RESEARCH ARTICLE OPEN ACCESS

USER RATING CONFIDENCE PREDICTION FOR PRODUCTS USING SOCIAL USER'S TEXTUAL AND TEMPORAL INFORMATION

[1]V Swedha , [2]A Pravi

[1]Master of Engineering in Software Engineering Department of ComputerScience
Anna University,Chennai-600025

[2]Assistant Professor Department of Information Technology
Sri Muthukumaran Institute of Technology,Chennai-600069

ABSTRACT:

Opinion Mining or sentiment analysis is a computation study and ongoing research field which aims to analyse the people's opinion, sentiment and emotions. E- commerce websites allows the people to share their feedback/opinion as textual reviews regarding the products/services. In traditional, many researches has focused on personalized recommendation ,which uses the historical data for recommendation .The historical data contains review rating and reviews for prediction and recommendation. Recommendation mainly focuses on User-item ratings. The problem arises when there is drift in user preferences/interest in purchasing the product. In the proposed work, the concept of User rating's confidence is used to denote the trustworthiness of user ratings. To predict the user confidence rating and analyse the best qualified product, entropy is calculated. In addition to find accuracy in the user rating confidence, constraints like sentiment and temporal features is used. Finally, all the three factors are merged to find each User item rating and overall item rating. The overall item rating suggests the quality of product. This system results in accurate performance accuracy and better prediction ofratings.

Keywords—Data mining, recommender system, sentiment analysis, review ratings, social networks.

1. INTRODUCTION

In recent years, enhancement of internet access helps the social users to upgrade the knowledge in various domains. E-commerce websites have been used to purchase the products share their opinions and gain much more information regarding the web services .Recommendation system helps to seek the user preference /interest over the rating of an item. Mostly recommendation system based on user –item ratings. There exists some difference while rating and reviewing the products posted by social users. To avoid the difference in rating the prediction, user confidence ratings is enhanced for trustworthiness of user-rating. To represent the inter user similarity between the social users , entropy value is calculated for user item ratings using the old user ratings. To constraint the user ratings confidence, sentiment and temporal features are used. Then the entropy values and constraints are fused to predict the exact user confidence ratings for their rated items. Thus the model uses the pre- processing in which the user confidence rating (entropy values) and low level constraints(sentiment and temporal features extraction using 4 th degree Gaussian method) calculation steps are done to find the user confidence rating and product quality. During thelow-level constraints calculation, the sentiment score is calculated for predicting the user confidence rating using sentiment features and curve fitting values are calculated using temporal features. Finally all three factors are fused, each user-item confidence rating is calculated and overall item rating is calculated. Thus the overall item rating suggests the quality of theproduct.

To address the problems in text analysis, the product reviews containing unwanted contents, special symbols problem must be handled. The constraint in the user confidence rating must be handled at regular interval. Since, the user's interests drifts over change in time. The overall rating of each product must be accurately calculated. The sparse matrix problem must be handled. The proposed work includes the user's confidence rating calculation and low level constraints calculation to handle the text analysis and overall rating of each product.

The major contributions of the proposed system are: i) filters the unwanted symbols, Uniform Resource Locator (URL)s, stemming the product description. ii)User Rating Confidence values are calculated with entropy value of the user and product rates iii) Low – level constraints values are calculated using the sentiment and temporal features. Curve fittings values are used for calculationiv)TheUserConfidence ratings are calculated with user confidence values and the low –level constraints values. Finally all three factors are fused ,each user-item confidence rating is calculated and overallitem rating is calculated.The experimental results shows better performance in rating prediction and in suggesting the quality of theproduct.

The remainder of this paper is organized as follows. In Section 2, we present the related work. In section 3, we describe the proposed approach .In section 4, the dataset introduction were provided. In section 5, Experiment results were discussed. In section 6, Conclusions were made. In section 7 , References are drawn.

