Hindawi
Mobile Information Systems
Volume 2018, Article ID 8071251, 13 pages
https://doi.org/10.1155/2018/8071251

*Research Article*

# Mobile Personalized Service Recommender Model Based on Sentiment Analysis and Privacy Concern

**Liang Xiao,[1,2] Fei-Peng Guo,[1,2] and Qi-Bei Lu[3]**

[1]*School of Management and E-Business, Zhejiang Gongshang University, Hangzhou 310018, China*
[2]*Modern Business Research Center, Zhejiang Gongshang University, Hangzhou 310018, China*
[3]*School of Science and Technology, Zhejiang International Studies University, Hangzhou 310023, China*

Correspondence should be addressed to Fei-Peng Guo; guofp@hotmail.com

Received 7 December 2017; Revised 23 February 2018; Accepted 19 April 2018; Published 3 June 2018

Academic Editor: Damianos Gavalas



The existing mobile personalized service (MPS) gives little consideration to users' privacy. In order to address this issue and some other shortcomings, the paper proposes a MPS recommender model for item recommendation based on sentiment analysis and privacy concern. First, the paper puts forward sentiment analysis algorithm based on sentiment vocabulary ontology and then clusters the users based on sentiment tendency. Second, the paper proposes a measurement algorithm, which integrates personality traits with privacy preference intensity, and then clusters the users based on personality traits. Third, this paper achieves a hybrid collaborative filtering recommendation by combining sentiment analysis with privacy concern. Experiments show that this model can effectively solve the problem of MPS data sparseness and cold start. More importantly, a combination of subjective privacy concern and objective recommendation technology can reduce the influence of users' privacy concerns on their acceptance of MPS.

## 1. Introduction

With the constant development of personalized recommendation technology and its wide use in mobile commerce, mobile recommender system crops up [1, 2]. It provides users with accurate and real-time mobile personalized service (hereafter abbreviated as MPS) [3, 4]. At the same time, users from online platforms like Weibo and Twitter generate an abundance of content (UGC). Platforms have become the main way for users to express their feelings and share information. Consequently, more and more online reviews of online purchases and transaction services emerge [5]. Yenter and Verma [6] said that the opinions given by online users on products and services contain user's emotional tendency. Mobile personalized recommender system can mine implicit interest of users in online reviews, which can fully provide user's preferences and sentiment polarity over attributes, functions, and experiences of products and services [7]. Since the information is stored in the form of behavior logs, tracks, and transaction data in network, opinion mining technology is required to extract and analyze user's sentiment tendency and business knowledge hidden in reviews, drawing many mobile commerce service providers' attention [8, 9]. Therefore, it is a hot topic to extract information about user's sentiment tendency and implicit preference on the basis of text opinion mining to assist personalized recommender system and provide quality mobile personalized service.

Pang and Lee [10] defined opinion mining as a process of analyzing, disposing, concluding, and reasoning on subjective texts with sentiment. However, some of the users' privacy information, such as location and preference, is often exposed through online reviews mining. The privacy leak in MPS is a big issue, jeopardizing the private life of users [11]. Users' privacy awareness and attention paid to the risk of privacy disclosure are gradually increasing [12]. Although many commercial websites implement online privacy policy mechanism, problems exist as follows: (1) the MPS providers unilaterally set up privacy policies but ignore privacy preferences of individuals. Users are only given two choices, either refusing to use the service or passively accepting all privacy policies in order to obtain the service.

They cannot subjectively choose the disclosed type of privacy information [13]. (2) Some researches start to use history preference data of mobile users and privacy protection technology to improve the quality of recommendation. However, they ignore the information about users' sentiment tendency and personality traits [14], hence reducing recommendation accuracy. The traditional researches show that personality traits of individuals have an important influence on how users shop online, their psychological preferences, and privacy awareness [15]. Therefore, how to integrate information, such as sentiment tendency, privacy concerns, and personality traits, into MPS is a research puzzle in this field.

To sum up, this paper analyzes influence factors of privacy concerns and online reviews and puts forward mobile personalized service recommender model based on sentiment analysis and privacy concerns. Firstly, it studies the sentiment analysis method at the level of opinion target and quantifies the sentiment tendency by opinion mining technology. Secondly, it integrates the quantitative influence of sentiment in recommend mechanism. Thirdly, it proposes a measurement method of personality traits integrated with privacy preference, mining personality traits, privacy preferences, and user preferences to obtain user groups with similar interests. At last, a novel hybrid collaborative filtering recommend method combining sentiment tendency and user's personality traits is put forward to achieve MPS.

This paper is organized as follows. After the introduction part, the paper discusses related work in the second part. In the third part, it proposes a mobile personalized service recommender model based on sentiment analysis and privacy concerns. In the fourth part, it evaluates the performance of the proposed model, and finally in the fifth part, the paper concludes with future work.

## 2. Related Work

*2.1. Sentiment Analysis Based on Opinion Mining.* Sentiment analysis of reviews on mobile commerce platform includes extraction, classification, retrieval, and induction of sentiment information [10]. There have machine-learning and knowledge-based approaches. The latter uses sentiment dictionary and syntactical rules for sentiment analysis. Hu and Liu [16] proposed a novel sentiment analysis method emphasizing reviews of product features by using traditional sentiment words library. Then, this method integrated context information with sentiment reviews to predict sentiment polarity of a product, increasing prediction accuracy. Wang et al. [17] analyzed the characteristics of ontology and Chinese online reviews and proposed a text mining model based on sentiment vocabulary ontology to build match of opinion target. This model attracts attention to mobile commerce enterprises. Similarly, Somprasertsri and Lalitrojwong [18] analyzed the sentiment tendency in the dependencies of commodity and opinion based on ontology model. This method integrated syntax with semantic information so as to quantify sentiment value, which had great business value in application. Ma et al. [19] constructed syntactic rules by adopting noun pruning and frequency filtering technology to extract review corpus. Then, they used syntactic rules to calculate the sentiment tendency of each sentence and boiled it down to four kinds of sentiment type.

Sentiment mining method based on machine learning, which is very different from the knowledge-based, requires more training time, and the model is too complex [20]. Therefore, Mi et al. improved the traditional sentiment mining method and built a text sentiment opinion mining model on the basis of binary language model and gray theory to achieve quality sentiment-oriented mining [21]. Additionally, considering the complexity of the machine-learning method, Somprasertsri and Lalitrojwong [22] extracted commodity features offline according to the maximum entropy and trained the model using corpus-tagged library. Finally, they extracted product features of online opinions by using traditional auxiliary commodity information. The model reduced learning time and achieved good prejudging results.

*2.2. Mobile Personalized Services Recommendation Based on Sentiment Analysis.* Some researchers have found that similar preferences of music can be matched by analyzing sentiment features. On one hand, Yang et al. [23] focused on sentiment characteristics of users who had common musical interests when calculating similarity of users. Similarly, Kuo et al. [24] found sentiment characteristics among different users by analyzing features of theme music in the movie and built a user interest of music model based on sentiment tendency to achieve accurate recommendations. On the other hand, Cai et al. [25] proposed cross-domain recommendation to predict users' sentiment tendency in music by analyzing opinions from Weibo in real time. Han et al. [26] presented a context-aware music recommendation system. They modeled and classified users' sentiment based on ontology language and solved the problem of recommendation data sparsity with context by using the nonnegative matrix factorization technique. Additionally, Mudambi and Schuff [27] integrated analysis of online reviews with users' behavior preference mining. They can get better recommendation results through analysis of users' sentiment tendency toward good opinions.

The research of sentiment mining in collaborative filtering recommendation also becomes a hotspot recently. Traditional collaborative filtering recommendation method relies on a matrix of "user-review" to calculate user similarity or item similarity. But, it covers limited information and often leads to user interest bias due to many factors such as context. She and Chen [28] introduced the sentiment analysis method based on topic model into collaborative filtering recommendation, which increased accuracy by using rich-text review information. Winoto and Tang [29] studied sentiment analysis in the film recommendation system. They found user groups of similar interests through sentiment tendency submitted actively by users and then proposed a sentiment-aware collaborative filtering method. Shi et al. [30] proposed a method based on decomposition matrix when calculating the films similarity on the basis of sentiment. They integrated sentiment analysis results in the process of collaborative filtering recommendation, collecting users' ratings data of IMDB in three stages, that is, before

a movie, during a movie, and after a movie. Then, this method utilized users' sentiment reviews published on Twitter for improving the algorithm of collaborative filtering recommendation and accurately forecasting the box office.

