# Postprint: An Aggregated Recommendation Algorithm Incorporating Item Fatigue and Diversity Preferences

**Authors:** Que Zhenghao, Deng Mingtong, Liu Xuejun, Li Bin

**Date:** 2018-08-13T00:00:00+00:00

## Abstract

To address the problem of recommendation lists being biased towards popular items and exhibiting poor diversity, we propose the ARIFDP algorithm (Aggregation Recommendation Algorithm for Embedding Item Fatigue and Diversity Preference). First, users' diversity preferences are captured by analyzing their historical feedback data to derive their diversity inclination degree. An item fatigue function that is negatively correlated with the number of ratings is then constructed. Finally, matrix factorization is aggregated with the item fatigue function, and the diversity inclination degree is incorporated to modulate the weight of the item fatigue function, thereby increasing the probability of recommending long-tail items. Experimental results demonstrate that the ARIFDP algorithm can effectively enhance the diversity of recommendation results while maintaining accuracy.

## Full Text

## Preamble

**Title:** An Aggregation Recommendation Algorithm Embedding Item Fatigue and Diversity Preference

**Authors:** Que Zhenghao, Deng Mingtong, Liu Xuejun, Li Bin
(College of Computer Science & Technology, Nanjing Tech University, Nanjing 211816, China)

**Abstract:** To address the problem of recommendation lists being biased toward popular items with poor diversity, this paper proposes the ARIFDP (Aggregation Recommendation Algorithm for Embedding Item Fatigue and Diversity Preference) algorithm. First, we analyze users' diversity preferences through their historical feedback data to derive their diversity tendency degree. We then construct an item fatigue function that is negatively correlated with the number of evaluations. Finally, we aggregate matrix factorization with the item fatigue function, incorporating the diversity tendency degree to adjust the weight of the item fatigue function, thereby increasing the probability of recommending long-tail items. Experimental results demonstrate that the ARIFDP algorithm can effectively improve recommendation diversity while maintaining accuracy.

**Keywords:** topic model; matrix factorization; diversity tendency; item fatigue function; recommendation diversity

## 0 Introduction

Broadening users' horizons has become an essential characteristic of modern recommendation systems. A system capable of expanding user perspectives can achieve a win-win outcome: users discover more interesting items, while website operators increase sales and improve user satisfaction. As an effective tool for solving information overload, recommendation systems help users find relevant items from vast item pools. However, existing recommendation systems often focus primarily on accuracy, making diversity improvement increasingly critical. For instance, when a user searches for the movie "Wrong Turn" on the movie database website IMDB, all movies from the "Wrong Turn" series are recommended, as shown in [Figure 1: see original paper]. While satisfactory from an accuracy perspective, this lack of diversity fails to broaden user horizons and may cause user fatigue.

## 1 Related Work

The primary goal of recommendation systems is to provide personalized lists that meet user requirements based on historical behavior data. Several approaches have been proposed to address different aspects of recommendation quality. Some researchers have introduced three discounting methods for user purchase or click behaviors—user-oriented, item-oriented, and time-oriented discounting—which integrate well with existing matrix factorization models to improve recommendation accuracy. Others have argued for collecting both positive and negative feedback, modeling user feedback as categorical variables in triplets with users and items, and employing third-order tensor decomposition techniques for recommendations. Additionally, classification frameworks based on various attributes (including cognitive effort, user models, measurement scales, and domain relevance) have been developed to leverage explicit and implicit user feedback, enhancing system performance.

While these methods improve accuracy to some extent, they share a common limitation: neglecting recommendation diversity. Current diversity research focuses on two aspects: aggregate diversity and individual diversity. Some studies utilize meta-path similarity measures in heterogeneous information networks to calculate item correlations, linking popular and niche items to improve aggregate diversity. Others propose tensor decomposition frameworks to reveal latent topics in multi-modal data including communities, users, and social tags for enhancing individual diversity. Heat conduction algorithms oriented toward diversity have been combined with accuracy-oriented energy diffusion methods to balance precision and diversity. Collaborative topic modeling approaches integrate traditional collaborative filtering with probabilistic topic modeling to provide interpretable latent structures for users and items, thereby increasing diversity. The UC-BCF model transforms user-item relationships into item-item relationships through item co-occurrence to discover novel items. However, these studies concentrate on item sets without considering users' individual preferences for diversity.