2. RELATEDWORK

The opinion mining is used to collect and categorize the opinions about the products and also find the product quality through rating and reviews as stated by the social user's .It generally analyse the user's preferences on the product .The purchasing of the products dependson the historical data about the user ratings and the product features. The existing system only calculates the user's rating according to the reviews and itsratingsgiven.Theproblemoccurswhenthereisdeviationin the user's confidence about the products from time to time. To overcome user's confidence rating with constraint (sentiment and temporal) and to predict the exact user- rating on the product and to check the overall product quality the proposed system isdeveloped.

A. *Entropy Calculation(User ConfidenceRating)*

User preferences and the overall rating for the products differ over time .Entropy is a measure of uncertainity. The entropy value represents the confidence value of user ratings .So the lower the entropy value is, more stable the product quality is ,and the more confident the user's rating is.

Bell and Volinsky [1] developed a personalized product recommendation which involves Matrix factorization method for analyzing the latent factors and nearest-neighbour techniques. In addition, conventional Singular Value Decomposition (SVD) is used for overcoming sparse matrix. However this system lacks in finding the sentiment score and spatial –temporal features. The dataset used was Amazon and Netflix.

Zhao & Quain [11] developed a framework to predict User-Servicerating.Thisframeworkfusesfourfactorsnamely, user personal interest (related to user and the item's topics), interpersonal interest similarity (related to user interest), interpersonal rating behavior similarity (related to users' rating behavior habits), and interpersonal rating behavior diffusion (related to users' behavior diffusions), into a unified matrix-factorized framework. But however, no temporal features prediction rating were involved to find exactness of user's interest. The dataset used is Yelp and Douban Movie dataset forprediction.

Zhang & Ding [9] developed collaborative recommender system and used the user-item rating matrix along with sentiment classification to overcome the sparseness of matrix. The system lacked in finding the spatio-temporal features which in-turn lacked in finding user's preference. Youko (Chinese dataset) was used for experiment.

B. *User confidence rating (sentimentfeatures)*

Mostly, in e-commerce websites, the users are allowed to rate the product, share their opinions and attitudes byreviewing.Thetextualreviewsanduserratingmaydifferin rare circumstances. This system lacks in the exactness of user rating. It is necessary to analyze the relevance between user confidence and textural review sentiment. So the sentiment score is calculated using the reviews and sentiment dictionary and finally the relevance user confidence sentiment value is predicted.

Chelaru et al. [2] developed a query based recommendation system which involved sentiment analysis using the SentiWordNet thesaurus and a lexical based analysis for resource containing sentiment annotations. Query recommendation employed sentiment detection and sentiment classification to discover the topics. Sentiment Analysis was done with the available reviews and queries.

Tan et al. [6] proposed a Latent Dirichlet Allocation (LDA) using Foreground and Background LDA (FB-LDA) concept. This model was used to distill foreground topics and filter out longstanding background topics between two sets of documents. Twitter Dataset was used for the experiment.

Lei et al. [5] developed a sentiment-based rating prediction to improve prediction accuracy in recommender systems. The user's sentiment scores and interpersonal sentiment score was calculated along with product reputation. The experiment was done using real-world dataset collected from Yelp. Even though, the system lacked in linguistic rules (and, but) while during sentiment analysis.

C. *User Confidence Rating (TemporalFeatures)*

User preference about the product may drift over change intime.Forpredictingthechanges,first,theaveragedifference between user ratings and overall rating of product is calculated within a period. Secondly, the curve fitting values i.e, peak values, and centroid and peak widths (centroid to other values) are calculated of products within certain period (months) using 4th degree Gaussian method. Finally, the user confidence ratings are calculated using these curve fittingvalues.

Yang et al. [7] developed a cross-domain feature learning (CDFL)algorithmbasedonstackeddenoisingauto-encodersto handlethetext,imageandvideoofthecontextsimultaneously. In this model, sentiment classification, spam filtering, and event classification was involved. The experiment used Amazon dataset. This system lacked in spatial features prediction.

