# Personalized Recommendation Based on User Decision-Making Mechanism (Postprint)

**Authors:** Lin Xin, Sang Yunxin, Long Cunyu

**Date:** 2023-07-26T00:00:00+00:00

## Abstract

[Purpose/Significance] For content-based personalized recommendation strategies, propose optimization strategies for resource feature selection and weight calculation to improve personalized recommendation effectiveness.

[Method/Process] Construct a personalized recommendation model based on user decision mechanisms. The model utilizes user decision mechanisms as background knowledge for resource feature selection, user interest model construction and semantic representation, and user decision function construction. To verify the model's effectiveness, experiments were conducted on movie viewing data from 4,748 users, using the vector space model as a reference model and P@N as the evaluation metric.

[Results/Conclusion] Experimental results demonstrate that for N values of 5, 10, 20, 50, 100, and 200, the personalized recommendation model based on user decision mechanisms significantly outperforms the vector space model, thereby validating the model's effectiveness.

## Full Text

## Preamble

**ChinaXiv Cooperative Journal**
Volume 63, Issue 2, January 2019

**Personalized Recommendation Based on User Decision-Making Mechanism**

Lin Xin[1,2], Sang Yunxin[3], Long Cunyu[2]
[1]Institute of Scientific and Technical Information of China, Beijing 100038
[2]School of Information Management, Central China Normal University, Wuhan 430079
[3]School of Information Management, Wuhan University, Wuhan 430072

## Abstract

[**Purpose/Significance**] This paper proposes an optimization strategy for feature selection and weight calculation in content-based personalized recommendation to improve recommendation effectiveness. [**Method/Process**] The study constructs a personalized recommendation model based on user decision-making mechanisms, which leverages user decision mechanisms as background knowledge for resource feature selection, user interest model construction and semantic representation, and user decision function formulation. To validate the model, an experiment was conducted using movie viewing data from 4,748 users, with the vector space model serving as the reference model and P@N as the evaluation metric. [**Result/Conclusion**] Experimental results demonstrate that the personalized recommendation model based on user decision-making mechanisms significantly outperforms the vector space model across all tested values of N (5, 10, 20, 50, 100, 200), thereby confirming the model's effectiveness.

**Classification Number:** G203
**Keywords:** Decision-making mechanism; Content-based recommendation; Personalized recommendation
**DOI:** 10.13266/j.issn.0252-3116.2019.02.012

In recent years, personalized recommendation has gained widespread attention as an effective solution to information overload [1-3]. Among various approaches, content-based recommendation represents a common implementation strategy, where the key to achieving good performance lies in appropriate resource feature selection and weight calculation [4]. From the fundamental principles of content-based personalized recommendation, ideal resource features should not only serve as reference factors in user decision-making but also enable differentiation between resources that interest users and those that do not. Similarly, an ideal weight calculation strategy must reflect the magnitude of influence that individual resource features exert on user decisions while aligning with user decision-making mechanisms during multi-feature fusion. However, previous research has overemphasized whether selected features and their weight calculation methods can distinguish between historically interesting and uninteresting resources, paying insufficient attention to their alignment with user decision-making mechanisms. This oversight may result in selected features that are not actually considered by users during decision-making, and may lead to improper handling of these features in weight calculation, consequently compromising recommendation effectiveness.

To address these issues, user decision-making mechanisms can serve as background knowledge for designing personalized recommendation strategies, thereby guiding resource feature selection, user interest model construction, and feature weight calculation and fusion. In recent years, scholars have conducted extensive research on user decision-making mechanisms across various contexts and have begun applying these mechanisms to content-based personalized recommendation. The following section provides an overview of relevant findings.

# 1 Related Research

## 1.1 User Decision-Making Mechanisms in E-commerce and Information Consumption

Research on decision-making mechanisms has a long history in psychology, economics, and related fields, yielding multiple theoretical models including rational decision-making models, bounded rationality models, prospect theory, and preference construction theory [5]. With the development of the internet and e-commerce, scholars have explored user decision-making mechanisms in online shopping and information consumption contexts, achieving numerous results. Li Zongwei et al. studied factors influencing consumers' online purchase decisions from a holistic perspective, identifying key factors such as product price, sales volume, seller credit rating, service rating, store age, and online review length [6]. Regarding smartphone purchase decisions, K.L. Lay-Yee et al. [7] and J. Sujata et al. [8] conducted separate studies, with the former identifying product features, convenience, brand, price, and peer influence as primary factors, while the latter categorized them differently into technology, hardware, basic factors, brand, and price. In digital library resource utilization decisions, S. Joo et al. found that users' perceptions of usefulness and ease of use, along with information resource quality (availability, credibility, scope, novelty, and format), significantly influence final decisions [9]. Cha Xianjin et al. divided influencing factors into direct and indirect categories, with direct factors including information usefulness and attachment to digital libraries, and indirect factors encompassing digital library information quality, source credibility, and reputation, which exert influence indirectly through information usefulness [10]. Wu Jiang and Zhou Lusha employed regression analysis to study factors affecting users' decisions to purchase online health information services, categorizing them into physician title, hospital level, review quantity, positive rating ratio, thank-you letter count, and number of post-diagnosis registered patients [11].