This paper proposes the ARIFDP algorithm, which embeds item fatigue and diversity preferences into an aggregation recommendation framework. Our main contributions are: (a) analyzing user diversity preferences from historical feedback to propose the concept of diversity tendency degree, calculated through clustering analysis; (b) constructing an item fatigue function constrained by long-tail distribution to increase weights for long-tail items; (c) proposing the ARIFDP algorithm that aggregates matrix factorization with item fatigue functions, using diversity tendency degree to adjust the weight of item fatigue functions; and (d) demonstrating through experiments on real datasets that ARIFDP achieves favorable results.

## 2 Problem Description and LDA Topic Model

### 2.1 Problem Description

We formulate the recommendation problem as follows: given item description information and a target user' s historical feedback, recommend a set of items that align with the user' s diversity preference degree.

### 2.2 LDA Topic Model

The LDA (Latent Dirichlet Allocation) model, proposed by D.M. Blei in 2003, is an unsupervised machine learning technique for identifying latent topic information in large document collections or corpora. The model structure is shown in [Figure 2: see original paper].

Given a document collection D with M documents containing K topics and N words w, where and are corpus-level parameters following Dirichlet distributions— is a vector parameter for generating topic distribution , and is the word probability distribution matrix for each topic—the text generation process can be described as:

a) For each document d, sample its topic distribution _d from Dirichlet distribution Dir( )

b) For the nth word in document d, sample its topic z_dn from multinomial distribution Multi( _d)

c) Sample the topic-word distribution   from Dirichlet distribution Dir( )

d) Sample the final word w_dn from multinomial distribution Multi( _z_dn)

The probability of generating M documents in collection D using LDA is given by:

$$p(D|\alpha,\beta) = \prod_{m=1}^{M} \int p(\theta_m|\alpha) \left( \prod_{n=1}^{N_m} \sum_{z_{mn}} p(z_{mn}|\theta_m) p(w_{mn}|z_{mn},\beta) \right) d\theta_m$$

The model contains two hidden variables: document-topic distribution   and topic-word distribution  , both obtainable through Gibbs sampling.

# 3 Aggregation Recommendation Algorithm Embedding Item Fatigue and Diversity Preference

## 3.1 User Diversity Preference

**Definition 1 (Diversity Tendency Degree).** Let K be the number of clusters after clustering all items in the item set, and L be the number of subcategories in a user’ s feedback history. The user’ s diversity tendency degree   is defined as:

$$\mu = \frac{L}{K}$$

**Definition 2 (Item Topic Distribution).** For each item v, we obtain its topic distribution from item description information using the LDA topic model, denoted as  _v.

Research on user historical feedback reveals that most users interact with many different types of items. The greater the variety of item categories, the stronger the user’ s preference for diversity. We employ K-means clustering to categorize items. As a centroid-based technique, K-means requires calculating dissimilarity between different item types. For item features, we extract topic distribution vectors from item descriptions using LDA and compute Euclidean distance between vectors a and b:

$$dist(a,b) = \sqrt{\sum_{i=1}^{n} (a_i - b_i)^2}$$

where n is the dimension of the topic vector. After setting the number of clusters K and completing clustering, we analyze validity using the average silhouette coefficient. For object o in cluster S_1, a(o) represents the average distance between o and other objects in the same cluster, reflecting compactness (smaller values indicate tighter clusters). b(o) represents the minimum average distance between o and objects in other clusters, capturing separation (larger values indicate better separation). The silhouette coefficient for object o is:

$$s(o) = \frac{b(o) - a(o)}{\max\{a(o), b(o)\}}$$

The average silhouette coefficient across all n items measures overall clustering quality:

$$\bar{s} = \frac{1}{n} \sum_{i=1}^{n} s(o_i)$$

Larger average silhouette coefficient values indicate better clustering quality. We find the K value that maximizes the average silhouette coefficient to determine the optimal number of clusters.

After obtaining K clusters (excluding outliers), we compare each item in a user' s historical feedback against these clusters to identify L subcategories (L K). The diversity tendency degree is then calculated using Definition 1.