Yang & Zhang [7] developed a users' spatial temporal activitypreference(STAP)toreducetheproblemcomplexityof user's social activities. This model separately considered the spatial and temporal features of user activities by introducing the notion of spatial specificity and temporal correlation. Finally, a context-aware fusion framework was designed and fused both features in activity preference inference. Three datasets where collected from two LBSNs, i.e., Foursquare and Gowalla was used forexperiment.

D. Product Quality RatingPrediction

Product Quality Prediction is done to find exact user confidence rating and overall rating of each product. First the user confidence rating about the product is calculated usingthe old product rating, entropy values and the low levelconstraints value (sentiment features and temporal features). Then the overall rating of the product is calculated. This overall ratingis used to suggest product quality and the user'stastes.

Zhao& Qian [4] proposed a model to solve service objective evaluation by deep understanding social users since users' tastes and habits are drifting over time. This model focused on user ratings confidence (trustworthiness of user ratings) in service objective evaluation. This model utilized entropy to calculate user ratings confidence and spatial - temporal features of user ratings to constrain confidence level oftheuserratings.TheexperimentwasbasedonYelpdatasets. The system lacked in finding the sentiments of social users' reviews.

Lei X. & Qian X. [4] developed a social user's reviews sentiment measurement system for calculating each user's sentiment score on items/services, service reputation and accurate rating prediction based on probabilistic matrix factorization. The experiment was conducted on Yelp dataset. The system lacked in calculating the depth of social users sentiment.

3. DESIGN AND IMPLEMENTATION

A. Proposed work

The proposed system framework predicts the new user confidence ratings for the given rating after the inclusion of low level constraint. This system starts with the preprocessing stage which filters the unwanted information from the reviews dataset; perform the stop words and stemming process to obtainpreprocessedcontent.Thepreprocessedcontentareused to calculate the entropy values and the overall user confidence rating of the products without the constraint information. Later the low level constraints, such as sentiment user confidence rating and the temporal user confidence ratings are calculated. Then the overall user confidence ratings for each user-item pairs are calculated with the entropy and the low level constraints values and finally the overall confidence rating of items are calculated, to find the product quality and likes among theusers.

Figure 1 describes the overall system architecture of the proposed user confidence rating prediction for products using social user's textual and temporal information.

B. ModuleDescription

3.1 Data Pre-processing(StepI)

Input: Product Review Dataset, Stop words list

Output: Pre-processed data

Steps:

1. Loop through the productsreviews.
2. Removal of URL, Special character anddigits.
3. Tokenization is done for each productreviews.
4. Remove stop-words from tokenizedcontent.
5. Stem the each word of thecontent.
6. Stemmed words for each sentence arecombined.
7. End the loop

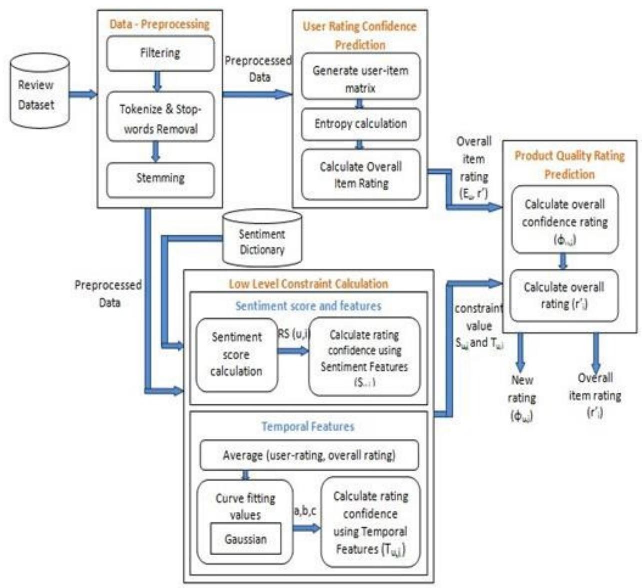

Fig. 1 Overall architecture of the proposed framework

3.2 User Rating Confidence Prediction(StepII)

Input: Preprocessed data

Output: Overall user confidence rating for each product.