*2.3. Mobile Personalized Services Recommendation Based on Privacy Concerns.* The term "privacy concerns" is used to measure consumers' worriedness, perceptions, and controls of information privacy [31]. Culnan and Armstrong [32] claimed that mobile internet activities often involved personal privacy information, such as payment information, motion, and geographical location. Obtaining user's personal information is crucial to the survival and development of MPS providers [31], which can be used to accurately recommend users' information, and to stimulate their continuous acceptance of the service by satisfying their needs. Thus, their satisfaction and loyalty are improved [33, 34]. However, the problem of privacy leak is increasingly serious, and more and more researchers propose new ways to solve privacy issues [35]. For example, some people use data encryption and anonymous protection in mobile personalized services to generate recommended content for reducing privacy worriedness. Li and Sarkar [36] studied individual privacy policy mechanism based on cluster and analyzed problems of privacy concerns in the process of user-based collaborative filtering recommendation. They presented a novel method of encrypting and anonymizing for individual privacy information to protect user privacy. McSherry and Mironov [37] put differential privacy technology into collaborative filtering recommendation system. They carried out differential privacy process on item-to-item covariance matrix. Through experiments in learning and forecasting stage, they found that introducing differential privacy protection technology in recommender system with high accuracy was feasible. Duckham and Kulik [38] proposed a privacy protection model of location based on noise model for encrypting the MPS in mobile environment.

Some scholars study privacy concerns and strategies from the subjective perspectives of users. Wang and Duan [39] synthesized the influence from individual and system privacy factors of online users to quantify these factors. Then, they designed a privacy quantification model based on universal vectors from the subjective perspective of user perception. Similarly, Sutanto et al. [40] took advertising application of personalized mobile as the example and designed a technology solution called Personalized, Privacy-Safe Application, trying to solve the paradox between personalization and privacy. This method was verified by field experiments. Chellappa and Shivendu [41] designed user's privacy protection model based on economics. They calculated users' utility and facilitated users' use by weighing the costs of not having a MPS service and being supervised by a service provider. In addition, Korzaan and Boswell [42] studied the influence of the five personality traits, which include extraversion, agreeableness, conscientiousness, neuroticism, and openness on consumer privacy, and found that agreeableness has an effect on privacy concerns. According to Junglas and Spitzmuller [43], openness and agreeableness had a negative influence on privacy concerns, and person with different personalities have different psychological tendency of privacy concerns. Choi and Choi [44] also showed that privacy concerns were different among users with different characteristics and thus affected their behaviors. For example, people who are open are more willing to experience and accept MPS.

## 3. Mobile Personalized Service Recommender Model Based on Sentiment Analysis and Privacy Concern

Since most sentiment analysis algorithms excessively rely on the size of data at present, it is hard to meet user requirements for fast response of MPS. Therefore, this paper adopts the text analysis method based on sentiment dictionary. But the traditional methods have a poor effect on emotion analysis, which is not included in sentiment dictionary. So, our model proposes a sentiment analysis algorithm based on sentiment vocabulary ontology to accurately obtain user interests. Firstly, users with similar sentiment tendency or privacy preference may have the same interest preference [13, 15]. Secondly, the key step is to find the $k$ users who have similar characteristics with the target users in collaborative filtering recommendation method [45]. This article uses "user-sentiment" vector matrix to find the $k$ users who have similar sentiment tendency. Considering the differences of users' privacy preference [13], this paper introduces the concept of "personality traits integrating with privacy preference." It uses "user-personality traits" vector matrix to find the most $k$ similar users who have similar privacy preference. Thirdly, as shown in Figure 1, our model provides recommendation based on sentiment analysis and privacy concern.

### *3.1. Sentiment Tendency Analysis Algorithm Based on Sentiment Vocabulary Ontology*

*3.1.1. Building the Sentiment Vocabulary Ontology Library.* Users usually give "positive" or "negative" opinions to online service. Firstly, users express opinions with words like "service quality is very good, value for money" after electric business transactions. Secondly, users use emotional words containing emotional polarity to describe the sentiments, such as "this shopping experience is very happy, I like it." Thirdly, the opinion of "I am not very happy today" shows a "negative" emotion tendency. However, the opinion of "I feel very happy" shows "active" emotion tendency. The example above shows that negative words can make emotion polarity of sentiments reverse. Adversative phrase and excessive adverbs of degree will also change user's emotion tendency, such as "mobile phone is very cool, but flashy" and "mobile phone keys are too small." Therefore, this section constructs three vocabulary ontology library of comment, emotion, and inversion for different descriptions of online opinions. Finally, this paper sums up users' sentiment tendency as negative, positive, and neutral in MPS. The negative is subdivided into evil, anger, fear, and grief. The positive is subdivided into good and happy. The neutral is surprise.

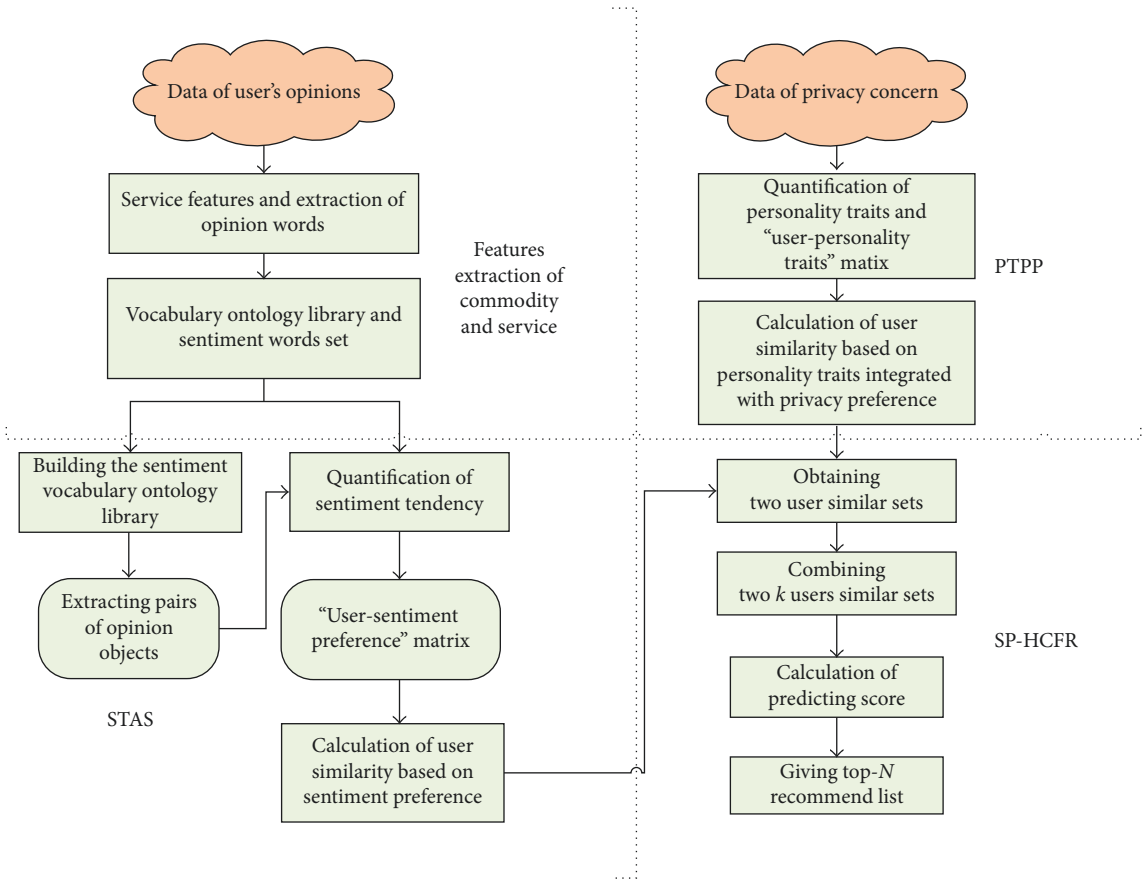


Figure 1: The model architecture.

*(1) Building the Comment Vocabulary Ontology Library.* Users use comments to directly express their preferences for goods and services. They reflect users' satisfied or unsatisfied attitude. Opinion mining methods calculate the sentiment type according to the comments. Considering the inconsistency and nonstandard use of the users' comments on the Internet, this paper puts forward comment vocabulary set based on ontology and uses triple ET = $(B, R, E)$ to describe the sentiment transform. The $B$ means the basic information of original words. The $R$ means the synonymous relations. The $E$ is the two-tuple of sentiment type and membership degree.

*(2) Building the Emotion Vocabulary Ontology Library.* Users use emotion words to directly express their emotion for goods or services. They reflect users' "like" or "dislike" attitude. Opinion mining methods calculate the sentiment type according to the emotion words. Considering the differences of emotion vocabulary classification, this paper studies eight kinds of emotion expression in social relationships [46] and twelve kinds of Chinese vocabulary classification [47]. On this basis, we select seven emotion types, which are happy, good, anger, sorrow, fear, evil, and surprise. Then, we use similarity algorithm to quantify the value of sentiment tendency.

