rating predictions. To differentiate from existing user influence measures [2]–[5], [49], [65] such as following, retweeting, mentioning and other interactive actions, we propose using user internal factors to measure his/her influence, including user sentimental deviations and review reliability. These internal factors are not from explicit behavior but could help us to understand users. Compared with previous works [1]–[6], [63], the main contributions of this paper are shown as follows.

1) We propose a review-based recommendation model by fusing users' internal influence into a matrix factorization to improve the accuracy of rating predictions.
2) We propose to leverage user sentimental deviations and a review's reliability to be the user internal factors on social influence. User sentimental deviations extracted from reviews help us to demonstrate that a user with clear and various opinions has more influence on others. In addition, a user who has a long review with many product features is more trustworthy. We fuse the user internal factors on social influence into a matrix factorization to enhance traditional methods.
3) Additionally, we utilize an attention mechanism that could auto-learn the weights of different factors. The attention mechanism enhances the scalability of our model because it provides an ability that could fuse multiple factors, even though the factor may be not substantial. We implement a case study on Yelp. Its experimental results demonstrate the effectiveness of our model on improving the accuracy of rating predictions.

The rest of this paper is organized as follows. We first provide an overview of related work in Section II. Then, we define our focus in Section III. Our review-based recommendation model is presented in detail in Section IV. Finally, we report the experimental results and analysis in Section V, and Section VI concludes this paper.

## II. Related Work

### A. Recommender System

With the ability to take advantage of the wisdom of crowds, Collaborative Filtering (CF) [10], [12], [24], [27], [40], [67] technique has achieved great success in personalized recommender systems, especially in rating prediction tasks. The task of CF is to predict users' preferences for unrated items. Item-based CF [12] produces the rating from a user to an item based on the average ratings of similar or correlated items by the same user. Cai *et al.* [40] investigated the collaborative filtering recommendation from a new perspective and present a novel typicality-based collaborative filtering recommendation. They improve the accuracy of predictions, and their method works well even with sparse training data sets.

Recently, Latent Factor Models based on Matrix Factorization [1], [23] have gained great popularity as they usually outperform traditional methods and have achieved great performance in some acknowledged datasets [28]. All kinds of MF algorithms have been proposed for solving different problems, such as Singular Value Decomposition (SVD) [43], Probabilistic Matrix Factorization [1], Non-negative Matrix Factorization [7], Max-Margin Matrix Factorization [8], and Localized Matrix Factorization [9]. They aim at learning latent factors from user-item rating matrices to make rating predictions, based on which to generate personalized recommendations. However, their latent characteristics suffer some problems when they faced with new users, and we define this situation as the "cold start" problem.

### B. User Influence

The flourishing of social media has promoted research on the social-based recommendation. Some Matrix factorization based social recommendations, e.g., Context MF [3], Social MF [4], and PRM [5] are proposed to solve the "cold start" problems. Besides, they also explore individual preferences. The basic idea is that user latent feature should be similar to the average of his/her friends' latent features with the weights of users' preference similarity. There are also some social trust-based works that try to calculate the interpersonal influence [2], [25], [26], [30], [39]. Most of them calculate the similarity between users to denote the trust value. Yang *et al.* [2] proposed the concept of "Trust Circles" in the social network based on probabilistic matrix factorization. Sato *et al.* [26] predicted the user's individual preference and influence from other users by applying the knowledge of probability and statistics. However, these methods are all restricted to the structured data, i.e., the used category information [2], [5] and the tag information [26] are not always available in some social networks.

It is important to notice that the increasingly growing amount of textual reviews users generated contain rich information about user preferences and item descriptions. There are some reviews based works for the task of recommendation. User topic based recommendation has attracted much attention to mine users' preferences [11], [13], [21], [23], [24], [36], [38]. Jiang *et al.* [24] proposed an author topic model-based collaborative filtering (ATCF) method to facilitate comprehensive points of interest (POI) recommendations for social users. There are also many works focus on user review details. Peng *et al.* [44] predicted the respective effects of review length and emotional intensity, which are used to emphasize reviewers' trustworthiness. Pan *et al.* [45] examined the effects of review characteristics, product type, and reviewer characteristics on perceived review helpfulness. They demonstrate that both review valence and length have positive effects on review helpfulness.

### C. Review Sentiment Analysis

Many sentiment analysis works are proposed to extract user preferences. Sentiment analysis can be conducted on three different levels: review-level, sentence-level, and phrase-level. Review-level analysis [18], [19], [20] and sentence-level analysis [17] attempt to classify the sentiment of a whole review or sentence to one of the predefined sentiment polarities, including positive, negative and sometimes neutral. Pang *et al.* [19] proposed a context insensitive evaluative lexical method. However, they cannot deal with the mismatch between the base valence of the term and the author's usage. While phrase-level analysis [6], [15], [16] attempt to extract the sentiment polarity of each feature word that a user expressed his/her attitude to the specific feature of a specific product. The main task of phrase-level sentiment analysis is the construction of sentiment lexicon [14], [31], [33], [42], [49]. There are some works attempt to leverage sentiment analysis to extract user features and product features for personalized recommendation [6], [42]. Lei *et al.* [49] also leveraged users' phrase-level sentiment analysis to make a personalized recommendation. They use user sentiment to infer a specific items' reputation, which has a good performance in a public dataset.

Generally, user's long-term interest is stable [57], [58], [59]. Extracting user's preferences with the content of reviews has received considerable attentions in recent years. Many topic models introduce user's interests as topic distributions according to their reviews contents [21], [29]. Most of them mine users' preferences based on a popular approach, LDA [11]. LDA is a Bayesian model. It is utilized to model the relationship of documents, topics, and words. Sentiment classification is one of the most popular works in user opinion mining. Existing works [19], [20] mainly focus on positive-negative binary classification. Normally, according to the psychological theories on sentiment, reviews are generally divided into two groups, positive and negative. To make a purchase decision, users not only need to know whether the product is good or not but also how good the product is [22].

TABLE I
SYMBOLS AND THEIR DESCRIPTIONS

| Symbol | Description | Symbol | Description |
|---|---|---|---|
| $U$ | a set of users | $P$ | a set of items |
| $M$ | user numbers | $N$ | item numbers |
| $\boldsymbol{R}_{M\times N}$ | the rating matrix | $\hat{\boldsymbol{R}}_{M\times N}$ | the predicted rating matrix |
| $\boldsymbol{U}_{M\times k}$ | the user latent feature matrix | $\boldsymbol{P}_{N\times k}$ | the item latent feature matrix |
| $E_{u,i}$ | user $u$'s sentiment on item $i$ | $R_w$ | initial score of the sentiment word w |
| $D(E_v)$ | user $v$'s sentiment variance | $F_v$ | the set of user v's friends |
| $\Psi$ | the objective function | $\lambda,\mu,\alpha,\beta,\gamma$ | the parameters in the objective function |
| $S_{u,v}$ | user $v$'s influence on user $u$ | $a$, $b$ | the coefficient of feature fitting formula |
| $H_v$ | user $v$'s popularity | $r_v$ | user $v$'s reliability |
| $V_w$ | word $w$'s feature vector | $F_{review}$ | the feature vector of a review |
| $K(\cdot)$ | kernel function | $y_{u,i}$ | the sentiment classification result of the review user $u$ to item $i$. |

Recognizing malicious users and reviews are important [53], [54], [55], [56]. Liu and Sun [53] proposed a scheme that identifies malicious users and recovered reputation scores by exploring the combination of temporal analysis and user correlation analysis. Zeng *et al.* [55] proposed an Equal Rating Opportunity (ERO) evaluation to detect dishonest reviews. Zeng *et al.* [56] proposed an algorithm that utilizes a detection method of the rating change interval. They leveraged the analysis of variance method to detect whether a product's reviews are manipulated.