Steps:

1. The user-item matrix isgenerated.
2. For the user-item match, particular user rating is populated else 0 is populated. Since entropy is a measure of uncertainity. The information entropy valueusedtorepresenttheconfidencevalueofuser ratings.
3. Theentropyvalueiscalculatedandthenormalized value for each user is calculated by cumulating each user ratings using Equation(i).

$$E_u = {}^{-1}/_{(\sum_i |d_i| * p(d_i) \log 2p(d_i))} \quad (i)$$

Where,

Euistheentropyvalueofeachuser,p(d_i) is the probability of itemrating.

The lower entropy value is, the more stable the system is, the more confident the users rating is.

4.Overall user confidence rating for each item is calculated using the entropy values and normalized values using Equation (ii)

$$Eu^* = \sum Eu(K) \qquad \text{(ii)}$$

Thus,theoveralluserconfidenceratingsforallitems withoutconstraintsarecalculatedusingEquation(iii) and obtained asoutput

$$\widehat{r_i} = \sum_{u=0}^{N} E_u^* * r_{u,i} \text{(iii)}$$

3.3 Sentiment Score Calculation and User Confidence Rating (SentimentFeatures)

Input: Preprocessed Data, Sentiment Dictionary terms

Output: Sentiment score and user confidence rating with sentimentfeatures.

Steps:

1) Loop through the pre-processedcontents.
2) Extract all the positive sentiment, negativesentiment, conjugative words and wordsdegree-levels.
3) Calculate the sentiment score using Equation(iv)

$$s(r) - {}^{1}/_{N_c}(\sum_{c\epsilon r} \sum_{w\epsilon c} Q.D_W.R_W) \qquad \text{(iv)}$$

Where,

S(r) - Sentiment Score,
Nc – Number of clauses
Q – Negation Check Coefficient ,
Dw – Degree level words (most, more …),
Rw – Initial Score of sentiment word.

4) Normalize sentimentscore

$$E_{u,j} = (10/(1 + e^{-S|r|})) - 5 \text{(v)}$$

5) Calculate the user confidence rating withsentiment features.

$$s_{u,j} = {}^{1}/_{(\sum_i a_i * (E_{u,i})^j)} \qquad \text{(vi)}$$

where,

Su,i - user confidence rating of user (u) to item(i)
aj - curve fitting value(high peak value of item(months))

3.4 User Confidence Rating (Temporal Features)

Input: Pre-processed Data

Output: User Confidence rating with temporal features

Steps:

1) Loop through the pre-processedcontents.
2) Calculate the average rating (user rating and overallrating).
3) Get the peak height (aj), centroid (bj) and peak width (cj) value as curve fitting values using Gaussian method foreach month anditem.
4) calculate the user confidence rating with temporalusing

$$T_{u,i} = {}^{1}/_{\sum_j a_j * \exp(-(\frac{R_{u,i} - b_j}{c_j}))} \qquad \text{(vii)}$$

where,

Tu,i - user confidence rating of user (u) to item(i),
aj - curve fitting value(high peak value of item(months)),
bj - scurve fitting value(centroid(months)),
cj - curve fittingvalue(peakwidth(months)),$R_{u,i}$
- Rating given by user (u) to item(i)

3.5 Product Quality Prediction

Input: Tu,i , Su,i , Eu, R

Output: Overall User Confidence Rating for Items and New user-item rating

Steps:

1) Loop through each item rating given by each userto calculate the overall user confidencerating.

$$\Phi_{u,i} = A_{u,t(u,i)} T_{t(u,i)} + B_{u,s(u,i)} S_{s(u,i)} + C_{u,(t(u,i),s(u,i))} \qquad \text{(viii)}$$

where,

$\Phi_{u,i}$- User Confidence rating of user(u) to item(i),
$A_{u,t(u,i)}$ - User Confidence rating using temporal features of user (u) to item(i) at time (t),
$B_{u,s(u,i)}$- User Confidence rating using sentiment featuresof user (u) to item(i) at time (t),
$C_{u,(t(u,i),s(u,i))}$- Entropy value with low – level – constraints user (u) to item(i) at time(t),