## 1.2 Application of User Decision-Making Mechanisms in Content-Based Recommendation

Aside from a few studies that select features such as frequency centroid, short-term average energy, zero-crossing rate, MFCC, and bandwidth for music recommendation [12-13], most content-based recommendation research has considered user decision-making mechanisms to varying degrees. These approaches can be broadly classified into two categories: (1) considering user decision mechanisms during feature selection, and (2) considering user decision mechanisms during weight calculation.

**(1) Feature selection considering user decision mechanisms.** This type of research selects factors influencing user decisions as modeling features in personalized recommendation systems but mixes various factors together without differentiation during weight calculation. A typical example is the vector space model. For instance, in book recommendations for college students, features such as user major, grade, book topic, and author are used as modeling features [14]; in movie recommendation, genre, director, and actor serve as features [15]; and in music recommendation, language, ethnicity, culture, location, style, and singer are considered [16].

**(2) Weight calculation considering user decision mechanisms.** This approach involves first calculating the weight of each feature and then fusing them according to their roles in user decision-making to generate final recommendations. For example, in context-aware personalized recommendation, user interest in resources and contextual factors are considered separately before being integrated to obtain comprehensive weights [17-18]. Li Jiang et al., in their study on expert recommendation for academic review, first calculated the expertise match, academic influence, and social relevance between candidate experts and review objects, then used the product of these three dimensional weights as the basis for recommendation [19]. Yang Cheng et al., in their research on open-source project recommendation for developers, calculated candidate project popularity, relevance to user technical capabilities, and social connections separately before generating recommendation lists through weighted summation [20].

Overall, scholars have conducted multifaceted research on user decision-making mechanisms across different contexts, and the integration of these mechanisms into personalized recommendation has gained recognition and demonstrated effectiveness. However, research focusing on user decision mechanisms often only notes their application value in marketing, rarely mentioning their significance for personalized recommendation. Meanwhile, the application of user decision mechanisms in personalized recommendation remains at a spontaneous stage, lacking systematic approaches: (1) Feature selection is often based on researchers' experience and observation rather than proceeding top-down from user decision mechanisms, potentially resulting in feature sets that fail to comprehensively cover user decision influences, such as the expert recommendation strategy proposed in [19] that neglects factors like work attitude [21]; (2) In analyzing relationships between features, existing research typically relies on experience or logical analysis rather than systematic investigation of user decision mechanisms, which may lead to biased analysis of factor relationships, particularly when many features are involved, thereby affecting feature fusion effectiveness.

To address these issues, we first construct a general personalized recommendation model based on user decision-making mechanisms, grounded in these mechanisms for feature selection, user interest model construction, and decision function fitting strategy design. We then validate the model using movie recommendation as an example.

## 2 Personalized Recommendation Model Based on User Decision-Making Mechanism

In personalized recommendation based on user decision-making mechanisms, the decision mechanism serves as background knowledge to guide resource feature selection, interest model construction and semantic representation, and decision function generation, enabling the constructed recommendation model to better approximate users' actual decision-making processes and thereby improve recommendation effectiveness. The model components and their relationships are illustrated in Figure 1 [Figure 1: see original paper].

### 2.1 User Decision-Making Mechanism Analysis

User decision-making mechanism analysis forms the foundation for subsequent modules and represents a critical factor affecting recommendation effectiveness. Its core task is to identify the main factors influencing user decisions, their mechanisms of action, and the relationships between these factors during the decision-making process for the product or service to be recommended. In terms of analytical implementation, psychology has proposed various effective methods, including meta-analysis, experimental methods, observation, questionnaires, interviews, and regression analysis [22], which can be selected based on the specific product or service context.

