



BRILL

ZEITGESCHICHTE

ISSN / eISSN: 0256-5250 / 2569-5304

Publisher: BRILL DEUTSCHLAND GMBH,
VIENNA UNIV PRESS, ROBERT BOSCH BREITE 6,
GOTTINGEN, GERMANY, D-37079

Volume 25-Issue 1-(2025)

OPEN  ACCESS

Research on Multimodal Recommendation Mechanism of Hybrid Collaborative

Filtering Algorithm in Mobile E-commerce



Dianyi Wu¹; Siwei Long¹; Yan Zhou¹, Mingxin Sun^{2*}

sunmingxin@btbu.edu.cn *

¹Advertising and Branding School, Communication University of China, No.1

Dingfuzhuang East Street, Beijing, 100024, China

²School of Languages and Communication, Beijing Technology and Business

University, Beijing, 100872, China

Fund: Innovation Centre for Digital Business and Capital Development of Beijing Technology
and Business University (SZSK202242)

Abstract: In the development of the artificial intelligence era, the e-commerce platform uses network information technology to provide an effective environment for online trading and negotiation for enterprises or individuals, which not only realizes primary online shopping, but also effectively structures a secure and easily scalable business framework system on the network, promoting the development of e-commerce in China. With the decline in the requirements of e-commerce platform merchant stationing and the simplification of the online review process of goods, the issue of risk related to content security has become one of the important contents of the review of collaborative filtering recommendation algorithms. In this paper, after understanding the multimodal recommendation model framework, according to the collaborative filtering recommendation algorithm oriented to e-commerce content security risk control, select the e-commerce platform's user ratings and content data to carry out experimental analysis, so as to clarify the application effect of the algorithm model.

Keywords: multimodal recommendation model; hybrid collaborative filtering algorithm; e-commerce; content security risk

1. Multimodal recommendation model analysis

According to the information characteristics in mobile e-commerce, create an end-to-end network for multimodal recommendation, the actual framework design, as shown in Figure 1 below:

