Postprint of Review Information Recommendation Integrating User Interest and Review Utility

Authors: Nie Hui, Qiu Yifei

Date: 2023-04-01T16:02:49+00:00

Abstract

[Purpose/Significance] In the Web 2.0 era, the problems of uneven quality and overload of online reviews are extremely serious, and the cognitive cost for people to extract valuable content therefrom is increasingly high. This study explores effective solutions to address review overload through information recommendation approaches, aiming to enhance network information utilization and information service quality. The comment ranking and recommendation scheme proposed in this paper focuses on comment information quality and places greater emphasis on satisfying users' personal information needs. [Methods/Process] The research employs probabilistic topic models and introduces word vectors to construct user models and comment models within a topic space. By incorporating these into a comment perceived utility evaluation system, it achieves comment recommendation that integrates user interests and comment quality. The recommendation effectiveness is tested through systematic experiments. [Results/Conclusion] Experimental results demonstrate that both comment information quality and users' individual information needs jointly influence user satisfaction with comment perceived utility; the recommendation strategy achieves an organic integration of the two. Evaluation results under three different recommendation modes show that, compared with pure "interest recommendation" and "utility recommendation", the "fusion recommendation" achieves the highest comprehensive satisfaction score.

Full Text

Preamble

Volume 65, Issue 10, May 2021
ChinaXiv Cooperative Journal

Integrating User Interests and Review Utility for Review Information Recommendation

Nie Hui, Qiu Yifei
School of Information Management, Sun Yat-sen University, Guangzhou 510275

Abstract:
[Objective/Significance] In the Web 2.0 era, the quality of online reviews is uneven and information overload is severe, leading to increasingly high cognitive costs for users seeking valuable content. This study explores effective solutions to review overload through information recommendation to enhance network information utilization and information service quality. The proposed review ranking and recommendation scheme focuses on review information quality while emphasizing satisfaction of users' personal information needs. [Method/Process] This research employs a probabilistic topic model and introduces word vectors to construct user and review models within a topic space. By incorporating these into a review perceived utility evaluation system, the study achieves personalized review recommendation that integrates user interests and review quality, with recommendation effectiveness tested through systematic experiments. [Result/Conclusion] Experimental results demonstrate that both review information quality and individual users' information needs jointly influence user satisfaction with review perceived utility. The recommendation strategy achieves organic integration of these two factors. Evaluation results under three different recommendation modes show that the "combined recommendation" approach yields the highest comprehensive satisfaction scores compared to pure "interest recommendation" and "utility recommendation."

Keywords: information recommendation; review utility prediction; user modeling; online reviews
Classification Number: P393.1 G202
DOI: 10.13266/j.issn.0252-3116.2021.10.008

In the mobile internet era, sharing opinions and posting reviews online has become part of daily life. The valuable content provided by online reviews helps people eliminate uncertainty in decision-making processes and profoundly influences individual behavior. However, in an environment where "everyone is a creator," problems such as review proliferation and uneven quality have become increasingly serious, raising the cognitive cost of obtaining valuable content. To reduce users' cognitive burden and improve information service quality, various websites have implemented information filtering mechanisms. "Taobao" uses criteria such as whether reviews contain images, whether they are follow-up reviews, and product rating levels as filtering standards; "Dianping" blocks untrustworthy content based on user feedback; platforms like "Douban" and "Amazon" employ user voting to rank reviews. These filtering strategies primarily target information quality, helping users quickly access high-quality information by placing top reviews prominently. Nevertheless, these approaches fail to address individual user needs.

Information adoption by individuals is influenced not only by information quality but also by personal information needs—people pay more attention to whether received information contains content of interest to them. Particularly when information volume exceeds cognitive load, users browsing quickly hope to find content they care about as soon as possible. Therefore, studying information utility value with the goal of reducing cognitive burden and improving information retrieval quality should incorporate user individual factors. Personalized information recommendation represents an important approach to solving information overload. This paper focuses on this research topic, exploring effective solutions to integrate users' personal information needs into review utility prediction models to construct a personalized review recommendation system that incorporates both personal interests and review quality. The study examines practical strategies for information service precision from a recommendation perspective, aiming to provide references for controlling review information proliferation and improving information service quality and utilization value.

2 Related Research

2.1 Review Ranking and Recommendation Based on Review Utility

The essence of review ranking is evaluating review utility to generate Top-N recommendation lists. In recent research, Guo Shunli et al. [?] employed fuzzy analytic hierarchy process and weighted grey relational analysis to predict review utility for ranking and selected reviews with high information volume for final recommendation. Zhang Yanfeng et al. [?] used K-means algorithm to classify review utility levels and subsequently optimize review sorting. Wang Zhongqun et al. [?] calculated review credibility based on the number of "feature-opinion" pairs in reviews for ranking, then invited users to evaluate Top-N reviews through questionnaires. Wu et al. [?] argued that reviewers' historical reviews could reflect their comment quality; they modeled reviewers based on previously published reviews and integrated this into review models, finding that models incorporating reviewer information better predicted review utility. These studies demonstrate that review ranking and recommendation primarily rely on calculated evaluation metrics, which consider information volume, content, credibility, reviewer expertise, and overall perceived utility by reading groups—all crucial for identifying high-quality reviews.