Ling *et al.* [60] proposed a unified model that combines content-based filtering with collaborative filtering, harnessing the information of both ratings and reviews. Ganu *et al.* [61] identified topical and sentiment information from free-form text reviews, and grouped similar users together using soft clustering techniques to improve user experience in accessing reviews. Chen *et al.* [62] provided a comprehensive overview of how the review elements have been exploited to improve standard content-based recommending, collaborative filtering, and preference-based product ranking techniques. Wang *et al.* [51] presented a simple model variant where an SVM is built over NB log-count ratios as feature values by combining generative and discriminative classifiers and demonstrated its effectiveness and robustness.

## III. PROBLEM FORMULATION

The symbols and notations utilized in this paper are given in Table I. Our approach aims to recommend user interested items based on their historical review records and interpersonal relationships in social networks. We have three key factors to infer the trust value in social networks. They are the user's internal factors for user influence, such as user sentimental deviations and review reliability, and the user's external factor, i.e., user popularity. We have a set of users $U = \{u_1, u_2, \ldots, u_M\}$ and

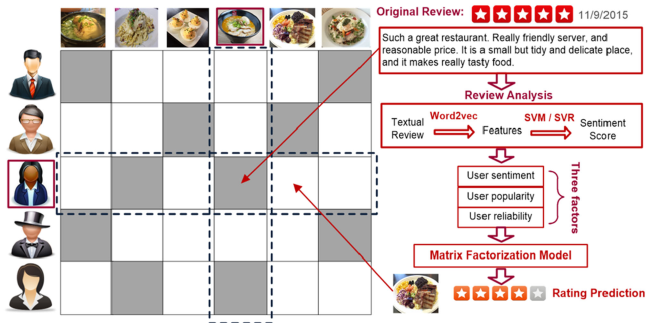


Fig. 2. An overview of our rating prediction method. Each shaded block is a review made by a user on an item. The features of these reviews are extracted by word2vec model. Then sentiment scores are calculated by SVM/SVR method. Finally, three factors, including user sentiment, user reliability, and user popularity, are fused to matrix factorization model for rating prediction.

a set of items $P = \{i_1, i_2, \ldots, i_N\}$. The ratings expressed by users on items are given in a rating matrix $R \in R_{M \times N}$. In this matrix, $R_{u,i}$ denotes the rating stars given by user $u$ on item $i$, which can be any real number but generally is an integer in the range of 1 to 5. In a social rating network, each user $u$ has a set of friends and $S_{u,v} \in [0, 1]$, which denotes the influence score of user $v$ to user $u$. Meanwhile, $D(E_u)$ denotes user $u$'s sentiment deviations. This helps us to judge whether the user has clear and various opinions or not. We use the sentiment deviations to calculate the user's sentimental influence. Then, $H_u$ denotes user $u$'s popularity, which we determine by calculating the sum of the number of users' friends and the number of users' ratings. $r_u$ denotes user $u$'s reliability. We leverage the average length of reviews to infer whether a user is responsible for writing a review.

The task of our recommender system is to predict user $u$'s ratings on an unknown item $i$. In addition, we explore the influence of user sentiment, user reliability, and user popularity. We employ matrix factorization techniques to learn the latent features of users and items. $\boldsymbol{U}_{M \times k}$ denotes the latent user feature matrix, and $\boldsymbol{P}_{N \times k}$ denotes the latent item feature matrix, with row vectors $U_u$ and $P_i$ representing $k$-dimensional user-specific and item-specific latent feature vectors of user $u$ and item $i$. Then, we use the learned latent factors of users and items to predict user's ratings on items.

The purpose of our method is to find the user internal influence and external influence from reviews and then fuse them into recommender systems. We give an overview of our rating prediction method in Fig. 2. We can see different users have different rating preferences, and each shaded block is a review made by a user on an item in the user-item review matrix. The review processing on the right side of Fig. 2 shows how to extract user sentiment from an original review. The features of these reviews are extracted by the word2vec model. Then, sentiment scores are calculated by the SVM/SVR method. Finally, we fuse the internal factors, i.e., user sentiment deviations and user reliability, and the external factor, i.e., user popularity, into a matrix factorization framework with an attention mechanism for rating prediction tasks.

## IV. Review-Based Recommendation Model

First, we introduce the internal factors for the user influence. Second, the attention mechanism is formulated. Finally, our model is expressed, and its training procedure is presented.

### A. Internal Factors on User Influence

In this subsection, we mainly focus on exploring the internal factors on influence, including user sentimental deviations and review reliability. However, first we present the sentiment calculation method.

*1) Sentiment Calculation:* In this work, we use a word2vec [50] -based sentiment analysis method to calculate the user's sentiment. We leverage the word2vec model to extract the feature vector of each word. The word $w$'s feature vector is $V_w = \{v_1, v_2, \ldots v_d\}$. The dimension $d$ of the word's feature vector $V_w$ is set as the default value 200. Then, the feature vector $F$ of a review is represented by

$$F_{review} = \frac{1}{|review|} \sum_{w \in review} V_w \tag{1}$$

Support Vector Machines (SVM) are often utilized for text classification [51]. Their performance varies greatly depending on the features. We leverage the features of reviews $F_{review}$ extracted by the word2vec model to train the SVM for sentiment classification as follows.

$$y_{u,i} = sign\left(\sum_{j=1}^{\ell} y_j \varepsilon_j K\left(F_j, F_{review}\right) + b\right) \tag{2}$$

where $y_{u,i}$ is the sentiment classification result of the review user $u$ to item $i$, $\ell$ is the count of the training data, $y_j$ is the classification of review $j$ in the training data $\varepsilon_j$ and $b$ are learned by SVM, and $K(\cdot)$ is the kernel function. More details about SVM can be found in [51].

In addition, Support Vector Regression (SVR) [52] can be used to predict the sentiment score $E_{u,i}$ from the feature vector $F_{review}$ of the review user $u$ gave to item $i$. The decision function is:

$$E_{u,i} = \sum_{j=1}^{\ell} \left(\varepsilon_j - \varepsilon_j^*\right) K\left(F_j, F_{review}\right) + b \tag{3}$$

where $\ell$ is the count of training data, $\varepsilon_i$ and $\varepsilon_i^*$ are Lagrange multipliers, and they are a set of dual variables. $K(\cdot)$ is the kernel, and we train the SVM and the SVR with the Radial Basis Function (RBF) kernel in this paper. More details about the SVR can be found in [52].

*2) User Sentiment for Interpersonal Influence:* When we search the Internet for purchasing, we are more concerned with the users who posted five-star reviews or critical reviews. In particular, critical reviews can truly reflect the deficiency of the product. In this case, we observe that reviewer sentiment will influence others. If a reviewer expresses polarized opinions, users may get an objective evaluation of items/products more easily. In this paper, we argue that the sentimental influence of the user's friends will offer more help when the user decides. If a

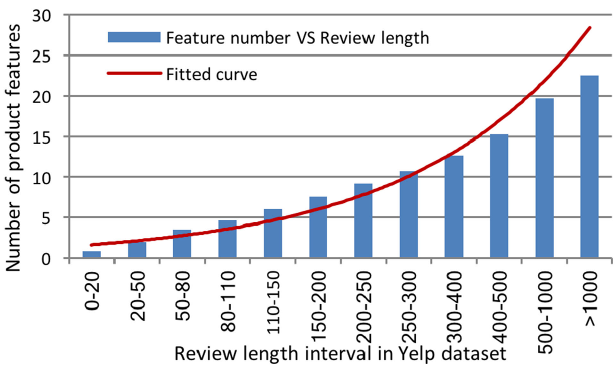


Fig. 3. A statistic of product features *VS* review length interval on the Yelp Website. The red curve is fitted by all Yelp review data, and according to the review length (total words number in a review), we divide all reviews into 12 intervals.

reviewer always has the same opinion on different items, we think he/she does not show any biases and he/she may have a low influence on others. Conversely, if a reviewer offers polarized and variant sentiments on different items, we think his/her ratings may have more influence.