*(3) Building the Inversion Vocabulary Ontology Library.* This paper complements the sentiment vocabulary ontology library of DUITR [48], which collects escape words to construct inversion vocabulary ontology library, including overdone degree adverbs, twist words, and negative words. But it does not consider general modifiers. Considering the complexity of the Chinese, this paper calculates the distance length in one statement between sentiment and inversion words (range [−4, +4]) to adjust the polarity of sentiment words.

*3.1.2. Sentiment Tendency Analysis Algorithm.* Sentiment tendency analysis algorithm based on sentiment vocabulary ontology (STAS) (Algorithm 1) divides opinion statements into pairs of words based on condition random fields (CRFs) [49] and domain ontology. Then, STAS extracts pairs of opinion targets based on CRFs model, including opinion target, opinion word, and phrase for sentiment tendency analysis, and improves the efficiency of the opinion mining.

*Definition 1.* Sentiment vocabulary library $S = \langle s_1, s_2, \ldots, s_m \rangle$; $s_i$ is the sentiment word, $i = \{1, 2, \ldots, m\}$.

*Definition 2.* Opinion text set $D = \langle d_1, d_2, \ldots, d_n \rangle$; $d_i$ is the opinion text, $i = \{1, 2, \ldots, n\}$. $d_i = \langle w_1, w_2, \ldots, w_k \rangle$; $w_i$ is the feature item of $d_i$, $i = \{1, 2, \ldots, k\}$.

*Definition 3.* Sentiment text set $D^s = \langle d_1^s, d_2^s, \ldots, d_t^s \rangle$ is the text collection containing sentiment words; $d_i^s$ is the sentiment text, $i = \{1, 2, \ldots, t\}$.

**Input**: Sentiment Opinions Object Set, Comment Vocabulary Ontology Library, Emotion Vocabulary Ontology Library, and Inversion Vocabulary Ontology Library.
**Output**: Triple<Opinions Service, Sentiment Type, Sentiment Tendency Value>.

(1) It extracts opinion objects and opinion words based on CRF model [50] and judges whether the opinion sentence has sentiment words. If there are no sentiment words, the STAS ends directly. Otherwise, it jumps to the next step.
(2) It matches each opinion text phrase through sentiment vocabulary ontology library and constructs the relationship between opinion objects and opinion phrases.
(3) It traverses opinion phrases to match sentiment words in opinion objects by using sentiment vocabulary ontology library. If sentiment word exists, STAS changes it to sentiment type, calculates the sentiment polarity according to inversion words, and is stored as Triple<Opinions Service, Sentiment Type, Sentiment Tendency Value>. If it does not exist, it outputs as Triple<Opinions Service, "neutral," 0>.
(4) STAS repeats step 1 to step 3, until it judges all the opinion objects.

The overall sentiment tendency of opinion objects is calculated by weight values of comment words, emotion words, and reverse words. STAS quantifies the sentiment type of opinion words based on artificial tagging and fuzzy set theory [50] and adopts the membership degree to predict its value. Besides, it calculates the similarity between $s$ and $s'$ based on Levenstein edit distance and predicts the sentiment tendency value of sentiment words. Thirdly, it calculates the sentiment tendency value of reverse words by the Triple<opinions object, reverse words, sentiment polarity> and PMI-IR algorithm [51]:

$$\mathrm{sim}(s, s') = \max(0, (\min(|s_i|, |s_i'| - \mathrm{ed}(s_i, s_i')))/\min(|s_i|, |s_i'|)),$$

where $\mathrm{ed}(s_i, s_i')$ is the Levenstein edit distance between $s_i$ and $s_i'$.

$$\mathrm{PMI}(\mathrm{word}) = \sum_{p\,\mathrm{word} \in P_{\mathrm{set}}} \mathrm{PMI}(\mathrm{word}, p\,\mathrm{word}) - \sum_{n\,\mathrm{word} \in N_{\mathrm{set}}} \mathrm{PMI}(\mathrm{word}, n\,\mathrm{word}),$$

where *word* is the target word whose sentiment type is unknown, and $P_{\mathrm{set}}$ and $N_{\mathrm{set}}$ are the positive sentiment vocabulary set and negative sentiment vocabulary set, respectively, in basic sentiment words.

Algorithm 1: Sentiment tendency analysis based on sentiment vocabulary ontology.

The recognition of opinion target can be viewed as a sequence labeling problem. It labels the opinion corpora in sequence based on CRF model. It inputs a word string $w_1, w_2, \ldots, w_k$ and defines $D = D_1, D_2, \ldots, D_n$ as observed sequence, outputs a labeling sequence with the highest probability $D^s$, and then defines $D^s = D_1^s, D_2^s, \ldots, D_n^s$ as the observation sequence. The chain conditional probability distribution of $D^s$ is shown in the following formula:

$$P(D^s|D) = \frac{1}{Z(D)} \exp \sum_i \sum_k \lambda_k f_k(d_{i-1}^s, d_i^s, d, i) + \sum_i \sum_k \mu_k g_k(d_i^s, d)), \quad (1)$$

$$Z(D) = \sum_d \exp\left( \sum_i \sum_k \lambda_k f_k(d_{i-1}^s, d_i^s, d, i) \right), \quad (2)$$

where $f_k$ is the transfer characteristic function from location $i$ to $i-1$, $g_k$ is the state characteristic function on location $i$, $\lambda_k$ and $\mu_k$ are weight values in the process of training, and $Z_X$ is a normalizing factor relaying on $D$. CRF model uses Viterbi method to find a tag sequence named $D^{s*}$ for getting the maximum $P(D^s|D)$, when the training uses the iterative algorithm based on maximum likelihood.

To the end, we calculate the sentiment tendency to judge users' sentiment preference of commodities or services by using sentiment vocabulary knowledge ontology library. The algorithm of STAS is shown as follows.

### 3.2. Measurement Method of Personality Traits Integrated with Privacy Preference

*3.2.1. Measurement of User's Privacy Preference.* Social network websites generally use privacy settings like "how to find me" (email or phone), "who are allowed to comment on me" (all, only fans, and only the persons who I care about), "recommend friends in your phone's address book," "who are allowed to give direct messages to me" (all, only fans, only the persons who I care about), "who are allowed to @ me" (only the persons who I care about, all), "binding with other accounts," "who are allowed to get my location" (all, only fans, only the persons who I care about), and so on. Henson et al. [52] found that users' active and personality traits have close relevance with their privacy behaviors. Mobile users' active in social network can be described with numbers of blogs delivered, photos uploaded, and opinions given. It also has significant correlation with "who are allowed to comment on me" (AC), "who are allowed to give direct messages to me" (AM), and "who are allowed to get my location" (AG). Thus, this paper chooses the above three items as a measurement index of privacy preference. Taking Weibo as an example, this paper uses multiple linear regression model and (3) to calculate the intensity of privacy preference. The three evaluation indexes are shown as follows, all setting half a year as the time period: (1) "who are allowed to comment on me" ($\mathrm{AC}_u$), (2) "who are allowed to give direct messages to me" ($\mathrm{AM}_u$), and (3) "who are allowed to get my location" ($\mathrm{AG}_u$). The privacy preference vectors of users are named "$P_u$":

$$P_u = (\mathrm{AM}_u, \mathrm{AC}_u, \mathrm{AG}_u), \quad (3)$$

**Input**: Mobile user $u$, recommend service set *Service(R)*, and score matrix of "user-sentiment."
**Output**: $\mathrm{sim}(u,v)_{\text{sentiment-pearson}}$.

(1) To calculate the average user preference for a certain sentiment type:
$r_{u,s'} = \left(1/|S_{u,s'}|\right)\sum_{s\in S_{u,s'}} r_{u,s,s'}$,
$S_{u,s'} = \left\{s|s \in S, r_{u,s} \neq \text{null}, ss' = 1\right\}$, where $|S_{u,s'}|$ is the number of items in $S_{u,s'}$ and $r_{u,s,s'}$ is the user preference for service $s$ with sentiment $s'$, and to construct a two-dimensional "user-sentiment" preference matrix.

(2) It proposes an improved user's similarity calculation method based on preference matrix of "user-sentiment":
$\mathrm{sim}(u_i,u_j)_{\text{sentiment-pearson}} = \left(\sum_{s'\in S'}\left(r_{u_i,s'} - \overline{r_{u_i}^{s'}}\right)\left(r_{u_j,s'} - \overline{r_{u_j}^{s'}}\right)\right)/\left(\sqrt{\sum_{s'\in S'}\left(r_{u_i,s'} - \overline{r}_{u_i}\right)^2 \sum_{s'\in S'}\left(r_{u_j,s'} - \overline{r}_{u_j}\right)^2}\right)$,

where $\overline{r_{u_i}^{s'}}$ is the average sentiment preference of $u_i$ in all service-relative sentiment. It can select the $k$ users similar set of $u_i$ based on $\mathrm{sim}(u_i,u_j)_{\text{sentiment-pearson}}$.