2) Calculate of overall rating of each item using (ix)

$$r_i' = \sum_{u=0}^{ni}((\frac{\phi ui}{\sum_{u=0}^{ni} \phi ui}) * r(u,i)) \textbf{(ix)}$$

where, ri'- overall user confidence rating ,ϕu,i - User Confidence rating of user(u) to item(i),r(u,i) - Rating of user(u) to item(i).

4. EXPERIMENT

A. DatasetDescription

The dataset used for the experiment is Amazon. It contains the product reviews according to user's interest or preferences. The product rating and reviews depict the overall interest of the consumers of Amazon products .List of features extracted from reviews and ratings information allows the system to handle the cold-start (new user and new item) problem. The table 1 contains the characteristics of the dataset which were utilised for the experiment. The important attributes used for the experiment are product id , Reviewer id, Overall rating, year.

Table 1 List of Attributes in Dataset

Attributes	**Description**
PID	ID of the product
ReviewerID	ID of the reviewer
Reviewer Name	Name of the reviewer
Year	Year in which the ratings and reviews are calculated.
Overall rating	Overall Star rating given by the social user's.

B. EvaluationMeasures

This section briefly describes about the evaluation measures .

1) Precision

Precision is known as positive predictive value. It is defined as the number of correct result divided by the number of the retrieved result. Precision returns the exactness (accurate) of the result. If the threshold value changes the precision may increase or decrease accordingly. The precision value is calculated after finding the new rating, to check whether the user had given the positive or negative ratings correctly.

$$\text{Precision}=\sum TP / (\sum TP + \sum FP)$$

Where,

TP = True positive, total number of true positive conditions ranked

FP = False positive, total number of false conditions incorrectly ranked.

2) *Recall:*

Recall is known as true positive value or sensitivity. It is defined as the number of correct result divided by the number of the relevant result. Recall returns the completeness of the result but doesn't depend on any threshold value. . The recall value is calculated after finding the new rating, to check whether the user had given the positive or negative ratings correctly.

$$\text{Recall} = \sum TP / \sum P$$

where,

TP = True positive, total number of true positive conditions ranked.

P = Total number of correctly ranked.

Table 2: Calculated Precision and Recall values for various products

system	seeds	fertilizer	crops	cereals
entropy + SF (precision)	0.9225	0.8267	0.9667	0.9715
entropy + SF (recall)	0.9174	0.9667	0.9176	0.9064
entropy + SF +TF (precision)	0.9776	0.9729	0.9669	0.9765
entropy + SF +TF (recall)	0.9325	0.9668	0.9177	0.9121

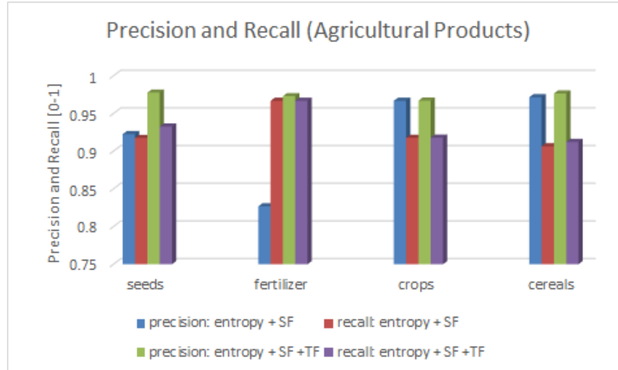

Fig. 2 Comparison of Precision and Recall for various agricultural products.

3) Accuracy:

The Accuracy or recognition rate measures the proportion of true results to the total number of the instances in the dataset. It specifies how many instances correctly predicted. This affects the system performance. Misclassification measures the error rate (ie) the proportion of false results to the total number of the instances for the given dataset. It specifies how many instances falsely predicted. This affects the systemperformance.

$$\text{Accuracy}= (\sum TP+ \sum TN)/(\sum P + \sum N)$$

Where,

TP=True positive, total number of true positive reviews is classified as positive.