Generally, in probability theory and mathematical statistics, the variance is used to measure the degree of deviation between a random variable and its mathematical expectation. Therefore, we calculate the sentimental influence, taking advantage of the concept of variance. The definition of variance is given by

$$D\left(E_u\right) = \frac{1}{n}\sum_{i=1}^{n}\left(E_{u,i} - \bar{E}_u\right)^2 \tag{4}$$

where $E_{u,i}$ denotes user $u$'s sentiment score on item $i$ and $\bar{E}_u$ is the average sentiment score of the items user $u$ has reviewed.

*3) User Reliability for Interpersonal Influence:* Users may have different attitudes when writing reviews after consumption; some may like to write a long and redundant review, while others may like to write a short review. Does a long review really reflect the quality of the product/service? We put forward an idea that a user who posts reviews with abundant descriptions and adequate features will be trustworthy. Hence, we believe he/she has high reliability. In this section, we study the relation between review length and product features. Note that product features are extracted by the method proposed in [49]. Fig. 3 shows the fitted curve using Yelp data. From Fig. 3, we can see that the fitted curve is approximately an exponential distribution. Based on different categories, we can set corresponding exponential coefficients to fit the curve. We define the reliability of user $u$ as follows:

$$r_u = a \times e^{bx} \tag{5}$$

where $a$ and $b$ are coefficients based on different categories and $x$ denotes a specific review length interval.

### B. An Attention Mechanism for the Combination of Factors

Generally, we deem that a user with many rating records and many friends is a popular user. In this paper, we define user popularity as follows:

$$H_u = |F_u| + |R_u| \tag{6}$$

where $F_u$ denotes the set of user $u$'s friends, $R_u$ denotes the set of user $u$'s ratings, and $|x|$ denotes the number of $x$.

After taking the three interpersonal influence factors above into consideration, we obtain user $v$'s influence on user $u$ in our recommendation model, which is defined as follows:

$$\begin{aligned} S_{u,v} &= \alpha\frac{D_v}{\sum_{v\in F_u} D_v} + \beta\frac{r_v}{\sum_{v\in F_u} r_v} + \gamma\frac{H_v}{\sum_{v\in F_u} H_v} \\ &= \alpha D^*_{u,v} + \beta r^*_{u,v} + \gamma H^*_{u,v} \end{aligned} \tag{7}$$

where $D(E_v)$ is user $v$'s sentiment variance, $r_v$ denotes user $v$'s reliability, $H_v$ denotes user $v$'s popularity, $F_u$ denotes the set of user $u$'s friends. In addition,$\alpha, \beta, \gamma$ are the coefficients of the three interpersonal factors, and $\alpha + \beta + \gamma = 1$.

To fuse the interpersonal influence factors into the matrix factorization model, we normalize $S_{u,v}$ as follows:

$$S^*_{u,v} = \frac{S_{u,v}}{\sum_{v\in F_u} S_{u,v}} = \frac{\alpha D^*_{u,v} + \beta r^*_{u,v} + \gamma H^*_{u,v}}{\sum_{v\in F_u}\left(\alpha D^*_{u,v} + \beta r^*_{u,v} + \gamma H^*_{u,v}\right)} \tag{8}$$

where $*$ denotes the normalization symbol, and each of the rows is normalized to unity, i.e., $\sum_v S^*_{u,v} = 1$.

### C. Review-Based Recommendation Model

In the matrix factorization framework, we predict the unknown ratings as follows:

$$\hat{R}_{u,i} = \bar{R} + U_u P_i^T \tag{9}$$

where $\hat{R}_{u,i}$ denotes the predicted rating given by user $u$ to item $i$ and $\bar{R}$ denotes the average value of all ratings.

We learn the coefficients of the three interpersonal factors and latent features of users and items on the observed rating data by minimizing the objective function. The objective function $\Psi$ is defined as follows:

$$\begin{aligned} &\Psi\left(\boldsymbol{R}, \boldsymbol{U}, \boldsymbol{P}\right) \\ &= \frac{1}{2}\sum_{u,i}\left(\hat{R}_{u,i} - R_{u,i}\right)^2 + \frac{\lambda}{2}\left(\|\boldsymbol{U}\|_F^2 + \|\boldsymbol{P}\|_F^2\right) \\ &\quad + \frac{\mu}{2}\sum_{u}\left(\left(U_u - \sum_{v} S^*_{u,v}U_v\right)\left(U_u - \sum_{v} S^*_{u,v}U_v\right)^T\right) \end{aligned} \tag{10}$$

where $R_{u,i}$ denotes user $u$'s real ratings on item $i$, $\boldsymbol{R}$ is the set of users' ratings on items, $R_{u,i} \in R_{M\times N}$, $M$ is the number of users, and $N$ is the number of items. $\boldsymbol{U}_{M\times k}$ and $\boldsymbol{P}_{N\times k}$ denote the user latent feature vectors and the item latent feature vectors, respectively. $U_u$ and $P_i$ are $k$-dimensional latent feature vectors of user $u$ and item $i$. They are obtained by the gradient descent method [2], [5]. The first term of (10) denotes the deviation between the actual ratings and the predictions, and the second term of (10) is a regularization term, which plays a role in case of overfitting. The factor of interpersonal influence is enforced by the third term. It means that if a user has three characteristics, e.g., popularity, reliability, and clear and various opinions, his/her friends may trust him/her more. Note that we utilize a learning method to fit the optimized weights $\alpha, \beta$ and $\gamma$.

### D. Model Training

We get the matrix factorization model as (10), from which we can obtain the user latent profile $U_u$ and the item latent profile $P_i$ by using the gradient descent method [2], [5], [49]. More formally, the gradients of the objective function with respect to the variables $U_u$,$P_i$, $\alpha$, $\beta$, and $\gamma$ are shown as follows:

$$\begin{aligned}\frac{\partial \Psi}{\partial U_u} = &\sum_i \left(\hat{R}_{u,i} - R_{u,i}\right) P_i + \lambda U_u \\ &+ \mu \left(U_u - \sum_{v \epsilon F_u} S^*_{u,v} U_v\right) \\ &- \mu \sum_{u \epsilon F_v} S^*_{v,u} \left(U_v - \sum_{w \epsilon F_v} S^*_{v,w} U_w\right)\end{aligned} \tag{11}$$

$$\frac{\partial \Psi}{\partial P_i} = \sum_u \left(\hat{R}_{u,i} - R_{u,i}\right) U_u + \lambda P_i \tag{12}$$

$$\begin{aligned}\frac{\partial \Psi}{\partial \alpha} = &\mu \sum_u \left(\left(U_u - \sum_v S^*_{u,v} U_v\right) \times \left(-\sum_v U_v\right.\right. \\ &\left.\left.\left(\frac{\left(D^*_{u,v} - H^*_{u,v}\right) \sum_{v \in F_u} S_{u,v} - S_{u,v} \sum_v \left(D^*_{u,v} - H^*_{u,v}\right)}{\left(\sum_{v \in F_u} S_{u,v}\right)^2}\right)\right)\right)\end{aligned} \tag{13}$$

$$\begin{aligned}\frac{\partial \Psi}{\partial \beta} = &\mu \sum_u \left(\left(U_u - \sum_v S^*_{u,v} U_v\right) \times \left(-\sum_v U_v\right.\right. \\ &\left.\left.\left(\frac{\left(r^*_{u,v} - H^*_{u,v}\right) \sum_{v \in F_u} S_{u,v} - S_{u,v} \sum_v \left(r^*_{u,v} - H^*_{u,v}\right)}{\left(\sum_{v \in F_u} S_{u,v}\right)^2}\right)\right)\right)\end{aligned} \tag{14}$$

where the initial values of $U_u$ and $P_i$ are sampled from the normal distribution with zero mean. We set $\alpha = \beta = \gamma = 1/3$ as the initial values of $\alpha, \beta, \gamma$. The user and item latent feature vectors $U_u$ and $P_i$ and the coefficients of influence factors are updated based on the previous values to ensure the fastest decrease of the objective function at each iteration. The update process is given by

$$U_u{}^{(t)} = U_u{}^{(t-1)} - \ell \frac{\partial \Psi^{(t-1)}}{\partial U_u} \tag{15}$$

$$P_i{}^{(t)} = P_i{}^{(t-1)} - \ell \frac{\partial \Psi^{(t-1)}}{\partial P_i} \tag{16}$$

$$\alpha^{(t)} = \alpha^{(t-1)} - \ell \frac{\partial \Psi^{(t-1)}}{\partial \alpha} \tag{17}$$

$$\beta^{(t)} = \beta^{(t-1)} - \ell \frac{\partial \Psi^{(t-1)}}{\partial \beta} \tag{18}$$

$$\gamma^{(t)} = 1 - \alpha^{(t)} - \beta^{(t)} \tag{19}$$

We set the step size $\ell = 0.0005$ and the total iteration number $\tau = 10000$ to ensure the decrease of the objective function while training.