**Algorithm 1: Diversity Tendency Degree Calculation**

Input: Dataset of all item descriptions, dataset of item IDs from user historical feedback
Output: Diversity tendency degree

1. For each item v, compute its topic distribution vector _v using LDA
2. Calculate pairwise dissimilarities between all vectors
3. Randomly select K objects as initial cluster centers
4. Repeat:
    - Assign each object to the nearest cluster based on mean values
    - Update cluster means until convergence
5. Vary K and repeat steps 3-4, computing average silhouette coefficients each time
6. Select K with maximum average silhouette coefficient
7. For each user u, compare their feedback items against K clusters to obtain L subcategories
8. Return = L/K

### 3.2 Diversity Recommendation Method

Traditional collaborative filtering struggles when new items lack sufficient user feedback. Inspired by previous work, we integrate item topic models into matrix factorization to compute initial user preference scores, then aggregate item fatigue functions to adjust these scores. This approach alleviates cold-start problems while improving recommendation diversity.

To enhance matrix factorization, we initialize the item latent feature matrix V using topic distributions rather than random initialization. Let represent the deviation between item feature distribution and LDA-learned topic distribution. The objective function incorporating user rating information becomes:

$$f(U,V) = \sum_{(i,j)\in R} (r_{ij} - u_i^T v_j)^2 + \lambda_u \sum_i \|u_i\|^2 + \lambda_v \sum_j \|v_j - \theta_j\|^2$$

where (i,j) R denotes user i' s rating on item j, r_{ij} is the rating, U and V are user and item latent feature matrices, and _u, _v are regularization coefficients.

Using alternating least squares, we obtain update formulas:

$$u_i \leftarrow (V^T V + \lambda_u E)^{-1} V^T R_i$$

$$v_j \leftarrow (U^T U + \lambda_v E)^{-1} (U^T R_j + \lambda_v \theta_j)$$

where E is the identity matrix. Through iterative updates, we obtain user latent features U* and item latent features V*, then compute user i' s preference score for item j:

$$r_{ij}^* = u_i^{*T} v_j^*$$

### 3.3 Item Fatigue Function Constrained by Long-tail Distribution

**Definition 3 (Long-tail Distribution).** Let S(x) be the cumulative distribution function of any distribution with complementary function c(x) = 1 - S(x). If for any > 0, lim_{x→∞} e^{ x} c(x) = ∞, then the corresponding distribution is a long-tail distribution.

To satisfy long-tail distribution conditions, previous work proposed describing it using linear combinations of exponential functions. For a dataset with h data points (x_i, y_i) where x_i is item ranking by rating count and y_i is the rating count, with n total items, we minimize the squared error:

$$F(p_j, \lambda_j) = \sum_{i=1}^{h} \left( \sum_{j=1}^{n} p_j e^{-\lambda_j x_i} + C - y_i \right)^2$$

where p_j > 0 are weight coefficients obtained by solving an n-th order nonlinear system for parameters _j that best fit the long-tail distribution. Analysis shows that when n = 2, the description is optimal:

$$P(x) = p_1 e^{-\lambda_1 x} + p_2 e^{-\lambda_2 x} + C$$

Since item rating counts follow a long-tail distribution, we construct an item fatigue function negatively correlated with rating count to increase long-tail item recommendations. Let  (x) map long-tail distribution function values to [0,1]:

$$\eta(x) = \frac{1}{1 + e^{-x}}$$

The item fatigue function is then:

$$W(x) = 1 - \eta(P(x)) = 1 - \frac{1}{1 + e^{-P(x)}} = \frac{1}{1 + e^{P(x)}}$$

### 3.4 Overall Description of ARIFDP Algorithm

Recommendation systems should aim for "long-term user retention." Algorithms focused solely on accuracy harm long-term performance and prevent long-tail items from being recommended. The ARIFDP algorithm addresses this by: (1) analyzing user historical feedback to determine diversity tendency degree; (2) integrating topic models with matrix factorization to compute initial preference scores; (3) aggregating initial scores with item fatigue functions; and (4) using diversity tendency degree to adjust the weight of item fatigue functions.

The final preference score is:

$$r_{ij}^{**} = u_i^{*T} v_j^* + \omega \cdot \mu \cdot \eta(x)$$

where  adjusts the value range (set to the maximum rating value) and  is the diversity tendency degree.