However, a recent research report points out that these evaluation metrics only reflect review information quality in terms of data reliability without emphasizing the suitability of review information for target users [?]. Researchers argue that evaluating perceived utility of online reviews is an information quality assessment based on user perspective, starting from users' subjective perceptions to explore information utility, requiring individuals to systematically evaluate information functional performance based on personal experience. Therefore, online user reviews should not only become high-quality information meeting standards but also focus on the degree to which review information satisfies user needs and expectations and the value it brings to users [?]. Many researchers share this view. Li Jian et al. [?] explored product recommendation, arguing that review valence should incorporate consumer individual preferences to find high-quality reviews matching consumer personal preferences. E. Ben-abdallah et al. [?] analyzed cloud service review quality across different platforms, achieving review recommendation by calculating similarity between reviewer personal information and cloud service platform information seekers' background information. These studies all examine review perceived value from a personalized perspective.

2.2 Review-Based Recommendation System Research

2.2.1 Classification of Review-Based Recommendation Systems Recommendation is an effective method for solving information overload. By exploring user information needs, recommendation systems can achieve personalized information push oriented toward individual interests, alleviating troubles caused by overloaded information [?]. The core of product recommendation systems is building effective user and product models. As review information is rich in user product evaluations, extracting user preferences from reviews to build user models and introducing them into recommendation systems has become a research hotspot in recent years. L. Chen et al. [?] categorized related research from the perspective of user modeling into three types: term-based recommendation, rating-based recommendation, and feature-based recommendation.

Term-based recommendation belongs to content recommendation, directly using review text to model users and products. For example, S. G. Esparza et al. [?] extracted terms from user-published reviews, used TF-IDF as term weights to generate user models, while product models were based on target product review sets, finally recommending based on content similarity between the two. Geng Lixiao et al. [?] built a literature recommendation system that modeled users based on literature they had read, used word vectors to represent terms, and elevated similarity calculation between users and recommendation targets (literature) to the semantic level.

Rating-based recommendation employs collaborative recommendation mechanisms requiring generation of "user-rating" matrices, but matrix sparsity has long been a bottleneck for improving collaborative recommendation system performance. One solution is using review text data to predict user product ratings, thereby improving the "user-rating" matrix and enhancing system performance. W. Zhang et al. [?] used sentiment analysis to predict user product ratings for product reviews, building a user model based on "predicted ratings" for product recommendation. C. Musat et al. [?] weighted user ratings by condensing product topic information contained in reviews, further improving model quality. Zhang Yihao et al. [?] proposed a hybrid recommendation method integrating user ratings, sentiment orientation, and product content, filling and correcting the sparse "user-rating" matrix to rank and recommend products.

Feature-based recommendation focuses on details of recommended objects (products) mentioned in review content. User preference descriptions point directly to product features, representing user modeling based on product feature evaluation. In H. Liu et al.'s [?] research, user interests involved two indicators: attention degree and demand degree, with attention degree related to product feature mention rate and demand degree related to product evaluation. H. Liu et al. extracted product feature words and evaluation words from review text, built feature-level user interest models, and achieved more accurate product recommendation effects. Zhang Yanliang et al. [?] focused on changes in user interests in research to improve user modeling, using user-published reviews to mine user evaluations of various product features and regularly updating reviews to predict user interests and changing trends. Feature-based recommendation also employs collaborative recommendation, requiring fine-grained analysis of review text.

2.2.2 Review-Based Recommendation System Modeling Methods Product reviews have become an important data source for recommendation system modeling. Extracting product features users care about from review content, conducting sentiment analysis, and predicting user product ratings are issues involved in related research. User models and product models mainly adopt implicit feature vectors, with probabilistic topic models and deep learning methods widely used for modeling. Y. Bao et al. [?] used non-negative matrix factorization to obtain implicit topics in review text, reflecting user preferences and product characteristics through topic distribution. S. Feng et al. [?] proposed a hybrid modeling approach combining probabilistic topics and random walks, where the probabilistic model was responsible for mining user latent preferences and product implicit features, and random walks were used to build global latent associations. Probabilistic topic models can extract semantic features of text and alleviate problems caused by sparse high-dimensional matrices through dimensionality reduction, but the shortcoming is ignoring important context information.

Deep learning models can effectively retain word order information, better extract text features, and improve modeling quality. L. Zheng et al. [?] proposed the DeepCoNN model, which used two parallel convolutional neural network models (CNN) to learn implicit vector representations of users and products from reviews, using a rating prediction function as the loss function for iterative training to predict user product ratings. S. Seo et al. [?] further introduced an attention mechanism (ATT) on the basis of CNN to model the correlation between different parts of reviews and user preferences and product features, optimizing the model. C. Chen et al. [?] introduced both review text and rating matrices simultaneously into modeling to make full use of information contained in users and products in the rating matrix. Feng Xingjie et al. [?] borrowed C. Chen's approach, replacing static word vectors with dynamically fine-tunable Bert pre-training models, combining bidirectional GRU structures and ATT mechanisms to extract deep feature vectors of users and products from user reviews and product reviews respectively, further improving recommendation algorithm prediction accuracy.

In summary, recent research on recommendation systems has emphasized online reviews as an important information source for mining user interest preferences. Using user reviews to generate user models or predict user product ratings to improve the "user-rating" matrix for collaborative recommendation strategies has been widely adopted. User and product models learned through review text are represented in implicit vector form, with probabilistic topic models and deep learning algorithms widely used to improve modeling quality.