It is worth noting that the analysis should particularly focus on: (1) Distinguishing between direct and indirect influencing factors; (2) Concretizing decision factors, ensuring at minimum the concretization of all indirect influencing factors and direct factors without indirect influences; (3) In analyzing factor mechanisms of action, special attention should be paid to multi-value influencing factors—for example, in the case of movie actors, when a film includes multiple actors that a user likes or dislikes, whether the effect is positive or negative and how the degree of influence changes. Additionally, factor interrelationships must be considered, i.e., the form of users' utility functions during decision-making, which are most commonly linear but may also take nonlinear forms such as U-shaped functions; (4) Due to individual differences in decision-making, multiple decision patterns may exist, meaning different users consider different factors and feature interactions may be diverse, so common decision patterns should be covered in the analysis.

### 2.2 Feature Selection

Resource features are the direct objects processed during interest modeling and recommendation generation. Therefore, to approximate user decision-making mechanisms in personalized recommendation, these mechanisms must be integrated during feature selection. First, feature selection should adopt a top-down approach, starting from factors influencing user decisions to identify features or feature combinations that reflect these factors. Second, the principle of independence should be maintained to avoid coupling between factors, but coupling should be determined based on user decision mechanisms rather than objective correlations. For example, a film's director and main actors are strongly correlated with its country/region, but users generally treat them as separate factors in decision-making, so all should be included in feature selection. Third, for the same direct influencing factor, users may employ different indirect factors (or combinations) for judgment, and individual features may have missing values, so appropriate redundancy can be maintained during feature selection.

To ensure reasonable feature selection, several principles must be followed: (1) Usefulness—selected features should be associated with at least one influencing factor; (2) Availability—data corresponding to selected features must be obtainable; (3) Usability—resource features need to be utilized in quantified form, so easily quantifiable features should be selected to reduce implementation difficulty. Additionally, some required features may not be directly available, necessitating the application of data mining methods such as data extraction, statistical analysis, classification, and clustering during data preprocessing.

### 2.3 Interest Modeling

This module's function is to externalize user interest preferences in the form of decision direct influencing factors—interest degrees—based on selected resource features and user decision mechanisms, using user historical behavior and related resource information, and to semantically describe them. Implementation can be divided into four stages: data preprocessing, interest degree calculation, decision function generation, and interest model semantic description.

**(1) Data preprocessing.** This stage processes user historical behavior and resource-related data into complete, standardized forms convenient for deep processing, incorporating selected resource features. Required operations include multi-source data integration, invalid data removal, data integrity verification, data normalization, and continuous data discretization [23].

**(2) Interest degree calculation.** To facilitate subsequent utilization, interest degree calculation should shift from a feature-granularity approach to using direct influencing factors identified through user decision mechanism analysis as the basic unit. Therefore, for direct influencing factors comprising multiple indirect factors or requiring multiple features for judgment, knowledge fusion methods must be employed in addition to common approaches such as frequency analysis, weighted statistics, Bayesian classification, decision trees, and neural networks.

**(3) User decision function generation.** In this stage, the basic function form is first determined based on relationships between influencing factors in the user decision mechanism, and regression analysis is then applied to fit the decision function using user historical behavior data. It is worth noting that: (1) To obtain the optimal user decision function, when multiple decision patterns exist, each pattern should be fitted and compared; (2) After determining the user decision function, irrelevant elements in the user decision mechanism should be adjusted or removed as appropriate to obtain a personalized decision model.

**(4) Interest model semantic representation.** Since various influencing factors play different roles in user decision-making, user interest models must be semantically represented to facilitate recommendation result generation. Specifically, a high-dimensional vector space is constructed using all direct influencing factors from the user's personalized decision model as the framework, with each dimension corresponding to a direct influencing factor in user decision-making. Each direct influencing factor can be represented as a vector composed of (feature value, interest degree) pairs.

### 2.4 Recommendation Result Generation

The core of recommendation result generation is calculating the matching degree between resources and user interest models based on user decision mechanisms, and selecting a subset to present to users to avoid creating new information overload. Implementation includes candidate recommendation object semantic representation, recommendation degree calculation based on user decision functions, and result screening based on recommendation degrees.

**(1) Candidate recommendation object semantic representation.** Similar to user interest model semantic representation, candidate recommendation object description must also use direct influencing factors from user decision-making as framework elements. If an element has a directly corresponding resource feature, its value is generated based on the feature value; if not, it must be processed according to the mapping relationship between original resource features and the element.

**(2) Recommendation degree calculation based on user decision functions.** In this stage, the vector space model is first used to calculate similarity between resources and each element in the user interest model. Based on this, element weights are fused according to the user decision function to generate final recommendation degrees.