## V. Experiments

We conducted a series of experiments to evaluate the performance of the proposed model on a specific dataset, i.e., the Yelp dataset. We have crawled nearly 60 thousand users' circles of friends and their rated items. Some previous works [5], [41], [42], [49], [64] are based on the Yelp dataset.[1] We have subsistent social relationships and user reviews to organize the experiments. The dataset contains eight categories: $\#^1$**Active Life**, $\#^2$**Beauty&Spa**, $\#^3$**Home Services**, $\#^4$**Hotel&Travel**, $\#^5$**Nightlife**, $\#^6$**Restaurants**, $\#^7$**Shopping**, and $\#^8$**Pets**. In total, there are 28,629 users, 96,974 items, and 300,847 ratings in our dataset, and we have every user's social relationships in our dataset. In addition, each item has been posted by at least one comment/review. In the following experiments, we first evaluate our sentiment algorithm by comparing it with the Lexicon Based Method [49] and NBSVM [51]. Then, we investigate how to use sentiment information for rating prediction with more accuracy. The compared approaches include BaseMF [1], CircleCon [2], ContextMF [3], PRM [5], EFM [6], IS [61], and RPS [49].

### A. Sentiment Evaluation

We leverage the word2vec model and SVM/SVR for sentiment analysis in this work. However, sentiment evaluation is performed generally by transforming each sentiment value $E_{u,i}$ into a binary value [33]. We transform the results of the SVR for sentiment evaluation. Namely, if the result is $E_{u,i} > 0$. the review will be regarded as positive. If the result is $E_{u,i} \leq 0$. the review will be regarded as negative.

When testing in a positive dataset, if $E_{u,i} \leq 0$, it means there was a misclassification. When testing in a negative dataset, if $E_{u,i} > 0$, it also means there was a misclassification. We first label all 5-star Yelp reviews as positive reviews and label all 1-star Yelp reviews as negative reviews. Then, we have 57,193 positive reviews and 9,799 negative reviews. We also utilize the Movie [48] and SFU [37] datasets for evaluation. The statistics and evaluation results are shown in Table II.

- **Lexicon Based Method** [49] extracts product features and utilized sentiment dictionary to calculate the user's sentiment.
- **NBSVM** is the compared method proposed in [51]. It identifies simple Naive Bayes and SVM variants which outperform most published results on sentiment analysis datasets, sometimes providing a new state-of-the-art performance level.
- **WV + SVM** and **WV + SVR** leverage the word2vec model and SVM/SVR to calculate the user's sentiment.

Table II shows that the NBSVM outperforms other methods in the Movie and SFU datasets. WV + SVM and WV + SVR have better performance on the Yelp dataset. We suppose that the review length significantly impacts the performance of sentiment classification. Note that the average length of reviews in

[1] http://smiles.xjtu.edu.cn/Download/Download_yelp.html

TABLE II
STATISTICS AND EVALUATION RESULTS

| Test Dataset | | Method | Precision of Positive | Precision of Negative | Average Precision |
|---|---|---|---|---|---|
| Movie | 2,000 reviews 0.78M words 394 words/review | Lexicon Based Method | 863/1,000 (86.3%) | 592/1,000 (59.2%) | 72.75% |
| | | NBSVM | 900/1,000 (90%) | 929/1,000 (92.9%) | 91.45% |
| | | WV+SVM | 853/1,000 (85.3%) | 812/1,000 (81.2%) | 83.25% |
| | | WV+SVR | 876/1,000 (87.6%) | 739/1,000 (73.9%) | 80.75% |
| SFU | 400 reviews 0.35 M words 876 words/review | Lexicon Based Method | 184/200 (92%) | 110/200 (55%) | 73.5% |
| | | NBSVM | 183/200 (91.5%) | 145/200 (72.5%) | 82% |
| | | WV+SVM | 158/200 (79%) | 137/200 (68.5%) | 73.75% |
| | | WV+SVR | 126/200 (63%) | 152/200 (76%) | 69.5% |
| Yelp | 66,992 reviews 9.02M words 134 words/review | Lexicon Based Method | 52,474/57,193 (91.75%) | 5,895/9,799 (60.16%) | 87.1% |
| | | NBSVM | 50,004/57,193 (87.43%) | 7,665/9,799 (78.22%) | 86.08% |
| | | WV+SVM | 54,276/57,193 (94.9%) | 7,682/9,799 (78.4%) | 92.49% |
| | | WV+SVR | 53,553/57,193 (93.64%) | 7,950/9,799 (81.13%) | 91.81% |

the Yelp dataset is 134 words, but in the Movie and SFU datasets, the average length is much longer. There are 394 words and 876 words per review, respectively, in Movie and SFU. From the feature vector of a review given in (1), we can see that this average mechanism of words' vectors is suitable for short texts, such as the reviews on Yelp. Meanwhile, the sentiment calculation is a basis of our model. The more accurate the sentiment calculation, the stronger the validation of our experiments.

With regard to the generalization to other domains or online services, WV + SVM is acceptable due to its good performance on short texts. For sentiment evaluation on short texts, we suggest audiences use WV + SVM. However, for long texts, NBSVM, which performs well on the Movie and SFU datasets, is a better choice than WV + SVM.

## *B. Rating Prediction Evaluation*

*1) Evaluation Metrics:* In each category of Yelp, we use 80% of the data as the training set and the remaining 20% as the test set. The evaluation metrics we use in our experiments are the Root Mean Square Error (RMSE) and the Mean Absolute Error (MAE). They are defined as follows:

$$RMSE = \sqrt{\sum_{i \in \Re_{test}} \left(\hat{R}_{u,i} - R_{u,i}\right)^2 \Big/ |\Re_{test}|} \tag{20}$$

$$MAE = \sum_{i \in \Re_{test}} \left|\hat{R}_{u,i} - R_{u,i}\right| \Big/ |\Re_{test}| \tag{21}$$

where $R_{u,i}$ is the real rating given by user $u$ to item $i$, $\hat{R}_{u,i}$ is the predicted rating and $\Re_{test}$ is the set of all items in the test set. $|\Re_{test}|$ denotes the number of items in the test set.

### *2) Compared Algorithms*

We conducted a series of experiments for comparison with our model. The compared algorithms are given as follows.