Algorithm 2: User similarity calculation based on sentiment analysis.

$P_u$ can be quantified by the following formula to show the value of privacy preference intensity:

$$P_u = \beta_1(\mathrm{AM}_u) + \beta_2(\mathrm{AC}_u) + \beta_3(\mathrm{AG}_u). \tag{4}$$

The probability value of $\mathrm{AM}_u$, $\mathrm{AC}_u$, and $\mathrm{AG}_u$ can be 0 or 1, respectively, meaning "allowed" or "not allowed." The measurement of $P_u$ means specific preference weighted value in privacy settings of mobile users. $P_u$ with lower value belongs to the group of low awareness in privacy concern, whereas the reverse is the group of high awareness.

*3.2.2. Measurement of User's Personality Traits.* Traditional user's personality traits are generally quantified with discrete values but yield no ideal effect. This paper adopts the "big-five personality" questionnaire and users' self-rated scores to get continuous personality traits data. Selfhout et al. [53] found that not all five dimensions of "big-five personality traits" could affect users' behaviors and social relations in network. Therefore, we select openness, extraversion, and agreeableness as the research objects and quantify the personality traits based on multivariate linear regression model. This model selects 15 independent variables, which are friends, visitors, blogs, photos, photo albums, direct messages, opinions, and the forwarded number of opinions, photos, hot topics, videos, music, and so on. This model extracts 400 users' behavior character data as a sample set from the questionnaire survey, Web crawler-like tools, and characterizes them in social network behavior data:

$$S_u = (O_u, C_u, E_u, A_u, N_u). \tag{5}$$

Among them, the vector $S_u$ is defined as user's personality traits, $O_u$ is the score of openness calculated by the number of friends $\mathrm{FN}_u$, forwarded hot topic $\mathrm{TC}_u$, comments $\mathrm{CN}_u$, and uploaded photos $\mathrm{PN}_u$. $C_u$ is the score of responsible. $E_u$ is the score of extrovert calculated by the number of friends $\mathrm{FN}_u$, blogs delivered $\mathrm{DB}_u$, and comments $\mathrm{CN}_u$. $A_u$ is the score of agreeableness calculated by the number of friends $\mathrm{FN}_u$, forwarded hot topic $\mathrm{TC}_u$, comments $\mathrm{CN}_u$, and uploaded photos $\mathrm{PN}_u$. $N_u$ is the score of neurological. This paper calculates the value of personality traits preference by quantifying the $S_u$. Therefore, the calculation method of user's personality traits integrated into privacy preference (PTPP) is shown as follows:

$$\begin{aligned} O_u &= \beta_1(\mathrm{FN}_u) + \beta_2(\mathrm{TC}_u) + \beta_3(\mathrm{CN}_u) + \beta_4(\mathrm{PN}_u), \\ E_u &= \beta_1(\mathrm{FN}_u) + \beta_2(\mathrm{DB}_u) + \beta_3(\mathrm{CN}_u), \\ A_u &= \beta_1(\mathrm{FN}_u) + \beta_2(\mathrm{TC}_u) + \beta_3(\mathrm{CN}_u) + \beta_4(\mathrm{PN}_u), \\ P_u &= \beta_1(\mathrm{AM}_u) + \beta_2(\mathrm{AC}_u) + \beta_3(\mathrm{AG}_u), \end{aligned} \tag{6}$$

where $S_{u,p} = (O_u, E_u, A_u, P_u)$, and $O_u$, $E_u$, $A_u$, and $P_u$ are, respectively, calculated by multivariate linear regression model.

### 3.3. Hybrid Collaborative Filtering Recommendation Method Based on Sentiment Analysis and Privacy Concerns

*3.3.1. User Similarity Calculation Based on Sentiment Analysis.* We change the score matrix of "user-service" to preference matrix of "user-sentiment" to calculate the similarity of users. The construct of preference matrix of "user-sentiment" defined as $R_{u,s}$ relies on the score matrix of "user-service" and the correlation matrix of "service-sentiment." Each low in $R_{u,s}$ is the sentiment preference vector: $R_{u,s} = \left\{r_{u,s}|u \in U, s \in S, r_{u,s} \in [0,100]\right\}$. $r_{u,s}$ is the value of user preference $u$ for specific sentiment $s$.

*3.3.2. User Similarity Calculation Based on Personality Traits Integrated with Privacy Preference.* This paper uses matrix of "user-personality traits" $R_{u,p}$ to construct the $k$ users similar set of target user $u$. The item score in $R_{u,p}$ is calculated by measurement method of PTPP. Each item score in $R_{u,p}$ is the weighted value computed by the score of "openness," "extraversion," "agreeableness," and "privacy preference." Each row of $R_{u,p}$ is the score vector of personality traits integrated with privacy preference, and the paper comprehensively quantifies $r_{u,p}$ in comprehensive quantification of personality traits in four dimensions so as to calculate the similarity among users.

*3.3.3. Collaborative Filtering Recommend Method Combining User's Sentiment Tendency with Personality Traits.* The core of SP-HCFR is the calculation of hybrid users' similarity that is weighted according to the $\mathrm{sim}(u,v)_{\text{sentiment-pearson}}$

**Input**: Mobile user $u$, privacy preference vector $P_u$, personality trait vector $S_u$, and score matrix of "user-personality traits" $R_{u,p}$: $R_{u,p} = \{r_{u,p} | p \in P, u \in U, r_{u,p} \in [0,1]\}$

**Output**: $\mathrm{sim}(u,v)_{\text{privacy-preference}}$.

(1) It obtains users' basic data of personality traits through the questionnaire of "personality assessment of mobile user" $S_u$: $S_u = (O_u, C_u, E_u, A_u, N_u)$

(2) It calculates the comprehensive value of users' personality traits under privacy concern through data of mobile network behavior $S_u = (O_u, E_u, A_u, P_u)$:
$O_u = \beta_1(\mathrm{FN}_u) + \beta_2(\mathrm{TC}_u) + \beta_3(\mathrm{CN}_u) + \beta_4(\mathrm{PN}_u)$,
$E_u = \beta_1(\mathrm{FN}_u) + \beta_2(\mathrm{DB}_u) + \beta_3(\mathrm{CN}_u)$,
$A_u = \beta_1(\mathrm{FN}_u) + \beta_2(\mathrm{TC}_u) + \beta_3(\mathrm{CN}_u) + \beta_4(\mathrm{PN}_u)$,
$P_u = \beta_1(\mathrm{AM}_u) + \beta_2(\mathrm{AC}_u) + \beta_3(\mathrm{AG}_u)$.

(3) Then, it uses the following equation to compute user similarity based on $R_{u,p}$. Among them, user $\vec{u}$ and $\vec{v}$ are the personality trait vectors integrating with privacy concern:

$$\mathrm{sim}(u,v)_{\text{privacy-preference}} = (\vec{u} \cdot \vec{v}) / (\|\vec{u}\| \|\vec{v}\|).$$

Algorithm 3: User similarity calculation based on personality traits integrating with privacy preference.

**Input**: Mobile user $u$, recommend service set *Service(R)*, and score matrix of "user-sentiment" and "user-personality traits."

**Output**: Top-$N$ recommend services and its score.

(1) The calculation of user similarity is based on sentiment analysis.

(2) The calculation of user similarity is based on personality traits integrated with privacy preference.

(3) It searches the $k$ users similar set of target use by using the composite user similarity, which is calculated by the following equation:

$$\mathrm{sim}(u,v) = \alpha \times \mathrm{sim}(u,v)_{\text{privacy-preference}} + (1-\alpha) \times \mathrm{sim}(u,v)_{\text{sentiment-pearson}}$$

It uses fifty percent of cross-validation method to determine the parameters $\alpha \in [0,1]$. When $\alpha = 0$, $\mathrm{sim}(u,v) = \mathrm{sim}(u,v)_{\text{privacy-preference}}$, and when $\alpha = 1$, $\mathrm{sim}(u,v) = \mathrm{sim}(u,v)_{\text{sentiment-pearson}}$.