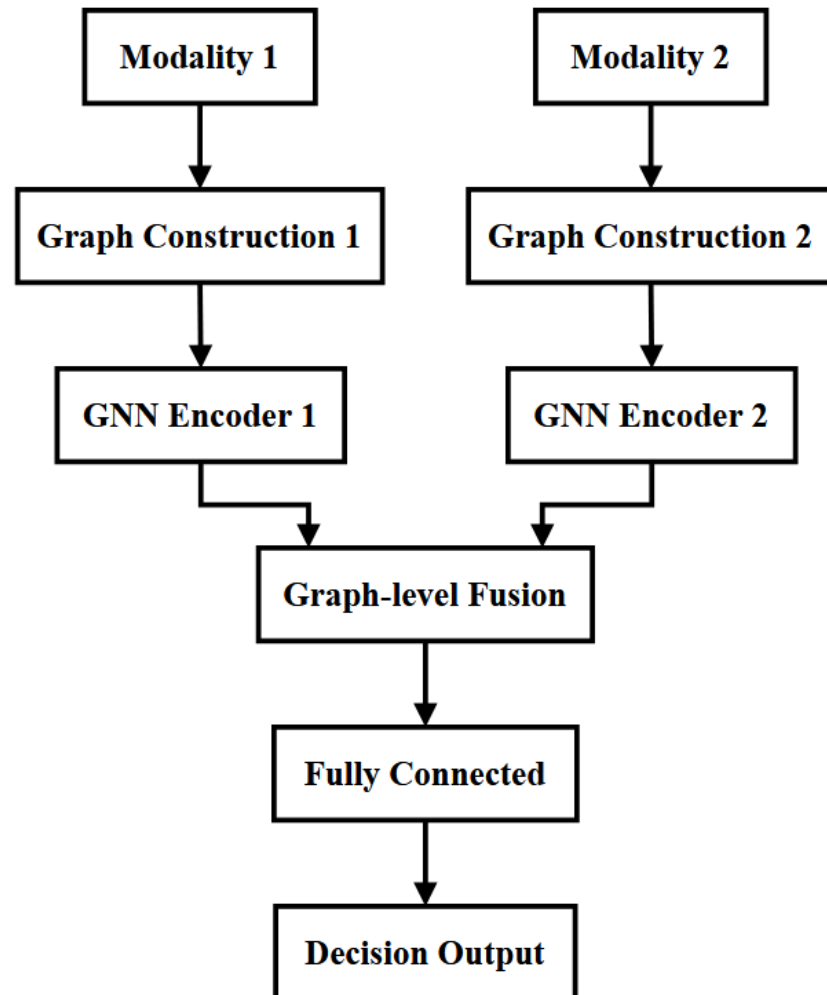


Figure 1 Flow chart of multimodal fusion module

Combined with the analysis of the above figure, it can be seen that the main task of the module is to build the interaction between the modalities, and to transform the higher-order matrix of the unimodal sparse representation into the lower-order matrix of the compact representation, and the whole process is divided into three steps: firstly, after adaptive dimensionality reduction processing of the sub-network, the use of the neural network to enhance the degree of nonlinearity of the model, and to reduce the number of generated high-dimensional feature dimensions; secondly, the use of the

three parallel branches to the three modalities interrelationships; finally, the three branches are fused. In the first stage, audio, text, and image are designed as three parallel branches respectively, and two features on a single branch are cascaded, thus constituting a cascaded feature vector, as shown below:

$$f_i = \text{Concat}(M_a, M_b)$$

In the above formula, f_i represents the cascade features of the i th branch, and different corner markers of M represent different modal features respectively. In order to have good interaction with each other, the P -order outer product of the cascade feature vectors is applied in the form of tensor product, which forms a high-dimensional tensor after a long period of time, thus realizing the modal complementarity, as shown below:

$$F_i = \underbrace{f_i \otimes f_i \otimes \dots \otimes f_i}_{P\text{-order}}$$

In the above formula, F represents the tensor after the P -order outer product, and \otimes represents the tensor direct product. On this basis, after a pooling layer, more compact fusion features can be obtained, at which time the single-branch fusion features are represented by the following formula:

$$\begin{aligned} & \sum_{r=1}^R a_r^h \prod_{p=1}^P w_a^{h(p)} \prod_{p=1}^P f_{i(p)} \\ & \&= \sum_{r=1}^R a_r^h \prod_{p=1}^P w_a^{h(p)} f_{i(p)} \end{aligned}$$

In the above formula, the dimension of W^H shows an exponential growth trend, at this time, the low rank network is used for dimensionality reduction, and the tensor decomposition based on rank-RCP is adopted, the specific formula is as follows:

$$\left\{ \{a_r^h, w_a^h\}_{r=1}^R \right\}_{h=1}^H$$

In the above formula, represents the parameter values during the fusion period, and the application of tensor decomposition can significantly reduce the parameter

dimensions during the whole connection period.

In the second stage, the actual steps are basically the same as the above, and only the first cascade step is performed on the operation results of the 3 branches as follows:

$$\hat{f}_i = \text{Concat}(M_x, M_y, M_z)$$

In the above equations, M_x , M_y , M_z refer to the modal features of the 3 branches during the fusion period, the collaborative filtering recommendation algorithm based on stacked noise reduction self-encoder shown in Fig. 2 below is an example, which is similar to the traditional collaborative filtering recommendation algorithm centered on matrix decomposition, and both will decompose the rating matrix R into the user potential feature vectors u and the item potential feature vectors v , in which the user's m predicted rating of item n is given by the formula:

$$\hat{r}_{mn} = u_m^T v_n$$

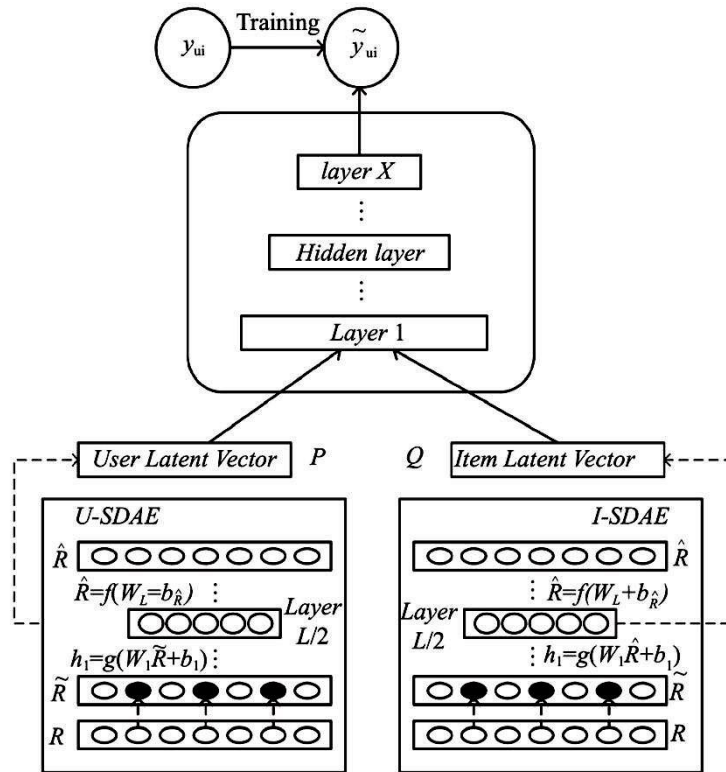


Fig. 2 Collaborative filtering recommendation algorithm based on stacked noise reduction self-encoder

In this model design, the overall selection of five levels, the first two levels belong to the encoder, the corresponding mapping function is as follows:

$$H_i = \text{dropout}(\text{ReLU}(W_i H_{i-1} + b_i))$$

In the above equation, i represents the layer encoding of the module, and H_{i-1} of the first layer adds the fusion feature of noise ; the last two layers belong to the decoder, and the corresponding mapping function is as follows:

$$Y_l = \text{ReLU}(W_l H_{l-1} + b_l)$$

In the running state of the model, the fusion semantic expression formulas of item fusion features are obtained from the output of layer 2 as follows:

$$v' = \text{dropout}(\text{ReLU}(W_2 H_1 + b_2))$$

In this paper, the research model uses the joint loss function, in order to avoid overfitting the model, it will add the normalization factor on the loss function, and each layer of the weight value and bias value regularized expression; in the matrix decomposition model, the regularized expression of the decomposed user potential vector is applied, as follows:

$$\begin{aligned} L_{\text{SDAE}} &= \sum_l \left(\frac{1}{2} W_l^2 + \frac{1}{2} b_l^2 \right) \\ L_{\text{MF}} &= \sum_m \|u_m\|^2 \\ L_1 &= \lambda_u L_{\text{MF}} + \lambda_w L_w \end{aligned}$$

The item feature in question contains the feature formula of the scoring matrix as follows:

$$L_2 = \lambda_v \sum_n \|v_n - v'\|^2$$

The optimization process for the manufacturing effect of the model encoder is aimed at reducing the error of the input fusion features and reconstructed features as follows:

$$L_3 = \lambda_n \sum_n \|Y_L\|$$

In the optimization of the rating matrix decomposition process, the squared error between the actual and predicted ratings is applied as follows:

$$L_4 = \lambda_w \sum_m \sum_n (r_{mn} - \hat{r}_{mn})^2$$

The final loss represents the sum between the regularization factor and the objective loss function as follows:

$$L = L_1 + L_2 + L_3 + L_4$$

2. collaborative filtering recommendation algorithm for e-commerce content security risk control

2.1 E-commerce content security risk

The traditional technical way of relying on search engines to obtain information has been unable to meet the needs of the development of mobile e-commerce in the current society, and with the continuous increase in the variety and number of platform products, it is increasingly difficult for users to use the browser or search to find products that meet their preferences, which prompted the emergence of the algorithmic recommendation model. Through practical investigation and research, it is found that the existing mainstream e-commerce platform's content security risk commodities refer to those commodities that violate the relevant national laws and regulations or platform operation norms, do not conform to the mainstream value orientation, and pose a threat to the sovereignty and interests of the state, the prosperity and stability of the social economy, the benign development of the e-commerce platform, and the legitimate rights and interests of the consumers, etc., for example, the commodities display contains one or more pictures with content security risks. The words used in the title of the goods contain illegal descriptions or undesirable information, and so on. Content security risk commodities are generally classified into seven categories: first,

pornographic and vulgar commodities and services; second, commodities and their derivatives that cause others to lose their ability to resist or endanger others' personal health; third, commodities such as military and police uniforms and related accessories, military and police equipment and products, or military and police special symbols and characters; and fourth, commodities affecting the unity of the country and social stability. Fourth, commodities that affect national unity and social stability, mainly including commodities and services in the three subcategories of national sovereignty, RMB, leaders and martyrs; fifth, commodities and services in the three subcategories of drugs and drug-making tools and food, drug-making processes and illegal descriptions; sixth, commodities or services in the three subcategories of ethnicity, religion and children; and seventh, commodities and services in the three subcategories of personal safety and privacy, illegal use and unsuitable for trading.