3 Research Design

This study proposes a ranking and recommendation strategy that integrates user interests and review information quality, starting from meeting users' personalized information needs and exploring effective methods to solve review information overload through information recommendation. The research posits that review utility lies in providing reference for user decision-making, and truly meaningful decision-making help should be oriented toward users' personal needs and preferences. Individuals' perceived utility of reviews varies, and review recommendation should pay more attention to users' personalized information needs. Li Jian et al. [?] deeply explored this issue from a theoretical perspective; Zhu Linlin et al. [?] conducted empirical research through questionnaire data based on statistical analysis methods. E. Ben-abdallah et al. [?] used similarity between information seekers and review authors as the basis for information recommendation but did not directly analyze target reviews. This paper employs review mining methods to achieve personalized review information recommendation by building a personalized review perceived utility evaluation model.

This research selects probabilistic topic models to condense product features in review text, generating feature description frameworks for users and recommendation objects (reviews). Review text is directly represented by implicit topic vectors. User interest models first obtain interested feature words through interaction, then map them to implicit topics. The entire study revolves around three questions: How to effectively model recommended reviews and user interests? How to build a review perceived utility prediction model that incorporates user interests? How to test recommendation strategy effectiveness and verify the scientificity and rationality of the recommendation scheme?

The study is divided into five modules, with the research architecture shown in Figure 1 [Figure 1: see original paper]. Data preprocessing: Clean and preprocess original review corpora to build review corpora. Build review r's topic model: Explore reasonable schemes for building review models. Introduce feature dictionaries and feature synonym dictionaries, use topic clustering to generate description frameworks for review content, and build review models r.topic_{profile} under topic space by mining product concerns. Build user u's interest model: Describe user interests based on topic models, introduce word vectors to achieve mapping from feature word sequence user model u.feature_{profile} to topic model u.topic_{profile}. Recommendation prediction model construction: Build a review perceived utility prediction model incorporating user interests and design review ranking recommendation strategies. Recommendation effect evaluation: Develop an online evaluation platform, adopt an interactive approach of "user selection-review push-user feedback" to obtain user ratings of various recommendation schemes, and test the actual effect of personalized review recommendation strategies. Among the five modules, and correspond to review modeling and user modeling, which are core components of the personalized review recommendation system; and introduce user factors into the review perceived utility prediction model, focusing on exploring and verifying personalized review ranking recommendation strategies and their implementation.

4 Model and Methods

4.1 Topic Modeling

4.1.1 LDA Implicit Topic Model The primary task is to model reviews and users under the same description framework, for which the LDA implicit topic model was selected. LDA is a multi-level generative probabilistic model with three layers: terms, topics, and documents. In the model, D (|D| = N) is the document set, W (|W| = M) is the term set, and ϕ ($|\phi|$ = K) represents implicit topics. Document d D is generated by random mixing of implicit topics, generally represented as a sequence pattern on W: d = {wi | wi W}; implicit topic ϕ is a polynomial distribution on W. α and β are important parameters of the LDA model. α reflects the relative strength of implicit topics in document set D, related to the prior probability distribution of topics Dirichlet; β characterizes the probability distribution of topics themselves. The topic layer ϕ is the "topic-term" distribution, parameterized by a K×M matrix β; the document layer characterizes the "document-topic" distribution, determined by parameter α; the term layer zi (i = 1, 2, …, M) represents the topic component assigned to each term in document d, obeying the polynomial distribution of d. With parameters α and implicit topic space ϕ set, d is generated by two processes: Randomly select a K-dimensional vector d from Dirichlet($|\alpha$) to generate document d's topic distribution. Generate document d's terms according to conditional probability p(wi| d, $ $1..K). The core of the LDA model is inferring the "document-topic" distribution α and "topic-term" distribution ϕ. Under the premise of known term distribution in the document set, zi (i = 1, 2, …, M) is derived in reverse to obtain and ϕ. The model involves multiple unknowns and generally uses approximate solutions, such as the commonly used Expectation Maximization (EM) algorithm and Gibbs sampling. The EM algorithm [?] achieves stability in maximum likelihood estimation through continuous iteration, realizing parameter inference by solving the maximum likelihood function. The LDA model in this study adopts the EM algorithm.

4.1.2 Topic Model Based on Standardized Features This study uses LDA to condense review content and generate a topic-based feature description framework. According to conventional topic modeling practice, nouns and verbs (substantive words) are generally extracted to describe documents (hereinafter referred to as "noun + verb modeling"). However, subsequent research found that topic semantics clustered based on substantive words were vague, especially for specific domains. Therefore, domain feature dictionaries were introduced, using only feature words for clustering. This optimization strategy enhances document topic representation, reduces feature space dimensionality, and improves computational efficiency. This scheme is named "feature word modeling."

However, "feature word modeling" remains a modeling approach based on independent terms. Synonyms such as "screen," "touchscreen," "display," and "large screen" are treated as different terms, unable to identify synonym relationships. Based on "feature words," this study standardizes synonyms, replacing synonyms like "touchscreen, display, large screen" with "screen." This is equivalent to semantic integration of document content. For example, if a review's feature word description is d = {screen, size, color screen, cost-effectiveness, price…}, after standardization it becomes d = {screen, size, screen, price, price…}. After integration, document expression rises to the conceptual level. This is named "synonymous feature word standardization" modeling. The study designed experiments to compare the application effects of three modeling schemes: "noun + verb," "feature word," and "synonymous feature word standardization," selecting the optimal scheme as the description framework for review and user models.