(4) It predicts the users' preference and sorts in Top-$N$ to give recommendation:

$$\begin{aligned} P_{u,i}' &= \overline{P_u} + \alpha \times \left(\left(\textstyle\sum_{v \in V} \mathrm{sim}(u,v)_{\text{privacy-preference}} \times \left(P_{u,i} - \overline{P_u}\right)\right) / \left(\textstyle\sum_{v \in V} \mathrm{sim}(u,v)\right)\right) \\ &\quad + (1-\alpha) \times \left(\left(\textstyle\sum_{t \in T} \mathrm{sim}(u,t)_{\text{sentiment-pearson}} \times \left(P_{t,i} - \overline{P_u}\right)\right) / \left(\textstyle\sum_{t \in T} \mathrm{sim}(u,t)\right)\right), \quad V \neq \varnothing \,\&\, T \neq \varnothing \\ &= \overline{P_u} + \left(\left(\textstyle\sum_{v \in V} \mathrm{sim}(u,v)_{\text{privacy-preference}} \times \left(P_{v,i} - \overline{P_u}\right)\right) / \left(\textstyle\sum_{v \in V} \mathrm{sim}(u,v)\right)\right), \quad V \neq \varnothing \,\&\, T = \varnothing \\ &= \overline{P_u} + \left(\left(\textstyle\sum_{t \in T} \mathrm{sim}(u,t)_{\text{sentiment-pearson}} \times \left(P_{t,i} - \overline{P_u}\right)\right) / \left(\textstyle\sum_{t \in T} p\mathrm{sim}(u,t)\right)\right), \quad V = \varnothing \,\&\, T \neq \varnothing. \end{aligned}$$

Algorithm 4: Collaborative filtering recommend method combining sentiment tendency with user's personality traits.

calculated by Algorithm 2 and $\mathrm{sim}(u,v)_{\text{privacy-preference}}$ calculated by Algorithm 3.

## 4. Performance Evaluation

### 4.1. Experimental Data and Evaluation Standards

*4.1.1. Data Sets of Practical Application.* Firstly, we tag the text opinion corpus and select group of opinion match, including opinion words and opinion targets. Experiments fetch 800 opinion corpus from Weibo and sorts out 500 users, 305 opinion targets, 541 opinion words, and two-tuple of sentiment opinions unit 416. This paper uses <opinion target, opinion words> to construct these sentiment opinion unit tagged by experts. Secondly, we use the sentiment opinion unit as a test data set and classify each unit into different sentiment types with artificial methods. Then, STAS uses these artificial results as contrast standards. On this basis, we obtain real recommended data sets with 500 users, 7 sentiment types, 800 score records, and 305 items.

Table 1: The data sets of recommend.

| Data sets | Score records | User | Item | Sentiment type |
|---|---|---|---|---|
| Data sets of practical application | 800 | 500 | 305 | 7 |
| Training data sets | 4,544,409 | 105,137 | 25,058 | 16 |
| Testing data sets | 19,506 | 160 | 3,396 | 16 |

*4.1.2. Standard Data Sets of Recommend.* It selects Moviepilot-mp.mood, which contains the information of users, movies, sentiment types, locations, opinion time, and so on. It divides the Moviepilot-mp.mood into training data set and testing data set. The training data set includes 4,544,409 score records, which is scored by 105,137 users in 25,058 movies under 16 sentiment types. The testing data set includes 19,506 score records, which is scored by 160 users in 3,396 movies under 16 sentiment types. It adopts five marks as a step size when users give scores in the range of 0–100. The data sets of recommend are listed in Table 1.

Table 2: The experimental result of STAS.

| Sentiment type | Precision | Recall | $F1$ |
|---|---|---|---|
| Sorrow | 0.62 | 0.65 | 0.63 |
| Happy | 0.64 | 0.66 | 0.65 |
| Anger | 0.66 | 0.67 | 0.66 |
| Fear | 0.68 | 0.71 | 0.69 |
| Surprise | 0.69 | 0.73 | 0.71 |
| Good | 0.74 | 0.77 | 0.75 |
| Evil | 0.75 | 0.79 | 0.77 |

Table 3: Influence of different conditions of STAS.

| Different condition combination | The results of sentiment analysis | | |
|---|---|---|---|
| | Recall | Precision | $F1$ |
| Comment words, emotion words, and inversion words | 0.66 | 0.64 | 0.65 |
| Comment words and inversion words | 0.61 | 0.56 | 0.58 |
| Emotion words and inversion words | 0.53 | 0.49 | 0.51 |
| Emotion words | 0.34 | 0.28 | 0.31 |

*4.1.3. Evaluation of standard of experiment.* It uses precision rate $p$, recall rate $r$, and $F$-measure $F_\beta$ for the effect analysis of STAS:

$$\text{Precision} = \frac{\text{the number of opinion units really belongs to correct sentiment type}}{\text{the sum of opinion unit numbers belongs to correct and incorrect sentiment type}},$$

$$\text{Recall} = \frac{\text{the number of opinion units really belongs to correct sentiment type}}{\text{the sum of opinion unit number judged to correct and incorrect sentiment type}}, \tag{7}$$

$$F_\beta = \frac{\left(\beta^2 + 1\right) \times p \times r}{\beta^2 \times p \times r}.$$

It uses mean average precision (MAP) for the effective analysis of SP-HCFR (Algorithm 4). The MAP measures the sort accuracy of Top-$N$. The higher the MAP, the higher the accuracy of the recommendation.

*4.2. Experimental Results of STAS.* In order to verify the superiority, STAS matches 416 two-tuple<sentiment opinion unit, sentiment type> with ontology of sentiment vocabulary library, including the ontology library of comment words, emotion words, and reverse words. The paper uses STAS to judge the sentiment type of each opinion unit in testing data set. The results of STAS experiment are compared with artificial tag standard, which are shown in Table 2.

It can be seen that STAS has a high rate of precision, recall, and $F1$ in seven sentiment types for the classification of sentiment. It also suggests that STAS can obtain accurate user preferences. At the same time, it divides text corpus into several sentiment opinion units, which are mined as 2-tuple<sentiment opinion unit, sentiment type> for more accurate prediction of sentiment analysis. The experimental results also show that three kinds of sentiment vocabulary ontology library have more accurate sentiment analysis than comments vocabulary. STAS can make up these problems, such as diversity of sentiment, deviation of sentiment similarity, and incorrect sentiment judgement.

To verify the influence of sentiment analysis based on comment words, emotion words, and inversion words, this paper does four experiments, respectively, by selecting comment words, emotion words, and inversion words, by selecting comment words and inversion words, by selecting emotion words and inversion words, and by selecting emotion words. It compares overall sentiment analysis through the indexes of precision rate, recall rate, and $F1$. The experimental results are shown in Table 3.

It can be seen that four kinds of words have a significant influence on sentiment analysis and improve the calculation accuracy of user preferences in MPS. Therefore, the STAS can semantically express the relationship among complex opinion targets by the ontology. It also improves the effectiveness of traditional opinion mining by combining with multisource information.

*4.3. Experimental Results of PTPP.* This paper adopts the stepwise multivariable linear regression model (SMLRM) to measure PTPP. It selects variables in SMLRM whose $F$-probability is less than 0.05. PTPP keeps variable with the highest significant level of coefficient, eliminates the non significant variables, and obtains the final significant regression equation of coefficient through several times of selection and elimination.

*4.3.1. To Quantify Openness Dimension of Personality Traits.* This paper finds the results in Table 4 through SMLRM that the dependent variable (openness) has linear regression relationship with the independent variable, such as the number of comments $CN_u$ (regression coefficient is −0.007), friends $FN_u$ (regression coefficient is 0.126), uploaded photos $PN_u$ (regression coefficient is 0.595), and forwarded hot topics $TC_u$ (regression coefficient is 0.088). At the same time, the openness has a positive linear relationship with the number of friends, uploaded photos, and forwarded hot

Table 4: Regression model in openness dimension.

| Variable | Regression coefficient $b$ | Standard coefficient $r$ | $T$ test | Significant degree $p$ |
|---|---|---|---|---|
| $FN_u$ | 0.126 | 0.807 | 11.105 | 0.000 |
| $TC_u$ | 0.088 | 0.122 | 2.491 | 0.015 |
| $CN_u$ | −0.007 | −0.145 | −2.458 | 0.014 |
| $PN_u$ | 0.595 | 0.919 | 2.163 | 0.031 |

Table 5: Regression model in extraversion dimension.

| Variable | Regression coefficient $b$ | Standard coefficient $r$ | $T$ test | Significant degree $p$ |
|---|---|---|---|---|
| $FN_u$ | 0.105 | 0.655 | 11.942 | 0.000 |
| $DB_u$ | 0.126 | 0.167 | 3.748 | 0.000 |
| $PN_u$ | 0.877 | 0.171 | 3.399 | 0.001 |

Table 6: Regression model in agreeableness dimension.

| Variable | Regression coefficient $b$ | Standard coefficient $r$ | $T$ test | Significant degree $p$ |
|---|---|---|---|---|
| $FN_u$ | 0.146 | 0.778 | 10.572 | 0.000 |
| $TC_u$ | 0.088 | 0.099 | 2.004 | 0.046 |
| $CN_u$ | −0.009 | −0.171 | −2.878 | 0.004 |
| $PN_u$ | 1.162 | 0.194 | 3.472 | 0.001 |

Table 7: Regression model in privacy preference dimension.

| Variable | Regression coefficient $b$ | Standard coefficient $r$ | $T$ test | Significant degree $p$ |
|---|---|---|---|---|
| $AM_u$ | −0.821 | −0.654 | −11.941 | 0.000 |
| $AC_u$ | −0.139 | −0.166 | −3.747 | 0.031 |
| $AG_u$ | −0.137 | −0.170 | −3.399 | 0.022 |

topics. The number of uploaded photos exerts most influence to openness. The composite correlation coefficient $R = 0.832$, and the decision correlation coefficient $R^2 = 0.767$. It means that regression model in openness dimension has positive correlation with these factors and fits testing data set well.