**Algorithm 2: ARIFDP Diversity Recommendation Algorithm**

Input: User rating dataset D, item set to recommend
Output: Top-N recommendation list for each user

1. For each item, compute topic distribution  _v from description
2. Initialize item latent feature matrix V with  _v
3. Incorporate rating dataset D to adjust user/item feature distributions and solve for U* and V*
4. Compute initial preference scores r_{ij}^* = u_i^{$T$} $v$_$j$^
5. Construct item fatigue function W(x) from dataset D
6. Compute final scores r_{ij}^{**} = r_{ij}^* +  ·  · W(x)
7. Sort scores descending and select top-N as final recommendations

# 4 Experiments

## 4.1 Experimental Dataset

We use the MovieLens dataset provided by the GroupLens research group at the University of Minnesota, containing approximately 1 million movie ratings from 6,040 users on 3,952 movies. Each user rated at least 20 movies on a scale of 1-5 (“1” = poor, “5” = perfect). We selected a subset of 943 users, 1,682 movies, and approximately 100,000 ratings (sparsity: 93.7%). Since the dataset lacks movie descriptions (required by our algorithm), we scraped Storyline information from IMDb URLs provided in the u.item file using a Python crawler. We randomly split 80% of data for training and 20% for testing.

## 4.2 Evaluation Metrics

**4.2.1 Accuracy Metric** We use Mean Absolute Error (MAE) to measure accuracy:

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|p_i - q_i|$$

where p_i are predicted ratings and q_i are actual ratings. Smaller MAE indicates better recommendation quality.

**4.2.2 Diversity Metrics** While high accuracy is the standard for most recommendation systems, long-term user retention requires diversity. We evaluate two diversity aspects:

**Aggregate Diversity:** Measured using information entropy:

$$H = -\sum_{i=1}^{n} p(i)\log p(i)$$

where p(i) is item i’ s popularity (its occurrence frequency across recommendation lists divided by total occurrences). Higher entropy indicates better long-tail item discovery.

**Individual Diversity:** Computed as the average pairwise dissimilarity between items in each user’ s recommendation list:

$$D(R_u) = \frac{1}{|R_u|(|R_u|-1)}\sum_{i,j\in R_u, i\neq j}(1 - sim(i,j))$$

$$\bar{D} = \frac{1}{|U|}\sum_{u\in U} D(R_u)$$

### 4.3 Experimental Results

**4.3.1 Cluster Number Selection** K-means clustering quality depends heavily on K. We use average silhouette coefficient to determine the optimal number of clusters. With 1,682 movies, we test K [2,50]. Due to K-means' randomness, we repeat each K value 30 times and compute the average silhouette coefficient AVG(s(o)). shows partial results, with K=13 yielding the maximum average silhouette coefficient of 0.6912, indicating optimal clustering quality. We use K=13 in subsequent experiments.

**4.3.2 Distribution of Diversity Tendency** In ARIFDP, balances accuracy and diversity—larger means greater influence from the item fatigue function. With K=13, we compute the number of movie categories each user rated across 30 clustering runs and calculate each user' s average . shows the distribution: most users' values concentrate in [0.2,0.3), indicating that most users have some diversity requirements that traditional accuracy-based algorithms cannot satisfy, validating our approach.

**4.3.3 Comparison Experiments** We compare ARIFDP against: (a) Traditional MF, which computes preferences via matrix factorization of user browsing behavior; and (b) CTR, which combines collaborative filtering with probabilistic topic modeling for interpretable latent structures and increased diversity.

With rating range 1-5, we set =5. We evaluate accuracy, individual diversity, and aggregate diversity while varying recommendation list length from 10 to 100. Results are shown in [Figure 3: see original paper]-[Figure 5: see original paper].

**Accuracy:** [Figure 3: see original paper] shows ARIFDP' s accuracy is slightly lower than CTR due to the item fatigue function enhancing diversity. However, ARIFDP outperforms traditional MF when more items are recommended because it incorporates both rating and topic information, while MF uses only ratings. ARIFDP' s accuracy remains acceptable.