4.2 Construction of Review and User Models in Topic Space

4.2.1 Review Model During LDA topic modeling, the "document-topic" probability distribution matrix is simultaneously obtained. We represent this as = Review_{MAX}i×k, where I corresponds to the number of documents in the review corpus and K is the number of topics. The row vector of Review_{MAX}i×k is the probability distribution description of review r in topic space, as shown in Formula (1). The review model r.topic_{profile} is a K-dimensional vector representing the probability distribution of review r across K topics.

Formula 1:
r.topic_{profile} = [p1, p2, …, pk]
Formula (1)

4.2.2 User Interest Model The user model is also built on the implicit topic space. First, a set of product feature words Interest_{set} describing user interests is established, from which users select terms of interest. The algorithm maps the selected term sequence to the implicit topic space. The modeling process consists of three steps:

Step 1: Set Interest_{set} and generate user interest description based on feature words according to user selection. Based on LDA clustering results and referencing e-commerce platform classifications of mobile phone features, feature words describing mobile phone performance are divided into 8 topics: "screen effect, network signal, appearance design, photography, entertainment, running performance, cost-effectiveness, battery life." Users select features they care about. For example, if user u cares about "appearance" and "battery performance," the corresponding topic description words characterize u, yielding u.feature_{profile} = {battery, life, appearance, exterior, screen, body, size, …}. The formal expression is Formula (2), where Topic(f) corresponds to the topic word set under user interest topic f. The u.feature_{profile} is then mapped to the LDA implicit topic space.

u.feature_{profile} = {ti | ti   Topic(f), f   Interest_{set}, i = 1, 2, …, m}
Formula (2)

Step 2: Word vector expression of user interest. Word vectors are distributed representations of words learned through shallow neural networks, representing terms as N-dimensional high-density real number vectors where each term corresponds to a point in N-dimensional space, with distances between points reflecting potential semantic relationships between terms. Before mapping the feature word-based user interest to topic space, word vectors are introduced to first convert u.feature_{profile} into a word vector matrix u.vec_{MAX}m×v (v is word vector dimension). The user interest model based on word vectors can convey semantics and improve recommendation accuracy. Using the matrix representation u.vec_{MAX}m×v also facilitates mapping the user model to topic space. The word vectors introduced in the study are an open-source Chinese pre-training model from Beijing Normal University [**?**]. The training corpus for these word vectors is "Baidu Baike" with a corpus size of 4.1G and vector space dimension of 300.

Step 3: User interest model in topic space. The topic t is expressed through the "topic-term" probability distribution generated by LDA clustering, as shown in Formula (3):

t.feature_{profile} = {<fi, wi>, i = 1, 2, …, n}
Formula (3)

where fi is a feature word describing topic t, wi is the weight of fi, and n is the number of feature words. Correspondingly, the word vector matrix t.vec_{MAX}n×v for topic t is established. In the word vector space, multiply u's interest matrix with the transpose matrix of topic t, while incorporating the topic feature word weight matrix W_n×v = [w1, w2, …, wn]^T. Finally, the maximum value of the matrix operation result is taken as the semantic relevance between u and t (see Formula 4). Calculate the relevance between user u and K topics according to Formula 4 to generate the user interest model in topic space, as shown in Formula 5.

Simt = Max(u.vec_{MAX}m×v × t.vec_{MAX}n×v W_n×l)

Formula (4)

$$u.topic_{profile} = [Sim1, Sim2, \ldots, SimK]$$
Formula (5)

4.3 Review Ranking and Recommendation Integrating User Interests

Based on the construction of user models and review models, user interest factors are introduced into the recommendation model. Three recommendation modes are proposed: recommendation based on user interest (hereinafter "interest recommendation"), recommendation based on review perceived utility (hereinafter "utility recommendation"), and recommendation integrating user interest and review perceived utility (hereinafter "combined recommendation").

4.3.1 Interest Recommendation The essence of content-based text information recommendation is calculating the relevance between content and user interest, and sorting recommended information by relevance for Top-N ranking. This study uses cosine distance to measure relevance, letting $\hat{u} = u.topic_{profile}$ and $\hat{r} = r.topic_{profile}$, calculated as shown in Formula (6). $Pref_{score}$ is named "interest score."

$$Pref_{score} = \hat{u} \cdot \hat{r} / (\|\hat{u}\| \times \|\hat{r}\|)$$
Formula (6)

4.3.2 Utility Recommendation The ranking metric for utility recommendation directly adopts the review perceived utility value based on website user voting, as shown in Formula (7). $Helpfulness_{score}$ is hereinafter referred to as "utility score."

$$Helpfulness_{score} = \ln(Num_{{{of}}{\mathit{helpful}}}{votes}/Total_{votes})$$
Formula (7)

4.3.3 Combined Recommendation The combined scheme incorporates both interest score and utility score into the calculation formula, i.e., a recommendation strategy integrating user interest and review quality factors. The comprehensive evaluation indicator $Combined_{score}$ is named "combined score" as shown in Formula (8).

$$Combined_{score} = wp \cdot Pref_{score} + wH \cdot Helpfulness_{score}$$
Formula (8)

In Formula (8), wp and wH are weights for interest score and utility score respectively, satisfying $wp + wH = 1$. In specific applications, the two weights are set as parameters, allowing flexible configuration of the combined recommendation scheme. This framework also incorporates interest recommendation and utility recommendation into the same framework.