TN= True negative, total number of true negative reviews is classified as negative.

P = Total number of correctly ranked.

Table 3: Calculated Accuracy and Misclassification percentages for various products

system	seeds	fertilizer	crops	cereals
entropy + SF (Acc)	88.54	83.6	84.2	85
entropy + SF (Mis)	11.46	16.4	15.8	15
entropy + SF +TF (Acc)	93.5	88.5	91	87.6
entropy + SF +TF (Mis)	5.5	11.5	9	12.4

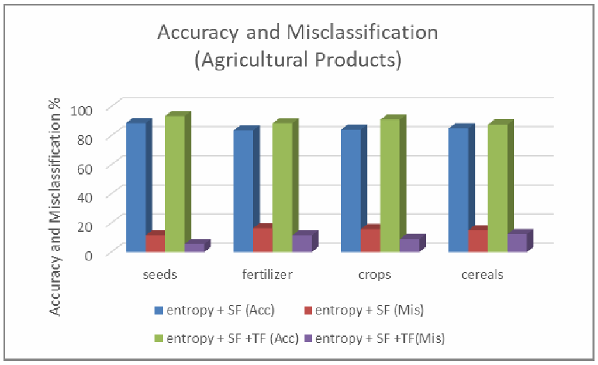

Fig. 3 Comparison of Accuracy and Misclassification for various agricultural products.

4) Root Mean Squared Error(RMSE)

The Root Mean Squared Error(RMSE) measures the average error rate calculated using the actual value and the predicted value. If the measure is low, the system performance is good else not good. If the measure is Zero. It proves that the system is best .The RMSE values is calculated after finding the new ratings for each review. Finally the difference is calculated between old and new ratings to check the error in thesystem.

RMSE= $$\sqrt{1}/_{m} \sum_{i=1}^{m} (ti - yi)^2$$

Where,
m=total number of reviews accessed,
ti=actual value predicted for the reviews.

4) Mean Absolute Error(MAE)

The Mean Absolute Error (MAE) measures the average deviation (error) between the predicted overall rating confidence and overall rating. This measure the forecasting error and misleading in the output predicted. If the MAE is 0 it is considered to best prediction result. The MAE value is calculated after finding the new ratings for each review.

Misclassification=1 – accuracy.

Table 4: RMSE and MAE value for various agricultural products

system	seeds	fertilizer	crops	cereals
entropy + SF(mae)	0.4798	0.588	0.783	0.635
entropy + SF (rmse)	0.814	0.795	0.864	0.851
entropy + SF +TF (mae)	0.2346	0.3346	0.645	0.425
entropy + SF +TF(rmse)	0.7533	0.7663	0.6953	0.641

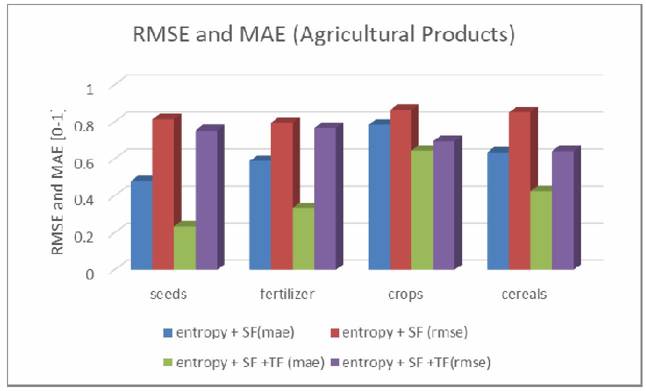

Fig. 4 Comparison of RMSE and MAE values for various products.

Figure 4 shows the snapshot of average values calculated for Root Mean Squared Error (RMSE) and Mean absolute error(MAE)for various products listed in Table 6. The Root Mean Squared Error (RMSE) and Mean absolute error (MAE) of the proposed system are low compared to the existing system. This proves the proposed system is performing better prediction compared to the existing system.