**Diversity:** [Figure 4: see original paper] and [Figure 5: see original paper] demonstrate that ARIFDP achieves superior performance in both aggregate and individual diversity. By increasing weights for long-tail items, ARIFDP recommends more varied items globally (improving long-term system performance) and provides richer lists to individual users (broadening horizons and increasing satisfaction). Thus, ARIFDP effectively recommends item sets matching users' diversity preferences.

## 5 Conclusion

To recommend item sets aligned with users' diversity preferences, we propose the ARIFDP algorithm. Through K-means clustering, we analyze user diversity preferences to compute diversity tendency degree. By aggregating matrix factorization with item fatigue functions and adjusting fatigue function weights using diversity tendency degree, we increase recommendation probabilities for long-tail items. Experiments on real datasets show ARIFDP produces accurate and diverse recommendation lists.

Future work will explore associating diversity tendency degree with user attributes (gender, age, occupation, education) and modeling its dynamic nature. Designing models that effectively combine behavioral and attribute information to measure diversity tendency degree represents our next research direction.

## References

[1] Cheng Peizhe, Wang Shuaiqiang, Ma Jun, et al. Learning to recommend accurate and diverse items [C]// Proc of the 26th International Conference on World Wide Web. International World Wide Web Conferences Steering Committee, 2017: 183-192.

[2] Kawai K, Kitagawa H. Collaborative filtering with implicit feedbacks by discounting positive feedbacks [C]// Proc of the 2nd IEEE International Conference on Multimedia Big Data. 2016: 41-48.

[3] Frolov E, Oseledets I. Fifty shades of ratings: how to benefit from a negative feedback in top-N recommendations tasks [C]// Proc of the 10th ACM Conference on Recommender Systems. New York: ACM Press, 2016: [4] Jawaheer G, Weller P, Kostkova P. Modeling user preferences in recommender systems: a classification framework for explicit and implicit user feedback [J]. ACM Trans on Interactive Intelligent Systems, 2014, 4 (2): 8.

[5] Liu Yezheng, Wang Jinkun, Jiang Yuanchun, et al. Utilize item correlation to improve aggregate diversity for recommender systems [C]// Proc of IEEE International Conference on Data Science in Cyberspace. 2016:

[6] Koochi M R, Hussin A R C, Dahlan H M. Improving recommendation diversity using tensor decomposition and clustering approaches [C]// Proc the 4th World Congress on Information and Communication Technologies. 2014: 240-245.

[7] Zhou Tao, Kuscsik Z, Liu JianGuo, et al. Solving the apparent diversity-accuracy dilemma of recommender systems [J]. Proceedings of the National Academy of Sciences, 2010, 107 (10): 4511-4515.

[8] Wang Chong, Blei D M. Collaborative topic modeling for recommending scientific articles [C]// Proc of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining. New York: ACM Press, 2011: 448-456.

[9] Niemann K, Wolpers M. A new collaborative filtering approach for increasing the aggregate diversity of recommender systems [C]// Proc of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. New York: ACM Press, 2013: 955-963.

[10] Blei D M, Ng A Y, Jordan M I. Latent Dirichlet allocation [J]. Journal of Machine Learning Research, 2003, 3 (1): 993-1020.

[11] Steinberger J. Update summarization based on novel topic distribution [C]// Proc of ACM Symposium on Document Engineering. New York: ACM Press, 2009: 205-213.

[12] Jain R, Koronios A. Innovation in the cluster validating techniques [J]. Fuzzy Optimization & Decision Making, 2008, 7 (3): 233-241.

[13] Park Y J, Tuzhilin A. The long tail of recommender systems and how to leverage it [C]// Proc of ACM Conference on Recommender Systems. New York: ACM Press, 2008: 11-18.

[14] Yin Guisheng, Zhang Yanan, Dong Hongbin, et al. A long tail distribution constrained recommendation method [J]. Journal of Computer Research and Development, 2013, 50 (9): 1814-1824.

[15] Sar Shalom O, Koenigstein N, Paquet U, et al. Beyond collaborative filtering: the list recommendation problem [C]// Proc of International Conference on World Wide Web. International World Wide Web Conferences Steering Committee, 2016: 63-72.

[16] Zhang QianSheng, Jiang ShengYi. A note on information entropy measures for vague sets and its applications [J]. Information Sciences, 2008, 178 (21): 4184-4191.

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