5 Experiments and Analysis

The experimental corpus is sourced from "ZOL Online" (http://www.zol.com.cn). An LDA model is constructed to condense product concerns in the corpus, generating a resource description framework based on clustering topics, i.e., establishing an implicit topic space. User interests are obtained through a developed data collection platform, and the application effect of the combined recommendation mode is tested through questionnaire feedback.

5.1 Data and Resource Construction

5.1.1 Review Data and Related Resources The corpus collection time was November 2018. The main extracted content included review text and review perceived utility, totaling 14,506 reviews. Related resources also include a mobile phone feature word table and synonym table, both outcomes of previous research. The feature word table summarizes 301 mobile phone feature words, and the synonym table contains 808 terms.

5.1.2 User Interest Data and Evaluation Data A user data collection and evaluation platform was developed according to the research design. The evaluation page lists a preset user interest word set. Invited users independently select product features they care about. The system builds user models based on selections, generates review recommendation lists, and pushes reviews to invited users. Users must evaluate reviews pushed by different schemes across two metrics: information appropriateness and information quality. Appropriateness measures satisfaction regarding individual needs, while information quality evaluates satisfaction with review content quality and format. Evaluation uses a five-point scale, with higher scores indicating higher satisfaction.

Research team members sent evaluation system links through various channels. Invitees entered the system through links, selected interest feature words according to instructions, and rated pushed reviews. The system backend aggregated and analyzed evaluation data. The entire evaluation experiment lasted one month (November 2019 to December 2019), during which platform functions were corrected and improved multiple times, ultimately obtaining 208 valid user evaluation data sets. A screenshot of the evaluation platform interface is shown in Figure 2 [Figure 2: see original paper].

5.2 Experiments and Results

Three experiments were conducted to address the proposed questions (see Section 3 Research Design). Experiment 1: Topic modeling, extracting product feature topics from review corpora, comparing LDA clustering effects of three modeling schemes ("noun + verb," "feature word," and "synonymous feature word standardization") to establish the model description framework. Experiment 2: User modeling, constructing user models under the LDA implicit topic space. Experiment 3: Review recommendation, comparing actual effects of three recommendation strategies ("interest," "utility," and "combined") to test the influence of individual factors on review perceived value. Experiments were implemented in Python, using jieba for Chinese preprocessing, sklearn's LatentDirichletAllocation class for LDA modeling, and pyLDAvis for clustering visualization.

5.2.1 LDA Topic Modeling Experiment and Results

(1) Model Parameters and Evaluation Metrics. To obtain an optimal model, the experiment first considered evaluation metrics and parameter settings for topic clustering. For LDA, when the average similarity between topics is minimized, the clustering structure is most stable and the model is optimal. Average similarity between topics can measure model stability. This study uses the mean of cosine distances between clustering topics, Avg_{similarity}, to measure clustering structure, where smaller values indicate more stable structure. Another metric is KL divergence (Kullback-Leibler divergence), calculated as shown in Formulas (9) and (10).

KL(p‖q) = Σp(xi)log(p(xi)/q(xi))
Formula (9)

KL(p,q) = KL(p‖q) + KL(q‖p)/2
Formula (10)

KL(p‖q) is the KL divergence between topic p distribution and topic q distribution, where p(xi) is the probability distribution of terms in p. Since KL divergence is asymmetric, the mean of KL(p‖q) and KL(q‖p) is generally taken as the KL divergence between p and q. KL divergence measures distribution differences, with larger values indicating greater clustering distinction and better clustering structure.

(2) Model Parameter Setting. For the LDA model, the number of topics K is a crucial parameter related to model parameters α and β. The study treats K as an optimization parameter determined through experiments. Figure 3 [Figure 3: see original paper] shows clustering effects for three modeling schemes with different K values. Overall, as K increases, Avg_{similarity} shows a downward trend, indicating decreasing similarity between topics and enhanced clustering structure stability. Conversely, KL divergence gradually increases, indicating larger differences between topics and enhanced internal cohesion. As K increases, both metrics gradually converge. For the three modeling schemes, both metrics show that the clustering effect of "synonymous feature word standardization" is significantly better than "noun + verb" and "feature word." This indicates that introducing feature dictionaries effectively reduces noise word interference, while feature word standardization further improves matrix sparsity, condenses topic content, and yields the best clustering model quality. Therefore,

subsequent experiments adopt the "synonymous feature word standardization" topic clustering scheme.

According to experimental results (see Figure 4 [Figure 4: see original paper]), both evaluation metrics begin to flatten at K = 13, with KL divergence = 8.267 and $Avg_{similarity}$ = 0.05. K = 13 is ultimately determined.