2.2 Collaborative Filtering Recommendation Algorithm

In order to reduce the content security risk during the application of collaborative filtering recommendation algorithm, this paper constructs a multimodal feature library for model training based on the above commodity identification, which is mainly based on deep learning technology and multimodal fusion technology to propose a collaborative filtering recommendation algorithm that meets the content security risk management needs of mobile e-commerce platform. This algorithm will identify products according to different feature libraries, which is mainly divided into four parts: product identification based on VGGNet16, product identification based on HTCBOV, multimodal model fusion and collaborative filtering recommendation. Among them, the collaborative filtering recommendation module will first clarify and delete the content safety risk commodities, commodity recommendation, user clustering and other steps, regard the corresponding coordinate of the largest element of the vector value $rsk_{\max} > 0.9$ as the content safety risk category of the current commodity, and delete the corresponding commodity columns in the scoring matrix M , and thus obtain the scoring matrix M' . In user clustering, the rating matrix is calculated and analyzed using the contour coefficient method to analyze the optimal value k for dividing clusters, and the

K-means algorithm is used to divide users into k clusters according to the rating distance. From the practical application point of view, the clustering effect of division clustering in high-dimensional space is significantly better than other algorithms. In the missing rating prediction, the reasonable use of Pearson correlation coefficient calculation to analyze the target user u 's nearest domain N_u , and the comprehensive consideration of the correlation coefficient and historical ratings, and the use of rating prediction formulas to predict the possible ratings of the target user u 's for the unsold goods, can guarantee the effectiveness and safety of the overall algorithm. The formula of the Pearson correlation coefficient is as follows:

$$sim(u,v) = \frac{\sum_{i=1}^n (r_{u,i} - \bar{r}_u)(r_{v,i} - \bar{r}_v)}{\sqrt{\sum_{i=1}^n (r_{u,i} - \bar{r}_u)^2} \sqrt{\sum_{i=1}^n (r_{v,i} - \bar{r}_v)^2}}$$

In the above formula, $r_{u,i}$ represents the rating of user u on the item, $\bar{r}_{(v)}$ represents the rating of user v on the item, which is the average value of all historical ratings of the two users, and n is the overall number of items in the rating matrix. The actual rating prediction formula is as follows:

$$r_{u,i} = \bar{r}_u + \sigma_u \frac{\sum_{v \in N_u} sim(u,v) \times r_{v,i}}{\sum_{v \in N_u} |sim(u,v)|}$$

In the above formula, σ represents the standard deviation of all historical ratings of user u , N_u represents the user's recent domain, and $sim(u, v)$ represents the correlation coefficient between user u and other users v in the same cluster.

3. Experimental Study

3.1 Experimental Design

This experiment utilizes the windows 10 operating system, configures the CPU as Core i7-5820K CPU@3.30GHz with 32 GB of internal storage, the GPU as GTX TITAN X with 12 GB of video memory, writes the code using Python 3.9, and the experimental framework is the Tensor Flow 2.6.0 as the core deep learning framework,

while using Cornac as a multimodal recommender system benchmark comparison framework parameters are set Table 1:

Table 1 Experimental parameter design results

Parameter Name	Parameter Value
Batch size	128
Learning rate	0.001
Dropout rate	0.1
Hidden dimensions	[64, 32, 128]
Out dimension	200
Rank	4
1	1
Ay	1
dw	0.1
An	1

According to the number of data sets in hand, on the basis of following the website protocol, the Scrapy framework is used to collect user ratings and content data of tens of thousands of products on an e-commerce platform, and generate three sets of experimental data sets as shown in Table 2 below, and then according to the time of user ratings, the first 70% of the user ratings are regarded as the training set, and the remaining 30% of the ratings are regarded as the test set.