Regression equation:

$$O_u = 0.126(FN_u) + 0.088(TC_u) - 0.007(CN_u) + 0.595(PN_u). \quad (8)$$

*4.3.2. To Quantify Extraversion Dimension of Personality Traits.* It finds the results in Table 5 through SMLRM that the dependent variable (extraversion) has linear regression relationship with the independent variable, such as the number of friends $FN_u$ (regression coefficient is 0.105), uploaded photos $PN_u$ (regression coefficient is 0.877), and blogs delivered $DB_u$ (regression coefficient is 0.126). At the same time, the extraversion has a positive linear relationship with the number of friends, uploaded photos, and blogs delivered. The composite correlation coefficient $R = 0.780$, and the decision correlation coefficient $R^2 = 0.821$. It means that regression model in extraversion dimension has positive correlation with these factors and has a good fitting degree in testing data set.

Regression equation:

$$E_u = 0.105(FN_u) + 0.126(DB_u) + 0.877(CN_u). \quad (9)$$

*4.3.3. To Quantify Agreeableness Dimension of Personality Traits.* It finds the results in Table 6 through SMLRM that the dependent variable (agreeableness) has linear regression relationship with the independent variable, such as the number of friends $FN_u$ (regression coefficient is 0.146), forwarded hot topics $TC_u$ (regression coefficient is 0.088), comments $CN_u$ (regression coefficient is −0.009), and uploaded photos $PN_u$ (regression coefficient is 1.162). At the same time, the agreeableness has a positive linear relationship with the number of friends, forwarded hot topics, comments, and uploaded photos. Finally, the composite correlation coefficient $R = 0.886$, and the decision correlation coefficient $R^2 = 0.771$. It means that regression model in agreeableness dimension has positive correlation with these factors and fits testing data set well.

Regression equation:

$$A_u = 0.146(FN_u) + 0.088(SC_u) - 0.009(CN_u) + 1.162(PN_u). \quad (10)$$

*4.3.4. To Quantify Privacy Preference Dimension of Personality Traits.* The person who gives low scores in this dimension prefers to ignore privacy and security and shares the privacy to others. He or she does not care about tagging his or her place and social network information but begins to care about the privacy setting. Instead, he or she prefers to protect his or her privacy, rejects new service, and tends to protect his or her own autonomy. This paper finds the results in Table 7 through SMLRM that the dependent variable (privacy preference) has a linear regression relationship with the independent variables, such as "who are allowed to comment on me" $AC_u$ (regression coefficient is −0.821), "who are allowed to give direct messages to me" $AM_u$ (regression coefficient is −0.139), and "who are allowed to get my location" $AG_u$ (regression coefficient is −0.137). The regression coefficients of these three factors are less than 0, which means privacy preference is of negative correlation with the three factors, and the "who are allowed to give direct messages to me" has the most powerful impact on users' privacy preference. Finally, the composite correlation coefficient $R = 0.850$, and the decision correlation coefficient $R^2 = 0.765$. It means that regression model in privacy preference dimension has positive correlation with these factors and fits testing data set well.

Regression equation:

$$P_u = -0.821(AM_u) - 0.139(AC_u) - 0.137(AG_u). \quad (11)$$

From the experimental results, it shows that personality traits with privacy preference, openness, extraversion, and agreeableness can better reflect and quantify online behaviors

Table 8: Comparison of SP-HCFR under the influence of different $\alpha$ ($P@R$).

| SP-HCFR | $P@5$ ($k=10, 20, 30, 50$) | | | | $P@10$ ($k=10, 20, 30, 50$) | | | |
|---|---|---|---|---|---|---|---|---|
| | 10 | 20 | 30 | 50 | 10 | 20 | 30 | 50 |
| 0.0 | 0.49 | 0.54 | 0.56 | 0.59 | 0.45 | 0.50 | 0.54 | 0.55 |
| 0.2 | 0.51 | 0.57 | 0.58 | 0.60 | 0.50 | 0.53 | 0.55 | 0.57 |
| 0.4 | 0.53 | 0.58 | 0.59 | 0.60 | 0.52 | 0.55 | 0.56 | 0.58 |
| 0.6 (cutoff points) | 0.55 | 0.59 | 0.60 | **0.61** | 0.53 | 0.56 | 0.57 | **0.59** |
| 0.8 | 0.54 | 0.57 | 0.59 | 0.60 | 0.51 | 0.55 | 0.56 | 0.57 |
| 1.0 | 0.50 | 0.54 | 0.56 | 0.59 | 0.49 | 0.53 | 0.54 | 0.56 |

Table 9: Comparison of SP-HCFR under the influence of different $\alpha$ (MAP and DOA).

| SP-HCFR | MAP ($k=10, 20, 30, 50$) | | | | DOA ($n=80\%–20\%, 70\%–30\%, 60\%–40\%, 50\%–50\%$) | | | |
|---|---|---|---|---|---|---|---|---|
| | 10 | 20 | 30 | 50 | 80%–20% | 70%–30% | 60%–40% | 50%–50% |
| 0.0 | 0.53 | 0.57 | 0.60 | 0.62 | 0.79 | 0.82 | 0.84 | 0.85 |
| 0.2 | 0.56 | 0.59 | 0.62 | 0.63 | 0.81 | 0.84 | 0.86 | 0.86 |
| 0.4 | 0.57 | 0.60 | 0.63 | 0.64 | 0.82 | 0.85 | 0.87 | 0.87 |
| 0.6 (cutoff points) | 0.58 | 0.61 | 0.64 | **0.65** | 0.83 | 0.85 | 0.87 | **0.88** |
| 0.8 | 0.56 | 0.60 | 0.62 | 0.64 | 0.82 | 0.84 | 0.86 | 0.87 |
| 1.0 | 0.54 | 0.58 | 0.61 | 0.63 | 0.80 | 0.83 | 0.85 | 0.87 |

of mobile users. PTPP obtains more accurate and objective scores of personality traits, improving the accuracy of calculation of user similarity in follow-up collaborative filtering recommendation.

### 4.4. Experimental Results of SP-HCFR

*4.4.1. Comparison of Hybrid Collaborative Filtering Methods Based on User Similarity under the Influence of Different $\alpha$.* Different values of $\alpha$, which is set to $\alpha = 0, 0.2, 0.4, 0.6, 0.8, 1.0$, mean different weight influence of sentiment characteristic and privacy concern in collaborative recommend. SP-HCFR combines the method of user collaborative filtering algorithm based on personality traits (PP-UCF, $\alpha = 0.0$) with the method of user collaborative filtering algorithm based on sentiment analysis (SA-UCF, $\alpha = 1.0$). This paper has some comparison experiments to the influence in the weighted coefficient $\alpha$. The results are shown in the evaluating indexes of MAP, DOA, $P@10$, and $P@5$ and are listed in Tables 8 and 9. The $\alpha$ and $k$ are set to $\alpha = 0.2, 0.4, 0.6, 0.8$ and $k = 10, 20, 30, 50$. Firstly, several experiment results show that SP-HCFR has a higher accuracy than others, and it reaches the highest accuracy when $\alpha = 0.6$. Secondly, the value of $\alpha$ is of nonlinear relationship with the sorting accuracy. Thirdly, the results show that SP-HCFR can obtain more accuracy in $k$ similar users by using the combination of user similarity calculation and also improves the accuracy of recommendation.

*4.4.2. Performance Comparison of Different Collective Filtering Algorithms.* In order to analyze the influence of privacy preference, personality traits, and sentiment characteristics on MPS, this paper does some comparisons among different algorithms. The results are shown in Figure 2. Firstly, SP-HCFR is ahead of traditional collaborative filtering algorithm based on user (UCF) in the indexes of $P@10$, MAP while the importance weight is $\alpha = 0.6$. Secondly, although actual score data are sparse, SP-HCFR solves this problem by using the information of privacy preference and sentiment characteristics for calculation of user similarity. Thirdly, the results show that it has great significance for introducing privacy concern, personality traits, and sentimental characteristics in recommendation.