(3) Clustering Results. Figure 5 [Figure 5: see original paper] shows the topic clustering visualization effect when K = 13, generated by pyLDAvis. Overall observation shows balanced distribution across topics, with clear distinction for most topics and overlap for a few (topics 4 and 5, topics 1 and 2). Overlap can lead to vague topic semantics, so the following processing is applied: for each clustering topic, topic words are sorted by term distribution probability from large to small, and the top 8 terms are taken to describe topic semantics. If a term appears under multiple topics simultaneously, it is assigned to the clustering topic with the highest weight. For example, "battery capacity" appears under both topic 4 and topic 12, but its weight under topic 12 (0.052) is higher than under topic 4 (0.019), so it is assigned to topic 12. After adjusting clustering topic words, topic meanings can be better clarified. According to each topic's word list and referencing digital website settings for mobile phone feature indicators, the 13 topics correspond to 9 feature categories: "running performance, screen effect, network signal, appearance design, photography, entertainment, cost-effectiveness, battery life, and others," based on which the user interest selection feature word set $Interest_{set}$ is generated for user modeling.

5.2.2 User Interest Modeling Experiment Using the method in Section 4.2.2, users select interested product feature words from $Interest_{set}$. The algorithm first generates $u.feature_{profile}$ based on user selection, then maps $u.feature_{profile}$ to the topic space to generate $u.topic_{profile}$. Table 1 shows a user modeling example. Mapping the feature word sequence user model to topic space is equivalent to semantic expansion of feature words describing user interests, making user information need expression more complete, enabling matching with more reviews related to user interests, and improving recall. The recommendation effect of this model is tested in subsequent experiments.

5.2.3 Review Recommendation Experiment Integrating User Interests

(1) Analysis of Review Ranking Sequence Differences. The experiment first analyzes differences in review ranking under three recommendation modes. Setting user-selected interest feature words (e.g., "screen effect"), three recommendation modes are used to rank reviews for four mobile phones respectively. Friedman tests are conducted on the three ranking sequences for each product's reviews, as shown in Table 2 . Results show that for the three recommendation modes, the probability p of Friedman statistics

for each product's data is less than 0.05, indicating significant differences in review ranking under the three modes. That is, with the same interest word, the three recommendation schemes produce different ranking results. According to average ranks, except for product C's reviews, the "combined ranking" mode considering both user individual factors and review quality shows obvious differences from the "utility ranking" mode considering only review quality, indicating that user individual interest preferences have certain influence on review perceived value.

(2) User Satisfaction Analysis of Three Recommendation Strategies. An evaluation platform was built according to Section 5.1.2 (see Figure 2 [Figure 2: see original paper]), and links were sent to invite users to participate in the experiment. Invitees selected interested product feature words under anonymous product conditions. The platform pushed a set of reviews based on their selection. To ensure data quality, each recommendation scheme pushed three reviews according to ranking gradients (reviews ranked 1, 4, and 7 were pushed). To shield the influence of review length on review perceived utility, pushed review length was controlled between 30-460 characters. Evaluators read reviews, rated their appropriateness and quality, and submitted evaluation results. The experiment tested the influence of user individual factors on review perceived utility, with the expectation that recommendation schemes incorporating user interest factors would achieve higher satisfaction than pure "utility recommendation" schemes, and "combined recommendation" would achieve the highest overall satisfaction as it simultaneously considers information quality. Statistical results of user rating data are shown in Figure 6 [Figure 6: see original paper].

As shown in Figure 6, for information appropriateness (i.e., meeting user information needs), "combined recommendation" received the highest score; for information quality, "combined recommendation" scored basically the same as "utility recommendation." This recommendation mode can still ensure pushed review quality. Further mean difference significance tests show (see Table 4) that in terms of information appropriateness, there are significant differences between "combined recommendation" and "utility recommendation" ($t = -0.084$, $p < 0.95$) and between "combined recommendation" and "interest recommendation" ($t = -2.228$, $p < 0.05$), verifying the effectiveness of the "combined recommendation" scheme in meeting user information needs.

Review quality test results (see Table 5) show significant differences between "combined recommendation" and "interest recommendation" schemes ($t = -2.2634$, $p < 0.05$), but no significant difference between "combined recommendation" and "utility recommendation." This further indicates that the "combined recommendation" strategy performs well in information quality, achieving the same effect as the "utility recommendation" scheme. However, pure "interest" recommendation without considering information quality has limited recommendation effect.

(3) Case Analysis of Recommendation Effect. Based on the above experimental results, the "combined recommendation" scheme with the highest score is analyzed in detail. In terms of information appropriateness, "combined recommendation" significantly improved user ratings compared to "utility recommendation" (3.149 → 3.377). The reason is that "combined recommendation" considers semantic matching of user interest features, and these reviews contain more evaluation content about product features users are interested in, improving information appropriateness. Compared with "interest recommendation," "combined recommendation" effectively filters out reviews with high interest feature mention rates but low information volume, ensuring information quality to some extent, thus performing best among the three recommendation schemes. Recommendation examples are shown in Tables 6 , 7 , and 8 .

Taking Table 6 as an example, where the user selects "screen effect" as the interest word, the review pushed by "utility recommendation" contains no relevant description of "screen features"; the review pushed by "interest recommendation" contains related features such as "screen," "size," and "6 inches," achieving high appropriateness but not high quality rating due to limited content, indicating that information volume is an important factor affecting review perceived utility; the review pushed by "combined recommendation" considers quality factors, emphasizing mention rates of relevant interest features. Tables 6 and 7 both show that reviews pushed by this scheme are superior to the former two in both content richness and satisfaction of personalized information needs. In terms of review quality, "combined recommendation" significantly improved compared to "interest recommendation" (2.992 → 3.139) because "combined recommendation" excludes reviews with low information volume and utility, ensuring pushed information quality.