Table 2 Analysis results of three sets of experimental data sets

Dat aset ID	Rat ings (r)	Rat ings Proport ion (r/R)	U sers (m)	Us ers Propor tion (m/M)	It ems (n)	Ite ms Propor tion (n/N)	Spa rsity (r/(m×n))
1	57, 209	29. 5 %	86 2	1. 1 %	1, 612	86. 5 %	95.8 8 %
2	62, 671	26. 8 %	74 0	0. 8 %	1, 009	39. 9 %	91.6 1 %
3	93, 078	29. 9 %	1, 076	0. 9 %	2, 069	72. 2 %	95.8 2 %

In the experimental process, Precision and Recall are used to evaluate the recommendation effect of the algorithm, At the same time, according to the algorithm application research needs, the content security risk evaluation index Violation is used, which represents the proportion of content security risk commodities in the commodity recommendation list, the larger the value represents the higher security risk, the specific formula:

$$Violation = \frac{1}{|U|} \sum_{u \in U} \frac{|X(u)|}{|R(u)|}$$

In the above formula, U represents the set of users, R (u) represents the commodity recommendation list, and X (u) represents the list of content safety risk commodities in the commodity recommendation list R (u). It is important to note that all the evaluation index values military calculated and obtained according to the

commodity recommendation list.

3.2 Result Analysis

When analyzing the model recommendation effect, the user-centered collaborative filtering recommendation algorithm (UCFR), the K-means-centered collaborative filtering recommendation algorithm (KMCFR), the linear regression-centered collaborative filtering recommendation algorithm (LRCFR), the review sentiment mining-centered collaborative filtering recommendation algorithm, and the above CSCFR algorithm are selected for comparative study, and the experimental results shown in Table 3 below are finally obtained. the experimental results shown in Table 3 below.

Table 3 Recognition performance results of CSCFR algorithm

Data sets	Model	Precision 1st (%)	Precision 2nd (%)	Precision 3rd (%)	Precision Avg (%)	Precision Std (%)	Recall 1st (%)	Recall 2nd (%)	Recall 3rd (%)	Recall Avg (%)	Recall Std (%)
Test #1	VGGNet16	70.25	72.73	71.35	71.44	10.01	60.3	1.43	9.28	0.45	0.89
	HTCBOW	60.93	64.88	64.20	63.34	10.72	8.85	0.70	5.05	8.20	0.35
	MF	73.84	74.03	74.31	74.06	10.19	1.05	0.85	0.12	0.67	0.39
Test	V	7	7	7	7	1					

est #2	GGN et16	2.72	0.23	3.06	2.21	.46	9.1 0	9.8 1	0.4 2	9.7 8	.54
	H TCB OW	6 2.52	6 3.81	6 0.22	6 2.18	1 .48	6.2 4	0.2 5	7.4 0	7.9 6	.68
	M MF	7 4.55	7 2.98	7 4.19	7 3.91	0 .67	1.0 4	0.0 4	0.2 2	0.0 3	.44
est #3	V GGN et16	7 3.33	7 2.39	7 0.47	7 2.06	1 .19	1.0 7	0.5 6	9.9 3	0.5 2	.47
	H TCB OW	6 3.20	6 0.40	6 2.03	—	1 .19	5.8 3	6.9 6			.53
	M MF	7 2.04	7 2.57	7 3.73	7 2.78	0 .71	1.4 5	0.5 1	0.6 5	0.8 7	.41

Combined with the analysis in the above table, it can be seen that the experimental results of the image recognition model, text recognition model and modal fusion model in its proof algorithm on the corresponding test set, in which the multimodal fusion model has a very high precision and recall rate in the content security risk recognition problem, and the cross-validation results of the three are more stable. It should be noted that not every multimodal fusion model in every experiment outperforms the single-modal model, which is most likely related to the fact that the late fusion assumes that the information between different modalities is completely independent, and the correlation between different modalities is lost. Therefore, future

studies could try pre-fusion or mid-fusion, the former having more complete information and the latter improving experimental flexibility.

In the same test set, for example, test set 1, the CSCFR algorithm has a significantly lower Violation indicator value than the other baseline algorithms, the lowest risk of algorithm content security, and the highest stability of the algorithm results. In this process, there is a positive proportional relationship between the baseline algorithm performance and Violation indicator values. In other words, for the better performing recommendation algorithms represented by CSCFR, the larger the actual Violation metric value is, the higher the content security risk is, whereas for the worse performing algorithms, the smaller the Violation metric value is, the lower the content security risk is. Similar conclusions can be drawn for test collections 2 and 3. In addition, there is no clear correspondence between the Violation metric value of the CSCFR algorithm and the sparsity of the data set in different test sets, while there is a significant inverse proportionality between the CSCFR metric value of the baseline algorithm and the sparsity of the data set; in other words, the higher the sparsity of the data set, the lower the value of CSCFR metric of the algorithm, and the lower the content security risk. The lower the content security risk.