**(3) Result screening based on recommendation degrees.** After calculating recommendation degrees for all candidate objects, a subset can be selected as final recommendations based on specific application requirements. Common screening methods include Top-N and threshold-based approaches. Top-N involves ranking resources by comprehensive weight and selecting the top N or N%, while threshold-based methods set a comprehensive weight threshold and present all resources exceeding it to users.

## 3 Experiment

To validate model effectiveness, experiments were conducted using movies as the object and Douban Movie (https://movie.douban.com/), a well-known Chinese movie website, as the data source. To facilitate effectiveness evaluation, the commonly used vector space model was selected as the baseline for comparison.

### 3.1 Sample Data

The sample data consists of a dataset collected between November 20 and December 15, 2014, containing viewing records of 830,682 Douban users. The dataset includes: (1) all viewing records of these users, with fields including user ID, movie URL, viewing time, and added tags; (2) basic information for 101,486 film and television works involved in the sample, including URL, title, director, actors, screenwriter, genre, production country, release time, rating, number of viewers, episodes, and duration.

After obtaining the basic data, non-movie data such as TV dramas and variety shows, data lacking release time fields, and corresponding user viewing records were removed. The sample set was then constructed as follows: First, 5,000 users were randomly selected from those who watched at least one movie between May 3 and 14, 2014; second, to avoid excessive data sparsity, users who had watched fewer than 20 movies before May 3, 2014 were removed, leaving 4,748 users. Subsequently, these users' viewing records before May 3, 2014 were used as the interest modeling dataset (62,612 records), and their viewing records from May 3-14, 2014 were used as the test dataset (27,482 records).

### 3.2 Movie Recommendation Experiment Process Based on User Decision-Making Mechanism

According to the proposed recommendation model, the movie recommendation experiment based on user decision-making mechanisms primarily includes user movie-watching decision mechanism analysis, feature selection, interest modeling, and recommendation result generation.

**3.2.1 User Movie-Watching Decision Mechanism Analysis** Douban Movie users are primarily well-educated young people such as college students and young white-collar workers with relatively stable viewing interests. To analyze users' movie decision-making mechanisms, 15 students from Central China Normal University and Wuhan University were interviewed. Overall, the main factors considered in user movie-watching decisions were relatively consistent, including actors, directors, theme/genre, country/region, rating, popularity, and novelty. Among these, actors, directors, theme/genre, and country/region do not have obvious directional effects and depend on individual user preferences; rating, popularity, and novelty have clear directional effects—users prefer movies with higher ratings, greater popularity, and newer releases—though the magnitude of these effects is influenced by individual preferences.

It should be noted that a movie may have multiple values for theme/genre, actors, directors, and country/region. During decision-making, user interest degrees for each factor are often not simple sums of interest degrees for individual values but more likely take the maximum interest degree among all feature values. The mechanisms of interaction between these factors vary across different movies, with two common patterns: (1) If users particularly like a movie's director or main actors and the rating, popularity, and novelty are within acceptable ranges, they will choose to watch; (2) If the main actors or directors are unfamiliar or not particularly liked, users will comprehensively consider the five factors of theme/genre, country/region, rating, popularity, and novelty, often first filtering by theme/genre in practice.

**3.2.2 Feature Selection** Although the aforementioned factors influencing user decisions exhibit certain correlations—for example, directors and actors are significantly correlated with movie theme/genre, country/region, and popularity—users treat them as distinct factors during decision-making. Therefore, when selecting movie features, features related to all these factors were considered. Specifically, selected features include (the mapping between features and influencing factors is shown in Table 1 ): (1) Directors and actors: These two factors operate similarly and have comparable influence on users, with overlapping values, so they were merged into a single "creator" factor; (2) Theme/genre: Since movie themes and genres overlap, they were combined (e.g., action, family, crime); (3) Production country/region: In decision-making, users consider country/region based on a combination of main actors, filming locations, and characters' origins, but this information is difficult to obtain, so production country/region was used as a proxy; (4) Rating: Different populations may rate movies differently, so Douban user ratings were used; (5) Number of viewers: The popularity considered in user decision-making reflects whether a movie has been widely watched, which can be indicated by the number of viewers on Douban movie pages; (6) Release time: This feature is used to determine a movie's novelty at present or a specific time.

**Table 1** Mapping Between Selected Features and User Decision Influencing Factors

| Decision Influencing Factor | Selected Feature |
|---|---|
| Actors, Directors | Creator |
| Theme/Genre | Theme/Genre |
| Country/Region | Production Country/Region |
| Rating | Rating |
| Popularity | Number of Viewers |
| Novelty | Release Time |

**3.2.3 Interest Modeling** To construct user movie interest models, movie viewing records and basic movie data must first be preprocessed, followed by calculating user interest degrees for each feature item, generating user decision functions, and providing semantic descriptions.