5. CONCLUSION

In the proposed work, special symbols and the URLs are handled. Stemming of contents is done for obtaining the better sentimental score. The Entropy values are calculated using the user ratings and normalized. The overall user rating with and without low – level constraints are predicted separately and finally the each product quality is known using the new ratings predicted. The new ratings are predicted using the entropy and the low-level constraints like, sentimental user confidence ratings and the user confidence ratings using the temporal features. Gaussian curve fitting values are calculated for predictingthelow-levelconstraintsvalues.Theproductquality is finalized by calculating the differences of overall user rating confidence of with and without low-level constraints values respectively. This prediction and product quality determines the attitude of the users. Instead of using the actual ratings given by the users, the overall user confidence ratings are considered, since there are few contradictions between the reviews and its star ratings or the consumers thoughts may differ. The accuracy and the misclassification of the proposed system are evaluated. The proposed system shows better performance than the existing system. The four constituents of the framework are the Data Preprocessing, User Confidence Rating with Entropy values, Low-Level constraints calculation with sentiment and temporal features and finally the overall user confidence ratings prediction along with overall item ratingpredictionmoduleusingtheentropy,low-levelpredicted constraints.Theproposedapproachaddressestheissuesspecial symbols and URLs, sparse values, accurate overall user confidence ratings using the low-level constraints. The effectiveness of our approach is verified by various product datasets. Experimental results shows good performance of the proposed approach compared to existingapproach.

REFERENCES

[1] Bell R and Volinsky C, "Matrix Factorization Techniques for Recommender Systems", IEEE Transaction in Knowledge Management, Vol.1, No.3,pp.42-50,2009.

[2] Chelaru S., Altingovde I.S., Siersdorfer S. and Nejdl W. "Analyzing, Detecting, and Exploiting Sentiment in Web Queries", ACM Transactions on the Web, Vol.8, No.1,pp.6(1-6),2013.

[3] Dror G., Koenigstein N. and Koren Y. "Yahoo! Music Recommendations: Modeling Music Ratings with Temporal Dynamics and Item Taxonomy", ACM RecommenderSystems,pp.1-8,(2016).

[4] Lei X. and Qian X. ,"Rating Prediction via Exploring Service Reputation", IEEE 17th International Workshop on Multimedia SignalProcessing,pp.1-6,2015.

[5] Lei X., Qian X. and Zhao G. "Rating Prediction based on Social Sentiment from Textual Reviews", IEEE Transaction on Multimedia, Vol.10, No.2,pp.1-12,2016.

[6] Tan S., Li Y., Sun H., Guan Z., Yan X. and Bu J. "Interpreting the Public Sentiment Variations on Twitter", IEEE Transaction on Knowledge and Data Engineering,Vol.26,No.5,pp.1158-1170,2014.

[7] Yang D. and Zhang D. "Modeling User Activity Preference by Leveraging User Spatial Temporal Characteristics in LBSNs", IEEE Transactions on Systems, Man AND Cyber net Systems, pp.1-14,2014.

[8] Yang X., Zhang T. and Xu C. "Cross-Domain Feature Learning in Multimedia", IEEE Transactions on Multimedia, Vol.17, No.1, pp.64- 78,2015.

[9] Zhang W. and Ding G. "Generating Virtual Ratings from Chinese Reviews to Augment Online Recommendations", ACM Transactions on Intelligent Systems and Technology,Vol.4,pp.9(1-17),2013.

[10] Zhao G. and Qian X. ,"Service Objective Evaluation via Exploring Social Users' Rating Behaviors", Proc in IEEE International Conference on Multimedia Big Data, pp.228-235,2015.

[11] Zhao G. and Qian X. , "User-Service Rating Prediction by Exploring Social Users' Rating Behaviors", IEEE Transactions on Multimedia, Vol.17,pp.1-22,2016.