From the calculation results of the precision and recall index values, the CSCFR algorithm obtains the overall optimal precision and optimal recall, and the standard deviation of the results of the three repetitions of the test is significantly lower than that of other algorithms, which proves that the CSCFR algorithm has a high degree of stability. The precision and recall of various algorithms selected for the experimental study have a significant inverse proportional relationship with the sparsity of the data set in different test sets. In other words, the precision and recall of the algorithms become smaller and smaller with the increasing sparsity of the data set, while the CSCFR algorithm studied in this paper is less affected by the data sparsity problem.

Conclusion

In summary, in the operation of mobile e-commerce platform, in order to

effectively control the security risk of e-commerce content, this paper studies the use of deep learning and multimodal late fusion technology, integrates the relevant issues into the optimization process of the user collaborative filtering recommendation algorithm, and constructs a hybrid collaborative filtering algorithm model with multimodal features as the core, and its practical application performance is significantly better than that of traditional algorithms, and plays an important role in the construction of e-commerce platforms in the new era. It plays an important role in the construction of e-commerce platforms in the new era, effectively reduces the risk of algorithmic content security, and provides a new direction for the future research of technical algorithms.

References

1. **Shibei Meng, Rui Zheng, Liang Chang, et al.** Multimodal and diversity recommendation algorithm based on fully connected tensor network[J]. *Computer System Applications*, 2023, 32(2):63-74.
2. **Jialei Xu.** Application of intelligent recommender system in new media platform[J]. *Satellite TV and Broadband Multimedia*, 2024(24):49-51.
3. **Z. Cao.** Research on online learning resources recommendation method based on deep learning[J]. *Information Recording Materials*, 2024, 25(11):62-64.
4. **J. Xu, T. Fu.** A collaborative filtering recommendation method based on feature clustering and time series analysis[J]. *Industrial Control Computer*, 2024, 37(5):74-76. doi:10.3969/j.issn.1001-182X.2024.05.030.
5. **Ting-Yu Lin.** Research on Collaborative Filtering Algorithm Based on RFM Modeling--Taking E-commerce Platform as an Example[J]. *Industrial Innovation Research*, 2025(1):61-64.
6. **Renfeng Deng.** Research on collaborative filtering recommendation algorithm based on RFMRQ model--Taking clothing e-commerce platform as

an example[J]. *Textile Report*, 2023, 42(12):18-21.

7. **X. D. Zhou.** Development of agricultural products e-commerce platform based on collaborative filtering recommendation algorithm[J]. *Computer Knowledge and Technology*, 2025, 21(2):57-59.

8. **Xuejing Xu, Chenwei Lin.** Collaborative filtering e-commerce personalized recommendation algorithm based on K-means clustering[J]. *Journal of Ezhou University*, 2023, 30(6):102-104.

9. **Q. L. Ge.** Modeling and analysis of user behavioral characteristics of e-commerce website based on improved k-means cluster analysis algorithm[J]. *Automation and Instrumentation*, 2024(12):262-266.

10. **Jianfeng Yuan, Jia Liu.** Design of recommendation algorithm for e-commerce platform incorporating item evaluation keywords[J]. *Computer Knowledge and Technology*, 2024, 20(28):48-51.

11. **Fang Wang, Lilan Gong.** Construction of e-commerce user repeat purchase behavior prediction model based on collaborative filtering technology[J]. *Journal of Jiamusi University (Natural Science Edition)*, 2024, 42(8):173-176.

12. **Yunhe Ding, Hong Chai.** An intelligent recommendation method for vegetable product information on e-commerce platform based on collaborative filtering[J]. *E-commerce Review*, 2025, 14(1):1579-1586. doi:10.12677/ecl.2025.141195.

13. **J. B. Li, Z. S. Qian, X. Y. Cai, et al.** Recommendation algorithm based on product attribute-context under cross-border e-commerce[J]. *Journal of Systems Engineering*, 2024, 39(3):333-343,468.

14. **H. Chen.** Intelligent dynamic collaborative recommendation model for user-perceived points of interest on e-commerce platform[J]. *Mathematical Practice and Understanding*, 2024, 54(3):29-38.

15. **Chunguang Zhang, Xiaoshan Chen, Sha Sha.** Research on user labeling application based on hybrid recommendation of agricultural products features[J]. *Agricultural Outlook*, 2024, 20(1):117-125.