6 Research Conclusions and Discussion

6.1 Research Conclusions

This study explores effective methods to solve review information overload through information recommendation. Using probabilistic topic models, user interest models are built in topic space and incorporated into review perceived value calculation models. Based on these models, a review recommendation strategy integrating user interests and review utility is proposed, and the effectiveness of the recommendation strategy is tested through an online evaluation system. The entire research focuses on personalized review information recommendation, with contributions mainly reflected in four aspects:

(1) This study constructs topic models describing review resources and user interests. The modeling algorithm introduces domain feature dictionaries and feature synonym dictionaries, using standardized review feature words to describe review documents. This optimization strategy integrates synonymous terms, reduces document feature dimensions, improves matrix

sparsity, and generates topic models with good interpretability and stable structure, which are well used for review and user interest modeling.

(2) This study builds user interest models at the semantic level. User modeling introduces word vectors to semantically represent feature word sequence-based user models, then maps semantically represented user models to topic space through matrix operations. This modeling scheme elevates user interest expression to semantic space.

(3) This study introduces personalized user factors into review evaluation mechanisms and proposes a review recommendation strategy integrating user interests and review utility. Based on user models and review models, review interest scores are predicted for users through semantic calculation, while review perceived utility measuring review quality is introduced. Combining review interest scores and utility scores, a review ranking and recommendation strategy under the joint action of the two factors is proposed.

(4) This study develops a platform to evaluate recommendation effects, comparing user evaluations of pushed reviews under three schemes: "interest recommendation," "utility recommendation," and "combined recommendation." Results show that the "interest recommendation" scheme can better achieve personalized review recommendation, and the "combined recommendation" scheme can meet users' personalized information needs while ensuring information quality, achieving the best comprehensive performance. This indicates that both review information quality and users' personalized information needs jointly influence users' perceived satisfaction with review information. The review recommendation strategy proposed in this paper achieves organic integration of the two factors. From an information service perspective, "combined recommendation" is undoubtedly a better review recommendation mode.

6.2 Practical Implications

Review ranking and recommendation based on review perceived utility starts from group perception without focusing on individual user information needs. This study verifies the significant influence of individual user factors on review perceived value and proposes a review recommendation scheme integrating user interests. This research can help internet enterprises better build review quality evaluation systems and provide theoretical references and practical suggestions for enterprises to monitor and effectively utilize network information. For example: Given the significant role of information quality and personalized factors in review perceived value, an information recommendation approach integrating user interests should be adopted to build an information filtering system based on user needs while ensuring information quality, optimizing information service models. For target users, businesses should screen reviews with both high perceived utility and satisfaction of their information needs for presentation, which can improve users' information adoption rate, promote purchase behavior, and thus positively impact business operations. For review information services and information platforms in other fields, introducing retrieval mechanisms to obtain user information needs and building complete information recommendation systems to achieve precise information recommendation can effectively alleviate troubles caused by information overload, improve information service quality, and promote healthy development of information service communities.

6.3 Research Limitations and Future Directions

The modeling and recommendation scheme design in this study is relatively simple. For example, for user models, feature words characterizing user interests are processed with equal weights. However, testing found that users have different focuses on product performance, and subsequent research could set weights for feature words describing user interests to build more refined user interest models. Similarly, in the combined recommendation scheme, the influence of review quality factors and personalized factors is also treated as equal, but experimental results show that recommendation results differ significantly under different weight settings, and the influence of the two factors varies for different types of reviews. This issue requires collecting more types of corpora for further exploration. Future research also plans to introduce deep learning algorithms for in-depth exploration of user modeling, extracting user features from user reviews to improve personalized recommendation algorithms.

References

[1] Guo Shunli, Zhang Xiangxian, Li Zhongmei. Research on mobile O2O online review usefulness ranking model oriented to user information needs—Taking Meituan as an example [J]. Library and Information Service, 2015, 59(23): 85-93.

[2] Zhang Yanfeng, Li He, Peng Lihui, Hou Litie. Online review usefulness classification algorithm and application based on emotional semantic feature extraction [J]. Data Analysis and Knowledge Discovery, 2017, 1(12): 74-83.

[3] Wang Zhongqun, Wu Dongsheng, Jiang Sheng, et al. Research on review credibility ranking based on mainstream feature-opinion pairs [J]. Data Analysis and Knowledge Discovery, 2017, 1(10): 32-40.

[4] Wu , Wang Zhongqing, Zhou Xiabing, et al. Online review quality detection based on text and user information [J]. Journal of Chinese Information Processing, 2019, 33(9): 107-114, 140.

[5] Zhu Linlin, Li He, Liu Jincheng, et al. Research on construction of online review information quality perception evaluation index system [EB/OL]. [2020-12-12]. https://kns.cnki.net/KXReader/Detail?TIMESTAMP=637449586862656250&DBCODE=CJFQ&TABLE EKEjGs%3d.

[6] Li Jian, Zhang Jun, Yuan Qingmin, et al. Improved analysis of online product reviews' utility to consumers—From the perspective of information quality and consumer satisfaction theory [J]. Information Science, 2018, 36(7): 137-144.