- **BaseMF:** This method is the basic matrix factorization approach proposed in [1], which does not consider any social factors.
- **CircleCon:** This method is proposed in [2], which focuses on the factor of interpersonal trust in the social network and infers the trust circle based on matrix factorization.
- **Context MF:** This method [3] improves the accuracy of traditional item-based collaborative filtering in [12] and SoRec in [4]. It takes both interpersonal influence and individual preference into consideration.
- **PRM:** This method is proposed in [5], which considers three social factors, including interpersonal influence, interpersonal interest similarity, and personal interest. It is also based on matrix factorization to predict users' ratings.
- **EFM:** This method is proposed in [6]. It builds two characteristic matrixes (i.e., user-feature attention matrix and item-feature quality matrix) in its framework. Each element in the user-feature attention matrix measures to what extent a user cares about the corresponding product feature/aspect. Each element in the item-feature quality matrix measures the quality of an item for the corresponding product feature/aspect.
- **IS:** This method is proposed in [61]. It identifies topical and sentiment information from free-form text reviews, and groups similar users together using soft clustering techniques to improve user experience in accessing reviews.
- **RPS:** This method [49] fuses user sentiment similarity, interpersonal sentiment influence, and item reputation similarity into a unified matrix factorization framework for rating prediction.
- **RRM:** It is the proposed Review-Based Recommendation Model, which explores social users' reviews. We have mined three interpersonal influence factors from users' reviews and then fused them into the matrix factorization framework with an attention mechanism for rating prediction.

### *3) Performance Comparison*

We compare the performance of our method with the existing models including BaseMF [1], CircleCon [2], ContextMF [3], PRM [5], EFM [6], IS [61], PRS [49] on the Yelp dataset. In the objective function of our model, $k$ is the dimension of user and item latent feature vectors $U_u$ and $P_i.\lambda$ is a coefficient for preventing overfitting. $\mu$ is a coefficient of interpersonal influence for exploring users' trust. $\alpha, \beta, \gamma$ are weights of the three interpersonal influence factors, respectively. We set the dimension of user and item latent features parameter $k =$ Please be consistent in using either a space or no space before and after mathematical symbols or operators. 10, the over-fitting parameter $\lambda = 1$, and set the interpersonal influence parameter $\mu = 2$.

In Table III, we show the comprehensive performance evaluation in eight categories of the Yelp dataset [5]. We used the color

TABLE III
PERFORMANCE COMPARISONS FOR EIGHT CATEGORIES ON YELP. DATA IN BLUEBERRY INDICATE THE BEST PERFORMANCE IN EACH CATEGORY, AND DATA IN AQUA REPRESENT THE SECOND-BEST PERFORMANCE

| Category | RMSE | | | | | | | | MAE | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | BaseMF | CircleCon | ContextMF | PRM | IS | EFM | RPS | RRM (Improvement) | BaseMF | CircleCon | ContextMF | PRM | IS | EFM | RPS | RRM (Improvement) |
| Active Life | 1.633 | 1.477 | 1.285 | 1.265 | 1.190 | 1.215 | 1.119 | **1.036 (7.4%)** | 1.238 | 1.126 | 1.002 | 0.984 | 0.929 | 0.941 | 0.876 | **0.79 (9.8%)** |
| Beauty | 1.813 | 1.656 | 1.454 | 1.431 | 1.378 | 1.385 | 1.287 | **1.195 (7.1%)** | 1.39 | 1.272 | 1.147 | 1.128 | 1.103 | 1.086 | 1.015 | **0.928 (8.6%)** |
| Home Services | 1.981 | 1.844 | 1.624 | 1.611 | 1.515 | 1.583 | 1.458 | **1.364 (6.4%)** | 1.558 | 1.454 | 1.294 | 1.284 | 1.236 | 1.273 | 1.177 | **1.123 (4.6%)** |
| Hotel&Travel | 1.683 | 1.539 | 1.337 | 1.321 | 1.181 | 1.267 | 1.115 | **1.103 (1.1%)** | 1.318 | 1.201 | 1.055 | 1.042 | 0.927 | 1.024 | 0.951 | **0.87 (6.1%)** |
| Nightlife | 1.408 | 1.311 | 1.176 | 1.150 | 1.173 | 1.134 | 1.092 | **1.052 (3.7%)** | 1.099 | 1.026 | 0.93 | 0.913 | 0.923 | 0.899 | 0.874 | **0.853 (2.4%)** |
| Pets | 1.873 | 1.715 | 1.499 | 1.481 | 1.516 | 1.436 | 1.344 | **1.23 (8.5%)** | 1.440 | 1.329 | 1.195 | 1.181 | 1.225 | 1.146 | 1.069 | **0.957 (10.5%)** |
| Restaurant | 1.261 | 1.202 | 1.149 | 1.094 | 1.203 | 1.113 | 1.070 | **1.043 (2.5%)** | 0.983 | 0.944 | 0.909 | 0.873 | 0.962 | 0.886 | 0.853 | **0.843 (1.2%)** |
| Shopping | 1.600 | 1.479 | 1.321 | 1.302 | 1.294 | 1.278 | 1.203 | **1.120 (6.9%)** | 1.228 | 1.138 | 1.032 | 1.016 | 1.053 | 0.999 | 0.941 | **0.878 (6.7%)** |
| Average | 1.657 | 1.528 | 1.356 | 1.332 | 1.306 | 1.301 | 1.211 | **1.143 (5.6%)** | 1.280 | 1.186 | 1.071 | 1.053 | 1.045 | 1.032 | 0.970 | **0.905 (6.7%)** |

of Blueberry to indicate the best performance in each category and leveraged the color of Aqua to represent the second-best performance. Our method performs better than others in every category. Additionally, we show the detailed improvements compared with the second-best performance. The overall increment is 5.6% and 6.7% for RMSE and MAE, respectively. Only the increments on Hotel&Travel (1.1% on RMSE) and Restaurant (1.2% on MAE) are lower than 2%. Therefore, the experiment results show the high accuracy of our approach.

### C. Discussion

In addition to the performance comparison in Table III, here we discuss four aspects of our experiments: 1) the learned weights by the attention mechanism, 2) the impact of the iteration count on the learned weights, 3) the impact of the factors, 4) the impact of less training data.

*1) Learned Weight by Attention Mechanism:* The attention mechanism utilized in this work is one of the main contributions. It is an auto-learning method for the weights of different factors. Here, we present the learned weights by the attention mechanism as shown in Fig. 4. The x-axis is the value of the weight. Alpha $\alpha$ is the weight of the sentimental deviations factor. Beta $\beta$ is the weight of the review's reliability factor, and gamma $\gamma$ is the weight of the user popularity factor. We find that the absolute values of $\alpha$ and $\gamma$ are much larger than that of $\beta$. Therefore, we suppose the factors of sentimental deviations and user popularity are much more important to our model than the review's reliability, no matter whether the impact is positive or negative. It can be concluded that the length of the review content has little relevance in improving the accuracy of our model. That is why the weight of review's reliability is always approximately 0. However, the results that $\alpha$ is positive and $\gamma$ is negative show the factor of sentimental deviations is more important than user popularity. Thus, $\alpha$ becomes larger and $\gamma$ decreases due to the sum of the three weights being a fixed value.

*2) The Impact of Iteration Count on Learned Weights:* Then, we track the log of the model training and try to find the relevance between weights and learning iterations. As shown in Fig. 5, the x-axis is the learning iterations and the y-axis indicates the values of weights. Due to the sum of the three weights being a fixed value, the gradient orientations are opposite to each other. We find that for the first 10 iterations, the fluctuations of weights are much larger than that after 10 iterations. We suppose that at the beginning, the value of our objective function is large so that the gradients of weights also become large. Thus, the fluctuations of the weights are high. After several iterations, the gradients become small so that the fluctuations also become smooth, and the weights start being learned in the right orientations. We suppose the importance of the factor of sentimental deviations is larger than user popularity, so the weight of sentimental deviations is increasing with learning iterations after 10 iterations. We suppose that the length of review content has

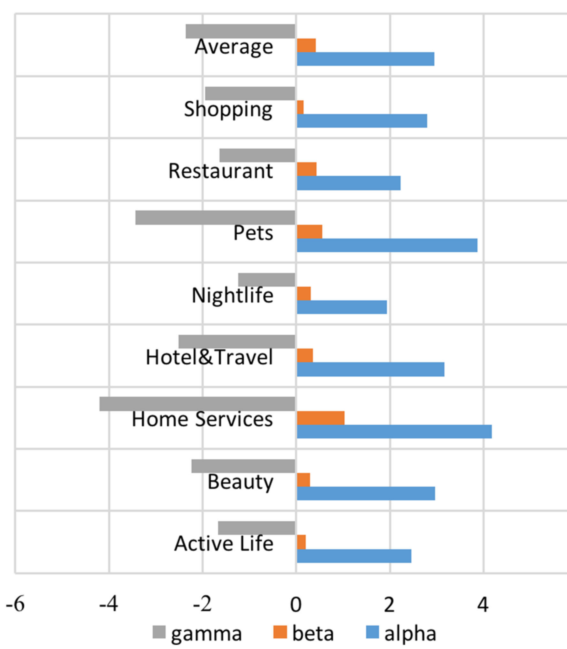


Fig. 4. The histogram shows the learned weights in different datasets.