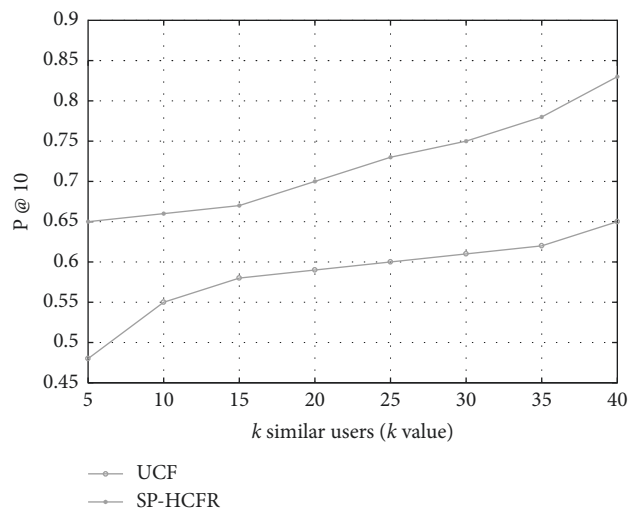


Figure 2: Performance comparison between SP-HCFR and UCF.

This paper does the performance comparison between collaborative filtering recommend method based on personality traits combined with privacy concerns (PP-UCF) and collaborative filtering recommend method based on sentiment analysis (SA-UCF) and UCF, while $\alpha = 0.6$. The results are shown in Figure 3. Firstly, SA-UCF has not considered the influence of privacy preference on users'

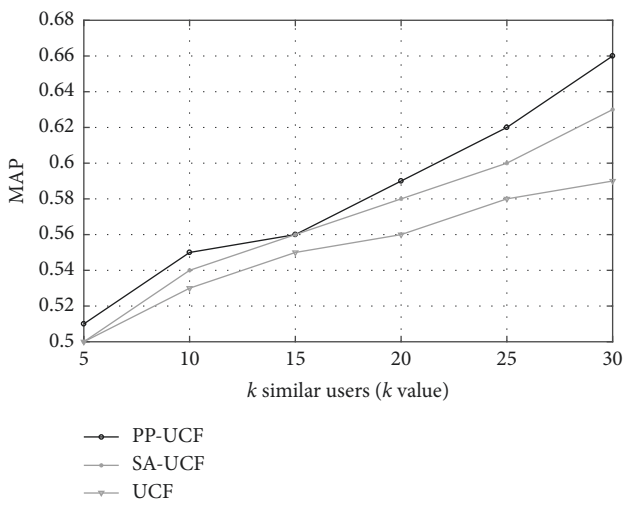


Figure 3: Performance comparison among PP-UCF, SA-UCF, and UCF.

interest, which cannot significantly improve the accuracy of recommendation. Secondly, PP-UCF has a greater influence on the quality of recommendation, which means privacy concern is more important than sentiment characteristics to obtain precision user's interest. Thirdly, PP-UCF and SA-UCF both have better performances than UCF. It also shows that using the information of privacy preference and sentiment characteristics for calculation of user similarity can solve the problem of data sparseness and cold start.

## 5. Conclusion and Future Work

With the wide use of MPS in mobile commerce, the problems in protecting user's privacy and obtaining user's interest with complicated sentiment are outstanding. A novel personalized recommendation technology should be proposed to address the privacy concern and sentiment analysis. Therefore, this paper proposes a novel recommendation model based on subjective privacy preference and objective recommend technology. The main contributions are as follows: (1) since it can better match user's preference through learning complex sentiment, this paper puts forward STAS to mine the sentiment preference, effectively solves problems of data sparse and cold start, and improves the accuracy of user interests. (2) User's interest is similar with persons who have common privacy concern and personality traits. Therefore, this paper puts forward PTPP to obtain the $k$ similar users. (3) This paper takes full use of both advantages of the above two contributions and puts forward a novel hybrid collaborative filtering recommendation method based on sentiment analysis and privacy concern to protect user's privacy and giving MPS.

A follow-up study may utilize methods of data mining to obtain user's dynamic privacy interest. At the same time, from the perspective of the importance of privacy protection, we will do the research of punctual personalized recommend services by using the control degree of users' privacy disclosure, intensity of privacy concerns, and so on.

## Conflicts of Interest

The authors declare that there are no conflicts of interests regarding the publication of this paper.

## Acknowledgments

This research was supported by Philosophy and Social Science Planning Project of Zhejiang Province (no. 18NDJC278YB), Natural Science Foundation of Zhejiang Province (no. LY18G020012 and LQ17G020003); Major Project of Social Science Foundation of China (no. 16ZDA053), Major Project of Humanities and Social Sciences in University of Zhejiang Province (no. 2016GH024), and Scientific Research Project of Zhejiang Provincial Education Department of China (no. Y201636584).

## References

[1] Z. Lin, K. Y. Goh, and C. S. Heng, "The demand effects of product recommendation networks: an empirical analysis of network diversity and stability," *MIS Quarterly*, vol. 41, no. 2, pp. 397–420, 2017.

[2] J. Bobadilla, F. Ortega, A. Hernando, and A. Gutiérrez, "Recommender systems survey," *Knowledge-Based Systems*, vol. 46, no. 1, pp. 109–132, 2013.

[3] D. Gavalas, K. Charalampos, M. Konstantinos, and P. Grammati, "Mobile recommender systems in tourism," *Journal of Network and Computer Applications*, vol. 39, no. 1, pp. 319–333, 2014.

[4] D. V. Sowmini, V. R. Kagita, A. K. Pujari et al., "Collaborative filtering by PSO-based MMMF," in *Proceedings of the IEEE International Conference on Systems, Man and Cybernetics*, pp. 569–574, San Diego, CA, USA, 2014.

[5] E. Cambria, "Affective computing and sentiment analysis," *IEEE Intelligent Systems*, vol. 31, no. 2, pp. 102–107, 2016.

[6] A. Yenter and A. Verma, "Deep CNN-LSTM with combined kernels from multiple branches for IMDb review sentiment analysis," in *Proceedings of the 2017 IEEE 8th Annual Ubiquitous Computing, Electronics and Mobile Communication Conference (UEMCON)*, pp. 540–546, New York, NY, USA, October 2017.

[7] P. Kefalas and Y. Manolopoulos, "A time-aware spatio-textual recommender system," *Expert Systems with Applications*, vol. 78, pp. 396–406, 2017.

[8] Z. Li, X. Fang, X. Bai, and O. R. Liu Sheng, "Utility-based link recommendation for online social networks," *Management Science*, vol. 63, no. 6, pp. 1938–1952, 2017.

[9] E. Cambria, B. Schuller, Y. Xia, and C. Havasi, "New avenues in opinion mining and sentiment analysis," *IEEE Intelligent Systems*, vol. 28, no. 2, pp. 15–21, 2013.

[10] B. Pang and L. Lee, "Opinion mining and sentiment analysis," *Foundations and Trends in Information Retrieval*, vol. 2, no. 1-2, pp. 1–135, 2008.

[11] P. Hawking, A. Stein, J. Zeleznikow et al., "Emerging issues in location based tourism systems," in *Proceedings of the International Conference on Mobile Business, IEEE*, pp. 75–81, Sydney, NSW, Australia, July 2005.

[12] J. Y. Tsai, S. Egelman, L. Cranor, and A. Acquisti, "The effect of online privacy information on purchasing behavior: an experimental study," *Information Systems Research*, vol. 22, no. 2, pp. 254–268, 2011.

[13] F. Provost, D. Martens, and A. Murray, "Finding similar mobile consumers with a privacy-friendly geosocial design," *Information Systems Research*, vol. 26, no. 2, pp. 243–265, 2015.

[14] P. Kefalas, P. Symeonidis, and Y. Manolopoulos, "New perspectives for recommendations in location-based social networks: time, privacy and explainability," in *Proceedings of the International Conference on Management of Emergent Digital Ecosystems, ACM*, pp. 1–8, Luxembourg City, Luxembourg, October 2013.

[15] J. Schrammel, J. Koffel, and M. Tscheligi, "Personality traits usage patterns and information disclosure in online communities," in *Proceedings of the 23rd British HCI Group Annual Conference on People and Computes: Celebrating People and Technology, British Computer Society*, pp. 169–174, Swinton, UK, September 2009.

[16] M. Hu and B. Liu, "Mining and summarizing customer reviews," in *Proceedings of the International Conference on Knowledge Discovery and Data Mining (KDD 2004)*, pp. 168–177, Seattle, WA, USA, August 2004.