[7] BEN-ABDALLAH E, BOUKADI K, HAMMAMI M. Personalized cloud service review ranking approach based on probabilistic ontology [C]//Proceedings of 22nd international conference on business information systems. Seville: Springer, 2019: 50-61.

[8] Jiang Lin, Zhang Qilin. Research on personalized recommendation strategy based on review sentiment analysis—Taking Douban movie reviews as an example [J]. Information Studies: Theory & Application, 2017, 40(8): 99-104.

[9] CHEN L, CHENG, WANG F. Recommender systems based on user reviews: the state of the art [J]. User modeling and user-adapted interaction, 2015, 25(2): 99-154.

[10] ESPARZA S G, P. O'MAHONY M, SMYTH B. Effective product recommendation using the real-time web [C]//Proceedings of the 30th SGAI international conference on innovative techniques and applications of artificial intelligence. Cambridge: Springer, 2010: 5-18.

[11] Geng Lixiao, Jin Gaojie, Li Yanhan, et al. Research on personalized recommendation of university library literature resources based on improved content filtering algorithm [J]. Library and Information Service, 2018, 62(21): 112-117.

[12] ZHANG W, DING G, CHEN L, et al. Generating virtual ratings from Chinese reviews to augment online recommendations [J]. ACM transactions on intelligent systems and technology, 2013, 4(1): 1-17.

[13] MUSAT C, LIANG Y, FALTINS B. Recommendation using textual opinions [EB/OL]. [2020-12-26]. https://www.researchgate.net/profile/Claudiu-Musat/publication/262348054_{{{Recommendation}}*{{using}}*}{{{textual}}_{{opinions}}}/links/57480c63 using-textual-opinions.pdf.

[14] Zhang Yihao, Zhu Xiaofei, Xu Chuanyun, et al. Hybrid recommendation method based on deep sentiment analysis of user reviews [J]. Chinese Journal of Computers, 2019, 42(6): 1316-1333.

[15] LIU H, HE J, WANG T, et al. Combining user preferences and user opinions for accurate recommendation [J]. Electronic commerce research and applications, 2013, 12(1): 14-23.

[16] Zhang Yanliang, Zhang Chao, Li Jing. KNN classification recommendation based on dynamic user portrait tags [J]. Information Science, 2020, 38(8): 11-15.

[17] BAO Y, FANG H, ZHANG J. TopicMF: Simultaneously exploiting ratings and reviews for recommendation [C]//Proceedings of the 28th AAAI conference on artificial intelligence. Quebec: AI Access Foundation, 2014: 2-8.

[18] FENG S, CAO J, WANG J. et al. Recommendation based on comprehensive exploitation of latent factors hidden in items' ratings and content [EB/OL]:

[2020-12-26]. https://dl.acm.org/doi/pdf/10.1145/3003728.

[19] ZHENG L, NOROOZI V, YU P S. Joint deep modeling of users and items using reviews for recommendation [C]//Proceedings of the 10th ACM international conference on web search and data mining. Cambridge, UK, 2017: 425-434.

[20] SEO S, HUANG J, YANG H, et al. Representation learning of users and items for review rating prediction using attention-based convolutional neural network [EB/OL]. [2020-12-26]. https://doogkong.github.io/2017/papers/paper8.pdf.

[21] CHEN C, ZHANG M, LIU Y, MA S. Neural attentional rating regression with review-level explanations [C]//Proceedings of the 2018 WWW conference on World Wide Web. Lyon, 2018: 1583-1592.

[22] Feng Xingjie, Zeng Yunze, Cui Guiying. Recommendation auxiliary network based on neighbor user reviews [J]. Application Research of Computers, 2020, 37(10): 2956-2960.

[23] BLEI D M, NG A Y, JORDAN M I. Latent dirichlet allocation [J]. Journal of machine learning research. 2003, 3(4/5): 993-1022.

[24] Chinese pre-training model [EB/OL]. [2020-12-20]. https://github.com/Embedding/Chinese-Word-Vectors.

Author Contributions:
Nie Hui: Research question proposal, research framework and process design, initial draft, revision and improvement;
Qiu Yifei: Data collection, evaluation platform construction, and data analysis.

Integrating User Interests and Review Utility for Review Information Recommendation

Nie Hui, Qiu Yifei
School of Information Management, Sun Yat-sen University, Guangzhou 510275

Abstract: [Purpose/significance] In the Web 2.0 era, the quality of online reviews is uneven and information overload is severe, leading to increasingly high cognitive costs for users seeking valuable content. This study explores effective solutions to review overload through information recommendation to enhance network information utilization and information service quality. The proposed review ranking and recommendation scheme focuses on review information quality while emphasizing satisfaction of users' personal information needs. [Method/process] This research employs a probabilistic topic model and introduces word vectors to construct user and review models within a topic space. By incorporating these into a review perceived utility evaluation system, the study achieves personalized review recommendation that integrates user interests and review quality, with recommendation effectiveness tested through systematic experiments. [Result/conclusion] Experimental results demonstrate that both review information quality and individual users' information needs jointly influence user satisfaction with review perceived utility. The recommendation strategy achieves organic integration of these two factors. Evaluation results under three different recommendation modes show that the "combined recommendation" approach yields the highest comprehensive satisfaction scores compared to pure "interest recommendation" and "utility recommendation."

Keywords: information recommendation; review utility prediction; user modeling; online reviews

Note: Figure translations are in progress. See original paper for figures.