**(1) Data preprocessing.** This includes: (1) Extracting creator information, including the director and top 4 actors (based on major film awards where Best Actor/Actress winners are typically among the top 3 billed actors, with top 4 accommodating ensemble casts), with deduplication within the same movie; (2) Extracting theme/genre information from user-added tags using the method from [24], removing cold theme/genre tags associated with fewer than 50 movies; (3) Extracting country/region information from the "production country/region" field; (4) Novelty extraction and discretization: Novelty is relative. For interest degree calculation, it is measured by the difference between viewing time and release time. Based on 2013 viewing data from 830,682 users, novelty was discretized as follows: Data was divided into half-year intervals, revealing that movies released within 6 months accounted for 30.7% of viewings, and those released 7-12 months accounted for 7.0%, significantly higher than other intervals. Therefore, "within 6 months" and "6-12 months" were defined as two novelty intervals. For 1-3 years, 3-5 years, and 5-10 years, data distribution across half-year intervals was similar (e.g., 1.45%-1.81% for the 5-10 year range), so these were consolidated into three intervals: 1-3 years, 3-6 years, and 6-10 years. Movies released over 10 years had relatively low viewership, so they were grouped into a single interval for simplicity; (5) Rating extraction and discretization: If the scraped rating field was not empty, its value was used; otherwise, the average rating of 6.9 was assigned. Douban's 10-point scale was simplified into five intervals: $ $5, 5.1-6, 6.1-7, 7.1-8, and >8; (6) Popularity discretization: Popularity spans a large continuous range. To facilitate measurement, movies were sorted by viewing frequency, with those accounting for the top 40% of total views classified as level 1. The remaining movies were processed similarly, resulting in 10 popularity levels. Preprocessed movie information is shown in Table 2 .

**Table 2** Partial Preprocessed Movie Data

| Movie Title | Creator | Theme/Genre | Country/Region | Novelty | Rating | Popularity |
|---|---|---|---|---|---|---|
| Example 1 | Director A, Actor B, Actor C, Actor D, Actor E | Action, Crime | Hong Kong | >10 years | >8 | Level 1 |

| Movie Title | Creator | Theme/Genre | Country/Region | Novelty | Rating | Popularity |
|---|---|---|---|---|---|---|
| Example 2 | Director F, Actor G, Actor H, Actor I, Director F | Crime, Police | Hong Kong | 6-12 months | 7.1-8 | Level 6 |
| Example 3 | Director J, Actor K, Actor L, Actor M, Actor N | Life, Comedy | South Korea | Within 6 months | >8 | Level 4 |

**(2) Interest degree calculation.** Interest degrees serve the recommendation generation stage by calculating matching degrees between candidate objects and interest models. In movie decision-making, users typically first make preliminary decisions based on creators or theme/genre, then finalize decisions based on country/region, rating, popularity, and novelty. Therefore, the interest degree for creators and theme/genre should represent the probability that a user will watch a movie containing that feature value, with a range of [0,1]. For the remaining four features, interest degrees should represent the extent to which a user prefers a movie with that feature value compared to other values, fluctuating around 1 with a theoretical range of [0,+∞). Accordingly, Formula (1) calculates interest degrees for creators and theme/genre, while Formula (2) calculates interest degrees for the other four features.

$$W(u_i, t_j) = \frac{F(u_i, t_j) - 1}{F(t_j)} \times \frac{F(all)}{F(u_i) \times (F(u_i, t_j) - 1)} \times \frac{F(u_i)}{F(t_j)} \quad \text{(Formula 1)}$$

$$W(u_i, t_j) = \frac{F(u_i, t_j)}{F(u_i)} \quad \text{(Formula 2)}$$

Where $W(u_i, t_j)$ represents user $u_i$'s interest degree in feature value $t_j$; $F(u_i, t_j)$ represents the number of movies watched by user $u_i$ that contain feature value

$t_j$ (requiring $F(u_i, t_j) \geq 3$ to avoid random effects); $F(t_j)$ represents the total number of movies containing feature value $t_j$; $F(all)$ represents the total number of movies; and $F(u_i)$ represents the total number of movies watched by user $u_i$.