[17] H. W. Wang, P. Yin, J. Yao, and J. N. K. Liu, "Text feature selection for sentiment classification of Chinese online reviews," *Journal of Experimental and Theoretical Artificial Intelligence*, vol. 25, no. 4, pp. 425–439, 2013.

[18] G. Somprasertsri and P. Lalitrojwong, "Mining feature-opinion in online customer reviews for opinion summarization," *Journal of Universal Computer Science*, vol. 16, no. 6, pp. 938–955, 2010.

[19] B. Ma, D. Zhang, Z. Yan, and T. Kim, "An LDA and synonym lexicon based approach to product feature from online consumer product reviews," *Journal of Electronic Commerce Research*, vol. 14, no. 4, pp. 304–314, 2013.

[20] J. Zhao, K. Liu, and G. Wang, "Adding redundant features for CRFs-based sentence sentiment classification," in *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, pp. 117–126, Honolulu, HI, USA, October 2008.

[21] C. Mi, X. Shan, Y. Qiang, Y. Stephanie, and Y. Chen, "A new method for evaluating tour online review based on grey 2-tuple linguistic," *Kybernetes*, vol. 43, no. 3-4, pp. 601–613, 2014.

[22] G. Somprasertsri and P. Lalitrojwong, "A maximum entropy model for product feature extraction in online customer reviews," in *Proceedings of the IEEE Conference on Cybernetics and Intelligent Systems*, pp. 786–791, Chengdu, China, September 2008.

[23] D. Yang, D. Zhang, Z. Yu et al., "A sentiment-enhanced personalized location recommendation system," in *Proceedings of the ACM Conference on Hypertext and Social Media, ACM*, pp. 119–128, Paris, France, May 2013.

[24] F. F. Kuo, M. F. Chiang, M. K. Shan et al., "Emotion-based music recommendation by association discovery from film music," in *Proceedings of the ACM International Conference on Multimedia*, pp. 7666–7674, Singapore, November 2005.

[25] R. Cai, C. Zhang, C. Wang, L. Zhang, and W. Y. Ma, "MusicSense: contextual music recommendation using emotional allocation modeling," in *Proceedings of the 15th International Conference on Multimedia (MULTIMEDIA'07)*, pp. 553–556, New York, NY, USA, 2007.

[26] B. J. Han, S. Rho, S. Jun, and E. Hwang, "Music emotion classification and context-based music recommendation," *Multimedia Tools and Applications*, vol. 47, no. 3, pp. 433–460, 2010.

[27] S. Mudambi and D. Schuff, "What makes a helpful online review? A study of customer reviews on Amazon.com," *MIS Quarterly*, vol. 34, no. 1, pp. 185–200, 2010.

[28] J. She and L. Chen, "TOMOHA: topic model-based HAshtag recommendation on twitter," in *Proceedings of the International Conference on World Wide Web Companion*, pp. 371-372, Seoul, Republic of Korea, April 2014.

[29] P. Winoto and T. Y. Tang, "The role of user mood in movie recommendations," *Expert Systems with Applications*, vol. 37, no. 8, pp. 6086–6092, 2010.

[30] Y. Shi, M. Larson, and A. Hanjalic, "Mining mood-specific movie similarity with matrix factorization for context-aware recommendation," in *Proceedings of the Workshop on Context-Aware Movie Recommendation, ACM*, pp. 34–40, Barcelona, Spain, September–October 2010.

[31] H. Cavusoglu, T. Q. Phan, H. Cavusoglu, and E. M. Airoldi, "Assessing the impact of granular privacy controls on content sharing and disclosure on Facebook," *Information Systems Research*, vol. 27, no. 4, pp. 848–879, 2016.

[32] M. J. Culnan and P. K. Armstrong, "Information privacy concerns, procedural fairness and impersonal trust: an empirical investigation," *Organization Science*, vol. 10, no. 1, pp. 104–115, 1999.

[33] H. Xu, T. Dinev, H. J. Smith, and P. Hart, "Examining the formation of individual's information privacy concerns: toward an integrative view," in *Proceedings of the 29th Annual International Conference on Information Systems*, pp. 14–17, Paris, France, December 2008.

[34] R. Zhang, J. Q. Chen, and C. J. Lee, "Mobile commerce and consumer privacy concerns," *Journal of Computer Information Systems*, vol. 53, no. 4, pp. 31–38, 2013.

[35] S. Menon and S. Sarkar, "Privacy and big data: scalable approaches to sanitize large transactional databases for sharing," *MIS Quarterly*, vol. 40, no. 4, pp. 963–982, 2016.

[36] X. B. Li and S. Sarkar, "Class restricted clustering and microperturbation for data privacy," *Management Science*, vol. 59, no. 4, pp. 796–812, 2013.

[37] F. McSherry and I. Mironov, "Differential private recommender system: building privacy into Netflix prize contenders," in *Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 627–636, New York, NY, USA, 2009.

[38] M. Duckham and L. Kulik, "A formal model of obfuscation and negotiation for location privacy," in *Lecture Notes in Computer Science*, pp. 152–170, Pervasive, Munich, Germany, 2005.

[39] B. Wang and Y. X. Duan, "Research on information privacy quantization method facing ubiquitous computing environment," *Computer Engineering and Applications*, vol. 47, no. 27, pp. 1–5, 2011.

[40] J. Sutanto, E. Palme, C. H. Tan, and C. W. Phang, "Addressing the personalization privacy paradox: an empirical assessment from a field experiment on smartphone user," *MIS Quarterly*, vol. 37, no. 4, pp. 1142–1164, 2013.

[41] R. K. Chellappa and S. Shivendu, "Mechanism design for "Free" but "No Free Disposal" services: The economics of personalization under privacy concerns," *Management Science*, vol. 56, no. 10, pp. 1766–1780, 2011.

[42] M. L. Korzaan and K. T. Boswell, "The influence of personality traits and information privacy concerns on behavioral intentions," *Journal of Computer Information Systems*, vol. 48, no. 4, pp. 15–24, 2016.

[43] I. Junglas and C. Spitzmuller, "Personality traits and privacy perceptions: an empirical study in the context of location-based services," in *Proceedings of the International Conference on Mobile Business*, pp. 36–47, Copenhagen, Denmark, June 2006.

[44] S. S. Choi and M. K. Choi, "Consumer's privacy concerns and willingness to provide personal information in location-based services," in *Proceedings of the International Conference on Advanced Communication Technology, IEEE*, pp. 2196–2199, Gangwon-do, Republic of Korea, February 2007.

[45] B. V. Roy and X. Yan, "Manipulation robustness of collaborative filtering," *Management Science*, vol. 56, no. 11, pp. 1911–1929, 2010.

[46] C. Q. Quan and F. J. Ren, "A blog emotion corpus for emotional expression analysis in Chinese," *Computer Speech and Language*, vol. 24, no. 4, pp. 726–749, 2010.

[47] Y. Zhang, Z. M. Li, F. J. Ren et al., "Semi-automatic emotion recognition from textual input based on the constructed emotion thesaurus," in *Proceedings of the International Conference on Natural Language Processing and Knowledge Engineering*, pp. 571–576, Wuhan, China, October 2005.

[48] L. Luo, "DUTIR at the BioCreative V.5.BeCalm tasks: A BLSTM-CRF approach for biomedical entity recognition in patents," in *Proceedings of the Biocreative V.5 Challenge Evaluation Workshop*, Barcelona, Spain, April 2017.

[49] J. D. Lafferty, A. Mccallum, and F. C. N. Pereira, "Condition random fields: probabilistic models for segmenting and labeling sequence data," in *Proceedings of the 18th International Conference on Machine Learning*, pp. 282–289, Sydney, NSW, Australia, 2002.

[50] W. Shi, H. W. Wang, and S. Y. He, "Study on construction of fuzzy emotion ontology based on Howne," *Journal of the China Society for Scientific and Technical Information*, vol. 31, no. 6, pp. 595–602, 2012.

[51] L. Yue, C. Malu, D. Umeshwar, and Z. ChengXiang, "Automatic construction of a context-aware sentiment lexicon: an optimization approach," in *Proceedings of the 20th International Conference on World Wide Web*, pp. 347–356, Hyderabad, India, March–April 2011.

[52] B. Henson, B. W. Reyns, and B. S. Fisher, "Security in the 21st century: examining the link between online social network activity, privacy, and interpersonal victimization," *Criminal Justice Review*, vol. 36, no. 3, pp. 253–268, 2011.

[53] M. Selfhout, W. Burk, S. Branje, J. Denissen, M. van Aken, and W. Meeus, "Emerging late adolescent friendship networks and big five personality traits: a social network approach," *Journal of Personality*, vol. 78, no. 2, pp. 509–538, 2010.