TABLE IV
THE IMPACT OF LESS TRAINING DATA ON PERFORMANCE IN NIGHTLIFE DATASET

| | 40% training | 50% training | 60% training | 70% training | 80% training |
|---|---|---|---|---|---|
| RMSE | 1.108984 | 1.107797 | 1.106621 | 1.105989 | 1.051773 |
| MAE | 0.877526 | 0.868535 | 0.861652 | 0.858347 | 0.853248 |

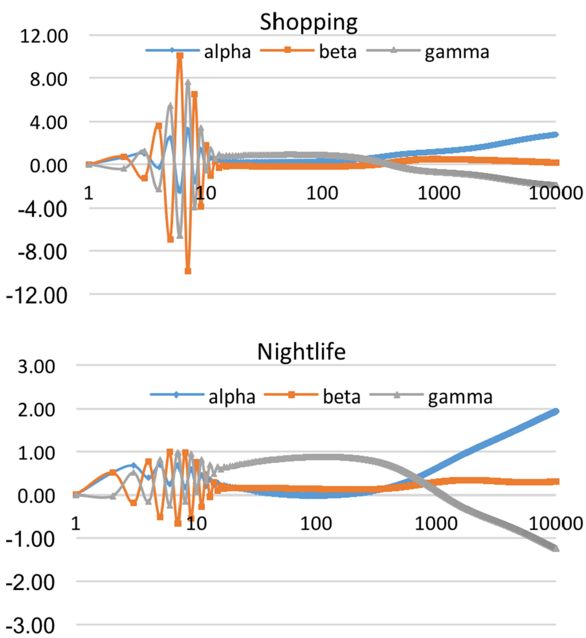


Fig. 5. The relevance between weights and learning iterations.

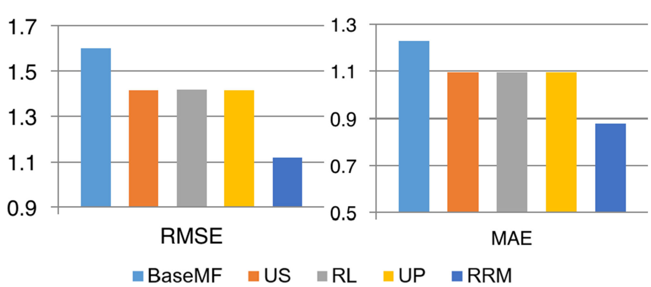


Fig. 6. The impact of each factor on performance in Shopping category.

slight relevance in improving our model so that its weight $\beta$ is always approximately 0. Due to $\alpha$ increasing and $\beta$ only slightly relevant, the weight $\gamma$ must decrease to fit the fixed sum of the three weights.

*3) The Impact of Factors:* In this part, we perform some experiments to validate the effectiveness of our hypotheses on performance, as shown in Fig. 6. We use the BaseMF [1] as a baseline and compare the results with the methods that consider different factors. US is the method considering the factor of user sentimental deviations. UR indicates the method of considering the factor of user reliability. UP denotes the method of using the factor of user popularity. RRM is the proposed model combining the three factors with an attention mechanism. We could find that each factor is effective, but the combination of them with an attention mechanism is the best way to improve the accuracy.

*4) The Impact of Less Training Data:* For the impact of less training data on the performance of our model, the data in the Nightlife category are used for this experiment. Table IV shows the impact of less training data on the performance of our model on Nightlife. In the process of model training, we randomly select some data from the complete dataset. 50% training data denotes that only 50% of our ratings are selected to train our model. It can be observed that there is little impact on performance. In addition, the performance of our model becomes worse with less training data.

## VI. CONCLUSION

In this paper, we implemented a case study on Yelp to explore social users' internal influence for recommender systems. We mined the fluctuation of user sentiments from textual reviews by utilizing the word2vec model and demonstrated that the users with higher sentimental variance may have more influence. We modeled the relationship between review length and product features with the exponential function and proved that the users who always post long reviews may have more influence. In addition, we showed the effectiveness of user popularity on performance. Finally, the combination of these factors with an attention mechanism was the final interpersonal influence to be used in the matrix factorization framework. Our model decreases the RMSE by at least 5.6% and the MAE by at least 6.7% compared with existing approaches.

One of the main contributions is that we utilize user internal factors to measure reviewer influence, such as the proposed user sentimental deviations and review reliability. Another contribution is that we utilize an attention mechanism that could auto-learn the weights of different factors. Through several experiments, we found that the utilized sentiment analysis methods based on word2vec are suitable for processing short texts. Additionally, considering more factors in the matrix factorization with the attention mechanism will obtain results that are more accurate.

In our future work, we will explore more factors and mine user social behaviors and reviews deeply for user attention mining. Textual reviews contain a large amount of information, and they can reveal users' actual needs and concerns, which are important to their decision making.

## REFERENCES

[1] R. Salakhutdinov and A. Mnih, "Probabilistic matrix factorization," in *Proc. 20th Int. Conf. Neural Inf. Process. Syst.*, 2008, pp. 1257–1264.
[2] X. Yang, H. Steck, and Y. Liu, "Circle-based recommendation in online social networks," in *Proc. 18th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, 2012, pp. 1267–1275.
[3] M. Jiang *et al.*, "Social contextual recommendation," in *Proc. 21st ACM Int. Conf. Inf. Knowl. Manage.*, 2012, pp. 45–54.

[4] M. Jamali and M. Ester, "A matrix factorization technique with trust propagation for recommendation in social networks," in *Proc. 4th ACM Conf. Recommender Syst.*, 2010, pp. 135–142.

[5] X. Qian, H. Feng, G. Zhao, and T. Mei, "Personalized recommendation combining user interest and social circle," *IEEE Trans. Knowl. Data Eng.*, vol. 26, no. 7, pp. 1763–1777, Jul. 2014.

[6] Y. Zhang, G. Lai, and M. Zhang, "Explicit factor models for explainable recommendation based on phrase-level sentiment analysis," in *Proc. 37th int. ACM SIGIR Conf. Res. Develop. Inf. Retrieval*, 2014, pp. 83–92.

[7] D. D. Lee and H. S. Seung, "Algorithms for non-negative matrix factorization," in *Proc. 13th Int. Conf. Neural Inf. Process. Syst.*, 2000, pp. 556–562.

[8] J. Rennie and N. Srebro, "Fast maximum margin matrix factorization for collaborative prediction," in *Proc. 22nd Int. Conf. Mach. Learn.*, 2005, pp. 713–719.

[9] Y. Zhang, M. Zhang, Y. Liu, S. Ma, and S. Feng, "Localized matrix factorization for recommendation based on matrix block diagonal forms," in *Proc. 22nd Int. Conf. World Wide Web*, 2013, pp. 1511–1520.

[10] X. Su and T. M. Khoshgoftaar, "A survey of collaborative filtering techniques," *Adv. Artif. Intell.*, vol. 2009, 2009, Art. no. 421425.

[11] D. M. Blei, A. Y. Ng, and M. I. Jordan, "Latent dirichlet allocation," *J. Mach. Learn. Res.*, vol. 3. pp. 993–1022, 2003.