**(3) User decision function fitting.** Based on the relationships between factors in user decision-making, user decision patterns can be abstracted as: (1) Preliminary decision-making based on movie creators or theme/genre and country/region—if interested, proceed to the next step; (2) Further decision-making based on novelty, rating, and popularity. Therefore, the user decision function can be expressed as Formula (3) (with logarithmic smoothing applied to the effects of country/region, novelty, rating, and popularity):

$$W(u_i, m_j) = \begin{cases} \max(\max w(u_i, cre_j), \max w(u_i, typ_j) \times (1 + \max \log w(u_i, cou_j))) \times (1 + \log w(u_i, rat_j)) \times (1 + \\ \quad \text{if } \log w(u_i, rat_j) > -1, \log w(u_i, pop_j) > -1, \log w(u_i, nov_j) > -1 \\ 0 \quad \text{if } \log w(u_i, rat_j) \leq -1 \text{ or } \log w(u_i, pop_j) \leq -1 \text{ or } \log w(u_i, nov_j) \leq -1 \end{cases}$$

Where $W(u_i, m_j)$ represents the matching degree between movie $m_j$ and user $u_i$'s interest model, and $w(u_i, cre_j)$, $w(u_i, typ_j)$, $w(u_i, cou_j)$, $w(u_i, rat_j)$, $w(u_i, pop_j)$, $w(u_i, nov_j)$ represent user $u_i$'s interest degrees in movie $m_j$'s creator $cre_j$, theme/genre $typ_j$, country/region $cou_j$, rating $rat_j$, popularity $pop_j$, and novelty $nov_j$.

**(4) Semantic description.** Since user movie-watching decision-making involves relatively consistent influencing factors, each user's semantic description framework consists of a high-dimensional vector comprising six elements: {creator, theme/genre, country/region, novelty, rating, popularity}. Each dimension consists of a series of (feature value, interest degree) pairs. For example, the creator dimension can be represented as [Creator: (Andy Lau, 0.34), (Li Bingbing, 0.13), …]. A partial example is shown in Table 3 .

**Table 3** Partial User Semantic Interest Model

| User ID | Creator | Theme/Genre | Country/Region | Novelty | Rating | Popularity |
|---|---|---|---|---|---|---|
| User 1 | Director A, 0.40; Director B, 0.25 | Black Humor, 0.039; Life, 0.028 | Hong Kong, 1.172; South Korea, 1.202 | Within 6 months, 1.022; 6-12 months, 0.768 | >8, 1.176; 7.1-8, 1.124 | Level 1, 1.830; Level 2, 1.642 |

| User ID | Creator | Theme/Genre | Country/Region | Novelty | Rating | Popularity |
|---|---|---|---|---|---|---|
| User 2 | Director C, 0.33; Director D, 0.25 | Life, 0.028; Humanity, 0.028 | Hong Kong, 1.218; Thailand, 1.572 | Within 6 months, 1.170; 6-12 months, 0.748 | >8, 1.228; 7.1-8, 1.176 | Level 1, 1.939; Level 2, 1.662 |
| User 3 | Director E, 0.25; Director F, 0.25 | Psychology, 0.017; Crime, 0.076 | Hong Kong, 1.152 | 1-3 years, 0.819; 6-12 months, 0.748 | 7.1-8, 1.021; 7.1-8, 1.176 | Level 1, 1.939; Level 3, 1.322 |

**3.2.4 Recommendation Result Generation** In the recommendation generation stage, all 75,694 movies were used as the candidate set, but for each user, movies already watched before May 3, 2014 were excluded. First, candidate movies were semantically represented. For the three features of creator, theme/genre, and country/region, binary weighting was applied (value = 1 if present, 0 otherwise). For novelty, rating, and popularity, values were mapped to corresponding intervals and then represented as vectors. Second, the recommendation degree for each candidate movie was calculated based on the user's personal interest model and decision function. Finally, candidate movies were ranked by recommendation degree in descending order, and the top N movies were selected as final recommendations. In the experiment, N was set to 5, 10, 20, 50, 100, and 200.

### 3.3 Movie Recommendation Experiment Process Based on Vector Space Model

The vector space model-based movie recommendation experiment served as a baseline for evaluating the previous experiment's effectiveness. In the experimental design, drawing on content-based movie recommendation research [25-27], director, actor, theme/genre, and country/region were selected as features, with popularity used to adjust results during recommendation generation.

In interest modeling, assuming $w(u_i, t_j)$ represents user $u_i$'s interest degree in feature value $t_j$, $F(u_i, t_j)$ represents the frequency of feature value $t_j$ in movies watched by user $u_i$, and $F(u_i)$ represents the total number of movies watched by user $u_i$, the interest degree for each feature value can be calculated using Formula (4):

$$w(u_i, t_j) = \frac{F(u_i, t_j)}{F(u_i)} \quad \text{(Formula 4)}$$

For popularity weighting strategy design, a similar approach to the previous experiment was adopted. Movies were divided into 10 popularity intervals using the same method. Assuming $w_{pop_j}$ represents the weight of popularity interval $j$, $F(all)$ represents the total number of movies, $F(pop_j)$ represents the proportion of total views accounted for by popularity interval $j$, and $N(pop_j)$ represents the number of movies in popularity interval $j$, the weight of popularity interval $j$ can be calculated using Formula (5):

$$w_{pop_j} = 1 + \log \frac{F(all) \times F(pop_j)}{N(pop_j)} \quad \text{(Formula 5)}$$

Based on the interest degree calculation formula and popularity weighting strategy, the final matching degree between a movie and user interest model can be calculated as the product of the user interest vector $u_i((t_1, w(u_i, t_1)), (t_2, w(u_i, t_2)), ..., (t_j, w(u_i, t_j)), ...)$, the movie feature vector $m_j((f_1, 1), (f_2, 1), ..., (f_k, 1), ...)$, and the corresponding popularity weight $w_{pop_{mj}}$, as shown in Formula (6):

$$W(u_i, m_j) = u_i \cdot m_j \cdot w_{pop_{mj}} \quad \text{(Formula 6)}$$

### 3.4 Experimental Results and Analysis

To measure experimental effectiveness, the typical P@N metric was selected as the evaluation index [28], and chi-square tests were conducted for significance testing. The results are shown in Table 4 .

**Table 4** Experimental Results and Significance Tests

| Model | P@5 | P@10 | P@20 | P@50 | P@100 | P@200 |
|---|---|---|---|---|---|---|
| User Decision Mechanism | 0.53%*** | 0.54%*** | 0.47%*** | 0.41%*** | 0.34%*** | 0.28%*** |
| Vector Space Model | 0.18% | 0.21% | 0.24% | 0.23% | 0.23% | 0.22% |

Note: *** indicates p < 0.001

Table 4 shows that the personalized recommendation model based on user decision-making mechanisms significantly outperforms the vector space model ($p < 0.001$). The improvement is substantial: for example, at P@10 and P@20, the user decision mechanism model achieves accuracies of 0.54% and 0.47%, respectively, compared to 0.21% and 0.24% for the vector space model. The superior performance stems from the model's better approximation of actual user decision-making mechanisms, specifically in three aspects: (1) Feature selection based on user decision mechanisms comprehensively and systematically covers the main factors users consider; (2) Weight calculation methods tailored to each feature dimension's characteristics better reflect reality, such as considering the number of works when calculating user interest in creators; (3) The comprehensive weight calculation strategy based on multi-dimensional fusion better aligns with the relationships between dimensions in user decision-making.

## Conclusion

To improve content-based personalized recommendation effectiveness, feature selection, interest modeling, and recommendation generation should be grounded in user decision-making mechanisms. This study constructed a personalized recommendation model based on user decision-making mechanisms and validated it using movie data. Results show that this approach substantially improves recommendation effectiveness compared to the vector space model, confirming the validity of the user decision mechanism-based recommendation approach.

Based on the prototype movie recommendation experiment, future research will focus on: (1) Analyzing and applying complex relationships between factors in actual user decision-making, such as features with both positive and negative values and competitive relationships among multiple features; (2) Automatically identifying user decision styles and adaptively adjusting recommendation models, as different decision styles may involve significantly different factors and utilization methods; (3) Expanding application domains to academic information resources such as books and academic papers to further validate model effectiveness.

## References

[1] LU J, WU D, MAO M, et al. Recommender system application developments: a survey[J]. Decision support systems, 2015, 74(C): 12-32.

[2] RUAN Guangce, XIA Lei. Internet recommendation system research: a survey[J]. Journal of the China Society for Scientific and Technical Information, 2015, 34(9): 999-1008.

[3] KUNAVER M, POŽR T. Diversity in recommender systems—a survey[J]. Knowledge-based systems, 2017, 123(9): 154-162.

[4] HARIRI N, MOBASHER B, BURKE R. Context adaptation in interactive recommender systems[C]//ACM conference on recommender systems. New York: ACM, 2014: 41-48.

[5] WANG Dongshan. Consumer purchase decision theory review and prospect[J]. Commercial Economic Research, 2017(21): 43-46.

[6] LI Zongwei, ZHANG Yanhui, LUAN Dongqing. Which factors affect consumers' online purchase decisions[J]. Management Review, 2017, 29(8): 136-146.