[12] B. Sarwar, G. Karypis, J. Konstan, and J. Reidl, "Item-based collaborative filtering recommendation algorithms," in *Proc. 10th Int. Conf. World Wide Web*, 2001, pp. 285–295.

[13] S. Bao *et al.*, "Mining social emotions from affective text," *IEEE Trans. Knowl. Data Eng.*, vol. 24, no. 9, pp. 1658–1670, Sep. 2012.

[14] W. Zhang, G. Ding, L. Chen, C. Li, and C. Zhang, "Generating virtual ratings from Chinese reviews to augment online recommendations," *ACM Trans. Intell. Syst. Technol.*, vol. 4, no. 1, pp. 1–17, 2013.

[15] Y. Lu, M. Castellanos, U. Dayal, and C. Zhai, "Automatic construction of a context-aware sentiment lexicon: An optimization approach," in *Proc. 20th Int. Conf. World Wide Web*, 2011, pp. 347–356.

[16] T. Wilson, J. Wiebe, and P. Hoffmann, "Recognizing contextual polarity in phrase-level sentiment analysis," in *Proc. Conf. Human Lang. Technol. Empirical Methods Natural Lang. Process.*, 2005, pp. 347–354.

[17] T. Nakagawa, K. Inui, and S. Kurohashi, "Dependency tree-based sentiment classification using CRFs with hidden variables," in *Proc. Human Lang. Technol., Annu. Conf. North Amer. Chapter Assoc. Comput. Linguistics*, 2010, pp. 786–794.

[18] P. Lou, G. Zhao, X. Qian, H. Wang, and X. Hou, "Schedule a rich sentimental travel via sentimental POI mining and recommendation," in *Proc. IEEE 2nd Int. Conf. Multimedia Big Data*, 2016, pp. 33–40.

[19] B. Pang, L. Lee, and S. Vaithyanathan, "Thumbs up? sentiment classification using machine learning techniques," in *Proc. Conf. Empirical Methods Natural Lang. Process.*, 2002, pp. 79–86.

[20] D. Tang, Q. Bing, and T. Liu, "Learning semantic representations of users and products for document level sentiment classification," in *Proc. 53rd Annu. Meeting Assoc. Comput. Linguistics*, 2015, pp. 1014–1023.

[21] X. Ma, X. Lei, G. Zhao, and X. Qian, "Rating prediction by exploring user's preference and sentiment," *Multimedia Tools Appl.*, vol. 77, no. 6, pp. 6425–6444, 2018.

[22] F. Li *et al.*, "Incorporating reviewer and product information for review rating prediction," in *Proc. 22nd Int. Joint Conf. Artif. Intell.*, 2011, pp. 1820–1825.

[23] M. Jiang, P. Cui, F. Wang, W. Zhu, and S. Yang, "Scalable recommendation with social contextual information," *IEEE Trans. Knowl. Data Eng.*, vol. 26, no. 11, pp. 2789–2802, Nov. 2014.

[24] S. Jiang, X. Qian, J. Shen, Y. Fu, and T. Mei, "Author topic model based collaborative filtering for personalized POI recommendation," *IEEE Trans. Multimedia*, vol. 17, no. 6, pp. 907–918, Jun. 2015.

[25] G. Zhao, X. Qian, and C. Kang, "Service rating prediction by exploring social mobile users' geographical locations," *IEEE Trans. Big Data*, vol. 3, no. 1, pp. 67–78, Mar. 2017.

[26] T. Sato, M. Fujita, M. Kobayashi, and K. Ito, "Recommender system by grasping individual preference and influence from other users," in *Proc. Int. Conf. Adv. Social Netw. Anal. Mining*, 2013, pp. 1345–1351.

[27] W. Wang, Z. Chen, J. Liu, Q. Qi, and Z. Zhao, "User-based collaborative filtering on cross domain by tag transfer learning," in *Proc. 1st Int. Workshop Cross Domain Knowl. Discovery Web Social Netw. Mining*, 2012, pp. 10–17.

[28] T. Wang, X. Jin, X. Ding, and X. Ye, "User interests imbalance exploration in social recommendation: a fitness adaptation," in *Proc. Conf. Inf. Knowl. Manage.*, 2014, pp. 281–290.

[29] H. Lin, J. Jia, Q. Guo, Y. Xue, and Q. Li, "User-level psychological stress detection from social media using deep neural network," in *Proc. 22nd ACM Int. Conf. Multimedia*, 2014, pp. 507–516.

[30] P. Bedi, H. Kaur, and S. Marwaha, "Trust based recommender system for semantic web," in *Proc. Int. Joint Conf. Artif. Intell.*, 2007, pp. 2677–2682.

[31] T. N. Phan and Y. Myungsik, "The lexicon-based sentiment analysis for fan page ranking in Facebook," in *Proc. Int. Conf. Inf. Netw.*, 2014, pp. 444–448.

[32] H. Ma, I. King, and M. Lyu, "Learning to recommend with social trust ensemble," in *Proc. 36th Int. Conf. Res. Develop. Inf. Retrieval*, 2009, pp. 203–210.

[33] Y. Lu, M. Castellanos, U. Dayal, and C. Zhai, "Automatic construction of a context-aware sentiment lexicon: An optimization approach," in *Proc. 20th Int. Conf. World Wide Web*, 2011, pp. 347–356.

[34] Z. Wang *et al.*, "Joint social and content recommendation for user-generated videos in online social network," *IEEE Trans. Multimedia*, vol. 15, no. 3, pp. 698–709, Apr. 2013

[35] X. Wang *et al.*, "Semantic-based location recommendation with multimodal venue semantics," *IEEE Trans. Multimedia*, vol. 17, no. 3, pp. 409–419, Mar. 2015.

[36] X. Ding, B. Liu, and P. S. Yu, "A holistic lexicon-based approach to opinion mining," in *Proc. Int. Conf. Web Search Data Mining*, 2008, pp. 231–240.

[37] M. Taboada, C. Anthony, and K. Voll, "Methods for creating semantic orientation dictionaries," in *Proc. Conf. Lang. Res. Eval.*, 2006, pp. 427–432.

[38] V. Suresh, A. Veilumuthu, and A. Krishnamurthy, "A Non-syntactic approach for text sentiment classification with stop words," in *Proc. 20th Int. Conf. Companion World Wide Web*, 2011, pp. 137–138.

[39] R. Forsati, M. Mahdavi, and M. Shamsfard, "Matrix factorization with explicit trust and distrust side information for improved social recommendation," *ACM Trans. Inf. Syst.*, vol. 32, no. 4, pp. 17:1–17:38, 2014.

[40] Y. Cai *et al.*, "Typicality-based collaborative filtering recommendation," *IEEE Trans. Knowl. Data Eng.*, vol. 26, no. 3, pp. 766–779, Mar. 2014.

[41] G. Zhao and X. Qian, "Service objective evaluation via exploring social users' rating behaviors," in *Proc. IEEE Int. Conf. Multimedia Big Data*, 2015, pp. 228–235.

[42] X. Lei, and X. Qian, "Rating prediction via exploring service reputation," in *Proc. 17th Int. Workshop Multimedia Signal Process.*, 2015, pp. 1–6.

[43] Y. Koren, R. Bell, and C. Volinsky, "Matrix factorization techniques for recommender systems," *IEEE Comput.*, vol. 42, no. 8, pp. 30–37, Aug. 2009.

[44] C. Peng, D. Yin, C. Wei, and H. Zhang, "How and when review length and emotional intensity influence review helpfulness: Empirical evidence from epinions.com," in *Proc. 35th Int. Conf. Inf. Syst.*, 2014, pp. 1–16.

[45] Y. Pan and J. Zhang, "Born unequal: A study of the helpfulness of user-generated product reviews," *J. Retailing*, vol. 87, no. 4, pp. 598–612, 2011.

[46] Y. Chen, A. Cheng, and W. Hsu, "Travel recommendation by mining people attributes and travel group types from community-contributed photos," *IEEE Trans. Multimedia*, vol. 15, no. 6, pp. 1283–1295, Oct. 2013.