[7] LAY-YEE K L, HAN K S, FAH B C Y. Factors affecting smartphone purchase decision among Malaysian generation Y[J]. International journal of Asian social science, 2013, 3(12): 2426-2440.

[8] SUJATA J, YATIN J, ABHIJIT C, et al. Factors affecting smartphone purchase among Indian youth: a descriptive analysis[J]. Indian journal of science & technology, 2016, 9(15): 1-10.

[9] JOO S, CHOI N. Factors affecting undergraduates' selection of online library resources in academic tasks[J]. Library Hi Tech, 2015, 33(2): 114-117.

[10] CHA Xianjin, LI Li, YAN Yalan, et al. Research on influencing factors of information usefulness and information acquisition in digital library environments[J]. Journal of the China Society for Scientific and Technical Information, 2017, 36(7): 669-678.

[11] WU Jiang, ZHOU Lusha. Research on influencing factors of users' purchase decisions for online health information services[J]. Journal of the China Society for Scientific and Technical Information, 2017, 36(10): 1058-1065.

[12] BOGDANOV D, HARO M, FUHRMANN F, et al. Semantic audio content-based music recommendation and visualization: based on user preference examples[J]. Information processing & management, 2013, 49(1): 13-33.

[13] TAN Xueqing, HE Shan. Review of music personalized recommendation systems[J]. Modern Library and Information Technology, 2014, 30(9): 22-32.

[14] HONG Wen, NIE Yanping, QING Qiao. Optimization analysis of model structure and processing flow for automatic collection resource recommendation[J]. Information Studies: Theory & Application, 2016, 39(5): 130-133.

[15] SOARES M, VIANA P. Tuning metadata for better movie content-based recommendation systems[J]. Multimedia tools and applications, 2015, 74(17): 7015-7036.

[16] JUN S, KIM D, JEON M, et al. Social mix: automatic music recommendation and discovery[J]. International journal of human-computer interaction, 2017, 33(3): 165-179.

[17] CHEN C, CHEN C, CHEN H, et al. Towards context-aware social recommendation and mixing scheme based on social network analysis[J]. Knowledge-based systems, 2015, 49(3): 835-856.

[18] MAN P, LIIKANEN L A. Interacting with context factors in human-computer interaction[J]. International journal of electrical engineering & computer sciences, 2016, 24(3): 833-848.

[19] LI Jiang, LI Dong, FENG Peihua, et al. Research on expert recommendation model based on expertise match, academic influence, and social relevance[J]. Journal of the China Society for Scientific and Technical Information, 2017, 36(4): 338-345.

[20] YANG Cheng, FAN Qiang, WANG Tao, et al. Personalized recommendation method for open-source projects based on multi-dimensional features[J]. Journal of Software, 2017, 28(6): 1357-1372.

[21] ZHOU Jianzhong, XU Fang. Comparative study on peer review methods in national research institutions[J]. Studies in Science of Science, 2013, 31(11): 1642-1648.

[22] ZHUANG Jinying. Decision Psychology[M]. Shanghai: Shanghai Education Press, 2006: 12-23.

[23] HAN J, KAMBER M, PEI J. Data Mining: Concepts and Techniques[M]. 3rd ed. Translated by FAN Ming, MENG Xiaofeng. Beijing: Mechanical Industry Press, 2012: 55-79.

[24] LIN Xin, SHI Yu, ZHOU Zhi. Optimization research on tag relevance judgment based on relative frequency[J]. Library and Information Service, 2016, 60(17): 130-135.

[25] LYU Xueqiang, WANG Teng, LI Xuewei, et al. Research on movie recommendation algorithm based on content and interest drift model[J/OL]. [2017-11-21]. http://www.arocmag.com/article/02-2018-03-049.html.

[26] ÇATALTEPE Z, ULUYAĞMUR M, TAYFUR E. Feature selection for movie recommendation[J]. Turkish journal of electrical engineering & computer sciences, 2017, 25(C): 58-66.

[27] SHI C, LIU J, ZHUANG F, et al. Integrating heterogeneous information via flexible regularization framework for recommendation[J]. Knowledge & information systems, 2015, 49(3): 835-854.

[28] HU Qian, LIN Xin. Comparative study on interest modeling based on tags and items in social tagging systems[J]. Journal of the China Society for Scientific and Technical Information, 2015, 34(12): 1296-1303.

## Author Contributions

Lin Xin: Responsible for data processing and paper writing
Sang Yunxin: Participated in data processing and responsible for paper revision
Long Cunyu: Responsible for paper revision

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