[47] X. Yang, T. Zhang, and C. Xu, "Cross-domain feature learning in multimedia," *IEEE Trans. Multimedia*, vol. 17, no. 1, pp. 64–78, Jan. 2015.

[48] B. Pang and L. Lee, "A sentimental education: Sentiment analysis using subjectivity summarization based on minimum cuts," in *Proc. 42nd Annu. Meeting Assoc. Comput. Linguistics*, 2004, pp. 271–278.

[49] X. Lei, X. Qian, and G. Zhao, "Rating prediction based on social sentiment from textual reviews," *IEEE Trans. Multimedia*, vol. 18, no. 9, pp. 1910–1921, Sep. 2016.

[50] T. Mikolov, I. Sutskever, K. Chen, G. Corrado, and J. Dean, "Distributed representations of words and phrases and their compositionality," in *Proc. 26th Int. Conf. Neural Inf. Process. Syst.*, 2013, pp. 3111–3119.

[51] S. Wang and C. Manning, "Baselines and bigrams: Simple, good sentiment and topic classification," in *Proc. 50th Annu. Meeting Assoc. Comput. Linguistics*, 2012, pp. 90–94.

[52] A. Smola and B. Schölkopf, "A tutorial on support vector regression," *Stat. Comput.*, vol. 14, no. 3, pp. 199–222, 2004.

[53] Y. Liu and Y. Sun, "Anomaly detection in feedback-based reputation systems through temporal and correlation analysis," in *Proc. IEEE 2nd Int. Conf. Social Comput.*, 2010, pp. 65–72.

[54] Y. Sun and Y. Liu, "Security of online reputation systems: The evolution of attacks and defenses," *IEEE Signal Process. Mag.*, vol. 29, no. 2, pp. 87–97, Mar. 2012.

[55] Y. Zeng, Y. Zhu, and Y. Sun, "Equal rating opportunity analysis for detecting review manipulation," in *Proc. IEEE China Summit Int. Conf. Signal Inf. Process.*, 2015, pp. 846–850.

[56] Y. Zeng, Y. Zhu, and Y. Sun, "Reviewsec: Security analysis tool for online reviews," *STCSN E-Lett.*, vol. 2, no. 4, 2014.

[57] L. Xiang *et al.*, "Temporal recommendation on graphs via long- and short-term preference fusion," in *Proc. 16th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, 2010, pp. 723–731.

[58] H. Kumar, S. Lee, and H. Kim, "Exploiting social bookmarking services to build clustered user interest profile for personalized search," *Inf. Sci.*, vol. 281, pp. 399–417, 2014.

[59] L. Li, L. Zheng, F. Yang, and T. Li, "Modeling and broadening temporal user interest in personalized news recommendation," *Expert Syst. Appl.*, vol. 41, no. 7, pp. 3168–3177, 2014.

[60] G. Ling, M. Lyu, and I. King, "Ratings meet reviews, a combined approach to recommend," in *Proc. 8th ACM Conf. Recommender Syst.*, 2014, pp. 105–112.

[61] G. Ganu, Y. Kakodkar, and A. Marian, "Improving the quality of predictions using textual information in online user reviews," *Inf. Syst.*, vol. 38, no. 1, pp. 1–15, 2013.

[62] L. Chen, G. Chen, and F. Wang, "Recommender systems based on user reviews: the state of the art," *User Model. User-Adapted Interact.*, vol. 25, no. 2, pp. 99–154, 2015.

[63] G. Zhao, X. Qian, X. Lei, and T. Mei, "Service quality evaluation by exploring social users' contextual information," *IEEE Trans. Knowl. Data Eng.*, vol. 28, no. 12, pp. 3382–3394, Dec. 2016.

[64] G. Zhao, X. Qian, and X. Xie, "User-service rating prediction by exploring social users' rating behaviors," *IEEE Trans. Multimedia*, vol. 18, no. 3, pp. 496–506, Mar. 2016.

[65] S. Huang, J. Zhang, L. Wang, and X. Hua, "Social friend recommendation based on multiple network correlation," *IEEE Trans. Multimedia*, vol. 18, no. 2, pp. 287–299, Feb. 2016.

[66] P. Zhou, Y. Zhou, D. Wu, and H. Jin, "Differentially private online learning for cloud-based video recommendation with multimedia big data in social networks," *IEEE Trans. Multimedia*, vol. 18, no. 6, pp. 1217–1229, Jun. 2016.

[67] G. Zhao *et al.*, "Location recommendation for enterprises by multi-source urban big data analysis," *IEEE Trans. Services Comput.*, p. 1, 2018, doi: 10.1109/TSC.2017.2747538.

[68] G. Zhao, X. Qian, and H. Feng, "Personalized recommendation by exploring social users' behaviors," in *Proc. 20th Anniversary Int. Conf. MultiMedia Model.*, 2014, pp. 181–191.


**Guoshuai Zhao** received the B.E. degree from Heilongjiang University, Harbin, China, in 2012, the M.S. degree from Xi'an Jiaotong University, Xi'an, China, in 2015, and he is currently working toward the Ph.D. degree, Xi'an Jiaotong University, Xi'an, China. He is with the Key Laboratory for Intelligent Networks and Network Security, Ministry of Education, and the SMILES Laboratory with Xi'an Jiaotong University. He was an intern with the Social Computing Group at Microsoft Research Asia from January 2017 to July 2017, and was a visiting scholar with Northeastern University, U.S., from October 2017 to October 2018. His research interests include of social media big data analysis, recommender systems, and natural language generation.

**Xiaojiang Lei** received the B.E. degree from the Chang'an University, Xi'an, China, in 2013, and the M.S. degree in July 2016 from Xi'an Jiaotong University, Xi'an, China. His research interests include social multimedia mining and recommendation.

**Xueming Qian** (M'10) received the B.S. and M.S. degrees from the Xi'an University of Technology, Xi'an, China, in 1999 and 2004, respectively, and the Ph.D. degree from the School of Electronics and Information Engineering, Xi'an Jiaotong University, Xi'an, China, in 2008. He was a Visiting Scholar with Microsoft Research Asia from 2010 to 2011. He was an Assistant Professor with Xi'an Jiaotong University, where he was an Associate Professor from 2011 to 2014, where he is currently a Full Professor. He is also the Director of the SMILES Laboratory, Xi'an Jiaotong University. His research is supported by the National Natural Science Foundation of China, Microsoft Research, and the Ministry of Science and Technology. His research interests include social media big data mining and search. Prof. Qian was the recipient of the Microsoft Fellowship in 2006 and the Outstanding Doctoral Dissertations of Xi'an Jiaotong University and Shaanxi Province, in 2010 and 2011, respectively.

**Tao Mei** (M'06–SM'11) received the B.E. and Ph.D. degrees from the University of Science and Technology of China, Hefei, China, in 2001 and 2006, respectively. He is a Technical Vice President with JD.com and the Deputy Managing Director of JD AI Research, where he also serves as the Director of Computer Vision and Multimedia Laboratory. Prior to joining JD.com in 2018, he was a Senior Research Manager with Microsoft Research Asia, Beijing, China, where he contributed 20 inventions and technologies to Microsoft's products and services. He is an Adjunct Professor with the University of Science and Technology of China. He has authored or coauthored more than 150 publications (with 11 best paper awards) and holds 20 US granted patents. Prof. Mei is or has been an Editorial Board Member of the IEEE TIP, IEEE TCSVT, IEEE TMM, and ACM TOMM. He is the General Cochair of the IEEE ICME 2019, the Program Cochair of ACM Multimedia 2018, the IEEE ICME 2015, and the IEEE MMSP 2015. He was elected as Fellow of IAPR and a Distinguished Scientist of ACM in 2016 for his contributions to large-scale video analysis and applications. He is also a Distinguished Industry Speaker of the IEEE Signal Processing Society.